

2016

Values, Ethics and Participatory Policymaking in Online Communities

Alissa Centivany

Western University, acentiva@uwo.ca

Follow this and additional works at: <http://ir.lib.uwo.ca/fimspub>



Part of the [Library and Information Science Commons](#)

Citation of this paper:

Centivany, Alissa, "Values, Ethics and Participatory Policymaking in Online Communities" (2016). *FIMS Publications*. Paper 80.
<http://ir.lib.uwo.ca/fimspub/80>

Values, Ethics and Participatory Policymaking in Online Communities

Alissa Centivany

Western University, Faculty of Information and Media Studies, North Campus Bldg., London, ON, N6A 3K7.
acentiva@uwo.ca

Drawing upon principles and lessons of technology law and policy, value-centered design, anticipatory design ethics, and information policy literatures this research seeks to contribute to understandings of the ways in which platform design, practice, and policymaking intersect on the social media site Reddit. This research explores how Reddit's users, moderators, and administrators surface values (like free speech, privacy, dignity, and autonomy), hint at ethical principles (what content, speech, behavior ought to be restricted and under what conditions), through a continuous process of (re)negotiating expectations and norms around values, ethics, and power on the site. Central to this research are questions such as: Who or what influences and/or determines social practice on Reddit? Who participates in decision-making and using what processes and mechanisms? Where do controversies arise and how are they resolved? Generating findings from a particular controversy surrounding the subreddit /r/jailbait, the author illustrates the complexities inherent in these questions and suggests that a participatory policymaking approach might contribute to future research and practice in this area.

Keywords

Information Policy, Social Media, Value-Sensitive Design, Anticipatory Ethics, Online Communities

Introduction

Recent scholarship highlights the need for emerging information technologies to account for human values and support ethical use and engagement (Friedman, Kahn, & Borning, 2013; Knobel & Bowker, 2011; Johnson, 2003, 2011; Shilton, 2014, 2015a, 2015b). A related but distinct strand of scholarship encourages more widespread and deliberative reflection on the integral role of policy, along with technical design and social practice, in the emergence and development of sociotechnical systems (Jackson et al, 2014; Centivany, 2016). A recently published work by the author explored the intersections of platform design, emerging practice, and participatory policymaking processes on the social media site Reddit focusing in

particularly on the importance of user loyalty in producing policy outcomes favorable to administrators, moderators, and members of the general user community. (Centivany & Glushko, 2016). This paper continues that line of inquiry but takes a slightly different tack, focusing less on policy outcomes than on the ways that values, ethics, and power are surfaced, discussed, debated, and negotiated as part of broader policymaking process. This research focuses on a particular controversy involving the now-defunct subreddit /r/jailbait as a window for understanding how and why values, ethics, power, and policies emerge, evolve, are disputed, reconstructed, and sometimes dissipate through a dynamic, ad hoc, co-constitutive participatory process.

This paper begins with a review of values, ethics, and policy literatures drawn from the fields of information science, human-computer interaction, and technology law and policy scholarship. It then provides a description of Reddit and the controversy surrounding /r/jailbait focusing in particular on the interplay between Reddit administrators, Reddit users, popular media and news reports, and members of the public in processing and evaluating the issues. The paper then introduces participatory policymaking as a potentially useful lens for understanding how and why values like freedom of expression are co-constituted and co-mediated by administrators and users of Reddit. This approach is intended to provide information science researchers with new tools for discussing and understanding the tensions and tradeoffs among values, ethics, and policies in emerging online communities.

Related Work

The relationship between technical design and social practice has been a significant focus of much research in information science and related fields concerned with the ways in which society shapes and is shaped by emerging technologies. Early leaders and disciplinary precursors to information science were deeply aware of and concerned with the value-laden potentials of new technologies. Great minds like Wiener, Kling, and many others understood that technological change did not equate with social, moral, or even technological progress. Innovative was not tracking a certain, if invisible, upward and optimistic trajectory. Rather, networked communications technologies were observed as simultaneously pregnant with promise and rife with peril. As Norbert Wiener (1964) (whom some have

called the father of information ethics (Bynum, 2001)) famously cautioned, increased technological ubiquity and interconnectivity is likely to present as many challenges as solves:

The world of the future will be an even more demanding struggle against the limitations of our intelligence, not a comfortable hammock in which we can lie down to be waited upon by our robot slaves (Wiener, 1964: 69).

Expanding upon this way of thinking, we might reason that it is ultimately up to us, the human factors in these technologies and systems, to ensure that we are designing, implementing, and using the *right technology* in the *right way*. Such a determination should naturally take into account concerns around design efficiency, effectiveness, and practicality of use, but also, and perhaps more importantly, concerns around human values like trust and dignity, and ethical principles like how we evaluate and ensure fairness, that are crucial to the overall health, well-being, and productivity of society.

Scholars have adopted various ways of framing and studying values and ethics emerging through the intersections of technology, practice, and law and policy. For example, scholars working in the law and technology space have long-recognized the challenges and promise inherent in Internet's open architectural design. Larry Lessig (1999) and others have argued that "code is law" – that technologies can effectively regulate behavior, sometimes in ways that are detrimental to human values and social welfare. Jonathan Zittrain (2006, 2008) has argued that the generative capacity of information technologies, most notably the Internet, to support distributed, unaccredited, and often uncoordinated audiences to build and distribute code and content across its vast network creates a serious and potentially troubling dynamic between new opportunities for innovation and progress and a new forms of regulation, control, closure. I have recently argued for a view of policy as source of embedded generativity, suggesting that not only can law and policy serve as a corrective response or gatekeeper — opening and closing space for technical design and social practice — but that it can also preserve and safeguard space for future unanticipated innovations, collaborations, and transformation might emerge (Centivany, 2016).

