

DOI:10.1145/2996181

User attitudes toward online intellectual property reveal how far social norms have strayed from legal notions of ownership.

BY CATHERINE C. MARSHALL AND FRANK M. SHIPMAN

Who Owns the Social Web?

USER-CONTRIBUTED CONTENT plays an increasingly important role in the Internet's evolution, overtaking professionally created and curated resources. Sophisticated recording technologies allow non-professionals to produce high-quality photos and videos. Improved editing and sharing applications facilitate other aspects of media creation, including larger-scale collaborative efforts. And social media venues give their users new opportunities to publish, curate, and recommend content. Every phase of the creative process—from recording to editing to publishing—has become more popular and interactive. At the same time, content ownership has become more complicated. Any distinct item may be associated with a virtual web of stakeholders.

USER-CONTRIBUTED CONTENT plays an increasingly important role in the Internet's evolution, overtaking professionally created and curated resources. Sophisticated recording technologies allow non-professionals to produce high-quality photos and videos. Improved editing and sharing applications facilitate other aspects of media creation, including larger-scale collaborative efforts. And social media venues give their users new opportunities to publish, curate, and recommend content. Every phase of the creative process—from recording to editing to publishing—has become more popular and interactive. At the same time, content ownership has become more complicated. Any distinct item may be associated with a virtual web of stakeholders.

» key insights

- Intellectual property law and social norms concerning content ownership are diverging in conspicuous ways; we find that legally contentious actions (such as downloading and saving content) may seem benign to most Internet users.
- Managing rights relies on content owners' ability to envision plausible reuse scenarios, including commercial reuse of their content as data, and predicting which are most likely.
- Everyday reuse of social media content is opportunistic, pragmatic, and highly contextual; users reason about the fairness of reusing other people's content but do not necessarily trust them to do the same.

contributed articles

DOI:10.1145/2900101

User attitudes toward online intellectual property reveal how far social norms have strayed from legal notions of ownership.

BY CATHERINE C. MARSHALL AND FRANK M. SHIPMAN

Who Owns Social Media?

an increasingly
ution, overtaking
sources.

low non-
otos and
ations
cluding
media
ish,

» key insights

- Intellectual property law and social norms concerning content ownership are diverging in conspicuous ways: we find that legally contentious actions (such as downloading and saving content) may seem benign to most Internet users.
- Managing rights relies on content owners' ability to envision plausible reuse scenarios, including commercial reuse of their content as data, and predicting which are most likely.
- Everyday reuse of social media content is opportunistic, pragmatic, and highly contextual; users reason about the fairness of reusing other people's content but do not necessarily trust them to agree.

A product review posted on Amazon might attract hundreds of comments that contribute substantively to the review's value and credibility. Videos on YouTube might respond to, excerpt, or satirize one another. Ongoing conversational threads on Twitter are held together by hashtags and @responses. Gamers use their avatars to interact with one another against the backdrop of a virtual universe and, in so doing, create new forms of data that build on the game's commercial content. Moreover, as individuals develop rich personal profiles, they publish new kinds of online representations of themselves.

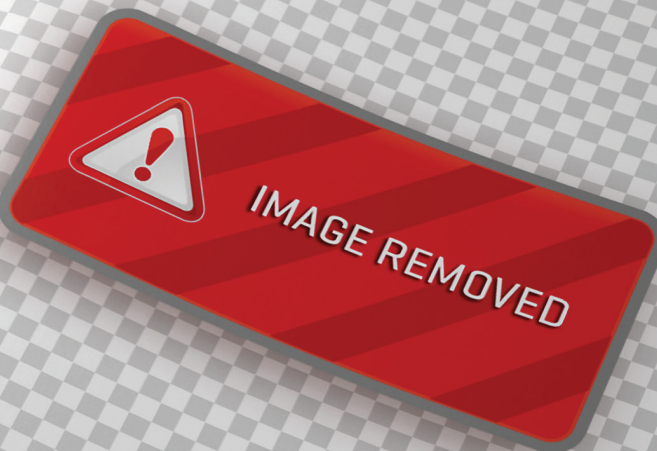
This complex non-professional digital-media landscape, along with new-found opportunities for copying, excerpting, and remixing professionally produced media, poses new challenges for managing intellectual property. Many social, legal, and technological forces shape our perceptions of who can do what with Internet content. As law professor and social activist Lawrence Lessig points out, in addition to legal notions of copyright, market forces and technologically enforced prohibitions constrain users' actions; additionally, emerging social norms make some online user behaviors seem acceptable to most people, while other behaviors are perceived as reprehensible.¹

In *Order Without Law*, property law scholar Robert C. Ellickson demon-

strated how people settle their disputes and regulate their behavior via these social norms; his analysis shows the importance of the norms and how they can be as effective as law.³ Legal scholar John Tehranian has highlighted how ordinary people (rather than legal scholars or jurists) now have a heightened awareness of the issues surrounding content ownership and, at the same time, the gap between

line user behaviors seem acceptable to most people, while other behaviors are perceived as reprehensible.⁶

In *Order Without Law*, property law scholar Robert C. Ellickson demonstrated how people settle their disputes and regulate their behavior via these social norms; his analysis shows the importance of the norms and how they can be as effective as law.³ Legal scholar John Tehranian has highlighted how ordinary people (rather than legal scholars or jurists) now have a heightened awareness of the issues surrounding content ownership and, at the same time, the gap between



the body of copyright legislation and the social norms that govern ownership behavior is growing.²² That is, although people are aware there are legal prescriptions for ownership and its reach, they are guided instead by social norms. By reflecting on his own online practices, Tehranian even showed how absurd copyright law has become relative to more sensible social norms.

In this article, we take a bottom-up view of content ownership and control, seeking to identify the norms and practices of everyday media users. Since 2010, we have conducted a series of surveys aiming to discover the social norms associated with content ownership and control and identify which media-specific user behaviors shape them. Once we arrive at a rich characterization of these norms, policy and technology can be designed accordingly and conflicts between practice and design can be anticipated and addressed when unavoidable.

