Are User-contributed Reviews Community Property? Exploring the Beliefs and Practices of Reviewers

Frank M. Shipman
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
shipman@cs.tamu.edu

Catherine C. Marshall
Microsoft Research, Silicon Valley
1065 La Avenida
Mountain View, CA 94043
cathymar@microsoft.com

ABSTRACT

User-contributed reviews form the cornerstone of many Web communities and online services. People rely on reviews as a source of information about products, services, creative efforts, and the reputation of other buyers and sellers. Although specific rights about the ownership and control of these reviews are spelled out in licensing agreements and by copyright law, most reviewers' actions are guided instead by evolving social norms. In this paper, we report on 203 responses to a questionnaire offered to reliable US-based Mechanical Turk workers who have written different types of online reviews. The questionnaire uses a series of realistic scenarios and specific questions about recent practice to probe participants about how online reviews may be reused, archived, re-purposed, deleted, and otherwise manipulated. We use these collective attitudes and behaviors to arrive at a picture of current social norms and examine user-contributed reviews as a counterpart to other types of online content, including photos and tweets.

Author Keywords

Product reviews, information rights, reuse, removal.

ACM Classification Keywords

H4.3 Information Systems: Communications Applications.

INTRODUCTION

User-contributed reviews are an essential mechanism for describing and organizing media, products, and services on the Web. For example, the weight of listener reviews on iTunes might put a podcast on the service's front page, thus distinguishing it from an undifferentiated sea of new audio content; a series of positive Yelp reviews might spell success for a neighborhood restaurant; and book and movie reviews represent an important literary genre that helps readers select and interpret creative works.

Thus within this broad category, reviews' specific function can vary. At one end of the spectrum, reviews of books, media, and music may be serious arbiters of taste or scholarly assessments of the creative efforts of others. At the other end of the spectrum, product and service reviews may have significant commercial impact; for example, CNet reviews can be as important as advertising in promoting a new technology product within a competitive marketplace. In essence, user-contributed reviews are an important commercial and intellectual component of what Lessig refers to as the Read-Write Web [17].

In the study we describe in this paper, we ask participants to consider a genre of reviews that is likely to have long-term value: those that help us identify the media (both on- and offline) worthy of our attention. Although these reviews may still play a commercial role (after all, people often buy books, movies, and music), they may also form the basis of large descriptive resources such as IMDB, and they may be important in determining how we parcel out our scarce attention as we navigate media-intensive sites.

In particular, because media reviews do seem to have intrinsic value, we are interested in peoples' perceptions of who owns and controls them, who can reuse them, and peoples' concomitant experiences with reviews they have written themselves.

Of course licensing terms and copyright law have important bearing on what can and cannot be done with this content. But people who contribute, save, and reuse content are often oblivious to (or, at best, only semi-knowledgeable of) law, policy, and the provisions of content licenses, and instead rely on their own nuanced sense of ethics and alignment with social norms [3, 11]. It is these culturally-specific attitudes and practices that we are after.

By eliciting these attitudes and practices, we hope to answer three research questions: (1) Which actions do people feel they and other members of the community can take with socially-contributed content like reviews? (2) Which factors influence peoples' intuitions about the ownership and control of reviews? (3) How does review ownership and control compare with that of other Web-resident digital belongings (e.g. tweets [19], photos [20], or possessions in general [23, 10])? The answers to these questions should help us understand not only how practices are likely to diverge from applicable law and policy but, when appropriate, how to design future systems and policy to better align with emerging social norms.

This paper first summarizes related work on online reviews and digital possessions. We then go on to describe the study and its participants, the steps we took to ensure the integrity of the results, and the dataset we collected. Finally, we present findings and discuss their implications.

RELATED AND PRIOR WORK

Research on online review systems and practices falls into three categories: mechanisms for identifying and promoting high-quality reviews and reviewers, review understanding and summarization, and methods for identifying review spam. This research is relevant to our work because it offers a mature characterization of phenomena associated with online reviewing. There is also an emerging body of related work on the ownership of digital belongings; our main contributions are in this area.

To promote trust, review systems must encourage members of the community (which could be as broad as the general public) to create real and valuable reviews. Gilbert and Karahalios explore the motivations of reviewers who echo earlier reviews, including the desire of amateur reviewers to simply make their own passionate views heard by the community [9]. Approaches to encouraging viable reviewing practices include allowing readers to evaluate the reviews of others. Such responses are then aggregated across reviews to create a reviewer reputation. Some sites allow users to create trust networks, e.g. identify reviewers they trust [29]. Opinion Space [8] differentiates between two rating dimensions—whether users agree with the comment and how they assess its quality—instead of conflating them in a single measure. When user evaluation of reviews is not possible, automatic approaches to identifying high-quality reviews have been explored [30].

Because products can receive large numbers of reviews, research has focused on applying text understanding to create overviews and summaries of the reviews [24, 4]. This approach includes work on identifying product features and analyzing the review's sentiment towards these features [6]. The results of such analyses have been used to design alternative shopping interfaces [27] and to estimate the value of particular product features [1].

A final area of research germane to online reviews attempts to identify review spam. Jindal and Liu identified three types of review spam: untruthful opinions, reviews based on brand only, and non-reviews (e.g. advertisements) [14]. Their resulting approach to identifying review spam makes use of duplicate and near duplicate detection combined with supervised learning. Instead of focusing on spam content, Lim and colleagues identify spammers by analyzing their rating behaviors [18]. To look more deeply at the strategies and practices of reviewers who reuse review content, David and Pinch identify different motivations for review reuse, including reuse to promote sales of an item or to bolster an opinion's strength and reuse to strengthen a reviewer's reputation or brand [5].