This shift away from viewing law and policy as static, monolithic obstacles operating in (often ineffective) parallel rather than concert with relatively faster moving sociotechnical systems is particularly crucial as we move into discussions of values and ethics in social media contexts where a continuously co-evolving and co-constitutive community of participants constantly (re)surface, (re)evaluate, and (re)configure what participation means, and what values and ethical principles are paramount. At this point it is worth noting that this research purposefully does not adopt a particular ethical framework or theory for

evaluating social media sites like Reddit. Instead the primary objective of this research is to describe, explain, and understand how the users, moderators, and administrators of the site, through their interactions, surface particular values as "core values" and adopt particular approaches to sensemaking and decision-making as the "right ones" for figuring out tough ethical dilemmas.

Researchers working in the fields of information science (Knobel & Bowker, 2011), human-computer interaction (Friedman, Kahn, & Borning, 2013), media and communication studies (Flanagan, Nissenbaum, & Howe, 2008), and science and technology studies (Johnson, 2003, 2011; Sclove, 1995; Winner, 1989) have taken a proactive and systematic approach to analyzing and/or incorporating human values in technological artifacts and systems. In particular, as networked technologies become increasingly ubiquitous and essential to our participation in the world, some researchers have begun to emphasize an emerging "platform society" where large-scale platforms like Google, Facebook, Twitter, and Reddit become the means through which we conduct business, communicate with family and friends, and learn about the world. Internet platforms are increasingly becoming the venues where integral aspects of contemporary life get played out. Understanding the ways computing platforms influence, monitor, and perhaps control our experience of the world is of utmost concern to researchers interested in values, ethics, and power as they relate to information and technology.

Amongst this community of researchers, it is observed that platforms like Reddit are not hollow technical infrastructures or neutral playing fields. Rather they are understood to be essentially political and contested, and comprised of dynamic overlapping zones of influence. A growing contingency of researchers, emerging from the field of human-computer interaction, have developed and adopted a Value-Sensitive Design approach which actively seeks to incorporate positive human values, such as privacy, trust, and autonomy, into the design of information technologies, systems, and platforms. This approach, championed by Friedman, Kahn, Borning, and others seeks to account for human values in a comprehensive and principled way through all stages of the design process. In addition, these scholars argue that measures of quality with respect to technological systems ought to include considerations of how the system promotes (or fails to promote) human values (Friedman, Kahn, & Borning, 2013).

Value-Sensitive Design therefore pushes designers and engineers of information technologies and systems to move beyond merely functional concerns such as speed, capacity cost, durability, and robustness. It recognizes that technologies directly and systematically influence the promotion and/or suppression of particular social, ethical, and political values (Flanagan, Howe & Nissenbaum, 2005). Moreover, this approach seeks to incorporate questions and concerns about the moral dimensions of the technology early

in its development (Friedman, Kahn, & Borning, 2013; van den Hoven, 2005). Focusing on moral values during the design process, rather than as an afterthought (or not at all), in a principled, systematic, and comprehensive way is, it is argued, provides the best chance that the technologies we create will promote moral, political, and socially beneficial ends. Ideally, ethics and policy researchers should be integrated into the design team at an early stage to provide perspective and guidance on the design process as it emerges.

In some ways this is the approach of the related Anticipatory Ethics researcher who seeks ways of incorporating ethics more explicitly into the design of emerging technologies. Anticipatory ethicists frame the approach as being primarily concerned with translating values – abstract, often ill-defined interests and goals – into actionable ethical practices. As described by Katie Shilton, “Anticipatory ethics is anticipation of how future technologies will be built, how they will be applied, and what their consequences might be” (Shilton 2015).

Similarly to the Value-Sensitive Design approach, Anticipatory Ethics stresses the R&D and introduction stages of technological development. One of the central challenges facing both of these approaches is the problem of uncertainty. Ethical issues relating to the use of information technologies that have not actually been fully developed or released into the wild are difficult to accurately and reliably predict. This uncertainty should not, in the view of these researchers dissuade technologists from considering the potential ethical implications of a proposed or plan design. As Philip Brey describes, uncertainty “come only be overcome through methodologically sound forecasting and future studies” (Brey, 2012). As a practical matter, however, the combination of uncertainty paired with the breakneck pace of technological development and the competitiveness of high technology markets, might suggest that, in practice, persuading developers to put functional concerns on hold in favor of important but largely inchoate ethical concerns may pose a significant barrier.

In addition, these approaches leave open questions for researchers studying the relationships among technical design, social practice, and values and ethics on *existing* large-scale Internet platforms like Reddit. Since we cannot turn back the clock and advise sites’ progenitors to consider the human values implicated by design choices, we need to think about how we might adapt and synthesize those methods to generate insights about sites already in existence. We might borrow useful insights from information policy and science of sociotechnical systems research.

Unlike the Anticipatory Ethics and Value-Sensitive Design approaches which focus primarily on the design process, a policy-oriented perspective offers a different, broader, approach to understanding and managing the intersection of emerging technologies and human

experience. At its essence, “policy” is rules and procedures designed to achieve particular goals. Policy includes forms of public law that regulate particular issue-spaces such as privacy, telecommunications, antitrust, security, intellectual property, and so forth where technologies may play an important or controversial role. Policy also includes a diverse range of rules, processes, mechanisms and procedures instituted by private firms to regulate users’ interactions with firms’ goods and services. (Jackson et al, 2014).

Policy, as well as design and social interaction, plays an important role in influencing which values may justify sensitivity or attention and how those concerns might translate into action and practice. This is not to suggest, however, that each of the approaches herein discussed compete or conflict with each other. Rather, Value-Sensitive Design, Anticipatory Ethics, and Information and Technology Policy approaches each offer distinct but related and complimentary opportunities for engaging with the moral and normative dimensions of our shared human experience with technology.