Uncovering Social Norms

Having acknowledged the role of social norms in reflecting and guiding online behavior, it seems like we might want to know what these social norms are. How situational are they? How much do they vary across media types or services?

To examine them, we must identify people’s current practice (what they do), both as media consumers and as media creators, and we need to get them to articulate their aspirations. What do they think is fair? What should they do? What should others do? What behavior do they object to? It is important to distinguish practice from aspiration; in a pinch or when

no one is looking, people’s behavior is highly situated and unselfconscious. It is not so much that study-participants might want to mislead us, but more that they are not always aware of what they usually do.

When we began these studies, we thus sought to pin down both practice and aspiration among a broad set of people who spend a substantial amount of time online and use specific types of social media services. We designed and ran a set of eight studies over the next four years, each focusing on the ownership and control of a different media type and service: tweets,⁹ photos,¹⁰ reviews,²⁰ podcasts, recorded videoconferences and educational videos,¹² recordings from multiplayer online games,²¹ and Facebook content.¹³

We screened participants for familiarity with the content type or social media service in question, then checked responses according to documented criteria to ensure the study had been completed in good faith, a process detailed in articles covering the individual studies.^{9,10,12,13,20,21} We collected valid responses from a total of 1,738 participants. Many had attended college; approximately one-third were students at the time of the study. Most were between 18 and 40, although close to 20% were over 40 at the time of the study. Participants’ individual interests and online activities varied, and they used a range of applications and services in addition to the one under investigation. Participants were generally both content consumers and creators. Two-thirds of them in the six studies after 2012 had been on the Internet more than 10 years. The

accompanying table summarizes key participant demographics.

In the studies, we adapted a technique that has been used successfully in legal education⁸ and legal argument analysis:¹⁶ scenarios (or cases) plus hypotheticals. Hypotheticals explore the features of a heuristic or rule, with the aim of discovering the edges of how and when it applies. For example, we might want to know if it is okay for a user to download and store a stranger’s photo if the user is the subject of the photo. To explore these rights, consider a scenario in which Sophia, a 25-year-old woman, encounters a photo of herself on Instagram that was taken at a wedding reception, and she does not know the amateur photographer who was sitting at her table and took the picture. Sophia likes the candid shot and wants to download it and store it locally. Hypotheticals can then be used to explore Sophia’s rights to the photo by varying different features, say, the action she is taking with the photo, storing it rather than re-uploading and using it as her Facebook profile photo; her relationship to the photo’s subject—herself—rather than if the subject was, say, another wedding guest; the status of the photographer, a fellow wedding guest, not a professional photographer; and where Sophia is keeping the photo, saving it to her hard drive rather than storing it on a cloud-storage service. Scenarios thus situate the hypotheticals in a story with concrete details derived from real-life situations.

The scenarios and hypotheticals are both more engaging than an abstract version of the question (such as “Should you be allowed to down-

Key participant demographics. Survey topics are listed in the order the studies were conducted. Values in the “Number of participants” column are the number of responses retained after screening. Parenthetical values are the total number of responses collected.						
Study topic	Number of participants	Percent attended college	Percent current students	Percent female	Percent born after 1980	Percent with >10 years Internet experience
Tweets	173 (190)	88%	NA	61%	75%	49%
Photos	242 (250)	91%	34%	71%	66%	55%
Reviews	203 (216)	92%	32%	59%	61%	66%
Podcasts	225 (239)	90%	31%	44%	69%	67%
Videos	200 (228)	93%	24%	47%	60%	67%
Educational videos	209 (250)	94%	36%	50%	66%	67%
Multiuser games	241 (251)	80%	29%	35%	77%	66%
Facebook content	244 (250)	92%	25%	45%	72%	68%
Total	1,737 (1,874)	90%	NA	51%	68%	63%

load any picture of yourself and save it?”) and less apt to leave the details to chance (such as “Are we asking about a posed picture or a candid picture?” and “Is the photographer a professional using a camera or a fellow party guest using a smartphone?”). These details can make a difference in how the hypotheticals are interpreted and what response(s) they might elicit.

We captured participant reactions using a seven-point Likert scale, from strongly disagree to strongly agree. Each study presented from 16 to 28 hypotheticals associated with from two to four related scenarios. We discussed methodological details and aggregate participant demographics for the first six studies in an earlier paper.¹¹

The studies explored three types of common user actions: saving, sharing, and removing:

Saving. We defined saving as the act of intentionally downloading content from a social media site or service and storing it to a place under the user's control. Saving user-contributed media has minimal effect on others; it neither affects the digital original nor will most legal copyright holders ever know the content has been saved elsewhere;

Sharing. We defined sharing as reposting existing user-contributed media on another site or service, possibly without attribution, along with varying degrees of content transformation and varying user intent. In essence, sharing tests the social norms that circumscribe the limits of fair use. Our first three studies—covering tweets, photos, and reviews—distinguished between sharing by, say, posting content on one's Facebook account so it can be accessed by a limited social group and publishing openly as, say, a public blog post, but participants did not seem to notice this distinction themselves without considerable explanation. Our later studies—covering podcasts, recorded videoconferences, educational videos, recordings from multiplayer online games, and Facebook content—did not make this distinction; and

Removing. We defined removing as deleting or limiting access to user-contributed content. Removal is an action that tests the limits of media ownership and control, since it is not usually supported if the remover is not the explicit content owner. Because removal in

Is saving someone else's photo appreciably different from saving someone else's tweet?

practice is often provoked by a desire to curate one's digital footprint or reflect a changing notion of privacy, we designed the removal scenarios to elicit feelings of shared ownership, as in “You have posted something I feel I should be able to control, as with, say, a picture of me.

We returned to saving, sharing, and removing throughout the studies.

