

Stanford **Technology** Law Review

NOTE

A Regulatory Proposal for Digital Defamation: Conditioning § 230 Safe Harbor on the Provision of a Site “Rating”

CAITLIN HALL *

CITE AS: 2008 STAN. TECH. L. REV. N1

<http://stlr.stanford.edu/pdf/hall-digital-defamation.pdf>

I. INTRODUCTION

¶1

Whatever lip service we may pay to those spaces “immemorially . . . held in trust for the use of the public,”¹ the Internet is operatively the most important public forum ever created.² Its vast interconnectivity far more nearly approximates the prototypical “marketplace of ideas”³ than do warring politicians duking it out on the op-ed pages or, for that matter, in opposing briefs. However, the very features that make the internet fertile ground for cultural and political discourse— anonymity and pseudonymity; intellectual symbiosis and parasitism; fractal sprawl, audience dispersal and many-to-many architecture—render it a treacherous landscape for its custodians. In recognition of that fact,⁴ Congress in 1996 passed the Communications Decency Act, which nearly eliminated the liability that website administrators face for third-party generated content.⁵

* © 2008, Caitlin Hall, Yale Law School.

¹ *Hague v. Comm. for Indus. Org.*, 307 U.S. 496, 515 (1939).

² *See, e.g., Denver Area Educ. Telecomm Consortium, Inc. v. FCC*, 518 U.S. 727, 802-3 (1996) (Kennedy, J., dissenting) (“Minds are not changed in streets and parks as they once were. To an increasing degree, the more significant interchanges of ideas and shaping of public consciousness occur in mass and electronic media.”).

³ *See Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

⁴ Communications Decency Act, 47 U.S.C. § 230(a)(3) (1996).

⁵ The law provides in relevant part:

(c) Protection for “Good Samaritan” blocking and screening of offensive material

(1) *Treatment of publisher or speaker*

No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

(2) *Civil liability*

No provider or user of an interactive computer service shall be held liable on account of—

(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or

(B) any action taken to enable or make available to information content providers or others the technological means to restrict access to material described in paragraph (1).

¶2 However, Congress may not have adequately anticipated how badly people behave when they know no one is watching⁶—a lapse in judgment that has grown increasingly conspicuous as internet users have become more sophisticated and successful in their efforts to avoid identification. While statistics on the phenomenon are unavailable, it is apparent that defamation in particular has begun to occur online with alarming frequency.⁷ However, there is reason to believe that we should be less concerned with digital libel than its print and broadcast analogs. This article will explore the critical differences between libel in traditional and digital media frameworks, and propose a scheme for the regulation of online defamation that can accommodate the internet’s unique structure.

II. LIBEL IN PRINT AND BROADCAST MEDIA

A. History

¶3 Given the propensity for irreparable reputational harm when an individual is publicly maligned, the English courts began punishing defamatory statements as early as 1500.⁸ At first, the truth or falsity of such a statement was considered immaterial to a defamation action;⁹ however, by the mid-sixteenth century, some courts started to recognize a number of defenses to libel, truth among them:¹⁰

As libel law matured, however, it lost its simplicity. Various defenses—justifications for punishing libels—sprang up. Truth was one. Privilege was another. Fair comment was a third. Even the publishers of untrue defamations, in some jurisdictions, were protected if they honestly believed the material they published to be true. In other jurisdictions, however, a publisher could be severely punished for publishing false libels, regardless of motives, regardless of the status of the person about whom the material was published.¹¹

¶4 The English common law tradition of libel was imported to the colonies and largely incorporated by the American judiciary, even after the passage of the First Amendment.¹² However, contrary to the common law English tradition,¹³ the political climate of early nineteenth-century America entailed that political libel—speech affecting governmental, rather than private, reputation—frequently went unpunished.¹⁴ When it was penalized, it was generally via civil, rather than criminal, sanctions¹⁵:

Id. § 230(c) (emphasis added).

⁶ See, e.g., Ellen Nakashima, *Harsh Words Die Hard on the Web: Law Students Feel Lasting Effects of Anonymous Attacks*, WASH. POST, March 7, 2007, at A01.

⁷ See, e.g., Douglas W. Vick, Linda Macpherson, & Sarah Cooper, *Universities, Defamation, and the Internet*, 62 MOD. L. REV. 58, 58 & n.2 (1999).

⁸ J.H. BAKER, AN INTRODUCTION TO ENGLISH LEGAL HISTORY 247-48 (1971) (citing Milsom H.F.C.L. p. 334).

⁹ W. WALT HOPKINS, ACTUAL MALICE: TWENTY-FIVE YEARS AFTER *TIMES V. SULLIVAN* 1 (1989) (“To show that one had been defamed by some published insult was a simple process. The defamed person showed publication and defamation, won a judgment, and the matter was settled.” (citing W. PROSSER, THE LAW OF TORTS 790-92 (3rd ed. 1964)); see also LAURENCE H. ELDREDGE, THE LAW OF DEFAMATION §§ 63-64 (1978); Marc. A. Franklin, *The Origins and Constitutionality of Limitations on Truth as a Defense in Tort Law*, 16 STAN. L. REV. 789, 790-805 (1964).

¹⁰ BAKER, *supra* note 8, at 253.

¹¹ HOPKINS, *supra* note 9, at 1 (citing Annot., 150 A.L.R. 358 (1944)).

¹² See, e.g., DONALD M. GILLMOR, POWER, PUBLICITY, AND THE ABUSE OF LIBEL LAW 50 (1992) (“Unable to accept the notion that the First Amendment had abolished preexisting liability at common law, especially where the reputations of public figures were concerned, almost all courts continued to adhere to the strict liability standards of English common law.”).

¹³ *Id.* at 73 (citing *De Libellis Famosis*, 77 Eng. Rep. 250 (1606), in which it was noted that “an offense against a private person should be punished because it could lead to a breach of the peace. An offense against a public official was far more serious because it was a scandal on the government.”)

¹⁴ *Id.* at 47.

