1-1-2010

IT'S MY NEWS Too! ONLINE JOURNALISM AND DISCRIMINATORY ACCESS TO THE CONGRESSIONAL PERIODICAL PRESS GALLERY

Ryan B. Witte

Follow this and additional works at: http://digitalcommons.law.yale.edu/yjolt

Part of the Computer Law Commons, Intellectual Property Commons, and the Science and Technology Commons

Recommended Citation

Witte, Ryan B. (2010) "IT'S MY NEWS Too! ONLINE JOURNALISM AND DISCRIMINATORY ACCESS TO THE CONGRESSIONAL PERIODICAL PRESS GALLERY," Yale Journal of Law and Technology; Vol. 12: Iss. 1, Article 5. Available at: http://digitalcommons.law.yale.edu/yjolt/vol12/iss1/5

This Article is brought to you for free and open access by Yale Law School Legal Scholarship Repository. It has been accepted for inclusion in Yale Journal of Law and Technology by an authorized administrator of Yale Law School Legal Scholarship Repository. For more information, please contact julian.aiken@yale.edu.
Despite its three hundred year existence, the American newspaper is being devastated as the Internet becomes the go-to source for news. Despite the rise in Internet journalism, the sharp increase in online readership, and the precipitous drop in the number of print newspapers, policymakers still have a dismissive attitude toward alternative news sources. Such attitudes must change. In particular, the government should give online-only journalists increased access to the Galleries of the House of Representatives, the Senate, and other state-owned facilities where mainstream journalists are permitted. With a world-wide audience of millions of readers, Congress and the courts can no longer afford to relegate Internet journalism to a second-class news medium.

In Consumers Union v. Periodical Correspondents’ Association, the plaintiff, the non-profit organization that publishes Consumer Reports, questioned the constitutionality of certain rules governing the issuance of press credentials to the Galleries when it was denied admission on ground that it was not an independent publication. Based on separation-of-powers concerns, the United States Court of Appeals for the District of Columbia Circuit avoided the constitutional issue with the political question doctrine, deeming the matter nonjusticiable. Since then, many courts have taken a similar path when faced with the exclusion of a journalist from an established press facility, completely skirting the constitutional issue of whether denial of access violates the freedom of the press protected by the First Amendment. Given the switch from traditional print media to websites and Kindles, the question of who has access to the places where the news is made becomes extremely important. If and when a court will be forced to decide the constitutional issue, it will need a set of principles that balance the constitutional concerns of Congress with the constitutional rights of the online journalist.

This Article will attempt to set forth those principals while at the same time explaining the history, the nature of the rights, and the state of the law as it exists today.
INTRODUCTION

On April 1, 2009, the Guardian announced that it would cease print publication after one hundred and eighty-eight years in business and begin publishing exclusively on Twitter via one hundred and forty character “tweets,” or instant messages. The newspaper cited the unprecedented challenge for all newspapers to begin harnessing the power of the Internet and social networking Web sites to maintain readership. While historical events would be condensed to the bare essentials—OMG Hitler invades Poland, allies declare war see tinyurl.com/b5x6e for more”—the newspaper was confident that brevity would be the key to its continued success.

* Ryan Witte is an associate in the Miami office of Boies, Schiller & Flexner, LLP. I would like to thank Judge Robert Sack of the U.S. Circuit Court of Appeals for the Second Circuit and Professor David Schulz for their comments on earlier drafts of this Article, and the students of the Yale Journal of Law and Technology for helping me make the most out of my work. I would also like to thank my wife, my family, and my friends for their love and support. Lastly, as long as he is on this earth, I will continue to dedicate my scholarship to Professor Steven G. Gey. Without your support and encouragement, I would not be where I am today.


2 Id.
With traditional print media giants going out of business over the last few years, the unobservant reader might have missed the fact that the above story was an April Fool’s Day gag by the Guardian itself. Since March 2007, dozens of newspapers have gone out of business, and still others, like the Seattle Post-Intelligencer, the Capital Times, and the Christian Science Monitor have moved to the Internet to stay afloat. Since 1990, a quarter of all American newspaper jobs have been lost. As Eric Alterman notes, “Few believe that newspapers in their current printed form will survive. Newspaper companies are losing advertisers, readers, market value, and, in some cases, their sense of mission at a pace that would have been barely imaginable just four years ago.” Despite its three hundred year existence, the American newspaper is being devastated in the span of a decade.

The Internet is becoming the go-to source for news, information, weather, movie reviews, and classified advertisements. A recent study by the Pew Internet and American Life Project found that the Internet is now the third most popular news platform behind television and radio, with about sixty-one percent of people turning to the Internet as a source of news, and around ninety-two percent of Americans are utilizing multiple news platforms such as Internet and print media on a typical day. With websites like Craigslist allowing anyone with a computer to post free classified ads, it is no wonder that users are refusing to shell out $325 to place an employment notice in the New York Times.

---


5 Id.


7 See Purcell et al., supra note 6, at 31.

IT'S MY NEWS TOO! ONLINE JOURNALISM AND DISCRIMINATORY ACCESS TO THE CONGRESSIONAL PERIODICAL PRESS GALLERY

cannot be said for the ink destined for its pages. News dissemination is just as vital to our democratic heritage as it has always been—even if the “ink” is now digital.

Despite the rise in Internet journalism, the sharp increase in online readership, and the precipitous drop in the number of print newspapers, policymakers still have a dismissive attitude toward alternative news sources. Such attitudes must change. In particular, the government should give online-only journalists increased access to the Galleries of the House of Representatives, the Senate, and other state-owned facilities where mainstream journalists are permitted. With a “world-wide audience of millions of readers, viewers, researchers, and buyers,” Congress and the courts can no longer afford to relegate Internet journalism to a second-class news medium. In *Consumers Union v. Periodical Correspondents’ Association*, the plaintiff, the non-profit organization that publishes Consumer Reports, questioned the constitutionality of certain rules governing the issuance of press credentials to the Galleries when it was denied admission on ground that it was not an independent publication. Based on separation-of-powers concerns, the United States Court of Appeals for the District of Columbia Circuit avoided the constitutional issue with the political question doctrine, deeming the matter nonjusticiable where the defendant acted in good faith and “pursuant to [its] express delegation of authority as aides . . . of Congress.” Since then, many courts have taken a similar path when faced with the exclusion of a journalist from an established press facility, completely skirting the constitutional issue of whether denial of access violates the freedom of the press protected by the First Amendment.

Given the switch from traditional print media to websites and Kindle, the question of who has access to the places where the news is made becomes extremely important. If and when a

---

11 515 F.2d 1341, 1342 (D.C. Cir. 1975).
13 See *Branzburg v. Hayes*, 408 U.S. 665, 728 (1972) (Stewart, J., dissenting) (“No less important to the news dissemination process is the gathering of information. News must not be unnecessarily cut off at its source, for without freedom to acquire information, the right to publish would be impermissibly compromised.”); see also *id.* at 723 (Powell, J., concurring) (“A popular Government, without popular information, or the means of acquiring it, is but a Prologue to a Farce or a Tragedy; or, perhaps both. Knowledge will forever govern ignorance: And a people who mean to be their own Governors, must arm

http://digitalcommons.law.yale.edu/yjolt/vol12/iss1/5
court will be forced to decide the constitutional issue, it will need a set of principles that balance the constitutional concerns of Congress with the constitutional rights of the online journalist.

It is important to note at the outset that the term “blogger” will not be found in very many places within this Article. This is because I believe there is a distinction between “bloggers” and online journalists that is rarely, if ever, clarified. For one, anyone with a computer and an Internet connection can be a blogger—online journalism requires the writer to engage in activities that are typically associated with traditional print reporting. Rather than simply posting commentary or opinion based off of someone else’s work, an online journalist would attempt to gather his own news information through investigation, sources, and access. Indeed, the right to access the source of the news is at the heart of this Article. These distinctions are important not only to First Amendment press jurisprudence, but to ensure that “commentators in pajamas” are not defining the right of access to the Galleries.\(^{14}\) The online journalist might be a former reporter for a major publication who decides to research, report, and publish his own news online. She might be a journalism school graduate. Lumping all online publication under the banner of “blogging” is sure to restrict protections that should be rightfully afforded to journalistic professionals who choose a “unique and wholly new medium of worldwide human communication.”\(^{15}\)

This Article will address several issues related to the freedom of access to the Periodical Press Gallery. Part I will briefly describe the history of the press and the history of congressional reporting leading up to the passage of the Periodical Press Gallery Rules. Part II will describe the rules that govern admission to the Gallery as they exist today. Part III will describe

\(^{14}\) This is not to say that bloggers are not necessarily real journalists, an issue which is highly debated. Typing the search term “Are Bloggers Journalists” into Google will reveal dozens of differing opinions on the subject. Rather, this Article presupposes that most bloggers do not engage in original and independent newsgathering, which is one of the hallmarks of traditional print journalism. To the extent a blogger does engage in this type of activity, then the protections discussed in this Article would be appropriate.

\(^{15}\) \textit{Am. Civil Liberties Union}, 521 U.S. at 850. Further validation for Internet journalism came on April 13, 2010, when Pro-Publica became the first online entity to win a Pulitzer Prize for investigative journalism. The prize was awarded to Sheri Fink for her article \textit{Deadly Choices at Memorial} about the choices faced by New Orleans hospital workers in the days after Hurricane Katrina. \textit{See Adam Goldman, New Media Recognized in Pulitzer Competition, ASSOCIATED PRESS, Apr. 13, 2010, available at http://abcnews.go.com/Business/wireStory?id=10358444.}
the nature of the right afforded to journalists to access government buildings which have already been opened up to the press. Cases discussed within this Part challenge the denial of access to the White House, the Gallery, and the Guantanamo Bay detention facility. Discussion in this Part will reveal how the D.C. Circuit Court has used the political question doctrine to avoid deciding these Gallery cases on their merits, while taking a deferential approach to denials from other established press pools. The final Part will argue that Consumers Union was wrongly decided, and address how the courts should decide cases which stem from a denial of admission to the Galleries. If the courts still refuse to decide these cases on nonjusticiability grounds, I will argue that Congress should refine its rules to cabin the discretion of the Executive Committee and permit more online journalists to enter the Gallery. I will describe how the current Periodical Press Gallery Rules act as an unconstitutional obstacle to the First Amendment rights of online journalists, as well as a Fifth Amendment violation of due process. Lastly, I will detail a few substantive amendments to the existing Rules which would provide for a fairer review of access for online journalists.

I. THE PERIODICAL PRESS GALLERY AND ITS RULES

A. History of the Periodical Press Gallery

What constitutes the press has changed dramatically since the First Amendment was adopted in 1791. “When the First Amendment was written, journalism as we know it did not exist.”

In the eighteenth century, the press was a trade of printers, not journalists, and “the press” meant “the printing press.” “Freedom of the press referred to the freedom of the people to publish their views rather than the freedom of journalists to pursue their craft.” At that time, the right was enjoyed by pamphleteers and individuals, rather than the media conglomerates of today.

In the later part of the nineteenth century, newspapers began hiring their own employees for the purpose of gathering news. Advances during the Industrial Revolution allowed mass production of newspapers, and years later gave rise to the titans of the newspaper industry that we know today. Industry pioneers such as Adolph Ochs, William Randolph Hearst, and Joseph Pulitzer all built their empires during the mid-to-late nineteenth century. Until

17 Id.
18 Id. at 446-47.
recently, “own[ing] the dominant, or only, newspaper in a mid-sized American city was . . . a kind of license to print money.” But, as the shift to online news continues, the press is in search of new models to make money. The online journalist, just like the pamphleteer of old, is in the business of publishing his own news and opinions without the allegiances to major newspapers—something the First Amendment originally envisioned.

The first congressional reporters were stenographers who attempted to publish congressional proceeding in the form of verbatim notes of speeches. These reports were often marred by incompleteness or inaccuracies leading to condemnation from lawmakers. In addition to the stenographers, a group of unaffiliated correspondents called “letter-writers” sent news and commentaries to out-of-town newspapers. Letter-writers were often critical of congressional members, drawing criticism from Congress, which retaliated by attempting to limit their access to congressional activities. In 1839, Congress debated a proposal to deprive out-of-town newspapers of access to congressional proceedings. The proposal infuriated a number of newsmen, leading to biting editorials. After being denied his usual seat in the Gallery, James Gordon Bennett of the New York Herald wrote of the “most outrageous, high-handed, unconstitutional act[] ever perpetrated by any legislative assembly in a free land—an act of despotism, tyranny and usurpation against the liberty of the press which the House of Lords of England . . . would not attempt against any newspaper in England.”

In response to Bennett’s protestations, a Whig majority in the Senate, led by Henry Clay, created the first “Reporter’s Gallery” for the press in July of 1841. Clay’s Senate resolution created “suitable accommodations to be prepared in the eastern gallery [for all bona fide reporters certified by the Editors of the papers for which they reported].” In 1879, the press itself took on the responsibility for monitoring the Galleries, drafting their regulations at the New York Times office in New York. The rules defined accreditable correspondents and barred lobbying by any member of the Gallery. The House adopted the New York Times

\[\text{19 See Alterman, supra note 4.} \]
\[\text{21 Id.} \]
\[\text{22 Id. at 35.} \]
\[\text{23 Donald A. Ritchie, Press Gallery: Congress and the Washington Correspondents 26 (1991).} \]
\[\text{24 Id.} \]
plan in 1879, and the Senate followed in 1884.25 This same system, drafted by the institutional press in the latter part of the nineteenth century, still controls admission to the Galleries today.