Highlights of Study Results

Our eight media-specific studies explored features that influence people's attitudes about the ownership and control of user-contributed content. Using a consistent set of actions—saving, sharing, and removing—supported comparing whether ownership rights are sensitive to expectations introduced by media type and nature of the actions users take upon them. That is, is saving someone else's photo appreciably different from saving someone else's tweet? Moreover, the scenarios helped us explore media-independent features, including:

Original context. This feature tests whether the content's original context influenced our study participants' perceptions of ownership rights; for example, do users have the same right to save a photo of a vintage picnic table they encounter on another user's public Flickr account as they do a similar photo that was used as an eBay product description?;

User's relationship to content. This feature tests some of the complexities of ownership. For example, if a person is the subject of a photo, as opposed to being the photographer, should this particular fact influence the person's right to save, reuse, or remove the photo from the service where it resides?;

Commercial concerns. This feature considers users' understanding of corporate ownership rights, as well as commercial use by individuals, apart from any terms and conditions spelled out by the service. For example, does the service owner have the right to save private communication that occurs within the service? And does it have the right to analyze the public communication it supports? Does it have the right to remove content it deems offensive?;

Genre-derived properties. Some content genres may have properties that raise specific expectations about associated rights. For example, media

considered ephemeral (such as an in-game chat session or other forms of communication) may influence perceptions of another user's right to save the content, regardless of whether that user was a participant in the conversation; and

Disaggregation. Disaggregation tests whether the rights to constituent parts of an item are different from the rights to the whole. This feature has allowed us to test whether, for example, the audio track of a recorded video inherits ownership rights from the video.

Here, we discuss the highlights of our findings, including social norms that emerged across studies and sometimes across actions. We also note media-specific norms and where norms break down.

Saving Social Media

To our participants, saving is the most benign, or least ethically contentious, action the scenarios explored. In the surveys, we define saving as an intentional act of downloading something—a photo, podcast, document, or video—to user-controlled storage to maintain a copy, rather than a side effect of performing some other action (such as viewing a webpage).

Users may save content (such as a tweet or a photo) because they fear its owner will delete it, because the site itself offers no guarantee of permanence (such as a story in a newsfeed may disappear and be difficult to re-find), or simply because they want to have a copy on hand. In the scenarios we spelled out in the surveys, saving was always motivated so participants would not imagine differing reasons for saving something; for example, guests interviewed on a podcast might want to save copies of the podcast for themselves.

The scenarios distinguished between saving for permanence and saving for reuse; the surveys considered reuse separately and are discussed in the next section. The scenarios also posited that the person was saving content without impediment; no tricks (such as screen captures of a Snapchat session) or special knowledge were necessary. That way, saving would not seem contrary to the media creator's expectations. In addition to testing the features outlined earlier, the hypotheticals checked two other aspects of saving—saving to cloud storage and explicitly imposed limits on saving.

Cloud storage. Cloud storage is often portrayed in the popular media and in

user interfaces as a seamless extension of local storage. Yet it is never fully under user control, and service-provider terms and conditions may apply. From a rights perspective, is saving downloaded content to local storage (such as on the person's hard drive) different from saving it to private cloud storage (such as in the person's Dropbox folder)?; and

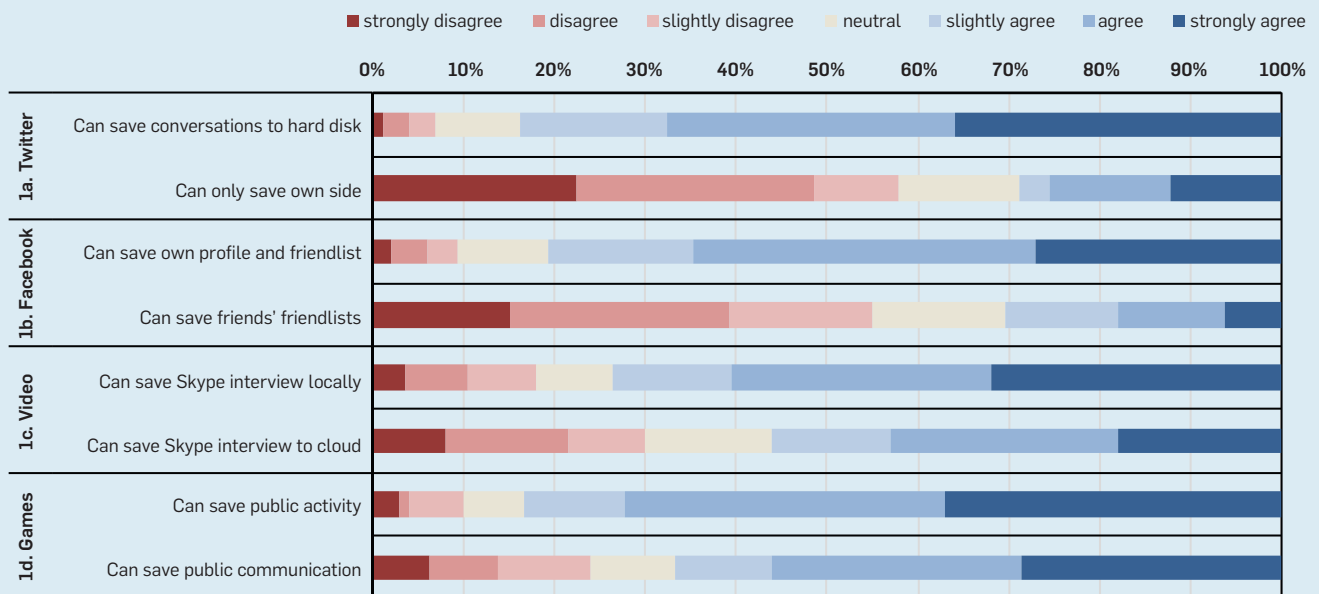
Limits. Responses to hypotheticals in the surveys suggest people expect to be able to download much of what they encounter online. This baseline may be tested by imposing artificial limits. For example, suppose people are permitted to save tweets they authored themselves but not the tweets other users wrote in response?