Our work examines peoples' attitudes and behavior about the ownership and control of reviews. In particular, we explore specific rights people feel they should (or shouldn't) have to save, reuse, or delete reviews. While duplicate detection assumes forms of reuse that are commonly considered inappropriate, our interest is in reuse in more ambiguous contexts. For example, can movie review authors post the same review to both IMDB and Amazon? Can a social media user copy an interesting or funny review and share it with her Facebook friends? In other words, instead of assuming such behavior is clearly permitted or prohibited by license terms or fair use, we look instead to the complexities of current behavior as a way of informing technology design or shaping content rights policies. In so doing, we extend the work reported in [19] and [20].

In [19] and [20], we examine the ownership and control of Twitter posts and personal photos, respectively. In these studies, we found that relationship to the material, along with the circumstances of reuse, strongly guided what participants felt they could do with user-contributed content. We also learned that participants feel that tweets and personal photos may always be saved and stored regardless of who created them and that removal is the most controversial action.

This research also serves as a counterpoint to Odom et al.'s digital possessions work [23]; while they investigate the materiality of digital content, and peoples' strategies for managing this content as a possession, we examine the social norms that have emerged to control its re-use as the content becomes increasingly owned by the community as well as the individual.

STUDY METHOD

The study used data collected from a three-part 48-question questionnaire that elicited participants' attitudes about the ownership and control of reviews and their experiences publishing and reusing reviews as well as other sorts of social media. We collected the data over a two-week period starting June 24, 2011. During this period, we received 216 responses. Participants spent an average of 14 minutes, 23 seconds completing the questionnaire.

Questionnaire structure

The first part of the questionnaire consisted of 12 demographic and background questions, including participants' online activities, social media experience, and what they had published on the Internet to-date.

The second portion of the questionnaire consisted of 29 questions about four realistic scenarios; the scenarios were based on actual reviews of a classic children's book, *Where the Wild Things Are*. We chose this book and several of its 500-plus reviews for three reasons: (1) The book is a well-known and enduring work, so reviews of it potentially have lasting value; (2) the reviews sparked actual conversations among reviewers that could be used in the scenarios; (3) our

participants were likely to be familiar with the book (either from their own childhoods or from reading the book to children or grandchildren), and to recognize it as a classic.

To investigate specific aspects of ownership, we used a technique borrowed from legal education: hypotheticals that systematically vary aspects of a situation's fact pattern [26]. First we presented a scenario that set up the basic situation; each scenario was then followed by a set of Likert-scale statements (using a 7-point scale) about hypothetical actions various people (including the author, specific reviewers, and others) could take with the reviews. By using detailed scenarios, followed by a series of 'what ifs' to elicit participants' attitudes, we reduced the opportunity for participants to envision key characteristics differently.

Finally, we asked 7 questions about participants' online reviewing experiences, including 3 open-ended questions: the types of reviews the participant generally writes; the last review the participant remembered writing; and the participant's overall views about reuse on the Internet.

Participant screening and data reliability
We administered the questionnaire on Mechanical Turk, recruiting English-speaking participants from the United States who performed reliably in past Human Intelligence Tasks (HITs) and who had experience writing online reviews; reliability was demonstrated by 95% or higher acceptance rate on past HITs. Reported experiences with similar questionnaires [19, 20], coupled with accounts of Mechanical Turk best practices [7, 13, 15], enabled us to take appropriate steps to ensure data quality. Responses to the open-ended questions suggested that participants not only took the questionnaire seriously, but also that they were genuinely engaged and thought carefully about their answers. The median length of responses to the open-ended questions about the last review written and their opinions about reuse of Internet content were 25 and 33 words respectively; generally, the answers were complete and surprisingly articulate. Some participants were moved to write longer answers (the longest was almost 300 words) to the reuse question. According to Ipeirotis, US-based Turkers accept work because it is interesting or entertaining instead of strictly for financial gain [12]; this ameliorates some concern we might have about participants' motivations or the veracity of their responses.

The recruiting limitations we placed on participants (and double-checked through demographic responses) were designed to minimize cultural interference with the results, and to ensure participants understood the questions with the same level of language comprehension. All participants who completed the questionnaire were paid at established rates for Mechanical Turk (50 cents for this questionnaire) whether or not we kept the data.

We further ensured data quality by applying a 2-point test to filter the completed questionnaires; if participants got two points, their questionnaires were disqualified and the data from them was discarded. Participants scored points for each of the following conditions:

- A wrong answer to any of three reading comprehension questions (the questions were from the material, and not simple attention checks);
- An unanswered question (or a nonsense response to an open-ended question); and
- Spending fewer than 8 minutes on the questionnaire.

After we had applied these conservative screening criteria, we discarded responses from 13 participants and were left with 203 responses for subsequent analysis.

Scenarios
The questionnaire's four related scenarios referred to Maurice Sendak's award-winning storybook *Where the Wild Things Are*. Originally published in 1963, this book is a staple of children's literature and has attracted over 500 Amazon customer reviews. We excerpted and modified actual reviews to use as a basis for our scenarios.

To ensure consistent interpretation of the statements in the questionnaire, at its outset we defined four user actions: (1) store, (2) share, (3) publish, and (4) remove.

The first scenario features a negative review purportedly written by a seven year-old girl. The review is charming and funny, but has elicited several critical comments questioning both its appropriateness for the Amazon website and its authenticity (e.g. is the author really 7 years old?). Three of these critical comments were used in the first scenario. The second scenario reveals that the review was written by the child's father, and explores who should be able to remove the review from Amazon. The third scenario introduces a positive review of Sendak's book written by a specialist in early childhood education who has a good reputation in Amazon's reviewing system. A fourth scenario parallels a premise used in the questionnaires reported in [19] and [20]: a public institution (the Library of Congress) creates an archive of user-contributed reviews to use as metadata for books and other library holdings. Because the final questions are focused on institutional archiving rather than on community ownership and reuse, the results were analyzed separately and reported in [21]. Table 1 presents the scenarios covered in this paper and the Likert-scale questions associated with them.