In addition, as this research will demonstrate policy and policymaking processes, particularly during moments of controversy, provide important clues of the design-practice-values-ethics interaction on existing platforms like Reddit. We can use the policymaking process as a touchstone for beginning to disentangle and trace the emergence and evolution of design choices, social practice, values, and ethics over time as way of understanding how and why users, moderators, and administrators make sense of and respond to the moral and ethical tradeoffs inherent in the life of social media platforms.

Methods

To understand the relationships among values, ethics, power, and policymaking on social media platforms this research undertook an in-depth qualitative study of a single controversy on Reddit involving the subreddit /r/jailbait. Data for this study, which consisted of publically available comments, discussions, and reports, were manually scraped from /r/jailbait and other subreddits engaged in discussions and debate on the controversy as well as popular news media outlets. Using an iterative, inductive process, data was coded for user name, subreddit, user role (e.g. moderator, user, administrator and so forth), are various qualitative measures signaling values and emerging ethical principles. In addition, comments with more “up votes” were interpreted as more representative of community values/beliefs than lower scored comments. We attempted to triangulate and anchor data wherever possible to foster increased credibility, validity, and trustworthiness.

The /r/jailbait controversy is studied, disentangled, and traced from its emergence to its resolution as a way of understanding how and why values, ethics, power, and policies emerge, evolve, are disputed, reconstructed, and

sometimes dissipate through a dynamic participatory process. In so doing this research hopes to shed light on the ways that design, values, ethics, and policy interact on social media platforms like Reddit.

Values, Design, and Policy Trade-offs on Reddit

Compared to other top social media platforms, Reddit's interface is simple, clean, and uncluttered: blue text on a white background. Few ads mar the user experience of what is essentially a list of headlines. On any given day, visitors to Reddit's homepage will be greeted by the so-called "front page of the Internet," a diverse, unrelated, and dynamic list of the top-voted user-supplied and/or user-generated posts pulled from Reddit's more than six thousand active subreddits. Subreddits are semi-autonomous niche forums that are created and moderated by Redditors and are typically dedicated to a particular topic. There are currently fifty default subreddits that tend to funnel the majority of content to Reddit's front page. The vast majority of subreddits remain in the shadows, far from the attention and glory of the FrontPage. Perhaps they languish from a lack of interest or purposefully remaining under the radar as meeting grounds for members espousing or exploring marginal, controversial, elicit, and perhaps illegal topics.

Reddit is therefore more than a link aggregator, more even than the "front page of the Internet," and it is distinguishable from other large-scale social media platforms like Facebook, Twitter, LinkedIn, Pinterest and others for at least three key interrelated reasons: first, Reddit *is* its users; second, Reddit was founded upon a techno-libertarian commitment to free speech online and this principle continues to be core to its identity; third, Reddit's technical design is based upon democratic ideals (voting mechanisms) with thin but sharp administrative oversight (admins have relatively few rules or policies but have ultimate control over content and user account data). This combination of unique features not only serves to distinguish Reddit from its peers but, as I will explain, also generates tensions that cause and exacerbate conflict and controversies around design, practice, values, ethics, and policy tradeoffs.

Reddit is its users

While the content of the homepage is diverse and appears somewhat random to an outsider, a seasoned observer or participant in Reddit might recognize a community ethos or culture reflected in the mishmash. A Reddit historian recently referred to it as a nest of "atheist, libertarian, porn-loving Ron Paul fans" (Fiegerman, 2015), but perhaps a more subdued description of its ethos or culture may be that it is youthful, predominately male, and thirsty for content that is humorous, erotic, informative, and newsworthy.

The culture of Reddit is reflected in the content on the front page. Beyond the front page, there are a growing number non-default subreddits created and moderated by users. At the time of this writing, there were over six thousand active subreddits covering a wildly diverse set of content.¹ Creators and moderators of subreddits enjoy a great deal of control over the content and the look and feel of the subreddit, subject to the (generally hands-off) discretion of Reddit administrators. As will be discussed in a following section, many of the conflicts and controversies at the intersection of values, ethics, and policy begin on non-default subreddits. Reddit's content and its community merge into a multifaceted but inseparable whole. Particularly as the instances of Reddit-originating content rises, through the wildly popular "IAMA" and "askreddit" subreddits for example, shared norms and expectations within the community begin to emerge. In the shadows of Reddit, amongst the more than one hundred thousand non-default subreddits, subcultures also emerge, coalesce, and break apart. In many of these subreddits, the value lies not in linking, link aggregation, or up votes. The value lies in the discussions and debates amongst users. Even where the initiating post begins as a link to external content, the discussions and debates that are sparked appear to be the aspect that redditors value most. Reddit has grown to depend on its users for much more than simply supplying links and votes. Reddit *is* its users.

This initial observation may not seem particularly insightful. Facebook, Twitter, LinkedIn, Pinterest and other social media sites are arguably *their* users as well. The significance of Reddit's relationship with its users becomes more apparent when considered in conjunction with these two additional key features of site.

Freedom of Expression is a Core Value

Freedom of speech has been a core value associated with Reddit since its earliest beginnings, when the company merged with Aaron Swartz's Infogami company in 2005 and its identity began to coalesce around content aggregation. Though not expressed as a positive, formal policy declaration, the commitment to freedom of expression may be fairly imputed through the actions, forbearances, and commitments of some of its creators, most notably Aaron Swartz, a technological wunderkind, hacktivist, and proponent of a free and open Internet. In addition, a negative inference in favor of free speech may be drawn from the express policies that do exist on Reddit. Reddit has relatively few content-based restrictions: illegal content (such as child pornography), spam, and personal and confidential information are deemed "unwelcome." More recently (post-/r/jailbait), the site has instituted prohibitions against involuntary pornography and content that

¹ Subreddits are identified by the text at the end of the slash at the end of the main reddit URL, e.g. [www.reddit.com/r/\(subreddit\)](http://www.reddit.com/r/(subreddit)). Additionally, registered users contributions are tracked and can

be viewed by going to the Reddit URL ending in their username, the convention for this is [www.reddit.com/u/\(name\)](http://www.reddit.com/u/(name)).

encourages or incites violence, or threatens, harasses, bullies or encourages other to do so. In addition to unwelcome content, Reddit prohibits certain behaviors including vote manipulation through “brigading”² or other means, “breaking Reddit,” doxing,³ and creating multiple accounts to evade punishment or avoid restrictions.