**B. Rules Governing Gallery Access**

The rules governing access to the Galleries come from several sources. Article I, Section 5, Clause 2 of the Constitution permits Congress to define the rules of its proceedings.26 Pursuant to that authority, both houses of Congress passed their own rules to delegate control of the Galleries. Senate Rule XXXIII permits the Committee on Rules and Administration to make all rules and regulations “respecting the reporters’ galleries of the Senate, together with the adjoining rooms and facilities, as will confine their occupancy and use to bona fide reporters of newspapers and periodicals, and of news or press associations for daily news dissemination.”27 Rule VI of the Rules of the House of Representatives provides that a “portion of the gallery over the Speaker’s chair as may be necessary to accommodate the representatives of the press wishing to report debates and proceedings shall be set aside for their use.”28 The rule notes that “[r]eputable reporters and correspondents shall be admitted thereto under such regulations” and that “[t]he Standing Committee of Correspondents for the Press Gallery, and the Executive Committee of Correspondents for the Periodical Press Gallery, shall supervise such galleries, including the designation of its employees, subject to the direction and control of the Speaker.”29

Pursuant to the authority of the rules, the Speaker of the House and the Senate Committee on Rules and Administration established Rules and Regulations which govern the Galleries.30 The rules give control of the Galleries to the Executive Committee of the Periodical Correspondents’ Association.31 In order to qualify

25 Id.
26 U.S. CONST. art. I, § 5, cl. 2.
29 Id.
31 Id. The current Executive Committee includes: Richard Cohen, National Journal, Chairman; Jay Newton-Small, Time Magazine, Secretary; Lauren Whittington, Roll Call, Treasurer; Heather Rothman, BNA News; Meg Shreve,
for access to the Galleries, prospective members must meet two criteria. First, prospective members must be “bona fide resident correspondents of reputable standing, giving their chief attention to the gathering and reporting of news.” Second, the applicants “must be employed by periodicals that regularly publish a substantial volume of news material of either general, economic, industrial, technical, cultural, or trade character . . . [which] require such Washington coverage on a continuing basis . . . .” The periodical itself must be “be owned and operated independently of any government, industry, institution, association, or lobbying organization” and must be “published for profit . . . supported chiefly by advertising or by subscription, or . . . published by a nonprofit organization [if additional criteria are met].” In addition, no member of the Gallery may be engaged in any form of lobbying. Despite the apparent clarity of the rules, standards such as “bona fide correspondent,” or “regular[] [publication of] a substantial volume of news” provide sufficient latitude for discriminatory and arbitrary interpretations. While it may seem the rules governing the admission of a journalist are clear, the manner in which the Executive Committee applies the rules is a real concern.

Members of the Periodical Correspondents’ Association enjoy a variety of advantages over reporters who are unable to obtain admission. First and foremost, they are provided with a seat in the Galleries without having to contend for space in the public galleries. In addition, Congress furnishes the accredited correspondents with support facilities and staff. Members are also permitted access to the House Speaker’s Lobby and the Senate President’s Room where they may seek and conduct interviews with members of Congress. Lastly, members of the Gallery are given exclusive access to attend the daily on-the-record press conferences held by the Senate leadership and the Speaker of the House.
These considerable perks provide lucky members with an advantage over all non-accredited journalists. The permanent presence of journalists in the Gallery allows them to foster relationships with members and cultivate valuable news sources. The constant presence and strong relationships allow Gallery journalists the opportunity to break stories and scoop the competition. Overall, the exclusion of a reporter from the Gallery “constitutes a permanent disadvantage with regard to the gathering of news and has a significant impact . . . both upon the publication excluded and others in similar situations.”\(^{40}\) At this point in history, it is the online-only journalist and her counterparts who bear the disproportionate burden of these disadvantages.

II. ACCESS TO ESTABLISHED PRESS POOLS

Several court decisions have addressed the issue of access to government institutions that have already been opened up to newsmen. The first Section of this Part reviews the nature of the right of access to these governmental institutions. The second and third Sections describe how courts have addressed decisions made by the Executive Committee of the Periodical Correspondents’ Association to exclude a particular journalist, analyzing decisions where access was denied and decisions where access was granted. The fourth section describes how “deference” has been shown to the legislative branch when it comes to determining access, while the executive branch has not been given the same judicial courtesy. The final section attempts to distinguish seemingly inconsistent holdings by the same court regarding access to governmental institutions. There are two lines of cases relating to the right of access involving established press facilities. In one line of cases, the courts refuse to address the First Amendment issue of access, relying instead on the nonjusticiability doctrine to determine that the issue is one not appropriate for judicial intervention. In another line of cases, courts have given protection to journalists who were excluded from access on the basis of unpublished, unclear, or arbitrary rules. The reasoning in the latter set of cases should ultimately prevail as courts will have to contend with these issues on a substantive basis and will have to give serious consideration to the rights of online-only journalists.

\(^{40}\) See id. at 26.
A. The Nature of the Right

The courts have come to varying determinations about the nature of the constitutional right of access afforded to journalists in Washington. This Article does not discuss the right to press access where the government has denied access entirely. Rather, the cases discussed in this Article are all examples of where the government “has voluntarily decided to establish press facilities for correspondents who need to report therefrom.” The courts have treated these cases quite differently than situations where all media access is denied to a particular event. All courts which have reached the merits of these individual exclusion cases have recognized a First Amendment right for journalists to access an already open press facility.

In Sherrill v. Knight, the Washington correspondent for the Nation was denied a White House press pass after the Secret Service deemed him to be a security risk. Here, the Secret Service did not publish guidelines governing the grant or denial of press credentials, did not establish procedures to appeal a denial of a press pass, and never informed the journalist of the precise reasons for his denial. In fact, Sherrill did not learn why he was denied access until five years after he applied, during the discovery phase of his action.

The government argued that because the White House was not open to the public and because the right of access due to the press is generally no greater than the right of access due to the public, there was no violation of the First Amendment unless the denial was arbitrary or based on the content of the journalist’s speech. The U.S. Court of Appeals for the District of Columbia Circuit acknowledged that arbitrary or content-based criteria are prohibited under the First Amendment, but noted that there were other considerations besides these. The court noted that “the First Amendment’s protection of a citizen’s right to obtain information concerning ‘the way the country is being run’ does not extend to every conceivable avenue a citizen may wish to employ in pursuing this right.” In particular, the First Amendment claim at issue did not demand that the President grant an interview to every journalist, nor that the White House open its doors to the press,

41 Sherrill v. Knight, 569 F.2d 124, 130 (D.C. Cir. 1977).
42 Id. at 127.
43 Id. at 126-27.
44 Id. at 127.
45 Id. at 129.
46 Id.
47 Id. (quoting Zemel v. Rusk, 381 U.S. 1, 17 (1965)).
IT’S MY NEWS TOO! ONLINE JOURNALISM AND DISCRIMINATORY ACCESS TO THE CONGRESSIONAL PERIODICAL PRESS GALLERY

conduct press conferences, or operate press facilities at all. Rather, the court held that where the doors are already open, and where press facilities are made publically available as a source of information for newsmen, “the protection afforded to news-gathering under the First Amendment guarantee of freedom of the press, requires that access not be denied arbitrarily or for less than compelling reasons.” Judge McGowen also found that “notice, opportunity to rebut, and a written decision are required because the denial of a pass potentially infringes upon First Amendment guarantees . . . [which] cannot be permitted to occur in the absence of adequate procedural due process.” The court observed that, in addition to the newsmen, “the public at large [has] an interest protected by the first amendment in insuring that restrictions on news-gathering be no more arduous than necessary, and that individual newsmen not be arbitrarily excluded from sources of information.”

Because of the important First Amendment rights implicated in the denial of individual access, the court determined that the refusal of a press pass should be based on a “compelling governmental interest.” In Sherrill, the court had no problem determining that the physical security of the President of the United States constituted a compelling—even overwhelming—interest, but the standards and process used to deny the press-pass did not pass constitutional muster. For one, the standard for denial of a press-pass was never formally articulated or published. In addition, informing journalists that they were denied for “reasons of security” was unnecessarily vague and subject to ambiguous interpretation. In clarifying the constitutional requirements, the court noted that while the specific interest in that case (i.e., presidential safety) did not lend itself to detailed articulation of narrow and specific standards or clear-cut factors, the standard must provide a meaningful way for journalists to be labeled a security risk and be sufficient to allow for “meaningful judicial review.” While Judge McGowen implored lower courts to be appropriately deferential, he still opined that notice, opportunity to respond, and a written statement of the

48 Id.
49 Id. (citations omitted).
50 Id. at 128.
51 Id. at 129-30.
52 Id. at 130.
53 Id.
54 Id.
55 See id.
56 Id. (emphasis added).
reasons for denial were “compelled by the . . . determination that the interest of a bona fide Washington correspondent in obtaining a White House press pass is protected by the first amendment . . . [which] undoubtedly qualifies as [a] liberty which may not be denied without due process of law under the fifth amendment.”

The court decided that at a minimum, an unsuccessful applicant must be informed of the factual basis for denial and provided an opportunity to rebut the denial. As in Sherrill, where applicants were only told that they were a “security risk,” the Periodical Correspondents’ Association often denies applications by reference only to a rule, setting forth no factual findings to be appealed. In light of the First and Fifth Amendment interests articulated in Sherrill, the Periodical Correspondents’ Association’s scheme is likely to fail constitutional review, but, unfortunately, courts have been hesitant to decide Gallery cases on their merits.

B. Access Denied

The seminal case regarding access to the Galleries is Consumers Union of United States, Inc. v. Periodical Correspondents’ Association. In 1972, Gilbert Thelen submitted an application to the Executive Committee of the Periodical Correspondents’ Association for membership as a representative of Consumer Reports. The Committee rejected the application on the ground that Consumer Reports was “not an independent publication,” as required by Rule 2 of the Periodical Press Gallery Rules. The Committee offered no factual basis for rejecting the application. Following the remedial scheme authorized by the rules, Thelen asked the Executive Committee to reconsider its decision, but the Committee again rejected the application. Thelan then appealed to the Senate Committee on Rules and Administration, and to the Speaker of the House, to no avail. While Thelan pursued administrative remedies, the basis of the rejection was later clarified. The Executive Committee contended that Consumer Reports was published by Consumers Union, a nonprofit organization “which is a self proclaimed advocate of consumer interests and, among other activities, testifies before Congressional committees on behalf of the interests of consumers.” Because the parent company of Consumer Reports

57 Id. at 130-31.
58 515 F.2d 1341 (D.C. Cir. 1975).
59 Id. at 1345.
60 Id.
was classified as an "advocacy group," rather than a publishing organization, Thelen was denied his credentials.

In his action for declaratory relief, the plaintiff argued that the Rules Governing Periodical Press Galleries were unconstitutional both on their face and as applied to Consumer Reports. More specifically, the plaintiff contended that Rule 2 “constituted a prior restraint upon, and otherwise abridged, its rights to gather, have full access to, and report to its readers upon, the news concerning Congress and of a public nature, in violation of . . . the First Amendment . . . ”\(^{62}\) In addition, the plaintiff argued that in “denying accreditation to Consumer Reports[,] the Association acted in a discriminatory, arbitrary, capricious, and unreasonable manner, thus violating Consumers Union’s rights under the Fifth Amendment.”\(^{63}\)

The U.S. District Court for the District of Columbia found for Consumers Union and declared the Periodical Press Gallery Rules unconstitutional on First and Fifth Amendment grounds.\(^{64}\) Like the Sherrill case two years later, the court opined that where certain journalists are excluded from gaining equal access to facts of public consequence, limitations must be clearly justified by a compelling and demonstrable governmental interest.\(^{65}\) In addition, “means selected for furthering [the governmental] interest must be no more restrictive of individual rights than is reasonably necessary.”\(^{66}\) Finally, the district court concluded that the rules may not be so vague or overbroad as to unnecessarily chill the exercise of those rights or provide insufficient guidance to those who must administer the legislation.\(^{67}\) Of particular note was the finding that the Periodical Press Gallery Rules were too ambiguous, despite being unmistakably clearer than the unpublished Secret Service rules in Sherrill.

In applying the law, the district court held that the exclusion of some reporters from an area which had been voluntarily opened to other reporters for the purpose of news-gathering poses grave constitutional problems.\(^{68}\) The district court found that when access to news sources is “unreasonably or arbitrarily denied by congressional action or publishers meeting under congressional auspices, [that denial] constitutes a direct limitation upon the content of news as recognized in Branzburg v.

\(^{62}\) *Consumers Union*, 515 F.2d at 1346.

\(^{63}\) *Id.*

\(^{64}\) *Consumers Union*, 365 F. Supp. at 26-27.

\(^{65}\) *Id.* at 25.

\(^{66}\) *Id.*

\(^{67}\) *Id.*

\(^{68}\) See *id.*
The court recognized that “[a]ll types of news compete and all types of publications are entitled to an equal freedom to hear and publish the official business of the Congress,” and that “[t]he Constitution requires that congressional press galleries remain available to all members of the working press, regardless of their affiliation.” The court held that the broad and generalized grant from Congress to the Correspondents’ Committee permits the Committee to implement arbitrary and unnecessary regulations to exclude publications they consider objectionable without any means to check the abuse of their delegated authority. This, the district court held, violated the constitutional rights of Consumers Union. The court concluded that more definitive rules were needed to permit due process prior to exclusion and provide some opportunity for adequate impartial review whenever a journalist is excluded.

In reversing the district court, the U.S. Court of Appeals for the District of Columbia Circuit held that the issue was nonjusticiable under the political question doctrine. Refusing to address the matter on the merits, the court found that Article I, Section 5, Clause 2 permits Congress to “determine its rules of proceedings,” and so long as the rule does not ignore constitutional restraints or violate fundamental rights, it is no impeachment of a rule to say that some other rule would be more just. The court also noted that the rules need only have a “reasonable relation” to the results that they seek to attain.