¹⁵ *Id.* at 57 (noting that by the Jeffersonian administration, “[p]unitives damages in civil suits were replacing criminal indictments,” though they were at first no less burdensome than criminal sanctions: “In fact, money damages were just as effective in quelling criticism.”). Significantly, the United States twice flirted briefly and unpopularity with the notion of imposing criminal penalties for libel of the government. The first such incident, the so-called Sedition Act of 1798, made it a crime to publish false, defamatory speech about the nation’s law, president, or Congress. An Act for the Punishment of Certain Crimes against the United States, Ch. 74, 1 Stat. 596 (1798). See also GILLMOR, *supra* note 11, at 51-54. As judicial review was not established until *Marbury v. Madison*, 5 U.S. 137 (1803), the constitutionality of the act was never tested in court, and the law was allowed to expire peacefully in 1801. However, the second such act, the Sedition Act of 1918, passed under Woodrow Wilson during World War I, was upheld by the Court in *Schneck v. United States*, 249 U.S. 47 (1919). The act was far more zealous in its reach than its predecessor:

Whoever, when the United States is at war, shall willfully make or convey false reports or false

The early years of the republic clearly were accompanied by an acceptance of—and an eagerness for—political polemic, part of what some called the “politics of controversy.” Readers did not take libels literally. Unrestrained public debate was part of the price to be paid for liberty. And it was entertaining.¹⁶

¶5

American libel law’s focus on compensating harm to private persons was well-tailored to suits arising from traditional media because their tightly constrained markets created the propensity for enormous and irreparable reputational damage. In the case of print media, that constraint was the product of relatively high production costs and a finite readership. This meant that only a very small number of magazines and an even smaller number of newspapers, in one geographical area or on one subject, could be sustained. In the case of radio and television, an additional constraint was the product of the limited broadcast spectrum, a structural restriction of no small importance for libel law.¹⁷ Even so, the restriction became increasingly irrelevant with the proliferation of cable and digital television and satellite television and radio.¹⁸ In whatever form, a small number of large, powerful conglomerates with exclusive access to publication technologies and their large and typically dispersed audiences have dominated traditional media markets because of those restrictions. Because such media have the power to substantially influence a large audience otherwise difficult to locate, they have the power to do irrevocable harm to individuals’ reputations, and those who suffer such injury have little recourse to self-help (*see* discussion, *infra* II(B)), even when the information used to defame them is untrue. As a result, in a traditional media framework, the courts decided that the compensation of actual damages was the best proxy for repairing the harm done by libelous speech.

B. *Actual malice*

¶6

The courts also recognized, however, that in attempting to curb damage to individuals’ reputations, they risked unwittingly shutting down constitutionally protected speech as well.¹⁹ Any punishment for libel had the propensity to have a chilling effect on speech because it might prompt publishers to avoid responsibility for the consequences of its publication by preemptively eliminate controversial material.²⁰ Thus, in *New York Times Co. v. Sullivan*, the Court attempted to broker a peace between the First Amendment and contemporary libel law by instituting the “actual malice”

statements with intent to interfere with the operation or success of the military or naval forces of the United States, or to promote the success of its enemies, or shall willfully make or convey false reports or false statements, or say or do anything except by way of bona fide and not disloyal advice to an investor or investors, with intent to obstruct the sale by the United States of bonds or other securities of the United States or the making of loans by or to the United States, and whoever when the United States is at war, shall willfully cause or attempt to cause, or incite or attempt to incite, insubordination, disloyalty, mutiny, or refusal of duty, in the military or naval forces of the United States, or shall willfully utter, print, write or publish any disloyal, profane, scurrilous, or abusive language about the form of government of the United States or the Constitution of the United States, or the military or naval forces of the United States, or the flag of the United States, or the uniform of the Army or Navy of the United States into contempt, scorn, contumely, or disrepute, or shall willfully utter, print, write, or publish any language intended to incite, provoke, or encourage resistance to the United States, or to promote the cause of its enemies, or shall willfully display the flag of any foreign enemy, or shall willfully by utterance, writing, printing or products, necessary or essential to the prosecution of the war in which the United States may be engaged, with intent by such curtailment to cripple or hinder the United States in the prosecution of war, and whoever shall willfully advocate, teach, defend, or suggest the doing of any of the acts or things in this section enumerated, and whoever shall by word or act support or favor the cause of any country with which the United States is at war or by word or act oppose the cause of the United States therein, shall be punished by a fine of not more than \$10,000 or the imprisonment for not more than twenty years, or both ...

Sedition Act of 1918, 40 Stat. 553-54 (1918) (codified in scattered sections of 18 U.S.C.).

¹⁶ GILLMOR, *supra* note 12, at 60.

¹⁷ *See generally* Red Lion Broad. Co. v. FCC, 395 U.S. 367, 388-89 (1969) (“Where there are substantially more individuals who want to broadcast than there are frequencies to allocate, it is idle to posit an unabridgeable First Amendment right to broadcast comparable to the right of every individual to speak, write, or publish.”)

¹⁸ *But see* Harvey Jassem, *Different Strokes for Different Folks: The Intersection of Regulatory Principles and Technology*, in REAL LAW @ VIRTUAL SPACE: REGULATION IN CYBERSPACE 31, 43 (Susan J. Drucker & Gary Gumpert eds., 1999) (“The tenets of the distinction, that broadcasters rely on a limited public resource, the electromagnetic spectrum, have been repeatedly challenged, but the distinction continues to stand, even in somewhat modified form.”).

¹⁹ *See generally* HOPKINS, *supra* note 9, at 3; *see also* Harry Kalven, *The New York Times Case: A Note on The Central Meaning of the First Amendment*, 1964 SUP. CT. REV. 191, 197 (1964).

²⁰ *See* N.Y. Times Co. v. Sullivan, 376 U.S. 254, 278 (1964) (“Whether or not a newspaper can survive a succession of such judgments, the pall of fear and timidity imposed upon those who would give voice to public criticism is an atmosphere in which the First Amendment freedoms cannot survive.”).

standard—the requirement that, in order to be held liable in a defamation action brought by a public official, the publisher of a false and injurious statement must have published it with knowledge of or in reckless disregard of its falsity.²¹

¶7 The Court premised its distinction between public officials and private individuals first on an access differential between the two types of plaintiffs. Because private individuals largely lacked the ability to directly rebut false and defamatory statements made about them,²² self-help could not be considered an appropriate remedy for such plaintiffs. By contrast, self-help was far more frequently a viable option for defamed public officials, because as persons in positions of relative power, they often had either direct access to the media or the persuasive power necessary to obtain such access. Furthermore, as the Court made clear in a subsequent case, the distinction between defamation of public and private individuals was further justified by public individuals’ willingness to engage in public disagreement, as evidenced and necessitated by their decision to enter civic life.²³ By deciding to participate in public affairs, the Court reasoned, officials knowingly assumed the risk that they would be criticized brashly, and sometimes unfairly; as a result, they had less right to complain when they were maligned.