In practice, Reddit administrators, moderators, and users have demonstrated a general aversion to viewpoint-based censorship. A consequence of this leniency has been a proliferation of highly offensive subreddits including /r/picsofdeadkids, /r/coontown, /r/greatapes, /r/nsfl, /r/thefapping, /r/theredpill, and /r/sexyabortions, not to mention every imaginable mishmash of pornography, violence, racism, and sexism.

For some, including /u/hueypriest (2011), a former General Manager of Reddit’s, the offensive, heinous, and egregious content is part of the tradeoff of enjoying a free and open site:

Personally, I think they are gross. But let’s take the infamous picsofdeadkids example. The actual content of that subreddits is mostly autopsy photos. Obviously it’s a troll subreddits and created to get a reaction, and I’d guess 98% of redditors think it’s gross/offensive etc. But what if the name of the subreddits was /r/autopsyphotos or /r/doyoureallywanttogointocriminalforensics and they were sincere in their discussion of these images? What if it wasn’t kids but adults? Or historical autopsy photos only? The point is I don’t want to be the one making those decisions for anyone but myself, and it’s not the business reddit is in. We’re a free speech site with very few exceptions (mostly personal info) and having to stomach occasional troll reddit like picsofdeadkids or morally questionable reddits like jailbait are part of the price of free speech on a site like this” (sic) (/u/hueypriest, 2011).

This quote encapsulates the notion that a freedom of expression is a core value on Reddit and requires tradeoffs. Users might not like or agree with all of the content that is posted but a user who chooses to can avoid viewing offensive subreddits and that is a relatively small price to pay for ensuring a free and open social media platform.

Embedded Democracy

Related to the first two features discussed (Reddit *is* its users and holds a core value of free speech) is the third one: Reddit embeds democratic values in the technical design of the site, primarily through voting mechanisms and distributed authority vis-à-vis subreddits. Content on Reddit’s front page and across all of its various subreddits continuously changes as new posts are created and existing posts rise and fall in prominence through the operation of a

popularity algorithm based on the users’ votes. If a registered user likes a link or a comment they can “up vote” it, adding points to the content’s score which helps move it up the page, raising its visibility and prominence. By contrast, if a user dislikes a comment, they can “down vote” it, which effectively subtracts from its overall score, moving it down the page and thereby reducing its visibility and prominence. Content is thus vetted and organized by fairly simple but extremely effective voting mechanics. Nearly all of the content that reaches the front page, and a huge majority of Reddit content overall, is informative, funny, or relatively inoffensive to the majority of redditors. In other words, content from the most offensive subreddits will generally lack sufficient votes to ever get close to the front page.

Similar voting mechanisms existed on other social computing/link aggregating websites such as Digg. But, unlike Digg and other content-sharing sites, Reddit is distinguished by users’ ability to create semi-autonomous subreddits, individual instances where users could moderate their own mini-reddits, allowing or disallowing content as they saw fit. Subreddits were not an entirely new idea, having been considered by Reddit’s major competitor and conceptual big brother Digg. Digg ultimately rejected the idea of subreddits, with founder Kevin Rose nixing them because of the “potential chaos that would bring” (Fiegerman, 2015). Whether this insight proved to be short sighted or prescient depends on the time scale under consideration. Reddit did begin to steadily eat into Digg’s traffic and influence, culminating with Reddit taking over Digg in traffic in 2010 (Fiegerman, 2015; Tassi, 2015). However, since the beginning, it has always been the presence of the subreddits and their internal communities that have caused the greatest tensions and conflicts across the wider Reddit community.

These unique features of Reddit not only distinguish it from its peers and competitors, but also prime the way for tensions and controversies to emerge. Collisions have been increasingly populating the intersection of Reddit’s core commitment to freedom of expression and its technical/governance structure as Reddit’s administrators, moderators, and users grapple with what *it* is. One such collision, arising in the context of the subreddit /r/jailbait, is now discussed.

/r/jailbait

On September 29th, 2011, the American journalist, author, and television host Anderson Cooper delivered this chilling warning to living rooms all over the United States:

² Brigading: “when a group of people get together to down vote the same thing, be it a single person, or a group of people representing a dissenting ideology” (drafterman, 2014).

³ Doxxing: “the search for, and subsequent (sic) publication of private personal information of people.” (splattypus, 2014).

Somebody somewhere is looking at sexually suggestive photos of your teenage child and you might be just as surprised to learn where these pervy grownups find the images. They're posted on a site that's part of one of the most respected publishing empires on Earth. We're talking about a site called Reddit (Cooper, 2011).

Cooper was reporting on a particular subreddit called /r/jailbait. As its name suggests, and as Cooper stressed, this subreddit invited users to post, comment on, and discuss images of minors. The images were often surreptitiously obtained and posted to /r/jailbait without the consent or knowledge of the subject, as fodder for visitors seeking to sexualize minors.