Our results have confirmed the baseline condition that participants usually felt individuals should be able to save anything they encounter on the Internet to local storage, regardless of whether the content is published on the open web or shared on a social media service, as long as the content is public.

This result was reaffirmed by our own hypotheticals that tested the idea of imposing limits on saving; these limits were based on a strong interpretation of ownership rights. Participants

Figure 1. Social norms for saving online content.

Pair 1a compares saving all parts of a Twitter conversation with a conventional ownership limit imposed on saving tweets, thus saving only one's own tweets. Pair 1b shows the effect of social distance on saving Facebook content. Pair 1c compares saving content locally with saving content to the cloud. And Pair 1d compares saving publicly visible in-game activity—on-screen avatar presence and action—with saving public in-game communication.



often disagreed with these imposed limits. In an extreme case—a hypothetical that limited saving tweets to saving only one's own tweets—58% (100/173) of the participants disagreed at least somewhat. Figure 1a contrasts saving all tweets in a Twitter conversation and saving just one's own tweets. That is, participants like the way content is controlled now; for example, if users are downloading content and saving it to local storage, they should not be limited to just the content they clearly own (such as photographs they have taken and posted, bon mots they have typed, or their own side of a conversation); instead, our study participants feel the norm is unfettered saving.

There are exceptions to this rule that also characterize norms associated with saving content. The strongest effect stems from the introduction of social networks. Participants respect explicit boundaries set by their social connections. While it seems perfectly acceptable to save any content encountered on the open Web, once one is inside Facebook, for example, different rules seem to apply. Our survey participants expressed a strong negative reaction to the hypothetical that one has the right to save the profile of a friend of a friend, even given reasonable motivation for doing so. Figure 1b contrasts one's right to save one's own profile and friends list (197/244, or 81%, agreed) with the right to save the equivalent content for one's friends (only 74/244, or 30%, agreed). This reaction is surprising, given the laissez-faire attitude about saving in general.

Two weaker effects also appeared in our survey results. First, saving to the cloud is viewed differently from saving to local storage. As a test, study participants judged two hypotheticals that differed only in where the downloaded content was stored. In the first, a recorded job interview—a Skype-based video—was stored to the user's local hard drive and in the second to a cloud storage service. In the first hypothetical, 18% disagreed with an individual's right to record the video and save it; in the second, the disagreement jumped to 30% (see Figure 1c.) Study participants may feel local storage is more private than cloud storage or are perhaps concerned that terms and conditions give one less control over the ultimate

disposition of the content. This effect may diminish as people become accustomed to cloud storage but may also grow if privacy breaches continue to be reported in the news.

A second effect stems from users' expectations that certain media types associated with communication will remain ephemeral. Some of our study participants were uncomfortable with the idea that a conversation may be recorded and stored locally for an unspecified period, even without intimations of reuse. Hypotheticals in a multiplayer-gaming scenario in one of our surveys revealed that participants were generally undisturbed by the thought of players saving recordings of other players' public avatar appearance, gestures, and other movement we refer to as "activity"; only 24/241, or 10%, disagreed with the right to save this content. Comparable recordings of public in-game conversations were regarded more skeptically; 58/241, or 24%, disagreed with the right to save this content. Communication carried on in public still carries with it an expectation of ephemerality that could change as standards for recording others in public grow increasingly lax.

Reusing Social Media

Reuse is one of the more contentious aspects of current legal interpretation of copyright and fair use. Some major social media sites (such as YouTube) receive numerous take-down notices for content that copyright owners feel has been inappropriately reused or reposted. Meanwhile, as Tehranian predicted in 2007,²² nuanced social norms have evolved to handle reuse of different media types and genres in a variety of circumstances.

Our reuse scenarios and hypotheticals examined at least eight features: the five described earlier—original context, the user's relationship to the content, commercial concerns, genre-derived properties, and disaggregation of constituent content—plus three additional concepts salient to reuse:

Public good. Public good scenarios seek a balance between individual rights (such as to privacy and to be forgotten) and the countervailing public interest (such as the right to preserve, access, and reuse historical content). Each of our studies included an institutional-archiving scenario that posits

the creation of a media-type-specific collection (such as an archive of public Facebook content or YouTube videos); associated hypotheticals in our surveys tested varying limits on access;

Permission. In our early studies, open-ended questions revealed that some participants thought permission was the essential bridge to fair use, although a legal approach to fair use does not require one to seek or obtain permission. Our later studies tested the mitigating force of permission with hypotheticals; that is, if permission is sought or obtained, does it drastically change participants' attitudes about reuse?; and

Venue and purpose. Our surveys used hypotheticals to compare reuse of the same content in varying contexts. For example, do participants' attitudes change if an Amazon book review is republished on a blog, on Facebook, or on another online bookstore? Because purpose may be entwined with venue, hypotheticals we spelled out in our studies specified a similar purpose so participants would judge them against the same baseline. Of special note are the hypotheticals in which user-contributed content is reused as data. This practice is common, as personal information is analyzed to draw conclusions about users as a group or to target advertising. Reuse hypotheticals also distinguished between a positive or neutral purpose and a distinctly negative purpose.

Our study results confirm the widely held user expectation that attitudes toward reuse crucially depend on circumstances and may stray far from what is legally permissible under systematized U.S. fair-use provisions.¹⁷ Aufdeheide et al.'s work with journalists¹ shows that users' stated attitudes are often more conservative than the law dictates, not less. Nonetheless, our participants' attitudes also confirmed Fiesler's and Bruckman's observation⁴ of the emergence of a rich set of reuse heuristics, norms, and self-policing tactics within communities, as reuse becomes not only commonplace but lauded in the creative arena.⁵

One of the more notable of our results is derived from commercial reuse hypotheticals, especially as they propose social media content that is reused as big data. Big data is often used in aggregate to profile community behavior and individually in personalization mecha-

nisms. The strongest disagreement arose when we asked the participants to react to a hypothetical in which Facebook sold users' profile information to Amazon (to target advertising). Figure 2a shows that more than 84% of our survey's 244 participants disagreed with the premise that this reuse is within the service's rights, and 57% disagreed vehemently. This was the most contentious of all of the reuse hypotheticals. Yet the same basic hypothetical was palatable (less than 7% objected) if the account owner's permission was solicited. The two hypotheticals define opposite ends of a spectrum of reuse attitudes.