¶8 In addition to simply reducing the number of successful libel suits, the distinction between public and private figures addressed the problem of chilling effects in two ways: first, it made it difficult for plaintiffs to recover on the basis of politically motivated statements (those pertaining to official conduct), which meant that it would tend to lift publisher-imposed constraints on political speech. This change directly comported with the political expression principle of the First Amendment. Additionally, coupled with the simultaneous barring of punitive damages for public officials,²⁴ it eliminated many of the most powerful potential plaintiffs—those with the potential to run roughshod over publishers by way of huge defamation suits—and thereby eliminated much of the fear of catastrophic loss that drove publishers to censor their own publications.²⁵

¶9 While the Court at first applied the actual malice standard only to public officials, it quickly extended it to “public figures,”²⁶ those persons whose non-governmental positions of power rendered them practically indistinguishable from public officials with respect to their access to the media and, more importantly, the public’s legitimate interest in their affairs. Chief Justice Earl Warren explained the doctrinal expansion in his concurrence in *Curtis Pub. Co. v. Butts*:

[B]lending of positions and power has also occurred in the case of individuals so that many who do not hold public office at the moment are nevertheless intimately involved in the resolution of important public questions or, by reason of their fame, shape events in areas of concern to society at large. Viewed in this context, then, it is plain that although they are not subject to the restraints of the political process, “public figures,” like “public officials,” often play an influential role in ordering society. And surely as a class these “public figures” have as ready access as “public officials” to mass media of communication, both to influence policy and to counter criticism of their views and activities. Our citizenry has a legitimate and substantial interest in the conduct of such persons, and freedom of the press to engage in uninhibited debate about their involvement in public issues and events is as crucial as it is in the case of “public officials.” The fact that they are not amenable to the restraints of the political process only underscores the legitimate and substantial nature of the interest, since

²¹ *Sullivan*, 376 U.S. at 279.

²² The Court definitively closed off one avenue to self-help for private individuals when it held that guaranteeing a victim of libel a right of reply in the publication in which the original defamation appeared was constitutionally impermissible because it was operatively similar to prior restraint; in media with limited resources—in the case of print media, space, and in that of broadcast media, time—commanding a publisher to publish something had nearly the same effect as preventing it from publishing something. *Miami Herald Pub. Co. v. Tornillo*, 418 U.S. 241, 258 (1974).

²³ *Gertz v. Robert Welsh, Inc.*, 418 U.S. 323, 338 (1974); see also *Coleman v. MacLennan*, 98 P. 281, 291 (1908) (“a candidate must surrender to public scrutiny . . . so much of his private character as affects his fitness for office.”); *Commonwealth v. Clapp*, 4 Mass. 163, 169-79 (1808).

²⁴ *Gertz*, 418 U.S. at 349.

²⁵ See Thomas A. Schwartz and Tracie L. Maruiello, *Journalistic Behavior and the Law of Libel: Legal and Ethical Controls*, in *EMERGING ISSUES IN CONTEMPORARY JOURNALISM* 289, 289 (Bala A. Musa & Cindy J. Price eds., 2006) (“In effect, the Court created the press’ constitutional mission to report vigorously on public affairs by crippling the ability of powerful people or groups to successfully sue the press for libel.”).

²⁶ *Gertz*, 418 U.S. at 351.

it means that public opinion may be the only instrument by which society can attempt to influence their conduct.²⁷

¶10 The Court thus extended the actual malice standard to general public figures—leaders of business and culture, as well as limited-purpose public figures—who operate as private figures, except in the context of a particular controversy in which they are voluntarily involved.²⁸ In carefully calibrating the distinction between public and private in this way, the Court made clear that it was concerned not with singling out those who have traditionally been considered part of governmental processes, but also those who willingly involve themselves in media affairs in the same manner as do public officials.

III. LIBEL IN DIGITAL MEDIA

A. *New problems*

¶11 In many ways, victims of digital libel are far better situated than their traditional counterparts. First, near-infinite information capacity²⁹ and low transaction and market entry costs mean that online content is generated at a phenomenal rate. That proliferation contributes to a simultaneous and proportionate devaluation of content on all but the most widely viewed sites.³⁰ It also forces people to invent or adopt mechanisms for sorting credible from questionable information.³¹ As a natural consequence, audiences tend to converge on consensus sources,³² limiting the depth and dispersion of defamatory harm.³³

1. *Defendants*

¶12 However, though the magnitude of harm resulting from libel has become less worrisome of late, the likelihood of being able to hold someone accountable for it has simultaneously plummeted. Section 230 of the Communications Decency Act provides that “[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider”³⁴—essentially, that those who own and run sites and servers containing defamatory content cannot be held liable for its publication, provided that it has a third-party source. Moreover, absent a judicial determination that a site administrator has interacted with offending content to an egregious extent, Section 230 applies even in cases in which the owner or administrator knows of the defamatory content.³⁵ Of course, were Section 230 plainly irresponsible law, Congress could simply repeal it. However, there are important information policy concerns that counsel restraint when debating punishment for those who provide forums for defamatory speech (*see* discussion, *infra* III(B)(2)).

¶13 While the barriers to owner and administrator liability are legal, those to speaker liability are technological. Much of what is written on the internet is anonymous or pseudonymous. That makes identifying the creators of defamatory speech a daunting task, and one that grows more difficult with each successive generation of IP masking technology.³⁶ As a result, a sophisticated internet user can

²⁷ *Curtis Pub. Co. v. Butts*, 388 U.S. 130, 163-64 (1967) (Warren, C.J., concurring).

²⁸ *Gertz*, 418 U.S. at 338.

²⁹ While the internet can currently accommodate 4.3 billion unique universally scoped web addresses, that number will increase 3.4×10^{38} with the impending IPv6 rollout. Simson Garfinkel, *Internet 6.0*, *TECH. REV.*, Jan. 7, 2004, http://www.technologyreview.com/printer_friendly_article.aspx?id=13426.

³⁰ Power Laws, Weblogs, and Inequality, http://shirky.com/writings/powerlaw_weblog.html, (Feb. 8, 2003).

³¹ *See, e.g.*, S. Brin & L. Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, 30 *WWW7 / COMPUTER NETWORKS* 107-117 (1998) (proposing the creation of a search engine, Google).

³² *See* Clive Thompson, *Blogs to Riches: The Haves and Have-Nots of the Blogging Room*, *NEW YORK MAGAZINE*, Feb. 20, 2006, available at <http://nymag.com/news/media/15967/>; *but see* Shirky, *supra* note 10, *The Median Cannot Hold*.

³³ *But see* Matthew Collins, *THE LAW OF DEFAMATION AND THE INTERNET 1.02* (2001) (“The Internet also has the potential, however, to increase dramatically the incidence of multi-jurisdictional defamation. Defamatory material published via the Internet may have an audience of thousands, or millions. That material might have a devastating impact on the defamed person’s reputation and give rise to a justifiable desire for vindication.”).

³⁴ Communications Decency Act § 230(c)(1).

³⁵ Collins, *supra* note 33, 27.54.