We might fairly assume that the vast majority of Cooper's audience, and many in the broader Reddit community, would indeed find /r/jailbait highly offensive, even shocking perhaps. But many redditors also took issue with the sensationalism and false-equivocation implied by Cooper's report which ignored the community's core values, any sense of values tradeoffs, and any potentially mitigating ethical or design principles (i.e. as a non-default subreddit one does not typically stumble upon /r/jailbait). The sentiment expressed by /u/NightBane's was shared by many:

"I just watched a segment on Anderson Cooper 360, where he highlighted Reddit. Which at first I thought was a good thing. However, he then began to focus on the obscure points of Reddit, singling out /r/jailbait, and continuously bashed Reddit, without even looking at the rest of the website. I'm a little offended, Reddit. There's more to us than "Dead Babies" and "Kiddy Porn." Anderson Cooper just tainted us all." (NightBane)

/u/NightBane's point may be well-taken but, at the time, the subreddit was indeed popular. In its coverage of the /r/jailbait controversy, *PC Magazine* reported:

As one user noted, the 'jailbait' reddit was the first result that came up in Google when users searched for 'jailbait.' It was definitely one of the major placed people went when looking at underage girls in bikinis,' kafka201 wrote. (Hachman, 2012)

Additionally, it was well-recognized by the Reddit administrators who awarded the subreddit a prize for "worst reddit" in 2009 and gave its moderator, /u/Violentacrez a special "pimp hat" badge to display in his profile in honor of his demonstrated service as a moderator (Ohanian, 2009; Chen, 2012). It is difficult to know how to interpret this mixed attention. Was /r/jailbait really part of the community or was it an anomaly? On a site that contains, at last count, over one hundred thousand subreddits (although only about 6% are active), how should redditors and members of the broader public interpret and make sense of /r/jailbait's significance? And how should redditors, moderators, and administrators respond to Cooper's report which unleashed an influx of both negative press and traffic to /r/jailbait?

A number of considerations further complicate this already complicated scenario. First, although /r/jailbait may cater to a niche audience, it may not be entirely accurate or fair to reject or dismiss /r/jailbait as some sort of playground limited to "pervy grownups" without at least considering the possibility that that category may be far more inclusive than first suspected. Indeed, in a questionably tongue-in-cheek sentiment echoed by several others, redditor /u/windwalker said: "I don't normally go to /r/jailbait, but being reminded of it, I'm browsing /r/jailbait, thanks Anderson Cooper!" Raising awareness of the subreddit may have the unintended consequence of also promoting its membership.

Second, although /r/jailbait may have been offensive to Cooper and many in his audience, there was no obvious legal basis for objecting to the subreddit. Although it facilitates and encourages the sexualization of minors, the images did not amount to child porn; the subjects were clothed and engaged normal, i.e. not sexually explicit, daily activities. In addition, because the images are taken in public, the subjects would not have a strong claim to privacy. Furthermore, while /r/jailbait may have been unethical, as a United States-based website, the First Amendment provides strong protections for speech even (and some would argue especially) speech that is unpopular or offensive. Indeed, for the state to shut down /r/jailbait, it would have had to show that the subreddit incites or produces imminent lawless action, a high burden to prove particularly where the nexus between speech and action is made more tenuous by the digital environment in which it operates.

Third, although it may reflect some specious reasoning, some viewed /u/Violentacrez moderation as providing a useful service to a community that seemed inevitably bound to attract offensive and pornographic material. In a Gawker article on /u/Violentacrez, Chris Slowe, the lead programmer for Reddit at the time stated, "(w)e just stayed out of there and let him do his thing and we knew at least he was getting rid of a lot of stuff that wasn't particularly legal," Slowe said. "I know I didn't want it to be my job." (Chen 2012) Moderating content, particularly in offensive subreddits, is a difficult and draining chore. Moderators like /u/Violentacrez identified and deleted much of the illegal content, such as child pornography, that was posted to Reddit (Chen, 2012). If the assumption holds, that offensive content will invariably find a home in a subreddit, then the *laissez-faire* attitude of the administrations might seem reasonable. While there may have been some discomfort around /r/jailbait, under this view, it was still part of the Reddit family, even if only as the black sheep.

While, as a private company, Reddit is not obligated to adhere to the First Amendment, as previously discussed, the principles of free speech it embodies resonate as a core value of Reddit's founders. In the techno-libertarian worldview of Reddit's founders, offensive content contributed by some users may simply be part of the price you have to pay to ensure free speech for all. In an interview with *Forbes*

Magazine, Reddit co-founder Alexis Ohanian cited the founding fathers of the United States and mused that they would have approved of Reddit as a “bastion of free speech” on the Internet: “I bet they would like it...It’s the digital form of political pamphlets” (Hill, 2012).

This value was also shared by a substantial proportion of the community which is why, when the /r/jailbait was raging, the threat of censorship loomed large in the minds of many members of the community. /u/ Scary_The_Clown, for example, pondered:

“(i)s /r/jailbait illegal? If so, reddit should take it down (and I believe they would)

It's not illegal - it's morally objectionable. Should reddit shut down subreddits that people find morally objectionable? Kiss /r/atheism goodbye. Is there an anti-Scientology subreddit? You know that's gone. /r/NSFW? Gone.

I'll wager there are folks who would step in to shut down /r/Assistance because they'll argue that charities should be registered or some crap.