Participants
The participants represent a fairly diverse population who write different types of online reviews and have varying motivations for contributing this online content.

Of the 203 participants, 64 are current students and 139 are non-students; this matches the student/non-student mix of similar studies [19, 20]. Table 2 shows participants' age and gender distributions. These distributions are consistent with studies characterizing US Turkers: many participants were in their 20s and 30s and women are represented in greater numbers than men [12].

Scenario 1: The Amazon customer reviews for *Where the Wild Things Are* are generally positive, but there are some negative reviews too. The following is a "1 star" (the most negative rating) review. There is no link to verify the reviewer's identity; she has used her first name (Nicole) to sign the review.

A Kid's Review:

I am almost 7 and my teachre said we have to say why we like a lot of books or do not like a lot of books this summer on amazon and then print out them and give them to our new teacher next year So I am starting with this book.

My dad reelly likes this book because he said it was good when he was a kid. I dont like it. The pictures are boring and the story is not long. My dad reads this to me a lot and I like the books that are newer. New books have pictures that are pretty and the storys are funner and longer. This book has pictures that look old. I wish my dad would read this to himself and let me read something diferent.

- Nicole

Amazon not only supports customer reviews, but also comments and conversations about the reviews. A number of readers responded to Nicole's negative review. Some enjoyed it. Other readers commented on the review's appropriateness (since it claims that an elementary school teacher had assigned it as a class exercise). Three examples of these comments (by Matt, Ulyyf, and Quadradox) are shown below.

• Matt says:

This is the best review ever. I used to like this book, but not anymore. You have good insights and a strong conviction, keep up the good work.

• Ulyyf says:

Quoting from Nicole's review: "I am almost 7 and my teachre said we have to say why we like a lot of books or do not like a lot of books this summer on amazon and then print out them and give them to our new teacher next year So I am starting with this book."

Do teachers think this is a clever idea? DO NOT DO THIS. Reviews written for this purpose are usually VERY BAD. (This one is a bit of an exception - a surprise!) It skews the ratings of books, and is not helpful to people who are using this review system for its actual purpose - to find out if they should buy a book. If you want your students to review books, have them keep a book journal. If you want them to write reviews online, have them email it to you. Don't waste the time of people browsing Amazon for this.

• Quadradox says:

Ulyyf, You are reacting to a review that is over 6 years old!!!! And with 490 reviews, of which only 15 are 1-star -- I doubt it is skewing anything. There are certainly better ways to do this teacher's project, but just as you claim that the review should be saved for another place ... perhaps your own critique of it really fits better somewhere else than attached to the kid's review.

Q1.	The author (Maurice Sendak) should have the right to save the review to his hard drive.
Q2.	Ulyyf, who teaches an early childhood education class, should have the right to save the review and his comment s to his hard drive.
Q3.	Ulyyf should have the right to republish the review and comments on the web site for a class he teaches on early childhood education.
Q4.	Nicole should have the right to save her own review to her family's PC, but not the right to save the comments.
Q5.	The author (Maurice Sendak) should have the right to post Nicole's review on his Facebook fan page. "Not so much a fan" he writes before quoting the review in its entirety.
Q6.	A non-profit website that promotes reading among children and adults should have the right to copy Nicole's review verbatim to support their description of Where the Wild Things Are, and to seed a discussion about the book.
Q7.	Years later, the now-adult Nicole sees her review on Amazon and is embarrassed. Nicole should have the right to remove her review and the comments others have made about her review.
Q8.	Nicole should have the right to remove her review and the other peoples' comments about her review only if the commenters give her permission to remove their comments.
Q9.	Maurice Sendak's publicist doesn't think that it's appropriate for a seven year old to rate the book. Sendak's publicist should have the right to remove the review from the Amazon site.
Q10.	Maurice Sendak finds evidence that Nicole's review was actually written by an adult. Sendak should have the right to remove the fraudulent review from Amazon's website.

Scenario 2. Suppose Maurice Sendak is right, and Nicole did not write the review. Instead, Nicole's father wrote it. He exaggerated Nicole's point of view to make the review funnier, and omitted the fact that Nicole was comparing the book to the Captain Underpants series, which she thought were hilarious and more modern.

Q11.	Nicole, who is now 13, should have the right to remove the review.
Q12.	Nicole's father, who suffers qualms of conscience, should have the right to remove the review and its comments.
Q13.	An Amazon customer discovers that Nicole didn't review the book herself. This customer have should have the right to remove the fraudulent review.
Q14.	Amazon discovers the misrepresentation. Amazon should have the right to remove the fictitious review.

Scenario 3. Most of the reviews of Where the Wild Things Are are not negative. On the contrary, they are overwhelmingly positive. The following 5-star review was written by an academic researcher from a university; like Ulyyf, the researcher specializes in early childhood education. Not only is the reviewer qualified to assess the book's quality; she also has a good reputation rating in Amazon's reviewing system.