³⁶ *See generally* Michael G. Reed, Paul F. Syverson, & David M. Goldschlag, Naval Research Laboratory, *Anonymous Connections and Onion Routing*, 16 *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS* 482, 482-84 (1998) (describing onion routing, one form of IP concealment); Anonymizer, Wikipedia (Feb. 14, 2008), <http://en.wikipedia.org/wiki/Anonymizer> (describing a

avoid identification in all but the most extreme circumstances.³⁷ The problem is further exacerbated by site administrators who enable user anonymity by deleting or disabling IP logging. Matthew Collins eloquently sums up the difficulty in tracking down digital audiences, and, analogously, speakers:

Before the Internet, communications emanated from determinate locations, and had audiences which were capable of being defined in geographic terms. Even mass communications, such as television and radio broadcasts, have a defined and geographically limited audience reach. By contrast, communication on the Internet is characterized by the transfer of signals from computers in indeterminate locations, to other computers in indeterminate locations, via routes which are indeterminate.³⁸

Given, then, how difficult it is to hold either site administrators or users accountable for defamatory speech online, it’s not surprising that malicious speech has proliferated wildly.

2. Plaintiffs

¶14 Moreover, in the unique landscape of the Internet, it is becoming harder to decide what standard of liability should apply to a particular plaintiff. This difficulty is largely a product of the way in which the structure of the Internet differs from that of traditional media. Because production, distribution, transaction and market entry costs for Internet media are exceedingly low (in many cases zero), anyone who wishes to can publish online. As a result, the access differential that served as one half of the traditional basis for the distinction between public and private figures³⁹ does not inhere in new media formats: not only do both types of plaintiffs—if such a distinction can even be drawn—have access to the relevant publication technologies, they also have access to the audiences of those technologies, and hence to the audiences relevant to the libelous speech they might seek to rebut.

¶15 Of course, the audience which a victim of libel will be able to access will not always be identical to that of the original defamatory speech; the proximity of the audiences will depend on a number of factors, such as a Web site’s moderative policies, as discussed in section 2(B)(3), the time lapse between the original and the rebutting speech, and audience transience and size (as one might expect sites with larger audiences to also have a larger proportion of sporadic visitors.)⁴⁰ Moreover, even if audience reproduction were far more perfect, “an opportunity for rebuttal seldom [alone] suffices to undo harm of defamatory falsehood.”⁴¹ Nonetheless, the approximate, *de facto* right of reply created by open, virtually universal⁴² access to Internet publication technologies is a vast improvement over the lack of access common to traditional media. Libel law certainly would have developed along a widely different trajectory had victims of traditional defamation been assured such a right of response.

¶16 The disappearance of the access differential between public and private figures (and, as a corollary, between victims of libel and their defamers) is not the only possible justification for the elimination of the distinction between the two. In a digital environment, it is also difficult to apply the “willingness” standard that served as the second justification for the division in *Gertz v. Robert Welsh, Inc.*⁴³ The willingness criterion was already undermined to some degree by the extension of the actual malice standard to public figures; rather than maintain a relatively bright line for its application based on a plaintiff’s choice to enter the political arena, the Court instituted a blurry line somewhere along the spectrums of success and visibility beyond which private persons are subjected to the same

class of IP masking programs).

³⁷ See Collins, *supra* note 33, at 2.25. *But see* Marisa Lagos, *UC Berkeley law student made threat against Hastings, officials say*, SAN FRANCISCO CHRON., April 20, 2007 (describing the FBI investigation and subsequent arrest of a pseudonymous poster on a pre-law website, AutoAdmit, <http://www.autoadmit.com>, after he joked that he intended to gun down U.C. Hastings students in the wake of the April 2007 Virginia Tech. shootings), available at <http://sfgate.com/cgi-bin/article.cgi?f=/c/a/2007/04/20/BAGEHPCJCP6.DTL>.

³⁸ Collins, *supra* note 33, at 3.02.

³⁹ *Gertz v. Robert Welsh, Inc.*, 418 U.S. 323, 338 (1974).

⁴⁰ See generally Shirky, *supra* note 30.

⁴¹ *Gertz*, 418 U.S. at 344.

⁴² See *The Harris Poll: Poll Shows More U.S. Adults Are Going Online at Home*, THE WALL STREET JOURNAL ONLINE, May 24, 2006, http://online.wsj.com/public/article/SB114840389678260791-IREjYVgN_rGLcE3_6Djin1jeJZc_20070523.html?mod=rss_free.

⁴³ *Gertz*, 418 U.S. at 338-39.

treatment as public officials. That evolution acknowledges a shift in our national conception of “valuable” speech⁴⁴ which reflects the growing importance of different types of governance—not only political, but informational, technological, and financial—and the difficulty in delineating what can properly and non-arbitrarily be called political or governmental speech;⁴⁵ it less overtly also calls into question our particular allegiance to “political” speech, as opposed to non-explicitly governmental cultural speech.

¶17 The imprecision in deciding who has willingly entered public discourse—as well as in defining what “public discourse” is, exactly—is exacerbated by the nature of online communities: the rapid proliferation of niche groups, each with its own experts, entails that an increasing proportion of Internet users could likely be properly treated as limited purpose public figures. At some point soon, if not already, there will be enough such users, and the rules for distinguishing them from “ordinary” private users will be sufficiently hard to select, that drawing the line between public and private figures will become wholly impractical.

¶18

3. Chilling effects

¶19 Even if we could reliably discern what standard to apply to Internet plaintiffs and surmount the technological and legal barriers to finding suitable defendants, Internet defamation would still present another problem: the potentially disastrous chilling effects of libel actions. The digital age, and its open invitation to publish, presents not only an entirely novel opportunity for the creation and dissemination of content by vast multitudes who were never before able to do so, but also options for forms of speech that have never before been available. There is no real non-digital analog to social software, for instance—and it would be hard to overestimate the cultural significance of, say, YouTube or MySpace.⁴⁶

¶20 More germane, however, are the opportunities afforded by anonymous speech, given that it is uniquely relevant to a discussion of online libel. Prior to the internet, with few exceptions,⁴⁷ anonymous conversations were nearly impossible; there was no technological infrastructure to enable such contact between two people, let alone networks of people carrying on interlocking conversations.⁴⁸ Now there are populous communities in which people can, unidentified, solicit advice,⁴⁹ make friends, and—for lack of a better phrase—form identities.⁵⁰ Even First Amendment reductionists must appreciate the expansion of anonymous speech: the value of such speech has

⁴⁴ See *Curtis Pub. Co. v. Butts*, 388 U.S. 130, 163-64 (Warren, C.J., concurring) (providing an early indication of this shift).