To help tease apart the effects of the different concepts—commercial reuse, selling data, and permission—our survey proposed a third hypothetical—that Facebook can analyze internal communication among users to target advertising. This hypothetical elicited a strongly negative reaction (over 65% of our survey's 244 participants were at least somewhat negative), although their reaction was milder than the reaction to the initial hypothetical—commercial reuse of personal content—described earlier (see Figure 2a).

Commercial reuse tends to be regarded by our study participants in

a cynical light. Even fairly justifiable commercial reuse—an author reusing a reader's positive review on the author's own website—elicited negative reactions from approximately one-third of survey participants (68 of 203 responses). Yet participants recognized the value of their personal data and exhibited a willingness to monetize it themselves; for example, a hypothetical that posited the sale of one's own Facebook data for personal gain tested relatively positively; 59%, or 145 of 244 participants, agreed they should be able to sell their personal content themselves.

Reusing information in a way that changes the information's veracity (that is, so it becomes deceptive, false, or is recast in an unintended way) elicited significantly more disapproval than benign reuse. Contrast the negative responses to a hypothetical in which a podcast guest re-edited a recording of himself and vetted audience comments to eliminate damaging, but valid, material (57% of 225 participants disagreed with the guest's right to create this remix) with a comparable positive response to republication of the podcast to seek a broader audience (81% favorable).

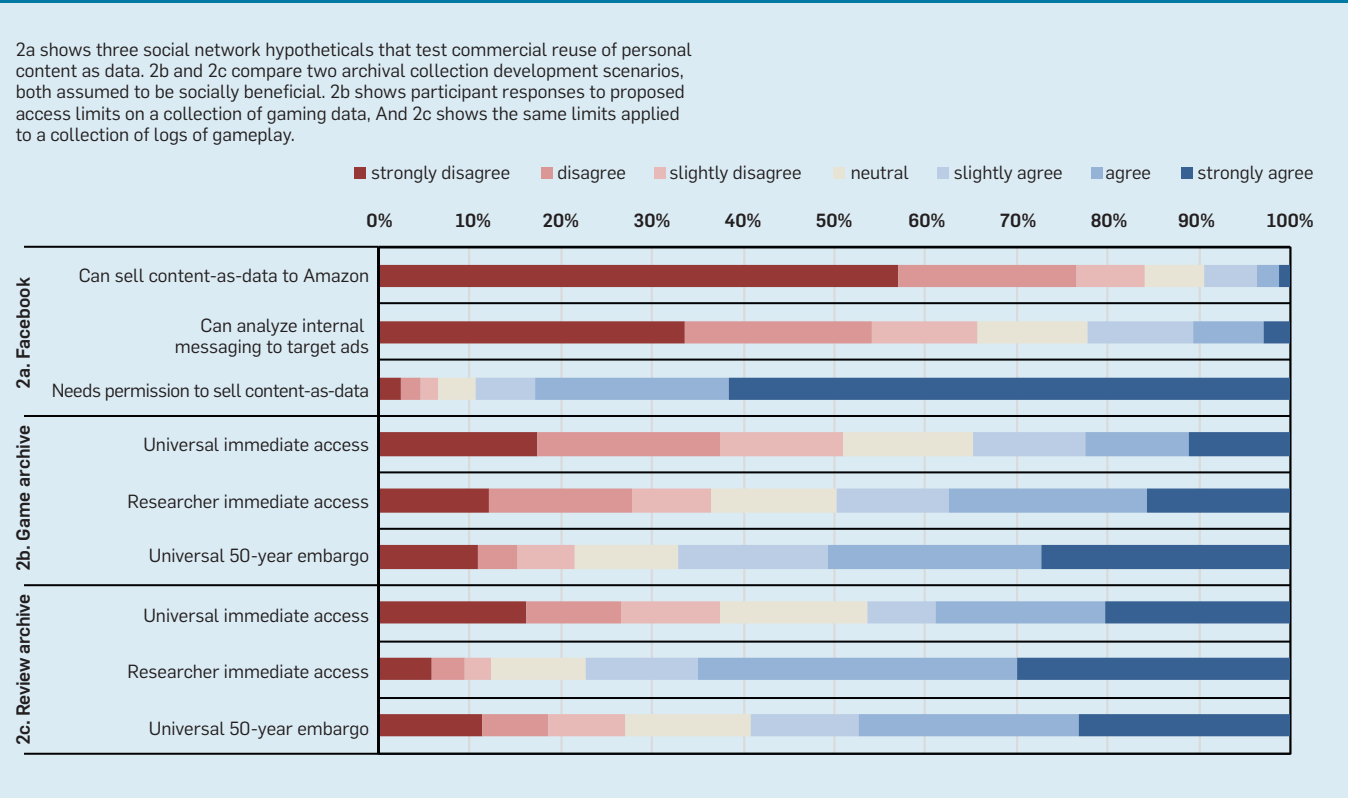
Humor is important when content is reused. Our study participants

seemed to feel that relatively few ownership restrictions should be placed on lighthearted content that falls within prescribed social norms. On the other hand, reuse that is “mean,” unwarranted, or offensive was judged more harshly. A “do no harm” heuristic provides a rough guide for how people propose to reuse other peoples' material.

Reuse for social good is viewed more skeptically than reuse of nominally humorous content. In each of our studies, a final scenario explored the idea of the U.S. Library of Congress acquiring public content from a type-specific social media service (such as Amazon Book Reviews, Facebook, Flickr, and YouTube) to create an historical collection. Three or more associated hypotheticals included in each survey tested different access restrictions on these archives. Results showed participants were more satisfied if collection access was restricted to researchers or embargoed for a 50-year period. Our own further examination of results revealed participants wanted to maintain long-term control of their public content.