Cable companies would love to get rid of /r/cordcutters /r/trees "only exists to advocate illegal activity"

And so it goes. Open the door to shutting down subreddits based on moral objections and it will never end” (Scary_The_Clown)

The commitment to the values of free speech have been tempered somewhat over time, under the pressures caused by controversies around /r/jailbait and through structural and organizational shifts and stabilization (Centivany & Glushko, 2016). Despite the bold proclamations of Alexis Ohanian, administrators have drawn lines in the sand to signal when redditors go too far. Reddit’s administrators, for example, ultimately did shut down /r/jailbait. Perhaps it was simply a byproduct of the necessary shift from generativity to stabilization, perhaps it was just an unavoidable consequence of “growing up.” /u/Khiva observed that, while Reddit did, at one time, hold steadfast to core values:

“these lines distinguishing between permissible and impermissible behavior are prone to shift as Reddit evolves and the overall terrain in which networked communities operate changes. People are just mad because he's airing reddit's dirty laundry in public. /r/jailbait has been a creepy embarrassment for some time, though it's an open question as to what (if anything) should be done about it. At a certain point in a site's growth it's going to have to choose between chaotic, free-form openness (a la 4chan) and all of the detritus that attracts (a la 4chan) or shed its nastier excesses for mainstream acceptability. Reports like this are simply byproducts of that friction” (Khiva, 2011).

While some debated (potentially unavoidable) outcomes, others focused on the means undertaken to effect a ban on /r/jailbait. Concerned about an apparent lack of transparency and consistency regarding content policies, /u/chanoop wrote:

“how are you banning /r/jailbait for Mod Drama but have ignored the Mod drama in many other subreddits (e.g: /r/Marijuana)? Reddit admins have always been VERY open about everything and it feels like this is being swept under the rug to never be spoken of again.

I don't visit /r/jailbait and am not really butt hurt about it being banned but this is a pretty big deviation from how reddit admins handles subreddits. I think other mods have a right to know "what not to do to get banned" (chanoop, 2012).

Still others objected to the implication that a moderator’s behavior could result in the disappearance of an entire subreddit. /u/patternfall wrote:

“while the situation itself involved specific individuals, the actions taken affect the entire community. And while control of subreddits belongs to the creators, I think that once a subreddit becomes large enough, there should be protections in place so the creator can't sabotage the entire thing out of spite.

I don't even care about r/jailbait specifically, but it upsets me that any of the subreddits I know and love can be destroyed simply because of a creator vs. admin pissing match. There's lots of good people I'm sure would be willing to keep things running if the current management becomes intransigent” (patternfall, 2012).

As these quote illustrate, there were substantial concerns not only around free speech and viewpoint censorship, but also issues of governance and decision- and policy-making process. The next section draws out some of the key observations from the /r/jailbait example and offers a new framework, based on participatory policymaking, to inform our understanding of the complex tradeoffs among values, ethics, policies, technical design elements and various stakeholders in an online social media community.

Discussion

The /r/jailbait controversy offers an entry point for beginning to think about the complex ways in which values, ethics, policies and technical design features interact, shape, and are shaped by participants of Internet platforms and social media sites (including creators, administrators, moderators, and users) and outside observers and commentators. In particular, this example raises a number of key observations about how different stakeholders navigate differently the various tradeoffs between shared values like freedom of expression, privacy, and dignity and principles related to governance such as fairness, transparency, and “getting a vote.” This discussion will highlight several observations and considerations drawn

from the /r/jailbait controversy and, by adapting a participatory policymaking approach, hopes to contribute to deeper understanding of the complex interplay among values, ethics, policy, and design and guide future explorations in this area.

Reddit can be understood as an example of Value-Sensitive Design in the wild. It also highlights the important differences between values and ethics. Arguably the core values of freedom of expression and democratic participation did not cause the /r/jailbait controversy or others like it but rather it was the lack of a set of principles for guiding sensemaking and decision-making around values tradeoffs that caused the most significant problems for the community. As any professor of ethics will tell you, the value of ethics is not that it tells you *what* to think but rather *how* to think, particularly in cases involve complex high-emotion tradeoffs between core values. In terms of Reddit's early design, its development was directed by a series of core values without the aid of a VSD specialist, Anticipatory ethicist, or information policy analyst. We can only speculate as to whether the contributions of these sorts of experts would have ameliorated some of the tensions and challenges that emerged as the Reddit community grew and norms and expectations emerged and evolved. Setting aside that speculation for a moment, we can draw a number of other lessons from the /r/jailbait controversy.

A first order effect of the /r/jailbait controversy is that it draws attention to potentially offensive content. As some redditors noted, Cooper's report may not have qualified as investigative journalism but, by shining a light on the dirty underbelly of Reddit it forced redditors to confront issues that had otherwise been largely "out of sight, out of mind." It raised serious and difficult to answer questions about the relationship between values, ethics, norms, policies, and laws, and it provoked discussion, debate, and engagement with the issues.

Drawing attention to /r/jailbait also produced a number of somewhat paradoxical follow-on effects. It increased viewpoint scrutiny which led some to marginalize and distance the content. It also increased curiosity and interest in the "morally questionable" content; /r/jailbait saw an uptick in traffic and participation following Cooper's report. Was this a fracturing of the Reddit community? Was this divergence of response a threat to the core value of "Reddit is its users"? Does traffic necessarily imply that those users are ratifying the viewpoint expressed in the content? The /r/jailbait controversy led some redditors and moderators to observe and question the tradeoffs between potentially competing values such as freedom of expression and dignity. It caused others to question the fairness of allowing a minority viewpoint to carry a disproportionate impact on the community. It made momentarily more tangible the often intangible processes by which we conceptualize and act upon questions of morality. And it positioned those processes in relation to other decision-making processes

informed by express policies and laws. In this way the discussions and debates engaged in by users (and perhaps the creation of competitor sites like Voat) reflect an emerging participatory policymaking process.