⁴⁵ See Gillmor, *supra* note 12, at 28-29 (“There is another kind of libel action that falls just short of a resurrected crime of sedition. That is the suit brought by what the law calls the pervasive or all-purpose public figure, or public personality—in lay language, the celebrity, or to be cute, the ‘glitterati.’ Here, the plaintiff may or may not intend to influence political policy but, with an access to mass media outdistancing that of many public officials, celebrities influence that policy nonetheless by *helping to fashion our common cultural context or a style that catches the public imagination.*” (emphasis added)).

⁴⁶ See generally Jack M. Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1 (2004).

⁴⁷ Examples that spring to mind include suicide and tipster hotlines—cases in which conversations could not take place without the benefit of an anonymous format.

⁴⁸ See, e.g., Law School Discussion, <http://www.lawschooldiscussion.org> (last visited Mar. 18, 2008) an Internet discussion-based forum for law school applicants. On the board, anonymous and pseudonymous “regulars” generate most of the content, carrying on conversations at times for months, through thousands of individual posts and multiple threads; forming friendships and cliques; and developing distinctive reputations. See also Slashdot, <http://slashdot.org> (last visited Mar. 18, 2008), a more popular technology news site with a similar structure.

⁴⁹ In March 2007, in response to defamatory attacks and threats aimed at some of its female students and originating from AutoAdmit, *supra* note 37, members of the Yale Law School community proposed attempting to uncover the identities of the site’s law school users. The importance of anonymous advice-seeking was adeptly communicated by one such user in an open letter to the law school posted on AutoAdmit: “I’m a Yale Law student who posts regularly on [AutoAdmit]. YLS students might be surprised to find that many of their classmates post here, and that many—and hopefully most—of us have never ranked our classmates, posted racist comments, or threatened anyone. Still, we post with the expectation that our comments will never be traced back to us. We ask medical questions, complain about partners, criticize professors, and rant about politics. None of this content is hurtful, but all of it is made with the expectation of privacy and should remain private.” YLSStudent, An Open Note to Yale Law Students, AutoAdmit (March 9, 2007), http://www.autoadmit.com/thread.php?thread_id=592810&mc=38&forum_id=2.

⁵⁰ See Danah Boyd, *Identity Production in a Networked Culture: Why Youth Heart MySpace*, American Association for the Advancement of Science (speech) (Feb. 19, 2006), available at <http://www.danah.org/papers/AAAS2006.html>; Danah Boyd, *Friends, Friendsters, and Top 8: Writing community into being on social network sites*, 11 FIRST MONDAY (Dec. 2006), available at http://firstmonday.org/issues/issue11_12/boyd/index.html.

figured prominently in our nation’s history since its inception⁵¹ and continues to visibly do so in the digital age.

B. New solutions

¶21

If we hope to preserve the Internet’s full communicative and generative capacity, a regulatory infrastructure designed to address digital libel will need to incorporate two features. First, in order to maintain the Web’s structural diversity, it must manifest a great deal of deference to site administrators’ authority to create code according to the community norms they hope to propagate.⁵² Second, in order to maintain the Web’s substantive diversity, its provisions must be procedural, rather than categorical—even within the indiscriminating bounds of unprotected speech. Cass Sunstein makes a compelling case for the preservation of that substantive diversity on two grounds. First, heterogeneity such as might be found on a street corner or in another traditional public forum is necessary for the continuing evolution of deliberative American democracy, and cannot be maintained if Internet content is tailored too narrowly to what people think they want to see: “[P]eople should be exposed to materials that they would not have chosen in advance. Unanticipated encounters, involving topics and points of view that people have not sought out and perhaps find quite irritating, are central to democracy and even to freedom itself.”⁵³ Second, Sunstein treats as dangerous the possibility that, in the absence of some measure of forced exposure to varied viewpoints, group polarization effects will lead individuals to ever more distorted—and in many cases, dangerous—viewpoints:

[W]hen group discussion tends to lead people to more strongly held versions of the same view with which they began, and if social influences and limited argument pools are responsible, there is legitimate reason for concern. Consider discussions among hate groups on the Internet and elsewhere. If the underlying views are unreasonable, it makes sense to fear that these discussions may fuel increasing hatred and a socially corrosive form of extremism. This does not mean that the discussion can or should be regulated in a system dedicated to freedom of speech. But as we have seen, it does raise questions about the idea that “more speech” is necessarily an adequate remedy—especially if people are increasingly able to wall themselves off from competing views. The basic issue here is whether something like a “public sphere,” with a wide range of voices, might not have significant advantages over a system in which isolated consumer choices produce a highly fragmented speech market. The most reasonable conclusion is that it is extremely important to ensure that people are exposed to views other than those with which they currently agree, in order to protect against the harmful effects of group polarization on individual thinking and on social cohesion.⁵⁴

If, as Sunstein argues, it is important to make sure internet users are exposed to different niches of information, it is certainly necessary to maintain a wide array of content online.

¶22

Even Robert Post, who views the development of modern defamation law as the process of codifying expectations of civility between citizens,⁵⁵ has recognized that it has been the policy of the courts to tend away from judicially enforcing these “civility rules” wherever possible:

[*Hustler Magazine v. Falwell*, 485 U.S. 46 (1988)] illustrates the depth of the Court’s commitment to preserving the neutrality of public discourse from the imposition of such norms. Although the Court assumed that the *Hustler* parody would “doubtless [be] gross and repugnant in the eyes of most,” it nevertheless refused to permit the parody to be

⁵¹ See, e.g., Federalist Papers, Wikipedia, http://en.wikipedia.org/wiki/Federalist_Papers (last visited Mar. 18, 2008).

⁵² Yochai Benkler, *There Is No Spoon*, in *THE STATE OF PLAY: LAW, GAMES, AND VIRTUAL WORLDS* 180, 186 (Jack M. Balkin & Beth Simone Noveck eds., 2006).

⁵³ Cass R. Sunstein, *The Future of Free Speech*, in *ETERNALLY VIGILANT: FREE SPEECH IN THE MODERN ERA* 285, 285 (Lee C. Bollinger & Geoffrey R. Stone eds., 2002).

⁵⁴ *Id.* at 298-99.

⁵⁵ Robert Post, *CONSTITUTIONAL DOMAINS* 128-29 (1995) (“Defamatory communications may be defined as those whose content is not civil, because their meaning violates the respect which we have come to expect from each other. They thus threaten not only the self of the defamed person . . . but also the continued validity of the rules of civility which have been violated The common law’s regulation of defamation contains numerous features that attempt to preserve the integrity of these rules of civility.” (citing Robert Post, *The Social Foundations of Defamation Law: Reputation and the Constitution*, 74 CAL. L. REV. 691, 711-15, 735-39 (1986))).

penalized. This result comports with the reasoning of *Cantwell*: if public discourse is constitutionally protected because it is the medium for the formation of future communities, its structural independence from all civility rules must be guaranteed, even if such rules in fact are accepted by every contemporary community. The “marketplace of communities” must be understood as extending in time as well as in space. The individualist methodology of First Amendment doctrine ultimately means that individuals must be free within public discourse from the enforcement of all civility rules, so as to be able to advocate and to exemplify the creation of new forms of communal life in their speech.⁵⁶

This line of thought—the emphasis on the importance of preserving the avenues of expression for even the most objectionable viewpoints—figured prominently in *Sullivan*⁵⁷ and its predecessors.⁵⁸ In order to ensure such an unrestricted environment, a response to the problem of digital libel should refrain from categorically barring particular types of content to the greatest extent possible.