Classic story, classy art from a classy illustrator:

Not much can be added to so many peoples' praise of Maurice Sendak. I grew up with his work; I read his books to my children; and I am reading them now to my grandchildren. When children ask to read a book over and over, you know that the book has found a place in their imaginations. Where the Wild Things Are is one such book.

|	
Max is a typical little boy, who dons his wolf outfit and becomes a 'wolf' in his active imagination. When he is sent to bed without dinner, he uses his imagination to turn his room into a wild forest inhabited by humongous monsters. He asserts his powers over the monsters, and becomes their rightful king, for he is surely a monster himself! Yet, his mother loves him in spite of his monstrous behavior—eventually he comes home to find that mother has relented and brought in his dinner. *This winter, get this classic book and sit down in a big chair with a child so you can pore over this book together. Our children have a right to be read to by someone they love and someone who loves them. Turn the television and computers off, and use a rainy evening to spend time with your own monsters. Maybe someday they will become the next Maurice Sendak!*	
Q15.	The author's publicist likes the 5-star review. She should have the right to save the 5-star review to her hard drive.
Q16.	The author's publicist should have the right to republish the 5-star review on her public KidLit blog, a blog in which she talks about new and classic books for kids, focusing on authors she represents.
Q17.	Sendak's publisher should have the right to extract a blurb, "When children ask to read a book over and over again, you know that book has found a place in their imaginations..." to use on the book's dust jacket.
Q18.	Barnes and Noble should have the right to reuse the review on their site, since they are selling the same book, and the review's author was not paid for the Amazon review.
Q19.	Sendak is republishing and selling his books on a website he manages himself. Sendak should have the right to copy the reviews of his own book, and post them on his website.

Table 1. Overview of the scenarios and the nineteen rights statements

| year born	before 1950	1950-1959	1960-1969	1970-1979	1980-1989	after 1990	total
female	1	5	16	29	56	12	119
male	1	0	9	18	45	11	84
total	2	5	25	47	101	19	203

Table 2. Participant age and gender.

We also asked participants about their reviewing experience (in part to confirm our screening criterion); online reviewing experience was limited to reviews that include textual content (not just ratings or thumbs-up recommendations). Figure 1 shows that 42 participants are avid reviewers who report having published more than 50 reviews, but it is more common to have published fewer.
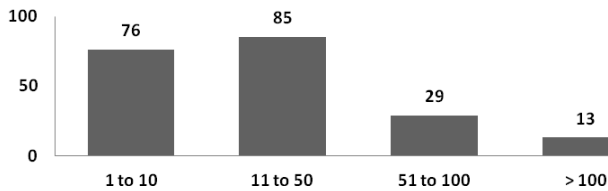
Figure 1. Reported number of on-line reviews written.

Responses to open-ended questions revealed that some reviews were written for pay, but these cases were exceptions, and did not reflect the dominant motivations. We discuss these motivations in the next section.
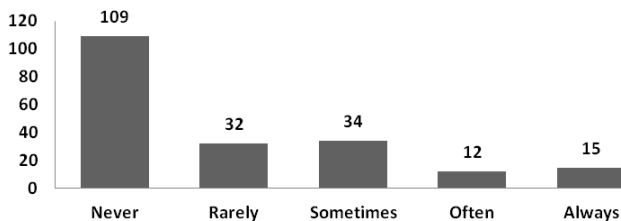
Figure 2. Participants' reported frequency of saving copies of reviews on their own computers

Figure 2 shows that approximately half of the reviewers occasionally save local copies of their reviews. Only 35/203

participants (17%) reported posting the same review on multiple on-line locations.[1] This number will become more important as we begin probing participants about their attitudes towards reuse. In a comparable questionnaire about photo reuse, pragmatic concerns that arose from personal reuse experience often led participants to a more nuanced view of the circumstances under which reuse is permissible [20].

Limitations

The results we discuss rely on responses gathered via Mechanical Turk HITs. Turkers tend to be young and technically savvy. They are also better educated than the population at large (as is evidenced by the fact that 62% self-report as having a college degree, and 92% have attended at least some college). Thus, some results may not generalize to the US Internet-using population at large.

RESULTS

Reviews of products, media content, or services are an essential part of the modern Web. Reviews not only allow people to contribute to significant online resources (e.g. IMDB or Amazon); they also strengthen the social fiber of communities (e.g. buyer and seller reputations on etsy or eBay), and provide a forum for people to share their good and bad experiences with Web content (e.g. a podcast in iTunes); physical products (e.g. a vacuum cleaner); and items that might be both (e.g. books and movies).

The last review that participants reported writing fell into three general categories: (1) reviews of media and creative works; (2) reviews of products; and (3) reviews of services (these three categories are ours; participants described what they reviewed in their own terms). Reviews of media and

---

[1] Similarly, Gilbert and Karahalios report that 10-15% of reviews on Amazon resemble previous ones, although (in keeping with the objectives of their study) this figure does not differentiate between self-plagiarism and reuse of others' reviews [10].

creative works dominated participants' efforts. Figure 3 shows the relative frequencies of the three types of reviews.
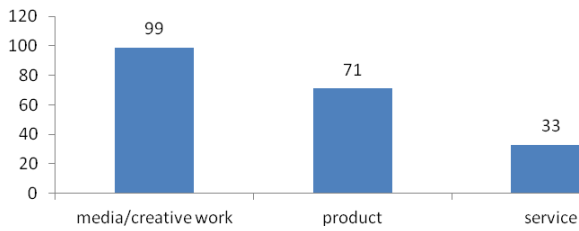

Figure 3. Subject of last review written.

What motivates participants to contribute reviews (often to a website that does not guarantee them any rights to the review)? We might suppose that since the study participants were recruited through Mechanical Turk that they are predisposed to review for pay. In practice, reviewing-for-pay is only a minor part of the story. Participants mainly reviewed for other reasons, most of which stem from a desire to share an experience or information. As Gilbert and Karahalios have noted, reviewers are often motivated by "*an almost visceral reaction to a product*" [9]. Five common reasons participants wrote reviews include:

- **A desire to share a positive experience.** For example, P002 said that he had reviewed "*Thurston Moore's album* Demolished Thoughts*. I wrote the review because I loved the album very much.*"

- **A desire to warn people about a negative experience.** For example, P026 wrote, "*I wrote a review about a pre workout supplement. I wrote it because it felt like people were being conned, and I wanted to share the truth.*"

- **A desire to contribute to community knowledge.** For example, P101 wrote, "*The last review I wrote was of an exceptional jazz recording that was being discussed on a music forum that I frequent. I wrote it because others hadn't listened to the recording yet, and it was long. Also, there weren't any other reviews, so I wrote mine to give potential listeners an idea of the content of the recording.*" P176 explained that she posted a review of a baby's car seat to the Target.com website, "*… since that is where we do most of our shopping and I seem to get the most useful reviews when researching that site.*"

- **A perceived need to bolster or detract from an individual's online reputation.** For example, P104 wrote, "*[My last review] was actually just yesterday on ebay—I needed to review the selling transaction between myself and a customer.*" Similarly, P178 wrote that she had reviewed "*a seller on eBay. I received extremely prompt service from them, and the article was sold for a 'used' price but was as good as new, and I was very impressed. The seller was new to eBay, so I went out of my way to post a review and help her*

get a positive rating. I talked about the quality of the product, and the dependability of the seller."