Study participants are increasingly aware of what they give up when they publish profiles that describe themselves. In addition to preventing privacy

Figure 2. Social norms for content reuse.



loss, participants also seek control over their digital identities. An important reason participants do not want the Library of Congress to maintain social media collections is that their old, unrevised, public selves may be on display in such collections; ordinary citizens wish to control the retrospective and future versions, not just the current version, of themselves. Media types more closely associated with one's sense of self—Facebook profiles, photos, tweets, gaming data, even reviews—provoke a stronger anti-collection response. For media types more aligned with personal privacy—photos, tweets, gaming data—survey participants reported that limiting access to researchers—even without a definition of what constitutes a researcher—mitigates anticipated harm; for media types strongly aligned with one's identity, the preferred mitigation strategy is to place the collection under a long-term embargo restricting access to the collection for a specified period. Figure 2b and Figure 2c compare reactions to retaining a collection of gaming data as opposed to a collection of online book reviews. Gaming data provokes a strong anti-collection response if everyone is given immediate access, mitigated somewhat by limiting access to researchers, and even more by putting the collection under a 50-year embargo. On the other hand, the harm associated with collecting online book reviews is better mitigated by limiting access to the collection to researchers, and less by the proposed 50-year embargo.

This finding returns the discussion to the issue of intent; to garner popular support, reuse for the public good must be weighed against individuals' ability to develop and maintain a sense of digital identity. Likewise, notions like veracity, which normally do not enter into the legal calculus of fair use, do play a part in defining social norms.

Figure 2 summarizes important reuse concepts, as they give rise to social norms by providing contrasting responses to pairs or triples of hypotheticals.

Removing Social Media

Removing social media content by anyone besides the person who posted it is the most speculative of the three actions we have investigated. Our surveys refer to this action as “removal” rather

than “deletion” because it is intended to be nondestructive. Removal targets the copy in a particular place or context, not the content itself. Through their responses to open-ended questions about content removal, participants' revealed they usually remove material for three curatorial reasons: as a personal information management task (such as “cleaning up” one's account); in service of online identity management (such as untagging an unflattering photo of oneself); or to reflect one's changing understanding of privacy or some other aspect of online life (such as removing one's birthday from a profile).

Most social media services do not allow users to remove content created or posted by someone else. We thus derived hypotheticals from what participants in other studies mentioned they wanted to do.^{7,19} Our surveys' removal hypotheticals primarily tested three variants: changes in the remover's relationship to the content (such as Should social media users be able to remove published content according to their own self interest?); circumstances in which a neutral non-owner can remove material (such as Should non-owners be able to remove published content they believe is demonstrably wrong); and situations in which removal requires requesting and receiving permission.

Participants generally did not support the idea of removing someone else's content in one's own self-interest, regardless of whether other mitigating circumstances were introduced.

What about the case of fraud detected by a neutral non-owner? Wikipedia has inured its users to the idea that content will be reviewed and removed if it does not pass the acid test. The results of our studies have revealed that certain entities are imbued with sufficient authority to support this hypothetical type of removal. As a social norm, veracity is apparently balanced with the first factor—self-interest. If self-interest is involved, a content non-owner is not entrusted with the public welfare.

Our study participants have often felt a custodial relationship to the content (such as a website owner, service provider, or podcast producer) should give users the authority to remove content. In fact, the common expectation is that commercial service providers

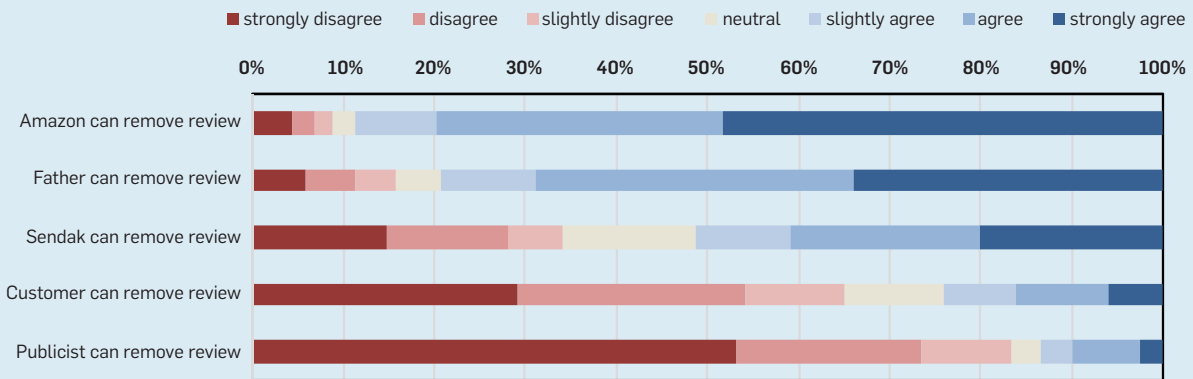
should not only have the authority to remove false or inaccurate content, they should also be charged with the responsibility for doing so.

A book-review scenario provides good examples of this test. In it, a seven-year-old girl has posted a negative book review that turns out to have been written by her father. Who has the authority to remove it? Hypotheticals we spelled out in our surveys tested removal by six different entities: the commercial host of the review website (Amazon); two potential content owners (the father or the girl); the book's author (Maurice Sendak); Sendak's publicist; and an Amazon customer who learned the review was posted under false circumstances. Who do participants believe should be given prevailing authority to remove the content, given a good reason to do so?

Study participants endorsed the idea of granting Amazon, the commercial host of the reviews, the responsibility of removing dubious content (89%, or 180/203, positive responses). The father or the girl (now a teen), as content owners, were secondarily given the authority to remove the review (77% and 79% positive, respectively). In spite of the question of self-interest, 51% of the participants thought Sendak (the author) should be able to remove the apparently fraudulent review. Only 24% thought a knowledgeable customer (aware of the circumstances under which the review was written) should normally be allowed to remove the review. Unsurprisingly, the least popular option was to allow the publicist, who was clearly acting in her client's self-interest, to remove the review; 83% thought this should not be allowed in normal circumstances. Figure 3 compares the action taken by different stakeholders.