As the community struggled with sorting through those issues, another, arguably more serious, set of questions emerged around governance and the *process* of decision-making and policy-making on Reddit. These questions were prompted, in large part, by the Reddit administrators' decision to ban /r/jailbait and, in so doing, seriously undercut, if not abandon the core values some members of the community had come to trust in and rely upon. Despite its long-standing commitment to free speech principles, the decision to ban /r/jailbait struck many as viewpoint-based and, moreover, unsupported by Reddit's express content policies as they existing at the time. In addition, the voting mechanisms and moderator discretion which embedded the system with a democratic sense of fairness, equity, transparency and legitimacy were sharply undercut by the administrators' swift and unilateral action. If Reddit *is* its community then administrators suddenly dictating an outcome seemingly without consulting its constituency, its lifeblood, was a serious blow. Alternatively, administrators may have been responding to the needs and expectations of a changing user demographic that prefers a more nuanced take on freedom of expression by, for example, balancing it when necessary against other key human values, such as dignity, privacy, and liberty. Or, they may have been responding to external pressures by actively seeking to preference particular viewpoints (and users) over others (like /u/Violentacrez).

It certainly seems plausible as //u/Khiva said, that controversies like /r/jailbait are the byproducts of a social media site's growing pains or, borrowing from Zittain, of the transition from generativity to stability. A key point that this controversy around /r/jailbait suggests is that the outcome (i.e. banning content, becoming more "mainstream," etc.) may be less important than the process used to reach the outcome. This point touches both on the need for values *and ethics* and also careful attention to how policy is made on a social media site like Reddit.

We can learn from researchers investigating the role and impact of citizen participation in local government policymaking in brick and mortar context for some useful signposts. Researchers in that space have found that the primary role of citizen participation may be to provide information which the local government officials can then use to make decisions. But they also found numerous other positive spillover effects on participatory democracy: citizens feel more responsibility for matters affecting the public; it increases public engagement; it encourages people to listen to a diversity of opinions; and contributes to a higher degree of legitimacy of decisions (Michels & De Graf, 2010). Those authors concluded that, at the local level, ensuring aspects of democratic citizenship (by facilitating

processes and procedures facilitating the sense of meaningful participation) may be more important than having a direct say or vote in decisions (Michels & De Graf, 2010).

This research is echoed by others who emphasize the *policy process* -- “processes of making policy, of decision-making, and ways of putting issues on the agenda as matters of public concern, along with often rather intangible processes of the way issues are thought of and talked about” -- over particular policy statements or outcomes (Keeley, 2001, p. 5). Policy processes are “distinctly nonlinear, inherently political and contested, and more incremental and haphazard” than traditional linear policy-making procedures. In the real world, how policy is made depends on context. Models of policy process can include trial and error, debate and negotiation between stakeholders, small incremental changes to existing policies, and/or political struggle between interest groups (Sutton, 1999).

These findings resonate with many of the findings of the /r/jailbait example. Users emphasized concerns regarding a perceived breakdown in the process by which administrators evaluate and respond to offensive content and moderator missteps. The unilateral decision to ban /r/jailbait appeared to many to contradict the core values of the community: Reddit *is* its community, freedom of expression, and adherence to principles of democratic governance. The ban prompted users to articulate and engage with some of the intangible aspects of their membership and in some cases even seek alternatives such as Voat. Questions of morality, trust, fairness, transparency, control, autonomy were discussed and debated amongst the community. The community engaged with issues, participated in the policy process, and Reddit administrators responded, imperfectly but affirmatively.

Where a community is the product and the company cannot administer the site without the users, basic assumptions about process and control are necessarily called into question. Participatory policymaking in the case of Reddit may be an unavoidable consequence of the mutual operation of the community’s core values and the platform’s design.

Conclusion

This research provides an example of value-sensitive design in the wild. It demonstrates how values, ethics, and policies interrelate in a dynamic co-constitutive manner. Even when technologists attempt to promote certain values and anticipate ethics outcomes, the distributed collaborative nature of platforms like Reddit are largely beholden to the shared expectations and norms of the user community.

In addition, this research observes how policies pertaining to content and expectations around use and behavior do not emerge in a strictly linear, top-down fashion. Policymaking on collaborative Internet platforms like Reddit unfolds through a far more dynamic, participatory, and ad hoc

process prone to near-instantaneous revision and critique. As the /r/jailbait example illustrates, the lines between permissible and impermissible behavior, and assessment of the tradeoffs among core values, are prone to shift as platforms like Reddit evolve and the overall terrain in which networked communities operate changes. Participatory policymaking might contribute a new and useful perspective on research and practice in this area.

References

- Brey, P. A. (2012). Anticipatory ethics for emerging technologies. *NanoEthics*, 6(1), 1-13.
- Bynum, T. (2001). “Computer Ethics: Basic Concepts and Historical Overview” in the *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- Centivany, A. (2016, February). Policy as Embedded Generativity: A Case Study of the Emergence and Evolution of HathiTrust. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 926-940). ACM.
- Centivany, A. and Bobby Glushko (2016, May). ‘Popcorn tastes good’: Participatory policymaking and Reddit’s ‘AMAgeddon’. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI’16)*, San Jose, CA.
- chanoop, (2012) Response to “Dear reddit, why did /r/jailbait disappear?”
https://www.reddit.com/r/AskReddit/comments/jkmx7/dear_reddit_why_did_rjailbait_disappear/
- Chen, A. (2012). *Unmasking Reddit's Violentacrez, The Biggest Troll on the Web*. Gawker. Retrieved 9 September 2015, from <http://gawker.com/5950981/unmasking-reddits-violentacrez-the-biggest-troll-on-the-web>
- Cooper, A. (2011). Transcript from CNN’s Anderson Cooper 360, available at <http://edition.cnn.com/TRANSCRIPTS/1109/29/acd.01.html>
- Deleted user, mod of /r/politics. Original Post in “An Announcement on Gawker Links in /r/Politics” https://www.reddit.com/r/politics/comments/119z4z/an_announcement_about_gawker_links_in_rpolitics/
- drafterman (2014). Response to “What’s vote brigading, and why is it illegal?” on Reddit, https://www.reddit.com/r/OutOfTheLoop/comments/24d8cj/whats_vote_brigading_and_why_is_it_illegal/ch5zbfv
- Fiegerman, S. (2015) “Aliens in the Valley: The complete history of Reddit, the Internet’s front page,” *Mashable*, <http://mashable.com/2014/12/03/history-of-reddit/>.
- Flanagan, M., Howe, D. C., & Nissenbaum, H. (2005, April). Values at play: Design tradeoffs in socially-oriented game design. In *Proceedings of the SIGCHI*