In evaluating the legislative purpose of the Periodical Press Gallery Rules, the court found that the intent was to ensure that the Galleries are only used for bona fide reporters who will not abuse the privilege by lobbying on behalf of private interests. The court found that the rules were reasonably related to the aforementioned purpose and that courts had no power to second guess Congress’s exercise of its Article I powers.
Invoking the Speech and Debate Clause under Article I, Section 6, Clause 1, the court also found that the Correspondents’ Association was entitled to immunity from Consumers Union’s challenge. The court opined that deciding the composition of the Gallery is a legislative function, as evidenced by Congress’s direct and historical control over the seating of the press in the nineteenth century. Furthermore, because the function was delegated to the Correspondents’ Association, the Association would be immune from suit so long as an individual member would be immune if the action was taken directly by him. The court was “content to rest [its] ruling... upon the ground that, performed in good faith, the acts of [the Correspondents’ Association] were within the spheres of legislative power committed to the Congress and the legislative immunity granted by the Constitution.”

The first and only case applying Consumers Union to an Internet journalist was Schreibman v. Holmes. In 1997, the U.S. Court of Appeals for the District of Columbia Circuit was again confronted with the question of whether denying a journalist access to the press galleries constitutes a violation of the First and Fifth Amendments. In this case, Schreibman was the sole owner, publisher, editor, and writer for Federal Information News Syndicate (FINS), which published a biweekly Internet news letter that reported on federal legislation and governmental policies. FINS had a number of paying subscribers and even more who read the publication online for free. The Executive Committee of Correspondents denied Schreibman’s request for accreditation on the grounds that his publication did not meet the requirements under Periodical Press Gallery Rules 1 and 2. The Committee failed to provide factual basis for the denial, but at the plaintiff’s

---

77 U.S. CONST. art. I, § 6, cl. 1; see Consumers Union, 515 F.2d at 1349.
78 See Consumers Union, 515 F.2d at 1350 (“Appellants were acting by virtue of an express delegation of authority as aides or assistants of Congress. If their actions would have been immune from inquiry... had they been performed by Members of Congress, the same immunity would attach to appellants.”). 79 Id. at 1351. Contra Consumers Union v. Periodical Correspondents’ Ass’n, 365 F. Supp. 18, 24 (D.D.C. 1973) (“[I]t is well established that a congressional rule which infringes upon the constitutional rights of persons other than Congressmen presents a proper question for the judiciary. ... [The conduct of the Correspondents’ Association] in barring the representatives of certain publications from the periodical press galleries and admitting others neither constitutes an integral part of nor has been shown to have a significant impact upon the proceedings on the floor of either House... In the absence of such a showing, it must be concluded that the Speech and Debate Clause does not shield the defendants from a challenge to their admission policies.” (citing Yellin v. United States, 374 U.S. 109, 143-44 (1963))).
81 See House Press Gallery Rules, supra note 30, Rs. 1, 2.
request, held a public hearing to reconsider the application. Again, Schreibman’s request was denied. He finally appealed to the Speaker of the House and the Senate Committee on Rules and Administration, but no action was taken on the appeal.

In the U.S. District Court for the District of Columbia, the Correspondents’ Committee finally asserted the factual basis for denying Schreibman’s request. They contended that in contravention of the Periodical Press Gallery Rules, FINS was not published for profit, Schreibman did not receive a salary from FINS, and Schreibman did not earn his livelihood as a journalist. The Committee also maintained that their interpretation of the Periodical Press Gallery Rules was immune from judicial review under the Speech or Debate Clause of the Constitution. Schreibman alleged that the Periodical Press Gallery Rules were unconstitutional on their face and as applied to his publication.

Relying on Consumers Union, Judge Urbina held that Schreibman’s challenge was nonjusticiable. The court again determined that the Speech or Debate Clause barred suit against the Correspondents’ Committee regarding its accreditation decisions “so long as the Committee was acting within the scope of its authority and in good faith.” Finding that Schreibman never pled that the Committee acted in bad faith, the court dismissed the suit.

C. Access Granted

Decided two years after Consumers Union and twenty years before Schreibman, the Sherrill court was the first court to venture into the constitutional issue underlying special access cases. Because Sherrill is still good law, it provides considerable weight to the argument that courts should not lightly abdicate their judicial responsibilities when it comes to reviewing Gallery access cases on the merits. Luckily, the D.C. District Court is slowly showing its willingness to address the issue.

The most recent case speaking to the issue of access is Getty Images News Services, Corp. v. Department of Defense. In this case, Getty Images alleged a violation of its First Amendment rights, due process rights, and equal protection rights when the DOD rejected the photo service from travel with certain press

---

82 See Schreibman, 1997 WL 527341, at *2.
83 Id.
84 Id.
85 Id. at *4 (emphasis added).
pools related to Operation Enduring Freedom of the Iraq War.\textsuperscript{88} Getty Images asserted that the DOD's method of selecting which media organizations could travel to Guantanamo Bay was arbitrary and capricious, and permitted the DOD to reject an application without any reasonable explanation, and without review. Dismissing all but one of Getty's claims, the U.S. District Court for the District of Columbia ultimately held that Getty was likely to succeed on the argument that inadequate and unpublished criteria relating to the selection of journalists for a press pool with finite space was a violation of their due process and First Amendment rights.\textsuperscript{89}

In 2002, the DOD began permitting journalists to travel to Guantanamo Bay on a military transport plane to cover the detention facility.\textsuperscript{90} Because there was only one way to Cuba, and one way home, the space allotted to journalists was limited, and the exclusion of some was necessary and inevitable.\textsuperscript{91} In order to aid in the selection process, the DOD crafted a set of six internal guidelines that would inform their decision. These guidelines were not published, and selection decisions were made by a DOD Public Affairs Officer based on his or her "general knowledge and expertise."\textsuperscript{92} On occasion, the Public Affairs Officer would elicit information from other members of the press pool to guide his or her decision.\textsuperscript{93} In support of its decision to exclude Getty Images, the DOD argued that not only is Guantanamo Bay not a public forum, but that "review of military regulations challenged on First Amendment grounds is far more deferential than constitutional review of similar laws or regulations designed for civilian society."\textsuperscript{94}

The district court agreed with the DOD that heightened deference was due to the military.\textsuperscript{95} In addition, the court found that Guantanamo Bay is a closed military base located on an island with no commercial air travel, dedicated to the housing of terrorist suspects in a military operation.\textsuperscript{96} But despite these important

\textsuperscript{88}Id. at 114.
\textsuperscript{89}Id. at 124.
\textsuperscript{90}Id. at 114-15.
\textsuperscript{91}Id. at 115.
\textsuperscript{92}Id. at 116.
\textsuperscript{93}See id.
\textsuperscript{94}Id. at 119 (quoting Goldman v. Weinberger, 475 U.S. 503, 507 (1986)).
\textsuperscript{95}See id. (“The Court agrees both that the Guantanamo Bay Navel Base is not a public forum and that consideration of Getty’s First and Fifth Amendment claims must be undertaken through the prism of the heightened deference due to military regulations and decision-making. Nonetheless, equal access claims by the press warrant careful judicial scrutiny.”).
\textsuperscript{96}Id. at 120.
policy concerns, the court found that components of the DOD selection process were unreasonable. For one, the court reasoned that the standard for due process is not met where “criteria . . . are either vague or completely unknown [so that] the party affected has no way of knowing how to achieve compliance . . . nor even of challenging them as being improper.” The court also noted—as in the situation of the Gallery—that competing journalists were permitted to inform the DOD’s decision on who is granted access. The court concluded that the “DOD must not only have some criteria to guide its determinations [about journalists], but must have a reasonable way of assessing whether the criteria are met.”

The court also held that “equal access claims by the press warrant careful judicial scrutiny.”

**D. Deference**

The *Sherrill* and *Getty* line of cases provides a proper standard for reviewing Gallery problems, as well as the appropriate measure of deference when reviewing access decisions on the merits. In addition, *Getty* is instructive when evaluating the reasonableness of the selection criteria for journalists.

First, both lines of cases involve the issue of deference. Whereas the courts in *Consumers Union* and *Schreibman* made the decision to give the ultimate deference to the Periodical Correspondents’ Association—rendering their decisions unreviewable—*Getty* and *Sherrill* carved out a more appropriate path for handling these types of cases.

Like in the Gallery cases such as *Consumers Union* and *Schreibman*, *Sherrill* involved the issue of deference to a coequal branch of government. The reasoning used to deny judicial review in *Consumers Union* was that the Constitution entrusted Congress with the authority to pass rules regulating access to the chamber, and that power was in turn delegated to the Correspondents’ Association. However, access to the White House and the President of the United States is within the scope of powers reserved to the Executive Branch. In turn, the Secret Service is the agency entrusted with deciding who may access the White House. But despite deference to a coequal branch of government and to the administering agency, the court still found that the constitutional concerns in *Sherrill* were within the purview of the

---

97 *Id.* at 121.
99 *Id.* at 119.
court. The court rejected the government’s argument that the case was “nonjusticiable . . . because protection of the President is vested within the sole discretion of the Executive.”\(^{101}\) The court found the argument “wholly without force,” and noted that this discretion “cannot be said to authorize procedures or actions violative of the Constitution.”\(^{102}\) The court found that once the White House made its press facilities available as a source of information for newsmen, “the protection afforded news-gathering under the first amendment guarantee of freedom of the press require[ed] that this access not be denied arbitrarily or for less than compelling reasons.”\(^{103}\) The court gave some weight to the government’s compelling reasons, but still found that decisions to exclude newsmen must be reasonable and guided by appropriate standards. *Sherrill* shows that courts can grant the appropriate measure of discretion to a coequal branch, while still safeguarding the rights of citizens.

If there were ever a case where public policy would counsel the courts to grant broad discretion to exclude journalists, it would be *Getty*. Unlike the Capitol in Washington, D.C., which houses the elected representatives of the people, Guantanamo Bay houses terrorist suspects and is accessible only by military transport. Unlike the Galleries, with over five hundred seats, space on the military base is much more restricted. Despite granting broad discretion to military decisions, the court in *Getty* still ruled that the decisions of the Department of Defense must be reasonable; a decision which is ultimately subject to judicial review.

There is no legitimate policy justification for holding the Department of Defense to a higher standard in excluding journalists than the Executive Committee of the Correspondents’ Association. Just as *Getty* demonstrates a new willingness to tread into access cases, courts should reconsider the district court opinion in *Consumers Union* and find the decisions made by the Executive Committee outside of the scope of immunity of the Speech and Debate Clause.

### E. Distinguishing Consumers Union from Sherrill

Arguably, the similarities between *Getty* and *Sherrill* may have led to their similar treatment, in contrast with the treatment of *Consumers Union* and *Schreibman*. In particular, the courts have

\(^{101}\) *Id.* at 128 n.14.

\(^{102}\) *Id.*

\(^{103}\) *Id.* at 129-30 (citations omitted).
only stepped in when the selection criteria for journalists were unpublished or when no appeal procedure had been established.

First, Getty and Sherrill involved selection criteria which were not published by the Secret Service or the Department of Defense. By contrast, the Periodical Press Gallery Rules have always been made public and available to any journalist prior to filing his application with the Committee. While this major difference is instructive on the matter of due process, the Periodical Press Gallery Rules are not so clear as to avoid constitutional scrutiny. As described above, although the Periodical Press Gallery Rules are long, much of the text deals with what Gallery journalists are prohibited from doing once they are admitted into the Gallery. The selection criteria themselves are subject to amorphous and arbitrary application. For instance, could a failing newspaper be denied admission because it is no longer turning a profit?\textsuperscript{104} As advertising dollars dry up and newspapers look toward alternative money-making schemes, can the newspaper be said to be “supported chiefly by advertising or subscription?”\textsuperscript{105} Is an online journalist categorically excluded from being a “bona fide resident correspondent[] of reputable standing, giving [his] chief attention to the gathering and reporting of news”?\textsuperscript{106} No matter how clear the rules are, judicial unwillingness to review decisions of the Executive Committee leaves the strong possibility of impermissible discrimination. All of these cases, including Consumers Union, provide that the Executive Committee is only entitled to immunity when its decisions are made in good faith. As Part III of this essay points out, the opportunity to make decisions in bad faith is too great to go unchecked.

Also important in Sherrill is that rejections were made without factual findings. The same is true in Schreibman and Consumers Union, where the Committee refused to give any factual basis for rejecting accreditation until after litigation commenced. This is most problematic when it comes to the appeal procedure. Although the Periodical Press Gallery Rules provide for “a right to a public hearing before the committee,”\textsuperscript{107} the right is meaningless when the factual basis for denial is not released until after the hearing. In addition, due process typically requires an impartial decisionmaker; in this case, the same committee that

\textsuperscript{104} See House Press Gallery Rules, supra note 30, R. 2 (“Applicants must also be employed by a periodical that is published for profit and is supported chiefly by advertising or by subscription.”).

\textsuperscript{105} Id.

\textsuperscript{106} Id. R. 1.

\textsuperscript{107} Id. R. 5.
made the decision is given the power to review it. Although the applicant can then appeal directly to the Speaker of the House, this method has been wholly unsuccessful in the past. The court in *Sherrill* made clear that the basis for rejection must be sufficient to allow for meaningful judicial review. In the absence of judicial willingness to review, the rationale for rejection may never be released.

### III. PROBLEMS WITH THE CURRENT SYSTEM

#### A. Problems with the Rules Themselves

As seats in the Gallery begin to open up, the question of who occupies the empty chairs will become a source of contention. Because *Consumers Union* and its progeny are on shaky ground jurisprudentially, the political question doctrine and the Speech or Debate Clause should be set aside, and cases involving discriminatory access to the Gallery should be decided on the merits. Even if the courts refuse to budge, there are still substantial policy reasons for Congress to redraft the rules or provide additional oversight to the accreditation process.

First, even if the existing rules are faithfully applied and are sufficiently distinguishable from those in *Sherrill* and *Getty*, the existing requirements infringe the First Amendment rights of the entire class of online journalists.

One piece of evidence of the discriminatory nature of the rules is the subsequent denial of Gallery accreditation to a reporter who was already accredited with a different organization. In *Schreibman*, the aggrieved plaintiff was accredited as a Gallery correspondent when he worked for the Electronic Public Information Newsletter. It was not until he struck out on his own and created the Federal Information News Syndicate that his accreditation was denied. This focus on the parent publication and the journalist’s income as a deciding factor stifles journalistic and entrepreneurial freedom and undermines competitive journalism.