- **Other (non-monetary) incentives are offered in exchange for reviews.** Often participants explained that they would have written the review anyway or that the reviews were negative. P181 responded that s/he had written a review "*For a restaurant on seamless.com; I wrote it because I feel strongly about the restaurant and I also got an incentive from seamless (a 10% discount on my next order).*"

Many participants reported that their reviews focused on creative efforts. These reviews provoked the most emphatic responses. For example, P111 wrote:

"*I wrote a review of a song for my music blog. For this particular post, I was spotlighting outstanding B-Sides. Naturally, I was overwhelmingly positive about the song. I went into great detail about the significance of its commentary, as well as the lyrical and musical structure of the song. It was of a decent length. I tried to make it as informative as possible without getting bogged down in details and while keeping it from being too long.*"

Often reviewers spoke about under-appreciated artists, or independent work that might not see the light of day without the efforts of the work's fans.

Less frequently, creative efforts evoked negative responses. For example, P160 said of the novel *The Accidental Tourist*: "*I hated that book with the fiery passion of a thousand suns.*"

Participants were most apt to report that the last review they wrote was positive; more than twice as many reviews were positive than negative (89 versus 43). A smaller number were non-committal (26) or balanced (46), written with other purposes in mind, such as educating the community or comparing multiple items. Later we explore whether the reviews' intended slants influences reuse attitudes.

Building on the results of our past studies, we hypothesize that people will generally take a liberal stance toward the reuse of positive reviews, especially those motivated by a desire to inform or recommend, rather than those with essentially commercial or self-interested aims.

Reuse in principle

At the conclusion of the questionnaire, we asked participants a general question about reuse on the Internet: *What do you think about the reuse of content on the Internet? When is it okay? When is it a bad idea?* Although the question was general, participants were likely to have reviews in mind, stemming from the scenarios and from the review-writing we had just asked them to describe. It was up to the participant to interpret the question and introduce notions of copyright, permission, plagiarism, intent, and other rights-related ideas. These responses were open-coded in an effort to capture their primary perspectives and to record distinctions that define secondary themes [28].

From the data, it was evident that participants approached reuse from four different angles, similar to those articulated in [20]. They were less apt to draw on specific reuse situations than they were in the earlier study (since they were less apt to have reused reviews), although they sometimes had specific types of reviews in mind.

We describe these four angles in an order that reflects their frequency in the data: (1) law- or license-driven considerations; (2) content properties; (3) situational factors stemming from how the review is reused; and (4) governing technology that can enforce specific restrictions.

Law or license. While fewer than 10% of the participants (18) referred explicitly to copyright law, the majority relied on it implicitly, invoking notions such as permission, plagiarism, public domain, attribution, free speech, and commercial vs. non-commercial use. Although the legal concept of fair use only arose specifically in 6 responses, it was frequently the underlying principle that participants relied on. It is important to note that fair use is not a well-understood principle outside of the legal profession, even in professions such as journalism where it may be central to the work [2]. In all, 149 responses (> 73%) were essentially appealing to the law; to the inability to enforce the law; or to licensing provisions. Other mixed responses used legal principles as a secondary perspective for guiding reuse.

Content. Most of the 22 participants who referred to content when they were addressing reuse were concerned about the literal veracity of the information, and potential imbalances stemming from reuse of positive or negative reviews on different websites. Some participants felt that there is intrinsic social harm in propagating incorrect information. Other participants constructed a similar argument to say that correct information should, in fact, be made ubiquitous to increase peoples' chances of finding it and that the propagation of correct information was in the public interest. As a counterpoint, other participants felt that verbatim reuse created deceptive redundancy. In other words, a review that is repeated multiple times may be given unfair weight and may interfere with a review's role in evaluating media, products, or services.

Situational factors. Only 11 participants appealed to the reuse situation. For example, was the reuse malicious, misrepresenting the author's original intent? By contrast, in our photo study [20], the most common type of argument relied on situational factors (how the photo was being reused). This isn't surprising, considering that only 35/203 participants (about 17%) claimed they had reused reviews, while 202/242 photo study participants (over 83%) said they had reused photos. Thus the review reuse arguments were apt to be entirely hypothetical, while the photo reuse arguments were apt to be based in real experience.

Technology. As was the case in our earlier photo study, a few participants (4/203) said they would rely on technology to guide reuse. On one hand, for example, review text might

be copy-protected, thus limiting the ability to reuse the text. That few participants suggest technological solutions may indicate that they regard this type of solution as ineffectual or that they simply dislike it.

Finally, several participants admitted that their views on reuse are evolving and that the questionnaire itself caused them to rethink implicitly-held attitudes. For example, P142 wrote,

> *"This is complex and difficult which is why I went 'neutral' on a lot of those questions. I think credit always needs to be given, that's for sure. I don't think it should be used commercially without permission. There are just too many variables, including what sites want to use it and why."*

Clearly ownership is a provocative topic, and it is never fully addressed until participants find themselves in a situation in which reuse is a plausible option.