1. Do nothing

¶23

Several commentators have seized on the need for a peculiar, medium-specific framework by advancing both legal and extra-legal proposals for the management of digital defamation. The simplest approach might in fact be viewed as the absence of management: the maintenance or expansion of current legal and technological barriers to prosecution. Declining to hold anyone liable for digital defamation would almost certainly accelerate its creation, and thereby speed its devaluation: “having reached some critical mass, it will be reduced to background noise and will no longer have the damning effect of singling out specific [victims].”⁵⁹ While this solution certainly has the advantage of ease,⁶⁰ it suffers from critical flaws: Short of a further expansion of liability exemption, it would do little but impotently urge victims of defamation to refrain from pursuing John Doe actions⁶¹—those that attempt to retroactively uncover the source of libelous speech—against anonymous content creators. An extension of immunity from civil liability to speakers, analogous to that granted to administrators by Section 230, would aid the process, but would for obvious reasons be a prospect unlikely to enjoy popular support. More importantly, as a matter of information policy, the value of flooding the Internet with bad information pales in comparison to that of a system capable of distinguishing good information from bad and thereby safeguarding the value of the former.

2. Solutions based on traditional libel law

¶24

As an alternative to letting defamation run its course, we might look to traditional libel law for ideas. For instance, we might seek to hold Web site administrators and owners responsible as publishers for their site content by repealing Section 230.⁶² Doing so would quickly and dramatically decrease the amount of defamatory content online; while it might be difficult to track down

⁵⁶ *Id.* at 151.

⁵⁷ See *New York Times Co. v. Sullivan*, 376 U.S. 254, 279 (1964) (“[e]ven a false statement may be deemed to make a valuable contribution to public debate”). *But see* *Gertz v. Robert Welsh, Inc.*, 418 U.S. 323, 340 (1974) (“Neither the intentional lie nor the careless error materially advances society’s interest in ‘uninhibited, robust, and wide-open’ debate on public issues. They belong in that category of utterances which ‘are no essential part of any exposition of ideas, and are of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality.’”) (citing *Sullivan*, 376 U.S. at 270; *Chaplinsky v. New Hampshire*, 315 U.S. 568, 572 (1942)).

⁵⁸ See *NAACP v. Button*, 371 U.S. 415, 444-45 (1963) (“[f]or the Constitution protects expression and association without regard . . . to the truth, popularity, or social utility of the ideas and beliefs which are offered”).

⁵⁹ See Caitlin Hall, *Swimming Downstream: Battling Defamatory Online Content Via Acquiescence*, 19 *YALE J.L. & FEMINISM* 287, 289 (2007)

⁶⁰ *Id.* (“This solution has the attractive feature of shifting the burden from those who seek to have their reputations protected to those who seek to destroy them. In other words, it renders attempts at online defamation self-defeating, and takes advantage of the inevitability of malice in unmoderated online spaces.”).

⁶¹ See generally Collins, *supra* note 33, at 27.10; John Schwartz, *Corporate Case in Ohio Raises Questions on Net Anonymity*, *N.Y. TIMES*, October 17, 2000, at C1 (noting the impact of defamation lawsuits on internet anonymity), available at <http://www.nytimes.com/2000/10/17/technology/17ONLI.html?ex=1180843200&en=336922f7bcd956f&ei=5070>.

⁶² For an example of such a proposal, see Robert T. Langdon, *Communications Decency Act Sec 230: Make sense? Or nonsense?—A private persons’ inability to recover if defamed in cyberspace*, 73 *ST. JOHN’S L. REV.* 829, 854 (1999). Langdon, however, does not foresee such dire consequences for administrators; he evidently anticipates they would be treated in much the same way as distributors, rather than publishers, of traditional media. See *id.* at 855 (“[I]here is little reason to believe that Congress would interfere with the vibrant and competitive free market that presently exists” for Internet providers by holding them liable in certain defamation actions. After all . . . we are yet to see bookstores, libraries, and newsstands closing up shop because of their potential liability in a defamation action.”) (citations omitted).

individual Internet users, the same is not true of website owners. However, such an effective solution would come at a high cost. Faced with liability, site administrators and owners would be required, at a minimum, to begin tightly moderating their content. More frequently, they would be forced to shut down entirely to avoid the time and expense of moderation. It seems unlikely that most website administrators have the resources to exercise responsible prior restraint in the same manner as newspaper editors, especially given the volume of prospective content; moreover, it seems unfair to ask them to do so. Even if administrators could reasonably be asked to act as strict information filters, placing such a burden on them would severely constrict the breadth of forum structures available online—it would necessitate, for example, serious limitations on anonymous and pseudonymous content—and would disappointingly reduce digital media to near-replicas of print and broadcast. A better solution would leave site administrators free to fashion their sites’ content rules based on the cultural values and social norms they wish to promote,⁶³ though such freedom would have the occasional consequence of allowing site users to disseminate uncorroborated information and outright lies.

¶25 A less unfair solution, though one just as critically flawed, would be to require site administrators to keep logs of the IP addresses of users by provisioning civil liability immunity on an agreement to maintain them. Keeping IP logs is not a technologically daunting task, but some administrators decline to do so anyway, preferring to proactively protect their users’ anonymity. It might seem that requiring those administrators to log IP addresses would make identifying their site users an easier task. However, there’s a problem with this reasoning. In most cases, even if site administrators agreed to keep logs of IP addresses, and to provide them on request, that would not necessarily make it possible to track down the identities of individual users, as discussed above in Part III(A)(1). Only in a limited number of cases would logging help civil litigants track down their defamers, and the proportion of cases in which it would help is sure to grow smaller as IP obscuring technologies grow more advanced.