Removal is generally the most controversial of the three actions. Although social media users want to be able to groom their own online self-presentation, there is a concomitant expectation that others should not remove content willy-nilly. The surprise in the study participants' responses to these hypotheticals was the degree of authority invested in the commercial service providers. The norm is to see commercial service providers as content guardians when such an action

Figure 3. Content-removal norms comparing five potential actors. Hypotheticals refer to a scenario in which an unfavorable book review submitted by a child was actually written by the child's father.



is necessary. Study participants also expressed intolerance for content removal motivated by self-interest.

Conclusion

We designed these eight studies, conducted over the course of five years, to elicit attitudes about how user-contributed content may be saved, reused, and removed by people other than the content's most obvious owner. The results reveal how far public attitudes have strayed from conventional legal concepts and how much they are tied to media type and other circumstantial factors. Yet these attitudes are surprisingly robust, regular, and predictable, suggesting emerging norms for the ownership and control of social media.

Among the highlights of our findings, which can be media-type-specific, are five recurring social norms:

Save anything but respect explicit social constraints. Study participants have felt they have the right to save almost anything they encounter on the open web. They reject notions of artificially imposed limits on the right to save content but also respect the explicit constraints introduced by a social network like Facebook. Social distance imposes a strong effect on whether a person can save personal information posted on Facebook. Even the powers of a friend of a friend are limited. Ownership effects are thus stronger inside social networking services than they are outside, on the open Web;

Concern for control of self-presentation. Participants object to proposals for the institutional archiving of public con-

tent on social media services like Twitter and Facebook. When probed, their objections stem less from conventional notions of privacy loss than from the loss of control of their own digital self-presentation. Digital self-presentation is subject to ongoing revision through deletion and curation of relevant online material.⁷ Congruent with Samuelson's analysis of personal data as intellectual property,¹⁸ our study participants have felt a strong right to own and control their own digital footprints, regardless of the source of this content;

Reuse norms reflect some aspects of fair use and ignore others. Social norms for reuse reflect many of the nuanced concerns of fair use (such as limiting commercial reuse, and encouraging creative or educational use) but are more pragmatic and often more conservative, relying excessively on the mitigating effects of permission. Original intent and original context are often part of a calculus of circumstantial fairness;

A right to veracity. Our study participants take information veracity seriously, though they put excessive responsibility for it into the hands of infrastructure providers. Content removed in blatant self-interest falls under this rubric, and participants generally deny others the right to remove content if self-interest is the only rationale. Fairness and accuracy are, however, seen by the participants as part of a right to remove content; and

Highly circumstantial reasoning about reuse. Differences among responses to the varying hypotheti-

cals demonstrate that participants' sense of media rights may be highly dependent on the actual reuse situation. This contextual sensitivity may interact with labeling systems like Creative Commons, since people may be unable to conceive of the full range of possible reuse scenarios or predict which are most probable. As mentioned earlier, "permission" is some participants' go-to way of mitigating unpredictable reuse. Not only does fair use case law make such a workaround unnecessary, it also does not scale to viral reuse, and experience suggests permission from the original content creator is often unobtainable.¹⁴

For example, Etsy artists may explicitly permit noncommercial use of their work, since they envision reuse that promotes their art. In so doing, they may fail to consider a popular crafts parody site that pokes fun at artisanal work. Although buyers flock to the artists' stores as a result of their work's exposure to a new audience (in line with their intent), artists may still feel indignant about the nature of the reuse. To complicate matters further, the parody site donates its proceeds to charity, so the sting is mixed with social good. Content creators are thus faced with complex trade-offs. Whether reuse restrictions are implemented through technology, policy, or a combination of the two, managing rights relies crucially on the ability of content owners to envision plausible reuse scenarios and predict which are most likely.

In *Code: Version 2.0*, Lessig⁶ identified four constraints that regulate

online behavior: architecture, law, market forces, and social norms. As seen in the overall responses in our studies, social norms seem to have an outsized effect on participants' perceptions of what they (and others) can do with user-contributed content. Even social-media-savvy participants have little understanding of the relevant legal guidelines. Software-based governance is easy to ignore or thwart. And much reuse is oblivious to market forces. Furthermore, social norms are often nonreciprocal in action; participants in our studies did not always apply the same standards to themselves that they did to others, especially in non-abstract practical situations. This lack of reciprocity is not uncommon in other aspects of online behavior and may be attributed to individual users' ability to reflect on their own motives and intentions but not those of others.

What are the design and policy implications of these results? For one, they signal certain design gaps when media creators use labeling schemes (such as Creative Commons²); study participants seemed more sensitive to actions like reuse when they are offered examples rather than abstract labels. Hypothetical examples of reuse, especially those based on the media being labeled, may be helpful for extending Web users' understanding of the abstract ideas expressed by labels. It is no accident that our final norm addresses highly circumstantial factors as the nature of the content (such as "Is it personal?"), the differential scope of the audience (such as "Is the content going viral or is it playing to an audience of 10?" and "How different is the scope from the original?"), the type of reuse (such as is the content used in a way that highlights the original intent?), and the way the implied (or explicit) social contract between all potential owners of both the original and derived work is handled (such as "Is attribution or anonymity desired?").

Note only one of these factors—the nature of the content—is known at publication time, or the time when content is usually labeled. Other factors depend on how the content is reused (such as changes in genre, audience, or publication venue). Still others are not revealed until time has

passed (such as the differential scope of the audience). That these factors are crucial to how a labeling scheme is used makes us think that supplemental mechanisms might be desirable; scenarios, hypotheticals, and mixed-initiative dialogs help content creators better envision many types of reuse or decide between attribution or anonymity or triggers that reveal when the scope or audience has changed. Still others depend on, say, the motivations for storing content. Past work tells us that individuals archive work that is not their own just as surely as institutions do.¹⁵

Ownership-driven questions need to be approached thoughtfully, lest we impose legal restrictions when none are necessary or fail to anticipate normal actions that will trigger reactions that could have been averted. Gaps between desired policy and current social norms may yet be bridged through education and thoughtful design.