Measured attitudes

We now turn our attention to the hypotheticals grounded in the scenarios reported in Table 1. Participants graded the hypotheticals using a Likert scale from 1 (strongly disagree) to 7 (strongly agree). All differences reported in this section were statistically significant (p<.03 for closest pair) except for the equivalence classes of Q1/Q2/Q15, Q3/Q6/Q16 and Q11/Q12 (Wilcoxon Signed Rank test).

Storing. From previous work, we expect saving reviews—writing them to local storage—to be uncontroversial unless there are limitations placed on what can be stored. The overarching principle is, anything that one encounters online can be saved; dissent only emerges when an obstacle is introduced (for example, one part can be saved and another part can't). We can refer to this as the *Digital Media Hoarding Principle*.
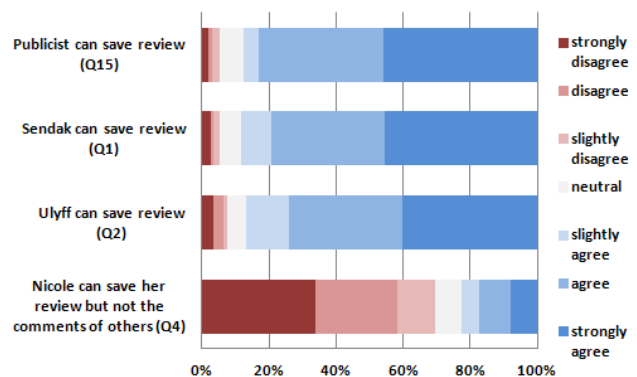
Figure 5. Saving an encountered review is an uncontroversial action unless limits are imposed.

Figure 5 illustrates how this principle manifests itself in the scenarios; refer to Table 1 to see the exact hypotheticals. The first three variations—the review is saved by the book's author, by a commenter, and by the author's publicist—are statistically indistinguishable: everyone with

any possible interest in the review can save a copy, even if they did not write the review themselves. The outlying statement in this case poses a situation in which Nicole (the review's author, a 7 year-old child according to the scenario's premise) cannot save other peoples' comments on her review, presumably because she does not own them in the same way that she owns her own review (apart from restrictions placed on the content by the site's license, which as we reported in the last section is generally ignored). This outlier trends negative.

Sharing and publishing. If the general question about reuse at the end of the questionnaire is a predictor, then participants will have a generally positive view of sharing someone else's review, particularly if the intent is not overtly commercial or promotional. Indeed, Figure 6 shows that this is is the case.
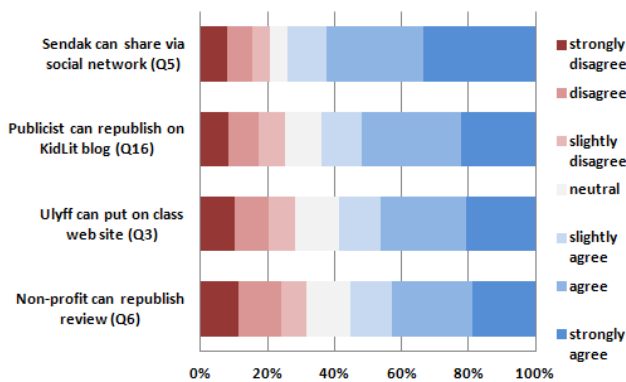

Figure 6. Few constraints are imposed on sharing reviews on a social network, especially if sharing is for public good.

Most of the responses—variations on public good—follow the same arc. For example, one of them is about a teacher sharing the review in his early childhood education course; another involves posting the review on a nonprofit organization's blog. The biggest outlier—which has met with an increased level of approval—poses that Sendak himself posts the little girl's negative review on his Facebook fan site. This positive reaction may be influenced by Sendak's self-deprecation; participants tend to approve of well-motivated reuse.
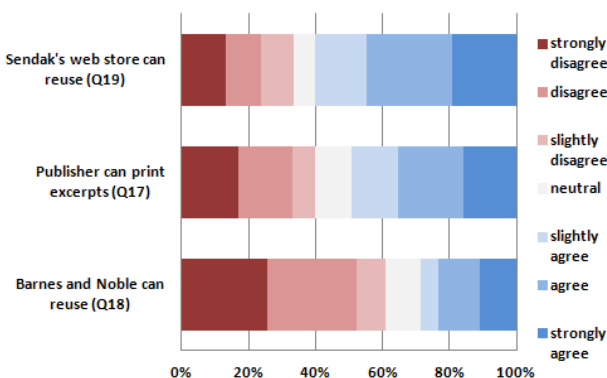

Figure 7. Views on commercial reuse vary with reuse details.

Our model of factors that influence the acceptance of reuse suggests that commercial reuse will be considerably more problematic than reuse that supports the public good; this aligns well with the trend in recent legal decisions [26]. Figure 7 explores this territory. The view that Barnes and Noble can reuse a review from the Amazon website is unsurprisingly negative. Sendak's publisher's ability to reuse excerpts is controversial, with a more bimodal distribution of responses. Sendak's ability to reuse the review on his own web store skews more positive. Although this is commercial reuse, it reflects an ownership conundrum we observed with photos: a photo's subject has almost the same perceived right to a photo as the photographer. Thus we might expect participants to think that authors have additional rights to reviews of their books.
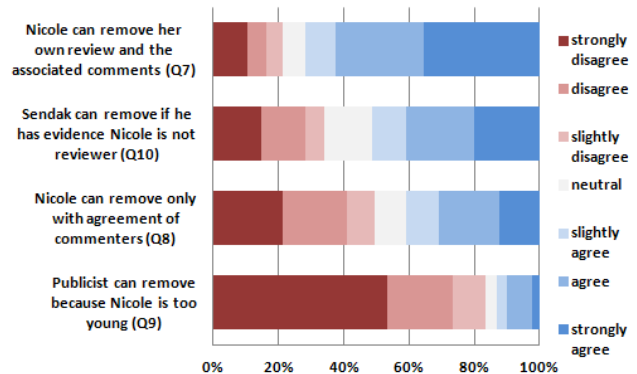

Figure 8. Attitudes toward the removal of a review depend on relation to review and reason for removal.