3. Solutions based on the structure of digital media

¶26 Another plan, inspired by the blogger.org community guidelines⁶⁴ and refined by Tim O’Reilly, calls for the creation of a “blogger’s code of conduct”⁶⁵ that licenses the deletion of abusive, threatening, libelous and invasive content. The code was devised for weblog owners, but is amenable to a wide array of forum structures. This solution is well-intentioned, but naïve. Bloggers have always been free to adopt such codes, and the fact that so few have demonstrates not an ignorance of their rights as moderators so much as a reluctance to codify and enforce community standards. Moreover, any solution that leans on administrators’ goodwill begs the question of the civility of the environment, and will thus fail to capture the most troublesome speech—that which is maliciously promulgated with knowing falsity under the condoning eye of site owners.⁶⁶

¶27 A more interesting and pragmatic approach, proposed by Mike Godwin and Katherine McDaniel, would be to enact a statute similar to Section 230 that would provision safe harbor for content distributors *and content creators* on a voluntarily imparted, site-specific right of reply.⁶⁷ Such a right of reply could take several forms:

For static, non-interactive web pages, the right of reply could take the form of external or internal links. An external link would allow the subject of the defamatory speech (the would-be plaintiff) to link to a webpage of his or her own creation that would contain a rebuttal of, or response to, the defamatory speech. An internal link would point to a part of the defendant’s website where the subjects of defamatory speech would be allowed to make a rebuttals or responses. For dynamic or interactive websites, such as blogs, or Myspace or Facebook pages, a publicly available comments section would suffice, so long as the

⁶³ See Benkler, *supra* note 52.

⁶⁴ BlogHer, What Are Your Community Guidelines? (Jan. 29, 2006), <http://www.blogger.com/what-are-your-community-guidelines>; see also Brad Stone, *A Call for Manners in the World of Nasty Blogs*, N.Y. TIMES, April 9, 2007, at A1.

⁶⁵ O’Reilly Radar, Call for a Blogger’s Code of Conduct (Mar. 31, 2007), http://radar.oreilly.com/archives/2007/03/call_for_a_blog_1.html; see also Stone, *supra* note 64.

⁶⁶ See, e.g., Nakashima, *supra* note 6.

⁶⁷ Mike Godwin & Katherine McDaniel, *A Modest Proposal: Libel Law Reform in the Internet Era 2* (2006), available at <http://crypto.stanford.edu/portia/pubs/articles/GM956689213.html>.

comment-space followed immediately after the post or section containing the allegedly defamatory speech.⁶⁸

The authors correctly reason that, given that their proposal involves an incentive—not a mandate—it would not likely run afoul of *Tornillo*’s print-based proscription of coerced publication.⁶⁹

¶28 However, the proposal is not without its problems. While it improves upon O’Reilly’s proposal by transmuting an ethical concern into a legal one, many Web site administrators might still resist such encroachment and resent the hassle of brokering conflicts. Furthermore, under Godwin and McDaniel’s plan, a response to defamatory speech would be unlikely to replace or be as prominent as the original content, and thus to suffice to undo the defamatory harm. On the other hand, if a response was required to replace the defamatory content, it would leave site administrators in the position of fact-finders, forced to figure out which speech is correct to avoid liability for defamation. That might lead to the same chilling effects the Court sought to avoid by way of its bar on rights of reply in *Tornillo*.⁷⁰ Moreover, if the response had to be placed on equal footing with the original content—either via an active or a passive mechanism—the scheme would interfere dramatically with administrative control of site conceit, which would inevitably retract the breadth of forum structures and content.

¶29 A better option—and a less restrictive one—would leave intact administrative control over site content and format and incorporate features of several of the above-mentioned proposals. It would limit the burden placed on libel victims by emphasizing self-regulation, facilitate the reputation-sorting process by prompting administrators to announce their moderation policies, and function via the dispensation of incentives, rather than penalties. I propose what I hope is such a plan: rewriting Section 230 so as to make safe harbor for site administrators conditional not on a right of reply but on the prominent display of a reasonably accurate administrator-maintained “rating” based on site moderation policies. Such a rating would be conceptually analogous to crude Internet content filters used to restrict material for consumption by minors⁷¹—though it would be premised on site structure, not substance. Also, it would be aesthetically similar to creative commons licenses,⁷² simple graphical labels employed by copyright holders to grant permissions to copy, share, and alter their works.

¶30 These tags would require site administrators to disclose four moderative stances: 1) whether anonymous and pseudonymous posting is permitted; 2) whether all, some, or no “offensive content” is tolerated; 3) whether or not incorrect information, when known to site administrators, is corrected or deleted; and 4) whether there are restrictions on who is permitted to contribute content to the site.⁷³ Of course, one might expect some difficulty in an attempt to explicitly define “offensive content” in a way that doesn’t violate this article’s proscription of substance-motivated regulation. It might be tempting, for example, to define “offensive content” as something akin to “racist, sexist, homophobic, or anti-Semitic material.” However, a “procedural” statement that the moderators of a site tolerate racist speech is in practice probably indistinguishable from a “substantive” statement that there is racist speech on the site. A definition such as “material that would be offensive to the average person” might seem to gloss over the problem. If we’re trying to stay away from selecting content based on “civility rules,” it won’t do to rely on the average person’s views as a guide. Nonetheless, such language might be the best available option. Making “offensive content” one prong of a moderative stance-based rating avoids delving problematically into substance in the sense described by Post. It doesn’t altogether restrict certain “impermissible” viewpoints; it merely warns viewers of their presence.

¶31 Though the moderative stances made prominent by the proposed ratings might otherwise become apparent to a third-party viewer on perusal of a site’s content, the ratings would nevertheless

⁶⁸ Katherine McDaniel, Let it Be, Let it Be: Accommodating Libel in the World of Digitally Distributed Speech 32 (April 2007) (unpublished manuscript, on file with author).

⁶⁹ See Godwin & McDaniel, *supra* note 67, n.11.

⁷⁰ See *Miami Herald Pub. Co. v. Tornillo*, 418 U.S. 241, 258 (1974).

⁷¹ See generally Paul Resnick & James Miller, *PICS: Internet Access Controls Without Censorship*, COMMUNICATIONS OF THE ACM, Oct. 1996, 87.

⁷² See generally Creative Commons, Share, Reuse, and Remix—Legally, <http://creativecommons.org/> (last visited Mar. 16, 2008).

⁷³ Implicit in this fourth prong would be the question of whether victims of defamation are given the opportunity to reply on-site as a matter of right.

serve a function. They would make immediately visible policies highly relevant to a site’s content credibility, and could thus serve as conspicuous proxies for reputation. In essence, the laxer a site’s moderative policies—the more information it tolerates—the less credibility its content should be afforded. At the same time, they would mitigate the potential harm of, for instance, requiring readers to review a large volume of vitriolic content in order to surmise a site’s position on defamatory material. While the ratings would doubtless be too coarse to serve as complete policy identification tools, they would nonetheless enjoy several advantages over the proposals described earlier in this article. Because they would be based on moderative policies, which would theoretically change only rarely, they could be created at the same time as a site and would require virtually no administrator investment over time to maintain. Furthermore, because they would offer only a few choices to reflect each stance,⁷⁴ they would require only minimal resource investment at first to create.⁷⁵ However, even though it would be advantageous to make the ratings as painless as possible for administrators, such a system would not rely on their goodwill (as would the blogger’s code of conduct). Instead, it would present very strong incentives for their compliance—freedom from responsibility where they could constitutionally be held liable in defamation actions. The system would thus be admittedly, but permissibly, coercive. Finally, a rating system would have the crucial advantage of leaving untouched site content. It would thus serve exactly the goal with which we are most concerned—the maintenance of the sanctity of administrative control of site content as a means to substantive and structural diversity on the Internet.