For instance, if an experienced *Washington Post* reporter decided to start his own news site, the Executive Committee could no doubt deny his application on several grounds. First, under Rule 2, the Committee could find that the journalist was not “employed” because he had yet to earn a salary. The proprietor of a new company might operate at a loss for some time, and if the owner is

---

108 In both *Consumers Union* and *Schreibman* the applicants appealed but cert was not granted.

also the sole employee, he might also forego a salary, choosing instead to reinvest his earnings back into the company until the business is financially secure. Rule 2 conditions accreditation on commercial viability, and allows only the instantly successful publishers—or the independently wealthy—to have access to the Gallery.

Not only is it improper to condition access to information on wealth, it ignores the underlying motivations of journalists.

Journalism has norms that often defy conventional wisdom about rational economic behavior . . . . Historically it has not been a high-paying occupation and . . . it still is not. As a consequence journalism is staffed largely by people who have rejected economic reward as their principal motivation. The rewards they seek come from their peers and their superiors, not the audience or the market.\(^\text{110}\)

By excluding journalists who have deeper motivations than economic concerns, the Executive Committee may be denying access to the most dedicated and altruistic reporters. While it is obviously necessary to set standards for access, money should not be the concern of the guardians of the Gallery. I do not suggest that every blogger with a website should have access, but hinging accreditation on readership rather than income might be more appropriate. In addition, the second part of Rule 2 operates as a catch-22 to prohibit newcomers from the Gallery. How does one demonstrate a necessity for “Washington coverage on a continuing basis” for an upstart publication? In the absence of a track record of congressional reporting, a new publication would likely be denied. Simply declaring its purpose would probably not be enough. While a history of congressional reporting with another publication could be instructive on this point, Schreibman demonstrates that the Committee might not take this fact into account. This rule favoring established media outlets further entrenches traditional media’s control of the Gallery.

Next, the administrative process to change the rules is undemocratic. Rule 6 of the Senate’s Rules Governing the Press Gallery states that “[t]he Standing Committee shall propose no changes in the [sic] these rules except upon petition in writing signed by not less than 100 accredited members.”\(^\text{111}\) However, the

\(^{110}\) Anderson, supra note 16, at 475.

dearth of accredited Internet-only journalists, and the strong economic motivation of traditional newsmen to maintain the status quo, ensures that the requisite number of online-only journalists to change the rules will not soon be reached. Thus, online journalists will not meet the accreditation standards until the rules change, and the rules will not change until the Executive Committee accredits a larger number of online journalists.

**B. Problems with the Administration of the Rules**

This is not the first time that the Correspondents’ Association has felt threatened by outsiders. When the rules were originally drafted at the New York Times Company headquarters in 1879, women and minorities were excluded from the Gallery. Only radio reporters who also reported for daily newspapers would be granted accreditation during the 1920s and 1930s. Nonprofit organizations, the “backbone of a civic society,” were excluded until 1979. Accreditation standards recognizing only print media also stood as an obstacle to the development of broadcast journalism. This history of discrimination demonstrates the need to address the problems faced by online-only journalists as quickly as possible.

First, the lack of genuine independence in the private press demands higher scrutiny of how the Correspondents’ Committee self-regulates the Galleries. Between advertisers, parent companies, and stockholders, the traditional press is constrained by market concerns that are less relevant to the independent online journalist. American newspapers, magazines, and broadcasters generate about $145 billion in revenue each year, more than double the revenue from oil and gas production or agriculture. Many conventional media outlets now have online operations and many more are betting their future on a switch to the Internet. With these bet-the-company strategies, it is easy to see how self-interested business can guide the supposedly impartial decision of who

---

113 See *id.*
114 *Id.*
115 *Id.*
116 Although one might point to the eventual incorporation of each of the excluded populations as a testament to the Executive Committee scheme, the case of Internet journalism is quite different. Unlike the admission of minorities or women, which posed no threat to the *institution* of the press, Internet journalism could deal a significant blow to the institution itself. In addition, newspapers were not attempting to break into the television or radio market. Nowadays, they are attempting to enter the world of Internet news.
should be accredited to the Gallery. In 1880, John Swinton, the managing editor of the *New York Sun* and former chief editorialist of the *New York Times*, made clear the effects that economics played on the newspaper and its employees:

> There is no such thing in America as an independent press, unless it is in the small towns. You know it and I know it. . . . [W]hat folly is this to be toasting an ‘Independent Press.’ We are the tools and vassals of rich men behind the scenes. We are the jumping-jacks; they pull the strings and we dance. Out talents, our possibilities and our lives are all the property of other men. We are intellectual prostitutes.\(^{118}\)

More recently, the “rich men” have become publically held conglomerates which must serve the expectations of investors, analysts, and fund managers. NBC is owned by General Electric, ABC by Walt Disney Co., and CBS by Viacom, and only a fraction of each parent company is dedicated to news gathering and news dissemination.\(^{119}\) News Corp., which owns the New York Post and 175 other newspapers, also owns television stations, sports teams, a book publisher, and a movie studio. Under the existing rules, it is debatable whether these companies actually engage in lobbying activities or whether their chief attention is the gathering and dissemination of news. For instance, should Walt Disney’s activities lobbying for an extension of the copyright protection period,\(^{120}\) prohibit ABC from joining the Gallery? Like *Consumer Reports*, the parent companies here may be engaged in activities which should prohibit membership into the Gallery, but, unfortunately, the rulemakers tend to avoid these issues when it comes to determining the accreditation of their own. These companies seem to have a fiduciary duty to their stockholders to stifle competition, and the current process for drafting the Periodical Press Gallery Rules and the unreviewable manner in which they are administered provides a perfect cover to do just that.

It is clear that the Executive Committee has the right to promulgate and enforce the rules relating to the Gallery as delegated by Congress. It must, however, draft clear rules and

\(^{118}\) Richard O. Boyer & Herbert M. Morais, Labor’s Untold Story 81 (1955).

\(^{119}\) Anderson, supra note 16, at 455.

\(^{120}\) See Lawrence Lessig, Copyright’s First Amendment, 48 UCLA L. REV. 1057, 1065 (2001).
administer these rules fairly. Because of *Consumers Union* and its progeny, the Executive Committee has no incentive to even-handedly administer the rules of accreditation. Well aware that the courts are unwilling to disturb its decision, the Committee is in a position to discriminate against online journalists at will. While proof of discrimination in any single case might be difficult to prove, the system as a whole should be evaluated for constitutional infirmities. As even the *Schreibman* court noted, decisions of the committee could not be predicated on bad faith. The confluence of problems in the rules themselves, the selfish interests of the current Gallery members in excluding Internet journalists, and the unreviewable administration of the rules provides a strong basis for inferring bad faith.

Ever since its inception in the late 1800s, the Correspondents’ Committee has been dominated by representatives of traditional print media. Of the seven members of the current Executive Committee, not a single individual represents an online-only publication. Given the competitive nature of journalism, this poses a significant problem. The press has been, and is now more than ever, a fluid and dynamic institution, with newcomers always contending for membership . . . . [There is a] troublesome risk in today’s environment [that] politically powerful media will capture the process to serve their own ends at the expense of the weaker or less politically engaged segments of the media.

The Committee is also well aware of the growing preference for individualized news tailored to readers’ specific interests and the advantage that smaller publications have in delivering the news to a niche market. In the area of television news, the past decade has proven that broadcast and cable networks which cater to specific interests, such as sports, finance, food, or entertainment, have outpaced the broadcast giants in terms of viewership. The same is true for print media, where metropolitan daily newspapers have lost readers to alternative newspapers, business publications, and national newspapers. Considering the low barriers to entry and low overhead for online journalists and publishers, the Committee’s impermissible desire to

---

122 *Supra* note 31.
124 *Id.* at 469-70.
125 *Id.*
hold on to a competitive advantage in congressional access can surely be inferred.

While the argument can be made that the Committee is discriminating based on the economic status of the Internet publications—a traditional rational basis category—it is the nature of the right that is most important. The court in Sherrill held arbitrary discrimination violated the First Amendment; the decision to discriminate based on market power is at best arbitrary. At worst, it is an example of bad faith because without access, market power is impossible to achieve.

In addition, the current accreditation scheme runs counter to the free marketplace of ideas theory of the First Amendment, in which a multiplicity of viewpoints is viewed as necessary to serve the best interests of the public. The First Amendment “rests on the assumption that the widest possible dissemination of information from diverse and antagonistic sources is essential to the welfare of the public, that a free press is a condition of a free society.” The value of the First Amendment provision for a free press lies in its protection of debate on public issues that should be “uninhibited, robust, and wide-open.”

The current scheme used by the Correspondents’ Association allows members of traditional print media to decide what the public needs to know, by selecting which journalists cover the news, instead of letting market forces determine which news outlets are ultimately successful. The existence of such a “self-appointed elite” is one source of popular dissatisfaction with the press.

Evidence of the failure of the free marketplace of ideas can also be seen in the rising dissatisfaction with the media, as well as the shrinking press coverage of the federal government. “On the whole, mainstream journalism seems to be edging away from the public-interest ideal. Coverage of foreign affairs, government, science, and business has been cut back in favor of coverage of

---

126 Associated Press v. United States, 326 U.S. 1, 20 (1945); see also Red Lion Broad. Co. v. Fed. Commc’ns Comm’n, 395 U.S. 367, 390 (1969) (“It is the purpose of the First Amendment to preserve an uninhibited marketplace of ideas in which truth will ultimately prevail, rather than to countenance monopolization of that market, whether it be by the Government itself or a private licensee.”).


lifestyle, consumption, sports, entertainment, and celebrities."\(^{130}\) Investigative reporting is suffering in the same way, with fewer than one in ten covering issues concerning education, economics, foreign affairs, the military, national security, politics, or social welfare, and over half focusing on lifestyle, behavior, consumerism, health, or entertainment celebrities.\(^{131}\) Given the media’s traditional role as a watchdog for the people, the lack of attention to matters of significance threatens that role. By allowing an injection of new blood into the Gallery, the public will have more options for finding their news; the addition of these new journalists should increase competition and lead to better coverage. Furthermore, congressional news which is important to only certain communities would benefit greatly from the increased variety of journalists in the Gallery. For example, an online publication dedicated to reporting on agricultural issues may be better suited to recognize and report on legislation like farm bills than some traditional print publications.

Currently, newspapers like the New York Times have dozens of seats in the Galleries. However, the marketplace of ideas theory of the First Amendment is best served by a diversified press pool. As Judge Learned Hand noted in United States v. Associated Press:

[The press] serves one of the most vital of all general interests: the dissemination of news from as many different sources, and with as many different facets and colors as is possible. . . . [That interest] presupposes that right conclusions are more likely to be gathered out of a multitude of tongues, than through any kind of authoritative selection. To many this is, and always will be, folly; but we have staked upon it our all.\(^{132}\)

While some may argue that the shift in reporting to soft news is simply a product of consumer demand, evidence suggests that the public seems less inclined to trust the press than ever.\(^{133}\)

\(^{130}\) See Anderson, supra note 16, at 477 (noting that only 513 reporters covered all of the state capitals in 1998, while 3000 reporters were accredited for one Super Bowl).

\(^{131}\) Id. (citing Bill Kovach & Tom Rosenstiel, Are Watchdogs an Endangered Species?, COLUM. JOURNALISM REV., May-June 2001, at 50, 53).


\(^{133}\) Anderson, supra note 16, at 480 ("From 1985 to 1999, the number of people who thought the news media usually get the facts straight dropped from fifty-
Perhaps, with such a wide array of independent journalists covering the same story, readers get one step closer to “the truth,” or at least “the truth” that they want to read. The ability of individual journalists to cover the Congress in their own way will help insure that institutional bias remains at a minimum. Without the crushing overhead of the institutional press, the future of independent online journalism is at a critical impasse. Provided with appropriate access, these journalists have the potential to change the way people receive their news. If the current system is allowed to discriminate against online-only journalism, the problems in coverage will persist and consumer confidence will continue to suffer.

IV. PROPOSED AMENDMENTS TO THE RULES

Of course, every denial of a journalist’s application should not turn into a federal case, but in order to ensure that the decisions of the Executive Committee are made in good faith, the existing Periodical Press Gallery Rules should be amendment. It is important to remember that once a journalist is admitted to the Gallery, his or her behavior still comes under the purview of the Executive Committee. If a journalist is ever found to no longer need access, there is nothing in the Rules which would prohibit the revocation of his or her credentials with appropriate notice and due process. The most important thing at this point, however, is to provide for the opportunity to be admitted in the first place. The proposed amendments are noted in italics.

First, Rule 1 should be amended to alter the selection criteria to permit more online-only journalists admission. The Rule should state that

Persons eligible for admission to the Periodical Press Galleries must be bona fide resident correspondents or independent journalists of reputable standing, giving their chief attention to the gathering and reporting of news. They shall state in writing the names of their employers and their additional sources of earned income. Independent journalists who are self-employed may be permitted

five percent to thirty-seven percent, the number who saw the press as ‘immoral’ rose from thirteen percent to forty percent, and the number who saw the press as lacking in professionalism tripled.”

access to the Press Gallery as a correspondent upon a showing that his or her publication is of the type that would benefit from ongoing Press Gallery access. Additional sources of income derived from sources not explicitly prohibited below should be irrelevant to determining Press Gallery admission.

By altering the Rule in this manner, the Executive Committee will have a textually demonstrable commitment to consider independent journalists. In addition, by allowing alternative sources of income, provided they are not from prohibited sources like lobbying activities, the proposed amendment ensures that access is never predicated on financial success.