Removing. In related studies, removal was deemed the most controversial action, particularly when self-interest came into play [19, 20]. The same pattern holds true with reviews. Figure 8 shows responses to the first 4 removal rights statements; each varies the circumstances under which Nicole's review is removed from Amazon's website.

The only removal scenario that elicited a generally positive response was one in which Nicole, now grown, can remove the review she wrote when she was a child; an effort to place an obstacle in her path (i.e. requiring her to ask permission of the people who had commented on her review) was deemed controversial; responses were bimodal. The hypothetical that met with the most negative reaction was the one in which Sendak's publicist removed the child's review (remember: the review is negative, and the publicist is acting in what she considers to be Sendak's best interest). Thus participants tend to be swayed by the particular features of the removal scenario: who is removing the material, and for what reason?

What happens if we introduce a notion of fraud? In the open-ended questions, we saw a number of participants react to the perceived correctness of a review: should a review be published if something is proveably wrong with it? In this case, we pose that the review allegedly written by a child was, in fact, written by her father. Will fraudulence

trump authority? Will a customer who detects the fraud be able to remove the bogus review? How about the site owner? How about the review's author, who now realizes the error of his ways? How about the child who is implicated by the fraudulent review? Figure 9 shows the responses to this slippery-slope relaxation of authority. Indeed, the most positive reaction is to situation in which the site owner is removing the review. There is virtually no difference between the authority of the purported author of the review and her father (the actual author) to remove the fraudulent content, but a customer who detects the scheme is not granted the authority to remove the bad content.
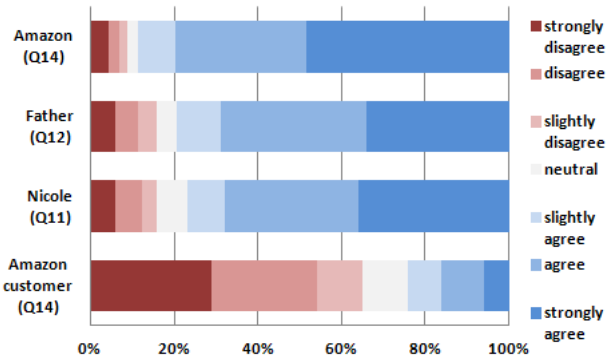

Figure 9. Who has the authority to remove a fake review?

DISCUSSION AND CONCLUSIONS

The results from the questionnaires paint a nuanced portrait of participants' views on the ownership of user-contributed reviews. They address our motivating research questions about socially acceptable user actions, the factors that influence their acceptability, and how these factors compare with those governing the reuse of other types of social media. Although the answers are intertwined, we tease them apart in this discussion.

In *Code version 2.0*, Lessig identifies four constraints that regulate online behavior: architecture, law, market forces, and social norms [16]. For our purposes, architecture is a shorthand for the rules and formalisms embedded in the software that mediates access to social media; market forces are realized in the terms and conditions that social media services impose on their users; and law is copyright—most notably fair use—as it is represented by legal cases and legislation. As we see in the overall responses to our questionnaire, social norms seem to have an outsized effect on participants' perceptions of what they (and others) can do with user contributed content: they have little understanding of the relevant legal guidelines; software-based governance is easy to ignore or thwart; and much reuse continues oblivious to market forces.

It is apparent from participant responses that keeping online reviews is readily tolerated (although it is not common practice); the uncomplicated storage scenarios were met with almost universal approval. On the other hand, republication is highly situated and depends strongly on

perceived motivation: if it is for the public good, or the republisher has a strong personal claim (it is about him/her), republication is well tolerated; if the republisher has commercial intent, it is less so. Removal is even more controversial, and depends on the strength of the remover's social distance from the review or the site where it is stored. Mitigating circumstances such as fraud complicate the situation further, with the site's owner granted the greatest authority to remove fraudulent content, followed closely by those with authorial claim.

To address our second question, we compare reviews as an emergent genre of digital content with amateur everyday photos and non-news tweets, two genres we have explored in the past. Properties that appear to influence ownership and reuse may be based on the content or on the nature of reuse and include:

- broad utility (including commercial value);
- relationship of reuser to the content (e.g. author, subject, external non-commercial, commercial);
- personal nature of content (i.e. how much it reveals about the author or the subject);
- specific reuse experience (i.e. whether the participant has ever reused this type of content);
- content reliability;
- content permanence (i.e. a tweet may be ephemeral and a review may be considered published);
- content sentiment (e.g. some participants are sensitive to what they refer to as 'meanness' or negativity);
- creative work;
- exogenous features (e.g. author or subject age); and
- potential for fraud (media manipulation and misattribution).

To-date, technology designers, publishers, and content owners have imposed a variety of technical and social constraints to control reuse. Technological mechanisms have included transclusion and micropayment (in alignment with Ted Nelson's original design for hypertext [22]); copy protection and DRM languages; and UI affordances that suggest the content is safe for sharing (e.g. "Share via" buttons for different social media sites). Social mechanisms have included labeling schemes such as Creative Commons; trusting peoples' familiarity with copyright and fair use provisions; and relying on social conventions (e.g. permission, attribution, voluntary donations). None have been wholly successful nor satisfying for authors or reusers.