C. Additional proposals

¶32 The rating system suggested above would be amenable to a number of ancillary proposals. I will briefly discuss two of them, search engine incorporation and content filters, as well as the idea of eliminating the distinction between public and private figures on the Internet.

1. Search engine incorporation

¶33 As a distant goal, search engines such as Google might be persuaded to include the ratings labels selected by site administrators directly into search result displays. While one might anticipate that such companies would be reluctant to participate in government created and supported ratings systems (not to mention to assume responsibility for working the systems into their displays and code), they might be interested nonetheless for two reasons. First, they might be reminded that under a different libel regime—one which imposed liability for any republication of libelous material, even an automatic republication—that they themselves might face sanctions. They thus stand to benefit from any system that disfavors the application of liability to faultless disseminators of libelous information, and might, as a matter of good faith, opt to assist in the administration of such a system. Additionally, and more importantly, search engines could use the ratings to aid their attempt to route around frustrating algorithm problems that sometimes buoy malicious information to the top of results lists.⁷⁶ Search engine participation would enhance the rating system’s sorting function by absolving unwitting viewers of the need to call up unmoderated content in the first place. Furthermore, search engine participation would dissuade readers from assigning search results too much credibility.

2. Content filters

¶34 While the aim of this article has been explicitly to avoid limiting access to content on any substantive basis, there are nonetheless instances in which such restrictions would be appropriate, even welcome. The rating systems here proposed could be easily incorporated into a content filter intended for use by minors. The use of such filters would be instituted either by users themselves or

⁷⁴ Site administrators would of course be free to elaborate on these policies; an internal link to a more complete description of stances, including a more specific description of what “offensive content” is permitted, would aid both site administrators and users. The idea, of course, is not to limit what information site administrators may provide about their content, but to limit what information they must provide.

⁷⁵ See Creative Commons License Creation Wizard, <http://creativecommons.org/license/> (last visited Mar. 16, 2008) (The site illustrates how little time the creation of such ratings could take. Using the Creative Commons License Creation Wizard, the code for such licenses can be generated in a matter of a few seconds.)

⁷⁶ See Google Bomb, Wikipedia, http://en.wikipedia.org/wiki/Google_Bomb (last accessed Feb. 15, 2008). (Google’s problems with this phenomenon have been substantial enough to spawn a new term: “Google-bombing.”); see also Tom Zeller Jr., *A New Campaign Tactic: Manipulating Google Data*, N.Y. TIMES, Oct. 26, 2006; An Explanation of Our Search Results, Google, <http://www.google.com/explanation.html> (last visited Mar. 16, 2008).

by their guardians, which would circumvent any worry of inappropriate censoring. Filters based on moderative stances would augment the function of filters based on specific content by recognizing the ways in which unmoderated sites can serve as gateways to profanity, hate speech, and even pornography.

3. *The private/public distinction*

¶35 Godwin and McDaniel’s proposal to eliminate the public-private figure distinction within the confines of the administrative safe harbor they propose (that is, to require plaintiffs to demonstrate actual malice, where they have been provided an automatic right of reply)⁷⁷ merits further mention. This article has already discussed at length the ways in which private individuals who are defamed on the Internet appear to be more like public figures than private from the standpoint of traditional media, as discussed above in Part III(A)(2). The access differential in Internet media between what would normally be considered public figures and what would be considered private is far smaller than that in traditional media, and both types of parties have access that closely approximates that of the (quasi-)media that are the source of defamatory speech online. Moreover, private persons who participate in digital media culture evince a willingness to be part of the public sphere that traditional private figures did not or could not. Finally, because of the growing importance of “cultural creation”—an information society focus on cultural, rather than political, speech and alternate forms of governance, it makes increasingly less sense to continue dryly insisting that the distinction between public and private figures should be premised on a vague notion of who has the capacity to influence “political” events.

¶36 As a result of this inevitable conflation of public and private figures, I would go one step further than Godwin and McDaniel. I propose that the distinction between public and private figures be abolished and with an eye mindful of chilling effects, the actual malice standard for compensatory damages applied to all plaintiffs in online defamation cases. It is worth mentioning that, even under such a standard, many Internet defamation cases could still succeed—specifically, those in which defamatory speech is published with reckless disregard for or knowledge of its falsity. Furthermore, as is also the case with public figures,⁷⁸ punitive damages should be eliminated for would-be private figures, in recognition of the fact that, even when one individual publishes defamatory information about another on the Internet with an intent to harm, limited audience size and a *de facto* right of reply will often narrowly limit the scope of damage to one’s reputation. Additionally, as a justificatory matter, one might find that punishing defamation on the Internet has less retributive or deterrent value than does punishing traditional libel. One might reasonably believe, for instance, that libeling someone in an open forum of limited scope is less wicked than libeling someone in a closed, widely viewed or read traditional forum. One might also expect that, because of the architecture of the Internet, punitive damages might do less to discourage defamatory speech not only because of its overwhelming volume but also due to the difficulty of associating it with its creators.

IV. CONCLUSION

¶37 Faced with the unique challenges of the Internet, it is tempting to want to jury-rig traditional libel law to make it “fit” in the context of digital media. However, as this article has attempted to establish, an attempt to do so would not only fail because of insurmountable technological and legal challenges but would in the process rob Internet media of many of the unique characteristics that make them so exciting as a new generation of communicative tools. A sound treatment of digital defamation will be minimally restrictive, both in the burden it places on site users and the extent to which it defers to administrators to maintain control over site content. It will be structurally based, so as to avoid constricting the breadth and depth of online content and forum structures. Finally, it will rely on reputation, rather than cumbersome legal procedures to sort good information from bad. By

⁷⁷ See Godwin & McDaniel, *supra* note 67, at 17.

⁷⁸ *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 350 (1974).

Caitlin Hall: A Regulatory Proposal for Digital Defamation: Conditioning § 230 Safe Harbor on the Provision of a Site “Rating”

providing Internet users with such accessible tools for gauging content’s worth, the damaging effects of defamation, no matter how epidemic, can be sharply mitigated.