Next, Rule 2 should be amended to provide more leniency to online journalists when evaluating their publication’s success. For instance, the Rule could read

Applicants must be employed by, or independently operate, periodicals that regularly publish a substantial volume of news material of either general, economic, industrial, technical, cultural, or trade character. The periodical may be published in print, on the Internet, or a combination of both. The nature of the periodical must be such that Washington coverage on a continuing basis would demonstrably improve the content of the publication. The publication must be owned and operated independently of any government, industry, institution, association, or lobbying organization. Applicants must also be employed by, or independently own and operate, a periodical that is published for profit and is supported chiefly by advertising or by subscription. . . . Online only publications which are distributed free of charge and without advertising revenue shall be permitted access so long as an alternative source of income is not prohibited below . . . 135

135 Pro-Publica is a good example of an online news organization which distributes its content free to other publications and relies on philanthropic contributions and foundation support to survive. See Pro-Publica, About Us, http://www.propublica.org/about (last visited May 28, 2010).
In addition, a few qualifications could guide the Executive Committee in its decisions. For instance, a section could be added that directs the Committee to consider certain factors when deciding whether to grant or deny the application of an online-only publication. Those factors should include:

(a) The number of subscribers to the publication, or the number of “hits” that the publication’s Web site receives on a daily basis;
(b) Whether the type of news material published by the applicant is substantially similar to that of a print publication which would be granted access to the Press Gallery;
(c) Whether the applicant has already been accredited by a different publication;
(d) Whether the applicant’s previous stories have contained independent research; and
(e) The number of publications which maintain multiple accredited journalists in the Press Gallery.

For the last factor, the Committee should consider the costs and benefits of allowing different publications to report the dealings of Congress rather than dozens of journalists accredited from the same institution. The Rules should also be amended to require the Committee to detail the factual findings associated with the decision to deny an applicant’s admission. These findings shall be reviewable and subject to appeal according to the current regulations. Lastly, the Rules should make clear that federal courts have jurisdiction to review decisions denying admission to a journalist after all of the administrative remedies authorized in the statute are exhausted.

By making these changes, the Executive Committee will be forced to give due respect to the emerging medium of online news. In addition, by making the decisions of the Committee subject to judicial review, there is a much smaller likelihood that the applications will be denied based on improper motives or bad faith.

CONCLUSION

While the Rules of the Gallery will probably eventually change to incorporate more online-only journalists, the change should not be so slow as to deny business opportunities to the very reporters that foresaw the shift from print media to the Internet. Cases like Consumers Union ensure that the Executive Committee has no incentive to administer the Periodical Press Gallery Rules fairly, and given the motivations to stifle expanding access, the
First Amendment rights of online reporters will be subjugated without judicial or congressional action. If access to the home of the President and to suspected terrorist suspects in a militarily controlled foreign land is within judicial oversight, so too should the Gallery be reviewable. In the absence of a judicial willingness to oversee this important matter, Congress should take steps to cabin control of the Executive Committee and ensure that online journalism can succeed or fail on its own merits, rather than its lack of access to primary news.
Quantifying the Cost of Substandard Patents: Some Preliminary Evidence

T. Randolph Beard*
George S. Ford, Ph.D.**
Thomas M. Koutsky***
Lawrence J. Spiwak****


Abstract

The purpose of patent policy is to balance the incentive to invent against the ability of the economy to utilize and incorporate new inventions and innovations. Substandard patents that upset this balance impose deadweight losses and other costs on the economy. In this paper, we examine some of the deadweight losses that result from granting substandard patents in the United States. Under plausible assumptions, we find that the economic losses resulting from the grant of substandard patents can reach $21 billion per year by deterring valid research with an additional deadweight loss from litigation and administrative costs of $4.5 billion annually. This brings the total deadweight loss created by our “dented” patent system to be at least $25.5 billion annually. These estimates may be viewed as conservative because they do not take into account other economic costs from our existing patent system, such as the consumer welfare losses from granting monopoly rents to patent holders that have not, in the end, invented a novel product, or the full social value of the innovations lost.

* Senior Fellow, Phoenix Center for Advanced Legal & Economic Public Policy Studies; Professor of Economics, Auburn University, Auburn, Alabama.
** Chief Economist, Phoenix Center for Advanced Legal & Economic Public Policy Studies.
*** Resident Scholar, Phoenix Center for Advanced Legal & Economic Public Policy Studies.
**** President, Phoenix Center for Advanced Legal & Economic Public Policy Studies. The views expressed in this paper are the authors’ alone and do not represent the views of the Phoenix Center, its Adjunct Fellows, or any of its individual Editorial Advisory Board members.
QUANTIFYING THE COST OF SUBSTANDARD PATENTS

TABLE OF CONTENTS

INTRODUCTION ..................................................................................241
I. SOURCES AND COSTS OF A "LOOSE" PATENT SYSTEM..............243
II. THE EQUILIBRIUM LEVEL OF VALID AND SUBSTANDARD
    PATENTS ..........................................................................................250
III. QUANTIFYING THE COSTS OF SUBSTANDARD PATENTS .........255
    A. Lost Patents ..................................................................................257
    B. Other Deadweight Losses .........................................................264
    C. Review of the Evidence ..............................................................268
CONCLUSION ......................................................................................268

INTRODUCTION

Patent policy necessarily involves a balance between encouraging inventors to create new products while simultaneously ensuring that innovations become diffused throughout the economy. Protecting intellectual property is a lynchpin of a vibrant, modern economy, and while the benefits of the patent system are undeniable, the system also imposes significant cost on the economy—even in the best of circumstances. Several high-profile patent disputes, such as the Blackberry\(^1\) and Microsoft MP3\(^2\) cases, have sparked a debate as to whether the U.S. patent law system adequately promotes the interests of inventors or whether the system is a legal quagmire that stalls new innovation in excessive litigation.\(^3\) When a patent system grants substandard patents or provides overly permissive legal remedies for patent holders, the protection of intellectual property can create substantial net loss of economic welfare. We envision a "substandard patent" as one that is not privately profitable to pursue in the absence of litigation opportunities afforded through inevitable imperfections in the legal system. Although such a patent does not protect an invention worth protecting from the social point of view, they may be privately profitable because the sometimes afford the opportunity to obtain payments from holders

\(^1\) The dispute between patent holder NTP Inc. and BlackBerry smartphone manufacturer Research in Motion Ltd. resulted in a settlement of $612.5 million. See Important Dates in BlackBerry Patent Case, http://www.msnbc.msn.com/id/11409695 (last visited Apr. 15, 2010) (providing a detailed timeline of the case).


of sufficiently similar or related “real” patents which are practiced. So, while a well-functioning patent system will necessarily balance the benefits of innovation with the costs of monopoly, a defective system adds to the social costs of patent monopolies the additional deadweight losses arising from reduced innovation and from the wasted resources directed at securing and protecting substandard patents, without providing any offsetting benefit. In economic terms, such substandard patents represent options permitting transfers of wealth from the holders of legitimate patent rights to those holding related, substandard patent claims. Their social value is low (or zero), yet their private option value rises with defects in the patent system.

The economic costs of substandard patents are highlighted by (but by no means limited to) “patent troll” litigation, to which a substandard patent regime can give rise. “Patent troll” litigation is one form of litigation arbitrage—it will exist in areas in which patents are relatively easy to obtain and the consequences to a defendant accused of infringement of losing a patent suit can be enormous and irreversible, such as an injunction against any future sales of a successful yet potentially infringing product. The presence of this arbitrage indicates that the current patent licensing and enforcement system are in need of reform and a thoughtful rebalancing of incentives.

In this Article, we attempt to quantify in a preliminary manner a portion of the cost to the United States’ economy of substandard patents granted by the United States Patent and Trademark Office (USPTO). In particular, we focus upon deadweight losses that result from the impact that a “loose” patent system that unduly grants “substandard” patents has upon innovation and the development of important, valid patents. These costs are deadweight losses and not merely transfers, so they reduce overall economic welfare. In the United States, we estimate that the deadweight loss of a “loose” patent system from lost innovation is approximately $21 billion each year in private costs alone, or nearly $200 per household per year. This sizeable deadweight loss constitutes approximately 7% of annual Research and Development (“R&D”) spending. Deadweight losses from litigation and administrative costs from substandard patents constitute an additional $4.5 billion annually, or 1.5% of the country’s annual R&D spending.

Our findings are described as preliminary, since there is very limited data upon which to base our estimates. However, we believe that our methods render conservative estimates because we do not take into account a number of other costs created by substandard patents. Most notably, in cases where a substandard patent allows a firm to enforce monopoly prices without truly
innovating, there is a welfare loss without commensurate benefit that our model does not attempt to quantify. We also ignore the fact that innovation has a greater social benefit than private benefit, so the social costs of lost innovation stand to be much larger than the $21 billion in annual private costs from lost innovation that we estimate.⁴

In Part I, we provide a brief description of the general problem of substandard patents and their causes and consequences. Our discussion is succinct, since there are many studies on this issue that are readily available to interested parties. In Part II, we explain an important component of our model, which focuses on the important interactions between the equilibrium level of “valid” and “substandard” patents. We show that substandard patents impose deadweight losses on the economy as a whole because they deter innovation and the development of important, valid patents. This idea serves as the basis for the estimation that we perform in Part III. Part III also contains a sensitivity analysis to allow the inputs to vary over the range of plausible values. Our findings are summarized in the Conclusion.

I. SOURCES AND COSTS OF A “LOOSE” PATENT SYSTEM

A well-functioning patent system engages in a delicate balance. In order to “promote the progress of Science and useful Arts,”⁵ a patent holder is granted a legal exclusive monopoly to an invention for a limited period of time. It is thought that granting monopoly profits to patent holders would direct societal resources towards scientific and useful innovations. Thomas Jefferson once wrote that patent law is about “drawing a line between the things which are worth to the public the embarrassment of an exclusive patent and those which are not.”⁶

What Jefferson calls the “embarrassment” of a legal patent monopoly, economists would call a social cost. By definition, the granting of a monopoly reduces output and causes a net loss in consumer welfare. The traditional justification for patent rights is predicated upon the assumption that without such monopoly rights, society will not achieve the optimal rate of innovation because innovations and scientific discoveries are, absent patent rights, often public goods that provide limited or no opportunity for the

---

⁵ U.S. CONST. art 1, § 8, cl. 8.
inventor to recover the costs of discovery. If every invention could be immediately copied, then few firms would invest the resources necessary to invent new products. Absent patent rights, an inventor also would have an incentive to prevent others from learning about any new discovery.⁷ A patent attempts to remedy these problems by giving the inventor the legal right to collect some portion of the social value attributable to the invention while inducing disclosure of the details of the invention to the public.⁸ This disclosure, in turn, likely increases innovative activity in that area due to increased information.⁹

At the same time, granting too much protection to inventors (or granting it too easily) can hamper the creation and diffusion of technology throughout the economy. Achieving an adequate balance of rights to compensate true innovators and fostering the use of patented technology is the goal of a well-functioning patent system. A patent regime that makes it too easy to obtain and enforce a patent could create too many of these monopoly "embarrassments" that would reduce economic welfare by virtue of their monopoly status yet not promote economic welfare because they do not reward true innovations.¹⁰ As the Supreme Court stated in 1950, the granting of patents for obvious and known methods "withdraws what is already known into the field of its monopoly and diminishes the resources available to skillful men."¹¹

There are several ways in which substandard patents can impose economic and welfare costs on the economy. As we describe in Part II below, a "loose" patent system—that is, a patent system that permits large numbers of substandard patents—causes deadweight economic losses because the presence of substandard patents diminish the overall level of innovation and development of valid patents. A valid patent, like a substandard patent, confers a

---

⁸ Kitch adds that patents also promote efficiency by deterring others from engaging in wastefully duplicative efforts of re-inventing the same technology. Edmund W. Kitch, The Nature and Function of the Patent System, 20 J.L. & ECON. 265 (1977); see also Dam, supra note 7, at 266-67.
QUANTIFYING THE COST OF SUBSTANDARD PATENTS

monopoly right to the holder, but a valid patent does so only in the case of genuine innovation. Valid patents then have a positive social value in addition to a private value, while substandard patents have only the private value. The deadweight losses occasioned by substandard patents will cause resources to be allocated inefficiently and therefore affect the entire economy. In addition, a “loose” patent system that grants large numbers of substandard patents also causes other inefficiencies and misallocations of resources, as such a system would:

- Cause consumers to absorb monopoly prices over “inventions” that were already effectively common knowledge;\(^\text{12}\)
- Direct resources away from productive research and instead towards strategic accumulation of patents already filed over innovations already deployed;\(^\text{13}\)
- Divert resources to “defensive patenting” or securing offensive “blocking patents;”\(^\text{14}\)
- Direct research away from areas of existing patents that should not have been granted;\(^\text{15}\)
- Direct resources toward acquiring and enforcing substandard patents and collecting royalties rather than productive fields of economic activity.

Given this potential for misallocating resources and the other costs, a well-functioning patent law regime should tailor the scope of the legal patent monopolies so that the harms described above are outweighed by the benefit to society from the economic innovation which results from those patent monopolies. As stated by Lévéque and Ménière, the “simple criterion” of economic welfare “helps define the elements of an optimal patent.”\(^\text{16}\)

Whether the United States patent system is “too loose” today is the subject of substantial debate. The claimed shortcomings of the USPTO and the United States court system are numerous and appear to stem primarily from a poor legal

\(^{12}\) “This deadweight loss reduces the total surplus created by the innovation at least during the lifetime of the patent.” FRANÇOIS LÉVÉQUE & YANN MÉNIÈRE, THE ECONOMICS OF PATENTS AND COPYRIGHT 21 (2004).

\(^{13}\) Jaffe & Lerner describe a number of such activities including the sealed crustless sandwich and the perpetual option pricing formula of Vergil Daughtery. JAFFE & LERNER, supra note 3, at 32, 145-47.


\(^{15}\) See Gallini, supra note 10.