Differences between the survey responses and an earlier photo reuse survey, coupled with reactions based on the exigencies of real reuse as opposed to reuse in the abstract, demonstrate that participants' sense of media rights are highly contingent on the actual reuse situation. This situatedness may interfere with labeling systems like Creative Commons, since people may relinquish (or hang onto) their rights in a non-productive way. For example, an etsy.com artist may allow non-commercial use of her work,

since she envisions reuse that promotes her store. In so doing, she may fail to consider a site like Regretsy.com (which pokes fun at etsy.com sellers). Although buyers may flock to the artist's store as a result, the artist may feel indignant about the nature of the reuse; to complicate matters further, Regretsy donates its proceeds to charity, so satire is mixed with social good. Hence the content labeler is actually faced with complex trade-offs rather than a single overarching constraint (non-commercial use is okay). Whether reuse restrictions are implemented through technology, policy, or a combination of the two, managing reuse (and other rights) relies crucially on authors' ability to envision a variety of plausible reuse scenarios.

Designers of review-driven web services and other content repository systems thus need to be aware that their users' understanding of what is acceptable is contingent on multiple features. Our results show that reuse and removal hinge on the user's relationship to the reviewed content and the review, the action being taken on the review, the nature of the review, and the action's context. Designs that affirm a single simple solution to all situations (e.g. no reposting is allowed) are likely to irritate users and lead them to identify work-arounds or to go elsewhere.

REFERENCES
1. Archak, N., Ghose, A., and Ipeirotis, P. Show me the money!: deriving the pricing power of product features by mining consumer reviews. *Proc. KDD '07*. 56-65.

2. Aufderheide, P, Jaszi, P., Bieze, K. and Boyles, J.L. Copyright, Free Speech, and the Public's Right to Know: How Journalists Think About Fair Use (July 30, 2012). http://ssrn.com/abstract=2119933

3. Boyle, J. *The Public Domain: Enclosing the Commons of the Mind*, Yale University Press, New Haven, 2008.

4. Danescu-Niculescu-Mizil, C. Kossinets, G., Kleinberg, J. and Lee. L. How opinions are received by online communities: a case study on amazon.com helpfulness votes. *Proc. WWW 2009*.

5. David, S. and Pinch, T. J.. Six degrees of reputation: The use and abuse of online review and recommendation systems. *First Monday*, July 2006.

6. de Albornoz, J.C., Plaza, L., Gervás, P., and Díaz, A. A joint model of feature mining and sentiment analysis for product review rating. *Proc. ECIR'11*, Springer-Verlag, 55-66.

7. Downs, J., Holbrook, M., Sheng, S., and Cranor, L. Are your participants gaming the system?: Screening Mechanical Turk workers. *Proc. CHI'10*. 2399-2402.

8. Faridani, S., Bitton, E., Ryokai, K., and Goldberg, K. Opinion space: a scalable tool for browsing online comments. *Proc. CHI '10*. 1175-1184.

9. Gilbert, E. and Karahalios, K. Understanding Deja Reviewers. *Proc. of CSCW '10*, 225-228.

10. Greengard, S. Digitally Possessed. *Communications of the ACM,* 55 (5), 2012, 14-16.

11. Hill, B., Monroy-Hernandez, A., and Olson, K. Responses to Remixing on a Social Media Website. *Proc. ICWSM 2010,* AAAI, 74-81.

12. Ipeirotis, P. Analyzing the Amazon Mechanical Turk Marketplace. *ACM XRDS 17*, 2, Winter 2010.

13. Jakobsson, M. Experimenting on Mechanical Turk: 5 How Tos. *ITWorld*, September 3, 2009.

14. Jindal, N. and Liu, B. Opinion spam and analysis. *Proc. WSDM '08*. ACM, 219-230

15. Kittur, A., Chi, E., Suh, B. Crowdsourcing user studies with Mechanical Turk. *Proc. CHI'08*. ACM, 453-456.

16. Lessig, L. *Code, Version 2.0*, Basic Books, 2006.

17. Lessig, L. *Remix: Making Art and Commerce Thrive in the Hybrid Economy*, Penguin, New York, 2008.

18. Lim, E.-P., Nguyen, V.-A., Jindal, N., Liu, B., and Lauw, H.W. Detecting product review spammers using rating behaviors. *Proc. CIKM '10*. ACM, 939-948.

19. Marshall, C.C., and Shipman, F.M. Social media ownership: Using Twitter as a window onto current attitudes and beliefs. *Proc. CHI'11*, ACM, 1081-1090.

20. Marshall, C.C., and Shipman, F.M. The ownership and reuse of visual media. *Proc. JCDL'11*. ACM, 157-166.

21. Marshall, C.C., and Shipman, F.M. On the institutional archiving of social media. *Proc. JCDL'12*. ACM, 1-10.

22. Nelson, T.H. The heart of connection: Hypermedia unified by transclusion. *Communications of the ACM*, 38 (8), 1995, 31-33.

23. Odom, W. Sellen, A., Harper, R., and Thereska, E. Lost in Translation: Understanding the Possession of Digital Things in the Cloud. *Proc. CHI'12*. 781-790.

24. Popescu, A.-M. and Etzioni, O. Extracting product features and opinions from reviews. *Proc. HLT '05*. Association for Computational Linguistics, 339-346.

25. Rissland, E. and Ashley, K. Hypotheticals as Heuristic Device. *Proc. Strategic Computing Natural Language*, 1986, 165-178.

26. Sag, M. Predicting Fair Use. Ohio State Law Journal, 73 (1), 2012, 47-91.

27. Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H., and Jin, C. Red Opal: product-feature scoring from reviews. *Proc. EC '07*. ACM, 182-191.

28. Strauss, A. and Corbin, J. *Basics of Qualitative Research*, Sage Publications, 1998.

29. Yeung, C.-M. A. and Iwata, T. Strength of social influence in trust networks in product review sites. *Proc. of WSDM '11*. ACM, 495-504.

30. Zhang, Z. and Varadarajan, B. Utility scoring of product reviews. *Proc. of CIKM '06*. ACM, 51-57.