Acknowledgments

We would like to thank Microsoft Research, Silicon Valley, for supporting the studies on which this article is based. Thanks, too, to the study participants for their patient and thoughtful completion of the lengthy surveys. ■

References

1. Aufderheide, P., Jasz, P., Bieze, K., and Boyle, J.L. *Copyright, Free Speech, and the Public's Right to Know: How Journalists Think About Fair Use*. SSRN, Elsevier, July 30, 2012; <http://ssrn.com/abstract=2119933>
2. Boyle, J. *The Public Domain: Enclosing the Commons of the Mind*. Yale University Press, New Haven, CT, 2008.
3. Ellickson, R. *Order Without Law*. Harvard University Press, Cambridge, MA, 1994.
4. Fiesler, C. and Bruckman, A.S. Remixers' understandings of fair use online. In *Proceedings of the 17th ACM Conference on Computer-Supported Cooperative Work and Social Computing* (Vancouver, BC, Canada, Mar. 14–18). ACM Press, New York, 2014.
5. Hill, B., Monroy-Hernandez, A., and Olson, K. Responses to remixing on a social media sharing website. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media* (Washington, D.C., May 23–26). AAAI, 2010, 74–81.
6. Lessig, L. *Code: Version 2.0*. Basic Books, New York, 2006.
7. Lindley, S., Marshall, C.C., Banks, R., Sellen, A., and Regan, T. Rethinking the web as a personal archive. In *Proceedings of the 22nd International World Wide Web Conference* (Rio de Janeiro, Brazil, May 13–17). ACM Press, New York, 2013, 749–760.
8. McCormick, D.N. and Summers, R., Eds. *Interpreting Precedents*. Ashgate/Dartmouth, Farnham, U.K., 1997, 528–9.
9. Marshall, C.C. and Shipman, F.M. Social media ownership: Using Twitter as a window onto current attitudes and beliefs. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada, May 7–12). ACM Press, New York, 2011, 1081–1090.
10. Marshall, C.C. and Shipman, F.M. The ownership and reuse of visual media. In *Proceedings of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries* (Ottawa, Canada, June 13–17). ACM Press, New York, 2011, 157–166.

11. Marshall, C.C. and Shipman, F.M. Experiences surveying the crowd: Reflections on methods, participation, and reliability. In *Proceedings of the Fifth Annual ACM Web Science Conference* (Paris, France, May 2–4). ACM Press, New York, 2013, 234–243.
12. Marshall, C.C. and Shipman, F.M. Saving, reusing, and remixing web video: Using attitudes and practices to reveal social norms. In *Proceedings of the 22nd International World Wide Web Conference* (Rio de Janeiro, Brazil, May 13–17). ACM Press, New York, 2013, 885–896.
13. Marshall, C.C. and Shipman, F.M. Exploring the ownership and persistent value of Facebook content. In *Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing* (Vancouver, BC, Canada, Mar. 14–18). ACM Press, New York, 2015, 712–723.
14. McDonough, J., Olendorf, R., Kirschenbaum, M., Kraus, K., Reside, D., Donahue, R., Phelps, A., Egert, C., Lowood, H., and Rojo, S. *Preserving Virtual Worlds Final Report. National Digital Information Infrastructure and Preservation Program*, Washington, D.C., Aug. 31, 2010; <http://hdl.handle.net/2142/17097>
15. Odom, W. Sellen, A., Harper, R., and Thereska, E. Lost in translation: Understanding the possession of digital things in the cloud. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Austin, TX, May 5–10). ACM Press, New York, 2012, 781–790.
16. Rissland, E.L. Dimension-based analysis of hypotheticals from Supreme Court oral argument. In *Proceedings of the Second International Conference on AI and Law* (Vancouver, BC, Canada, June). ACM Press, New York, 1989, 111–120.
17. Sag, M. Predicting fair use. *Ohio State Law Journal* 73, 1 (2012), 47–91.
18. Samuelson, P. Privacy as intellectual property? *Stanford Law Review* 52, 1125 (1999).
19. Sas, C. and Whittaker, S. Design for forgetting: disposing of digital possessions after a breakup. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Paris, France, Apr. 27–May 2). ACM Press, New York, 2013, 1823–1832.
20. Shipman, F.M. and Marshall, C.C. Are user-contributed reviews community property? Exploring the beliefs and practices of reviewers. In *Proceedings of the Fifth Annual ACM Web Science Conference* (Paris, France, May 2–4). ACM Press, New York, 2013, 386–395.
21. Shipman, F.M. and Marshall, C.C. Creating and sharing records of multiplayer online game play: Practices and attitudes. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media* (Ann Arbor, MI, June). AAAI Press, Palo Alto, CA, 2014, 456–465.
22. Tehrani, J. Infringement nation: Copyright reform and the law/norm gap. *Utah Law Review* 3 (2007), 537–550.

Catherine C. Marshall (ccmarshall@cse.tamu.edu) is an adjunct professor of computer science and engineering and affiliate of the Center for the Study of Digital Libraries at Texas A&M University, College Station, TX; she lives in San Francisco, CA, and volunteers at the Internet Archive (<https://www.archive.org/>).

Frank M. Shipman (shipman@tamu.edu) is a professor of computer science and engineering and associate director of the Center for the Study of Digital Libraries at Texas A&M University, College Station, TX.

© 2017 ACM 0001-0782/17/05. \$15.00



Watch the authors discuss their work in this exclusive Communications video. <http://cacm.acm.org/videos/who-owns-the-social-web>

Copyright of Communications of the ACM is the property of Association for Computing Machinery and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.