- conference on human factors in computing systems (pp. 751-760). ACM.
- Flanagan, M., Howe, D. C., & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. *Information technology and moral philosophy*, 322-353.
- Friedman, B., Kahn Jr, P. H., Borning, A., & Hultgren, A. (2013). Value sensitive design and information systems. In *Early engagement and new technologies: Opening up the laboratory* (pp. 55-95). Springer Netherlands.
- Hill, K. (2012). *Reddit Co-Founder Alexis Ohanian's Rosy Outlook On The Future of Politics*. *Forbes*. Retrieved 10 September 2015, from <http://www.forbes.com/sites/kashmirhill/2012/02/02/reddit-co-founder-alexis-ohanians-rosy-outlook-on-the-future-of-politics/>
- hueypriest, 2011. Communication in "IAMA reddit General Manager, AMA" on Reddit https://www.reddit.com/r/IAmA/comments/iuz8a/iama_reddit_general_manager_ama/c26ukq8
- Jackson, Steven J., Tarleton Gillespie, and Sandy Payette. "The policy knot: Re-integrating policy, practice and design in CSCW studies of social computing." *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 2014.
- Johnson, D. G. (2003). Computer ethics. *A companion to applied ethics*, 608-619.
- Johnson, D. G. (2011). Software agents, anticipatory ethics, and accountability. In *The growing gap between emerging technologies and legal-ethical oversight* (pp. 61-76). Springer Netherlands.
- Keeley, J. E. (2001). Influencing Policy Processes for Sustainable Livelihoods: strategies for change Lessons for Change in Policy & Organisations, No. 2. *Brighton: Institute of Development Studies*, 4.
- Knobel, C., & Bowker, G. C. (2011). Values in design. *Communications of the ACM*, 54(7), 26-28.
- Khiva, (2011). Response to "Anderson Cooper just bashed Reddit for /r/jailbait. What does Reddit think of this?" on Reddit, https://www.reddit.com/r/AskReddit/comments/kvzx4/anderson_cooper_just_bashed_reddit_for_rjailbait/
- Lessig, L. (1999). *Code and other laws of cyberspace*. Basic books.
- Michels, A., & De Graaf, L. (2010). Examining citizen participation: Local participatory policy making and democracy. *Local Government Studies*, 36(4), 477-491.
- NightBane (2011), Comment on "Anderson Cooper just bashed Reddit for /r/jailbait. What does Reddit think of this?" on Reddit, https://www.reddit.com/r/AskReddit/comments/kvzx4/anderson_cooper_just_bashed_reddit_for_rjailbait/
- Ohanian, A. (2009). *blog.reddit -- what's new on reddit: You've decided: here are the 2008 reddit award winners*. *Redditblog.com*. Retrieved 9 September 2015, from <http://www.redditblog.com/2009/01/youve-decided-here-are-2008-reddit.html>
- patternfall, (2012) Response to "Dear reddit, why did /r/jailbait disappear?" https://www.reddit.com/r/AskReddit/comments/jkxm7/dear_reddit_why_did_rjailbait_disappear/
- Scary_The_Clown ***
- Sclove, R. (1995). *Democracy and technology*. Guilford Press.
- Shilton, K. (2014). This is an intervention: Foregrounding and operationalizing ethics during technology design. In *Emerging Pervasive Information and Communication Technologies (PICT)* (pp. 177-192). Springer Netherlands.
- Shilton, K. (2015). Anticipatory Ethics for a Future Internet: Analyzing Values During the Design of an Internet Infrastructure. *Science and engineering ethics*, 21(1), 1-18.
- Shilton, K. (2015). "That's Not An Architecture Problem!": Techniques and Challenges for Practicing Anticipatory Technology Ethics. *iConference 2015 Proceedings*.
- splattypus (2014). Response to "What is 'doxxing', and why is it used?" on Reddit, https://www.reddit.com/r/OutOfTheLoop/comments/2e04k/r/what_is_doxxing_and_why_is_it_used/cjurge9
- Sutton, R. (1999). *The policy process: an overview*. London: Overseas Development Institute.
- Tassi, P. (2015). *Facebook Didn't Kill Digg, Reddit Did*. *Forbes*. Retrieved 10 September 2015, from <http://www.forbes.com/sites/insertcoin/2012/07/13/facebook-didnt-kill-digg-reddit-did/>
- van den Hoven, J. (2005). Design for values and values for design. *Information age*, 4, 4-7.
- van den Hoven, J. (2008). Moral methodology and information technology. *The handbook of information and computer ethics*, 49.
- Wiener, N. (1988). *God and Golem, Inc. A comment on certain points where cybernetics impinges on religion*. Cambridge: MIT Press.
- Windwalker (2011). Response to on "Anderson Cooper just bashed Reddit for /r/jailbait. What does Reddit think of this?" on Reddit, https://www.reddit.com/r/AskReddit/comments/kvzx4/anderson_cooper_just_bashed_reddit_for_rjailbait/
- Winner, L. (1986). Do artifacts have politics. *The Whale and the Reactor. A Search for Limits in an Age of High Technology*, 19-39.
- Zittrain, J. L. (2006). The generative internet. *Harvard Law Review*, 1974-2040.
- Zittrain, J. (2008). *The future of the internet--and how to stop it*. Yale University Press.