\(^{16}\) LÉVÉQUE & MÉNIÈRE, supra note 12, at 43.
framework and an understaffed and overworked agency. The debate often centers around the patenting of “inventions” such as a method for swinging on a swing, the sealed crustless sandwich, a financial technique developed four decades prior to patenting by academics unaffiliated with the patentee, and anti-gravity flying machines. As observed by Magliocca, the United States court system, in many ways, exacerbates the problem, brought to light by the explosion of “patent troll” litigation. Patent trolls engage in a very specific arbitrage opportunity and thrive in certain conditions in which patents are easy to obtain and keep, the costs of defending a patent suit are great, and the risk to a defendant of losing a patent suit are enormous because the defendant “cannot easily substitute away from the disputed technology.” Trolls thrive in situations in which patents are easy to get and damages uncertain. As Justice Kennedy observed in the eBay decision, a patent remedy such as a permanent injunction against an infringer “can be employed as a bargaining tool to charge exorbitant fees to companies that seek to buy licenses to practice the patent.”

As a result, almost unique among industrialized nations, United States companies face a plethora of patent suits brought by plaintiffs with arguably substandard patents. There are some

17 JAFFE & LERNER, supra note 3, at 145-47; U.S. Patent No. 6,960,975 (filed Mar. 14, 2005) (granting patent for a “space vehicle propelled by the pressure of inflationary vacuum”).
19 On the other hand, some commentators argue that “patent trolls” serve a useful purpose, most notably by providing liquidity to inventors as well as expertise in policing infringement. See, e.g., Steven Rubin, Hooray for the Patent Troll, IEEE SPECTRUM, Mar. 2007, http://spectrum.ieee.org/consumer-electronics/gaming/hooray-for-the-patent-troll (“[P]atent-holding companies provide another way, and sometimes the only way, for an inventor to monetize his patent. They foster innovation by making it possible for small companies and individual inventors to spend their time in research and development, knowing that if a patent does issue, they will not necessarily have to commercialize or litigate it. They can spend time doing what they are good at—enventing.”). But see Zachary Roth, Patent Troll Menace, WASH. MONTHLY, June 1, 2005, http://www.washingtonmonthly.com/features/2005/0506.rothsidebar2.html (“The reason this business is attractive to people such as Lockwood is simple: Trolling makes money. . . . [E]ven though his patent was overturned, Lockwood still got to keep the licensing fees he had extracted from other targets that chose not to fight.”).
21 Patent trolls are largely a U.S. phenomenon. See, e.g., Joe Brennan et al., Patent Trolls in the U.S., Japan, Taiwan and Europe, 13 CASRIP NEWSLETTER, Spring/Summer 2006, http://www.law.washington.edu/Casrip/Newsletter/default.aspx?year=2006&article=newsv13i2BrennanEtAl. We do not claim that most litigation by patent trolls is brought through substandard patents, only that trolls do bring such cases. That is why the pejorative “troll” is used.
signs that reform is brewing. Several recent Supreme Court decisions have addressed the standards for granting and challenging a patent\textsuperscript{22} and trimmed back lower court rulings that had increased the business risk and harm from losing a patent lawsuit.\textsuperscript{23} The Patent Reform Act of 2007, directed at improving patent quality and changing patent remedies, has been approved by the House and Senate Judiciary Committees. Analyzing and understanding the economic welfare costs of the current United States patent system is clearly of importance to policymakers as they consider these reform proposals.

The welfare costs of the current United States patent regime can be estimated empirically by comparing the valid patent output of our regime to the patent system in Europe. In contrast to the United States, the European patent system, while certainly not perfect, has a relatively “tighter” standard for granting patents and the process is administered and enforced differently as well. By this discussion we do not mean to imply that the European patent system is better than the United States system or that it should be adopted here, but only to assert that the two legal regimes are different in a way that allows us to perform an empirical analysis of the current United States patent regime.

To obtain a patent in the United States, the invention must be new, useful and non-obvious.\textsuperscript{24} In the United States, unlike some other countries, the process for granting a patent is usually confidential and solely between the applicant and the USPTO, and other parties are not permitted to intervene or oppose a patent application.\textsuperscript{25} Moreover, the USPTO cannot simply reject a patent

\textsuperscript{22} In \textit{KSR International Co. v. Teledex Inc.}, 550 U.S. 398 (2007), the Court tightened the Federal Circuit’s test for patentability, in particular the “obvious” standard. In \textit{Medimmune, Inc. v. Genentech, Inc.}, 549 U.S. 118 (2007), the Court overturned a Federal Circuit ruling that limited the ability of patent licensees to subsequently challenge the validity of a patent.

\textsuperscript{23} In \textit{Microsoft Corp. v. AT&T Corp.}, 550 U.S. 437 (2007), the Court overturned a Federal Circuit ruling that held Microsoft liable for computers manufactured and programmed abroad with software that infringed a United States patent. In \textit{eBay}, the Court ruled that traditional equitable principles should apply in patent disputes with regard to the granting of injunctions against infringing products; prior to that decision, lower courts had followed a “general rule” of always issuing such an injunction without considering the public interest.

\textsuperscript{24} 35 U.S.C. \S 101 (2006) (new and useful); \textit{id.} \S 103 (non-obvious). For a recent Supreme Court discussion of the obviousness test, see \textit{KSR}, 550 U.S. 398.

\textsuperscript{25} Patent applications in the United States are not necessarily made public until after a patent is issued. 35 U.S.C. \S 122(a). Applications are not made public until 18 months after filed. \textit{id.} \S 122(b)(1)(A). Within two months after publication, third parties may submit prior art related to patentability. Manual of Patent Examining Procedure 1134.01. Upon issuances of a patent, the protection has a term of twenty years from the date on which the application was filed, but only upon issuance of a patent does the information disclosed in the application becomes a matter of public record. 35 U.S.C. \S 154(a). Until 1995, the term of a
application; it also bears the burden of making a prima facie case that explains the reasons for rejection. Third parties do not have the right to participate in the patent application process and patents can only be challenged after a grant in limited instances, consisting of challenges based on prior art found in patents or printed publications. Moreover, in some instances, challenging a patent creates potential for the challenger to be estopped from asserting certain defenses in an infringement suit. Finally, in a suit for patent infringement, a plaintiff may obtain injunctive relief and damages, which may include lost profits due to the infringement or a reasonable royalty.

Pursuant to the European Patent Convention, which harmonizes the patent laws of its signatories, twenty-year patents are available for “any inventions . . . provided that they are new and which involve an inventive step and are susceptible of industrial application.” The standard for patentability in Europe, while similar to the United States in some respects, is different in other respects, particularly with regard to the European requirement that an invention be of a “technical” nature. In addition, patent applications in Europe are made public even if they have not been issued and the method for challenging a patent differs. As a result, a patent application in Europe is three times more likely to be opposed than a patent is to be reexamined in the United States. Of all the potential remedies to the United States patent system, remedying post-grant review process to adopt an

---

application; it also bears the burden of making a prima facie case that explains the reasons for rejection. Third parties do not have the right to participate in the patent application process and patents can only be challenged after a grant in limited instances, consisting of challenges based on prior art found in patents or printed publications. Moreover, in some instances, challenging a patent creates potential for the challenger to be estopped from asserting certain defenses in an infringement suit. Finally, in a suit for patent infringement, a plaintiff may obtain injunctive relief and damages, which may include lost profits due to the infringement or a reasonable royalty.

Pursuant to the European Patent Convention, which harmonizes the patent laws of its signatories, twenty-year patents are available for “any inventions . . . provided that they are new and which involve an inventive step and are susceptible of industrial application.” The standard for patentability in Europe, while similar to the United States in some respects, is different in other respects, particularly with regard to the European requirement that an invention be of a “technical” nature. In addition, patent applications in Europe are made public even if they have not been issued and the method for challenging a patent differs. As a result, a patent application in Europe is three times more likely to be opposed than a patent is to be reexamined in the United States. Of all the potential remedies to the United States patent system, remedying post-grant review process to adopt an

---


248
approach more consistent with the European Patent Office (EPO) is by far the most commonly mentioned.  

Remedies for patent infringement are also different between the United States and Europe, and these differences appear to encourage more litigation in the United States. In particular, European law tends to favor payment of license fees and damages instead of injunctions over the future sale of infringing products, which have been more common in the United States. For example, in the United States, patent holders do not have a duty to license and their licensing actions are limited only by antitrust law, but in the United Kingdom, France, and Germany, compulsory license statutes require patent holders to license their products. While the Supreme Court in 2006 took action to limit the scope of permanent injunctions in patent disputes, injunctive relief is still available to patent holders in the United States. Jury trials to enforce patent rights and establish damages are not guaranteed in Europe as they are in the United States. In the United States, patent litigators often get a second bite of the apple as well,

---


35 Stevenson, supra note 30, at 10.

36 In the eBay decision, the Court reversed the court of appeals “general rule” unique to patent disputes “that a permanent injunction will issue once infringement and validity have been adjudged.” eBay, Inc. v. MercExchange, 547 U.S. 388, 393-94 (2006). According to Chief Justice Roberts, “[i]n the early 19th century, courts have granted injunctive relief upon a finding of infringement in the vast majority of patent cases.” Id. at 394 (Roberts, C.J., concurring).

because the United States Court of Appeals for the Federal Circuit reviews all patent claim determinations under a de novo standard.\(^{38}\)

Patent suits in the United States are also generally more expensive than in Europe. Estimates indicate that the costs of a patent lawsuit through discovery are about $2 million for each side.\(^{39}\) These costs are substantially higher than that in several European countries. In Germany, for example, the cost of a suit ranges from approximately $30,000 to $80,000.\(^{40}\)

We outline these differences between the European and United States patent regimes not to imply that the European regime is somehow preferable to the United States system, but simply to demonstrate how the United States system maintains a relatively “looser” patent system than Europe, thereby permitting more substandard patents. (The looseness of the patent system is logically separate from the costs of litigating patent claims, yet the high costs evident in the United States may strengthen the hand of patent trolls in pretrial negotiations.) That distinction is important as it serves as the basis for our estimation of deadweight losses and other costs described in Part III below. As described below, a “loose” patent system discourages the development and filing of valid patents and creates a deadweight loss for the economy. We base this estimate on the fact that even with a tighter legal standard for patentability, Europe produces a higher share of “valid” patents, relative to substandard patents, than the United States. Therefore, while the European system has been criticized as being too “tight,” its relatively more stringent granting practices allow it to serve as a basis for our estimation approach.

**II. THE EQUILIBRIUM LEVEL OF VALID AND SUBSTANDARD PATENTS**

The basis of our argument is that substandard patents arising from a “loose” patent system reduce the number of valid patents by discouraging innovation. While this idea is generally accepted, we formalize it here by describing the “correct” level of patenting in the sense of the equilibrium values of valid and substandard patents. To begin, we divide total patents into two


QUANTIFYING THE COST OF SUBSTANDARD PATENTS

types: (1) valid patents \( (v) \) and (2) substandard patents \( (b) \). Total patents are just \( v + b \). Valid patents represent patents that are true inventions or discoveries in which the cost to society of granting a twenty-year monopoly to the patent holder are outweighed by the aggregate social benefit of the invention or discovery itself. Substandard patents are those which are granted that are of low quality (that is, for ideas that are not in fact new or non-obvious) and which create risks for valid patents through litigation and licensing. From the standpoint of our model, it is not necessary to assume that each and every patent of this sort is literally bogus, is created for a nefarious purpose, or is of no independent private value to the patentee. Our analysis simply assumes that such patents fall below an operative or ideal standard for approval and that they impose, on average, a harm to the economy as a whole and, specifically, a harm to so-called valid patents. More directly, the addition of a substandard patent reduces the private marginal benefit of a valid patent, and discourages the production of both valid and substandard patents (or inventions in general, whether patented or not).

The relationships between the flows of valid and substandard patents, and the assumed forms of the marginal benefits of these activities among agents in the economy, lie at the heart of our analysis, and therefore merit a brief discussion. First, valid patents, although they are differentiated by definition, are assumed here to be, broadly speaking, “competing” with one another. In other words, a valid invention will, on average, make money for its owner, but the product or service supported by the invention in question competes for the consumers’ attention with all other products offered in the market. In a market populated by many innovative products, any individual product is, on average, less likely to make a high return. Thus we assume, as is common in economic analysis, that the marginal benefit of an additional valid patent decreases as more valid patents are awarded. This feature of the return to innovative activity is also consistent with the observation that high value projects are pursued “first”, i.e., in almost all cases, while lower valued projects are funded only under more favorable conditions. The same logic applies to the diminishing marginal values of substandard patents.

The notion that increased numbers of valid patents will increase the marginal value of a substandard patent arises from the conceptualization of a substandard patent. Such a patent has value primarily (or solely) from its potential to support a patent claim against a valid, profitable patent. The more valid patents there are, the greater the probability such a claim can be constructed given an arbitrary substandard patent. Thus, the holder of a substandard patent would welcome increased numbers of valid patents, as this
would raise her opportunity for privately profitable litigation. In a similar way, an increase in the number of substandard patents will adversely affect the expected profitability of a valid patent, since the risk of costly patent claims will increase.

To describe the equilibrium, we assume that patents (or patentable inventions) are ordered from highest to lowest, with resources devoted first to those patents with the greatest value. With diminishing marginal benefits to patents, the equilibrium number of valid patents, \( v^* \), will solve:

\[
A(x) - a \cdot v - c \cdot b = 0
\]  

where the expression is the net private marginal benefit of an additional valid patent.\(^{41}\) The total benefit of valid patents is maximized where the marginal benefit of a valid patent is zero. The net marginal benefit includes a value \( A(x) \), which is a function of exogenous factors \( x \) such as the legal system for granting or challenging patents and the cost of enforcing patents. The parameter \( a \) measures the reduction in the net marginal benefit of valid patents given the addition of one more valid patent, and the negative sign indicates diminishing marginal benefits.\(^{42}\) The relationship between the number and creation of substandard patents and the marginal value of a valid patent is measured by the parameter \( c \). Substandard patents, on the other hand, reduce the net marginal benefit of valid patents. This consequence arises primarily from opportunistic litigation or licensing. Although they both represent a reduction in the value of a valid patent, both \( a \) and \( c \) are expressed as positive values.

The equilibrium number of substandard patents, \( b^* \), will solve:

\[
B(y) - d \cdot b + e \cdot v = 0
\]  

where \( B(y) \) is a scale factor for the net private marginal benefit of substandard patents, and its value is driven by a set of factors \( y \). The factors \( y \) will generally not be the same as \( x \), but some overlap is to be expected. Intuitively, \( y \) will encompass factors that measure the strength of the jurisdictional patent review process, the efficiency of the legal system, the generosity of patent

\(^{41}\) In both Equation (1) and Equation (2), we have linear marginal benefits, but this assumption is not required and is for convenience only.

\(^{42}\) This reduction in marginal benefit presumably occurs because patent opportunities are exploited in order of decreasing expected net value, although the formulation is not inconsistent with the existence of an additional effect that reflects an actual reduction in the economic value due to competition between products which is increased or facilitated by new inventions.
infringement awards, legal costs, and so on. The parameter $d$ is the effect on net private marginal benefits from the addition of one more substandard patent. And while $d$ is positive, the negative sign implies diminishing net marginal benefits. In contrast to Equation (1), an increase in the number of valid patents increases the net private marginal benefit of a substandard patent (because this increase creates more opportunities for litigation). The equilibrium number of both types of patents is determined by the condition that the marginal benefits of each are simultaneously equal to zero in the relevant jurisdiction.

Figure 1 is a graphical representation of the equilibrium. The figure has the number of substandard patents ($b$) on the vertical and valid patents ($v$) on the horizontal axis. The curves in the figures represent the loci of points where the marginal benefits of valid and substandard patents equal zero across the range of values of both $b$ and $v$ (i.e., iso-marginal benefit curves). The intersection of the two defines the equilibrium. In the figure, the equilibrium has $b^*$ and $v^*$ patents.

To demonstrate the comparative statics of the model, consider a court decision that makes the granting of substandard patents more difficult, such as the recent Supreme Court decision in *KSR* that overturned lower court’s permissive interpretation of the “non-obvious” test for patentability. In the model, this legal change is represented by a change in $y$ that reduces $B(y)$. Consequently, the number of substandard patents should diminish. In Figure 2, we illustrate this as a change in $y$ to $y'$, causing a shift in the upward sloping iso-marginal benefit curve for substandard patents down and to the right. The new equilibrium is $b^{**}$ and $v^{**}$.
where substandard patents fall and valid patents rise. Given our observation above that substandard patents diminish the value of valid patents and, therefore, reduce the incentive for firms to obtain such valid patents, upon the legal change that decreases the number of substandard patents, the number of valid patents will be expected to rise.

The comparative statics of the other parameters are similarly intuitive. Put simply, anything that increases the value of valid patents increases both $b$ and $v$. Any change that increases the value of substandard patents reduces $v$ and increases $b$. For example, a change in $x$ that makes valid patents more difficult to enforce will shift the iso-marginal benefit curve ($MB_v = 0$) down and to the left, thereby reducing both equilibrium valid and substandard patents.

The most important point about this analysis is that it illustrates an aspect of the problem of the United States patent system that has received insufficient attention. In particular, since the numbers of both types of patents affect the marginal values of each, any policy change that affects either relationship will, in equilibrium, affect the numbers of both types. Of special potential concern is the size of the effect of substandard patents on the values of valid patents. To the degree that valid patents, as described here, have much larger net social values, a patent system that allows too many substandard patents is likely to reduce the extent of innovation valid patents support, reducing economic welfare. This dampening effect may be far more important than the direct costs of litigation and licensing, much of which will necessarily involve transfers. While direct legal costs are surely not
de minimis, the discouragement of innovation, in the long run, will almost certainly swamp these more easily counted “direct” costs.

We also note that the underpinnings to our approach are conceptual. For example, we note that there is no need to assume that the socially optimal number of substandard patents need be zero, especially because there are costs associated with reducing substandard patents. For instance, adopting a patent regime that sets the bar high for granting any patent would certainly reduce or even eliminate the level of substandard patents, but that decision also could reduce the number of valid patents as well and therefore impose welfare losses on the economy. The administrative costs (and risk of mistakes) of sorting through valid and substandard patents may also be extraordinarily high. Like most things in economics and public policy, such a complete foreclosure of substandard patents would probably be too costly to be optimal in the real world. As a result, the efficient balance should be sought, and that appropriate balance is what a good patent policy must continually strive to achieve.

In addition, we are not assuming, and do not suggest, that substandard patents are intentionally created to use in opportunistic, socially destructive litigation or royalty seeking. It seems probable that very few patents are created with that primary end in mind. Rather, when the patent system is sufficiently “loose” in granting patents, and patents are had cheaply enough, firms and others will patent devices and procedures that are of limited commercial potential. In such cases, the possibility of obtaining an infringement award, or of licensing to others seeking legal defense, becomes a non-negligible consideration that encourages the patentee to proceed. Such expectations, of course, need to be accurate in equilibrium, so it must be the case that some opportunistic exploitation occurs. Since a patent is often an alternative to other means of protecting intellectual property, such as trade secret activity, one would expect that an increase in the number of potentially threatening patents would reduce the marginal benefit of a “valid” patent effort.

III. QUANTIFYING THE COSTS OF SUBSTANDARD PATENTS

As discussed above, the presence of substandard patents leads to a reduction in the number of valid patents. In this Part, we attempt to quantify the loss of valid patents in the United States due to substandard patents, and then put a monetary value on that loss. As a first step, we estimate the number of valid patents lost to substandard patents. To do so, we assume, as have others, that triadic patents—i.e., those in which the inventor seeks patent protection in the United States, Europe, and Japan—are “relatively
important" patents and are, to a large extent, generally regarded as "valid" patents. The validity of such patents is based on the fact that the patent must be granted by three patent offices: the USPTO, the EPO, and the Japanese Patent Office (JPO). By most accounts, the EPO is the most stringent in its requirements and evaluations, and we use that presumption below to specify some parameters of our estimation approach. We do not intend, however, to imply that the European patent system is the “correct” system. Rather, we assume, given the differences in the legal regime, that patents issued by the EPO are less likely to be substandard patents. This assumption, however, does not imply that the EPO system is in some sense ideal or perfect.

It is of course arguable that U.S. patents are, in some sense, qualitatively different from those in the European Union or Japan. This could be the case, for example, if the U.S. market were more important because of its size or profitability. Then, one might imagine valid inventions being patented only in the U.S. However, we find this conjecture unconvincing for several reasons. First, it seems unlikely one would have a patent for an invention that was profitable in the United States, but not in the European Union, for example. It is easier to imagine a profitable invention patented only in Japan, where cultural practices and relative prices are quite different (for example, it seems likely that a golf driving range device meant to be used atop tall buildings would be of more value in Japan). One might refer to this possibility as the "enculturation" of inventive activity.

More importantly, however, it seems undeniable that triadic patents are highly likely to be important, valid patents, even if there are valid patents that are not registered in all three jurisdictions. In this case, variations in the rates of triadic patents between jurisdictions will still provide a relative measure of lesser quality patents, although perhaps not universally substandard ones. Then, looking at the high rate of patents granted in the United States relatively to U.S. investment in R&D, one is forced to conclude either that: (1) the United States is highly efficient in R&D given its investments, or (2) the U.S. system generates a large number of dubious patents relative to the other jurisdictions. We base our calculations on the second of these interpretations.

Substandard patents are harmful in (at least) three respects. First, substandard patents may reduce future innovation

43 JAFFE & LERNER, supra note 3, at 143.
by discouraging research and development in a particular area for fear of infringing, or directing research away from valid to substandard opportunities.\textsuperscript{45} This reluctance to enter could affect market structure and prices. Second, substandard patents may induce unnecessary licensing royalties, distorting the incentives the patent system was designed to provide. Third, legal challenges to substandard patents can result in socially wasteful litigation costs. Our focus here is on the first harm, and we attempt to estimate the welfare losses from high numbers of substandard patents. We believe these costs will be the largest of the three, and our rough estimates of the other costs indicate that this is true.

A. Lost Patents

The presence of substandard patents clearly reduces the incentives for firms to innovate.\textsuperscript{46} Yet, there is no direct evidence of which we are aware on the precise extent of research deterrence. In an effort to approximate the number of lost “valid” or “relatively important” patents lost due to the presence of substandard patents, we assume the production of relatively important patents is a linear function of R&D expenditures.\textsuperscript{47} Thus,


\textsuperscript{46} Jean Lanjouw & Josh Lerner, Tilting the Table? The Use of Preliminary Injunctions, 44 J.L. & ECON. 573 (2001); Josh Lerner, Patenting in the Shadow of Competitors, 38 J.L. & ECON. 463 (1995). Hunt claims that the weakening of the non-obviousness requirement by domestic courts lead to more but weaker patents, and discouraged R&D activity. Robert Hunt, Nonobviousness and the Incentive To Innovate: An Economic Analysis of Intellectual Property Reform (Fed. Reserve Bank of Phila. Working Paper No. 99-3, 1999); see also T.S. Ellis, Judge, Distortion of Patent Economics by Litigation Costs, Address at the 1999 CASRIP Summit Conference, in 5 CASRIP PUBLICATION SERIES: STREAMLINING INT’L INT’L PRO. 22 (1999), available at http://www.law.washington.edu/casrip/Symposium/Number5/pub5atcl3.pdf. (“My thesis today is neither revolutionary nor abstruse. On the contrary, it is no more than a modest, straightforward, common-sensical observation that has likely already occurred to many veteran viewers of the patent scene. It is, simply put, that the escalating, indeed skyrocketing litigation costs of the 1970’s and 1980’s have distorted patent markets and patent economics. Put another way, it is my observation that the escalating costs associated with litigating patent infringement and validity issues discourage challenges to patents, thereby essentially equating the entry barriers for presumptively valid, but weaker patents with those entry barriers associated with strong or judicially tested patents.”).

\textsuperscript{47} We are also evaluating log transformations of the variables using comparable \( R^2 \) values, but the linear specification is superior to these alternatives.
the number of triadic patents filed by country \( i \) in period \( t \) is described by

\[
F_{i,t} = \beta \cdot RND_{i,t} + \theta \cdot DUS + \sum_{j=1}^{T} \alpha_j D_{j,t} + \epsilon_{i,t}
\]  
\text{(3)}

where \( F_{i,t} \) is the number of triadic family patents for country \( i \) in period \( t \), \( RND_{i,t} \) is the real research and development expenditures for country \( i \) in period \( t \), \( DUS \) is a dummy variable with a value of 1 for the United States (0 otherwise), the \( D_{j,t} \) are \( T (\leq \Sigma t) \) period specific dummy variables, the \( \alpha_j \) are estimated coefficients, and \( \epsilon_{i,t} \) is the econometric disturbance term.\(^{48}\) The coefficient \( \theta \) measures the extent to which the United States either under- or over-produces valid patents relative to other countries. Triadic patents are measured by the OECD using applications at the EPO and JPO and grants at the USPTO.\(^{49}\)

All of the data required to estimate Equation (3) is from the Organisation for Economic Co-operation and Development’s (OECD) Main Science and Technology Indicators. The variables are expressed in annual terms and cover the period 1995 through 2003. The variable \( RND \) is measured in two ways: civil R&D and total R&D (both civil and defense), and both are measured in real dollars (in millions).\(^{50}\) There is some evidence suggesting that defense spending on R&D generates few patents, and the United States spends far more on defense R&D than any other country in the sample.\(^{51}\) Thus, using only Civil R&D (i.e., total R&D less defense-related R&D) provides a more conservative estimate of the number of lost valid patents. To demonstrate the conservative nature of using only Civil R&D expenditures, we also present the results with the variable R&D measured using total R&D expenditures for comparison purposes. Given the large number of missing observations on the share of Civil R&D spending, the

\(^{48}\) Data required to estimate this equation is from the OECD’s Main Science and Technology Indicators (Subscription Service). The variables are expressed in annual terms and cover the period 1995 through 2003, and with missing variables the sample size is 299 observations. The model is estimated using least squares with period dummy variables to account for the time series nature of the data. The coefficients are highly statistically significant and are \( \beta_0 = -396.6, \beta_i = 0.10, \) and the year 2003 constant is -\textbf{105.2}. Overall, the model performs well, with an \( R^2 \) of 0.80. The linear specification fits the data quite well, much better than either the log-lin or log-log specifications.

\(^{49}\) Application data at the USPTO was not available prior to 2003, leading to this definition of triadic patents by the OECD. ORG. FOR ECON. CO-OPERATION & DEV. [OECD], PATENT STATISTICS MANUAL 71 (2009).

\(^{50}\) The GDP deflator is provided in the OECD data for those countries included in the sample.

share variable is assumed to be constant over the sample period (based on the average of available data). We do not suspect this will bias the results significantly, since the available data suggests the civil share across all countries is very stable over time.

The OECD data provides data on thirty-eight countries, although there are some missing observations. We present the results of the estimation using three sets of countries. Sample A includes thirty countries with 227 total observations. This sample includes all countries for which the necessary data is available. For Sample B, we include only countries in the European Union and the United States, since the patent and legal regimes in these countries are more likely consistent with that of the United States. Finally, in Sample C, we include all thirty-eight countries available. In this sample, however, we are limited to total R&D expenditures in nominal terms due to a lack of data. The results from this sample are provided for illustrative purposes only, and we do not discuss them in detail.

The model is estimated using least squares with period dummy variables to account for the time series nature of the data. Table 1 summarizes the results. Alternative procedures for computing the standard errors render no significant changes, so the t-statistics are based on the ordinary standard errors. All the variables are statistically significant. The model fits the data very well, with $R^2$ values of about 0.97 across all specifications (except for Sample C, with an $R^2$ of 0.92). The good fit is not surprising given the time series component of the data.

Turning to the number of lost patents (measured by $\theta$ in Equation (3)), across Samples A and B we observe similar estimates when using Civil R&D expenditures. The more conservative number, and probably the more sensible one given the comparison is across the United States and EU countries only, is

---

52 For every year the data is available, we compute the ratio of Civil to Total R&D (as a percentage of GDP), and then average these for each country.

53 We have 30 countries and 9 years of data for a total of 270 potential observations, but there are missing values. The countries in Sample A include Argentina, Australia, Austria, Belgium, the Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, New Zealand, Norway, Portugal, Romania, the Russian Federation, Slovak Republic, Slovenia, South Africa, Spain, Sweden, Switzerland, Taiwan, United Kingdom, and United States.

54 The countries in Sample B include Belgium, the Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Italy, Luxembourg, Norway, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Slovenia, United Kingdom, and United States.

55 Thus, Sample C includes all countries listed supra note 53, plus Canada, China, Hungary, Israel, Mexico, the Netherlands, Poland, and Turkey.

56 Using a time trend rather than dummy variables has almost no effect on the results.
7266 triadic patents, with 90% confidence interval boundaries of 6205 and 8327 [se(\(\theta\)) = 642.7]. In the larger Sample A, lost patents rises to 8447, with a 90% confidence interval bound by 7715 to 9181.

For both samples, the estimate of lost triadic patents is larger when using total R&D expenditures (12,004 and 9406, respectively). This difference and its direction were expected, given the higher percentage of defense related expenditures in the U.S. and the low patent productivity of such expenditures. To be conservative, we assume there are 7000 lost triadic patents due to the presence of substandard patents in the United States, a round number that is at the lower end of our approximation technique (i.e., \(\theta = 7266\)).
Table 1. Estimation of Lost Patents, Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample A</th>
<th>Sample B</th>
<th>Sample C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient (t-stat)</td>
<td>Coefficient (t-stat)</td>
<td>Coefficient (t-stat)</td>
</tr>
<tr>
<td>Constant</td>
<td>-404.35 (-8.23)</td>
<td>-417.79 (-7.14)</td>
<td>-249.60 (-4.63)</td>
</tr>
<tr>
<td>RND (Civil R&amp;D)</td>
<td>0.123 (60.19)</td>
<td>... (36.44)</td>
<td>... (50.19)</td>
</tr>
<tr>
<td>RND (Total R&amp;D)</td>
<td>... 0.115</td>
<td>... 0.105</td>
<td>... (35.50)</td>
</tr>
<tr>
<td>Period Constants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>-73.18 (-19.01)</td>
<td>-65.06 (-20.18)</td>
<td>-18.02 (-11.30)</td>
</tr>
<tr>
<td>1996</td>
<td>-24.38 (-13.81)</td>
<td>50.41 (63.51)</td>
<td>63.11 (67.97)</td>
</tr>
<tr>
<td>1997</td>
<td>50.93 (53.62)</td>
<td>63.11 (67.97)</td>
<td>135.30</td>
</tr>
<tr>
<td>1998</td>
<td>4.39 (5.69)</td>
<td>-5.49 (-1.89)</td>
<td>67.73</td>
</tr>
<tr>
<td>1999</td>
<td>101.16 (102.23)</td>
<td>24.40 (19.89)</td>
<td>82.75</td>
</tr>
<tr>
<td>2000</td>
<td>51.45 (45.43)</td>
<td>-1.93 (-10.63)</td>
<td>-10.12</td>
</tr>
<tr>
<td>2001</td>
<td>-9.92 (-16.89)</td>
<td>-55.26 (-64.45)</td>
<td>-109.85</td>
</tr>
<tr>
<td>2002</td>
<td>-67.83 (-71.36)</td>
<td>-50.51 (-52.54)</td>
<td>-213.70</td>
</tr>
<tr>
<td>2003</td>
<td>-37.68 (-43.37)</td>
<td>-43.92 (-52.56)</td>
<td>-234.70</td>
</tr>
<tr>
<td>R²</td>
<td>0.98</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>Cross Sections</td>
<td>30</td>
<td>30</td>
<td>19</td>
</tr>
<tr>
<td>Observations</td>
<td>227</td>
<td>227</td>
<td>148</td>
</tr>
</tbody>
</table>

Of course, a patent can be valid without being triadic (but we are assuming triadic patents are valid), since not all valid patents are worth filing triadically. As an approximation to the number of valid to triadic patents, we assume that more rigorous standards of the EPO render only valid patents (we relax this assumption later in our estimation procedure). The (average) ratio of valid patents to triadic patents can be approximated by

\[ P_{it} = \lambda \cdot F_{it} + \sum_{j=1}^{T} \alpha_j D_j + \epsilon_{it} \]  

(4)

where \( P_{it} \) is the number of patent applications by country \( i \) in period \( t \). Equation (4) is estimated in the same way as Equation (3) with period dummies \( D_j \). Sample B is used since it includes only EU countries (18 countries, 162 observations). The \( \lambda \) coefficient is estimated to be 3.0 (\( t\text{-stat} = 120.3 \)).\(^{57}\) So, the ratio of total valid patents to triadic patents in a jurisdiction is approximately 3.0. Triadics are measured in terms of grants by the USPTO rather than

\(^{57}\) The R² of the model is 0.99.
applications as in the case of the EPO and JPO. Given $\lambda = 3$ and a loss of 7,000 triadic patents due to the presence of substandard patents in the United States, both estimated above, the total loss of valid patents in the United States per year is estimated to be 21,000 patent grants (about 10% of patents granted annually by the USPTO). Over the period 1999 through 2003, applications in the US exceeded grants by about two-fold, so there is approximately a loss of about 40,000 applications for valid patents.

Assuming, for illustrative purposes, a 20% leakage in the EPO of substandard patents, reducing $\lambda$ to 2.4, we have 16,800 lost valid patents in the United States annually due to the research deterrence effects of substandard patents. If the EPO is too stringent, say leading to the rejection of 20% of valid patents filed, then $\lambda$ is 3.6 and lost U.S. valid patents is approximated by 25,200.

We can also vary the assumed loss of triadic patents, perhaps according to the estimated confidence interval, for even more approximations of the total loss of valid patents in the United States. Given the imprecise nature of all of these calculations, we believe it is prudent to consider a range of options when estimating the welfare loss from substandard patents.

Having set forth a method for determining the number of lost patents, the next step requires an estimate of how much each patent is worth. There exists a substantial literature on the economic value of patents. For our purposes, the most useful estimate is by Cockburn and Griliches, who estimate the average economic value of a patent as US$1 million (in current dollars) or $2.4 million adjusting for both inflation and economic growth.

Certainly, the distribution of value is highly skewed, but for our calculations the average is suitable. To be conservative, for the benchmark case we assume an average value per valid patent of $1 million, and we will also consider a range of potential values.

The calculation of the deadweight welfare loss from substandard patents is

$$DWL = \lambda \cdot P_{\text{lost}} \cdot V_{\text{valid}}$$  \hspace{1cm} (5)
where $P_{LOST}$ is lost valid patents and $V_{VALID}$ is the average value of a valid patent. In our benchmark case, we have

$$DWL = 3 \cdot 7,000 \cdot 1,000,000 = 21,000,000,000$$

so our “point” estimate of the deadweight loss from substandard patents is $21$ billion annually. Given annual R&D expenditures in this US of about $300$ billion, these losses represent about 7% of total R&D spending per year.

We do not wish to exaggerate the precision of our estimation approach. A rudimentary sensitivity analysis seems unnecessary given the simple form of the damage calculations (in Equation (5)). For example, if we assume any of the inputs to the calculation is understated by 10%, then the estimated cost increases by 10%.

We do think a simulation approach that estimates a distribution of plausible values may be useful. In this simulation, we take our “point” estimates of the three inputs to Equation (5) as mean values, and allow each to vary according to a specified distribution. From the econometric estimate of $P_{LOST}$, we observed a coefficient of variation of about 0.10 (i.e., standard error of the coefficient divided by the mean). For our simulation, then, we assume that $P_{LOST}$ is distributed normally with mean $7000$ and standard deviation $700$. We also assume that $V_{VALID}$ is distributed normally also with a coefficient of variation of 0.10 (so the 95% confidence interval is 0.8 million to 1.2 million). For $\lambda$, we also assign a coefficient of variation of 0.1, and this choice renders a 95% confidence interval bound by 2.4 and 3.6. Our simulation includes 10,000 draws of random numbers from these distributions, and these numbers are inserted into Equation (5) to compute the cost of substandard patents. The resulting distribution is illustrated in Figure 3.
Figure 3 is the histogram of welfare costs of substandard patents from the simulation. The simulated mean of costs ($20.989 billion) is essentially equal to the $21 billion from Equation (6), as expected. The standard deviation is about $3.6 billion (about 17% of the mean). The distribution has a slight positive skew, so it is not symmetrical. Repeating the simulation 100 times indicates the upper and lower bounds of the 95% confidence interval are $14.4 billion and $28.7 billion.\(^6\) For this particular simulation, the minimum value is about $10 billion and the maximum about $38 billion.\(^3\) We stress, however, that these calculations are illustrative since the nature of the random process is somewhat arbitrary. But, even with wide variation in the benchmark assumptions, the estimate of cost remains very high even at its smallest value ($10 billion annually).

**B. Other Deadweight Losses**

As mentioned above, we suspected that the research deterrence costs would be the largest of the deadweight losses from a loose patent system. There are, however, other costs. One direct cost of substandard patents relates to the typical administrative costs of pursuing substandard patents including legal fees, application fees, and the cost of the USPTO. These costs are

---

\(^6\) This confidence interval is not symmetric around the mean (-6.6 billion and +7.7 billion).

\(^3\) Given the very large number of simulations, the minimum, maximum, and confidence intervals are stable across runs.
deadweight losses. In the United States, the legal and filing fees are estimated to be a few thousand dollars for even a simple patent to upwards of $25,000 for more complex technologies. Offered estimates of costs from a variety of sources typically fall in the $3000 to $25,000 range per application.\textsuperscript{64} For our computations, we assume that the patent application costs $7500, on average, in legal and administrative fees.\textsuperscript{65}

These costs must be applied to some estimate of the number of substandard patents filed each year. According to OECD data, over the five-year period 1999 through 2003, there were 90,445 triadic patents filed from the United States (recall that triadic patents are counted by grants for the USPTO).\textsuperscript{66} From above, we estimated the ratio of valid patents to triadic patents to be 3.0.\textsuperscript{67} Applying our $\lambda$ to the United States, we would expect that there would be approximately 271,335 valid patent grants in the United States over this period. However, there were 594,827 patents granted the USPTO in this period, which suggests that approximately half of all U.S. patents granted are substandard.\textsuperscript{68}

While this percentage of substandard patents is high, it is consistent with other evidence. For example, Graham and Harhoff calculate that about 40% of U.S.-granted patents are rejected by the EPO, though the number is found to be much lower (about 4%) in Jensen et al.\textsuperscript{69} Not all United States patents are also filed at the


\textsuperscript{65} Over the period 2003 through 2005, the USPTO earned about $3.3 billion in revenue from 1.14 million applications, for an average application cost of about $3000. U.S. PATENT & TRADEMARK OFFICE, FY 2006 PERFORMANCE AND ACCOUNTABILITY REPORT, MANAGEMENT'S DISCUSSION AND ANALYSIS (2006).

\textsuperscript{66} Estimated according to OECD's \textit{Main Science and Technology Indicators}.

\textsuperscript{67} For that same period, European countries (EU25) filed 79,295 triadic patents and 250,275 applications at the EPO. Thus, dividing the two, we compute a $\lambda$ of 3.16, which is very close to our estimated $\lambda$ of 3.0.

\textsuperscript{68} Estimated according to OECD's \textit{Main Science and Technology Indicators}.

EPO (U.S. entities file about 15% as many applications annually at the EPO as they do the USPTO), but one would initially think that those filed at the EPO by American entities would be of relatively high quality.\(^{70}\) Allison and Lemley, in a study of patents litigated over the period 1989 through 1996, reveal that about half of litigated patents are invalidated at trial.\(^{71}\) Further, Trajtenberg argues that cited patents, and not simple patent counts, are correlated with patent value. In his data, about half of patents are not cited, again suggesting that about half of patents may be classified as substandard.\(^{72}\) Finally, Jaffe and Lerner summarize evidence from the OECD indicating that the growth rate of USPTO granted patents is twice that of "economically significant" (or triadic) patents.\(^{73}\)

Assuming 50% of filings are substandard and there are 400,000 filings per year, there are about 200,000 substandard patent filing at the USPTO annually.\(^{74}\) At an average cost of $7500 per application, the annual deadweight loss from administrative costs related to the acquisition of substandard patents is $1.5 billion. While this is certainly a large number and a significant cost of substandard patents, it is far below the costs of research deterrence caused by substandard patents.

Substandard patents also lead to litigation. While judgments are properly viewed as transfers, the costs of obtaining judgments (or royalties) are deadweight losses. In order to determine the expected cost of litigation from substandard patents, we need an estimate of the probability a patent is litigated and the cost of litigation. As for litigation rates, Lanjouw and Shankerman find a domestic litigation rate of about 1.6% during the early

Grant rates are highly contested figures. See, e.g., Cecil Quillen & Ogden Webster, *Continuing Patent Applications and Performance of the U.S. Patent Office*, 11 FED. CTR. BAR J. 1 (2001) ("The Grant Rate (allowances divided by total disposals, i.e., the sum of allowances and abandonments) for the USPTO for its fiscal years 1993-1998, corrected for continuing applications, ranges from 87% to 97%, depending on the extent to which prosecution of abandoned applications was continued in re-filed applications. Reported Grant Rates for 1995-1999 for the European and Japanese Patent Offices (averaged) are 67% and 64%, respectively.").

\(^{70}\) *FOUR OFFICE STATISTICS REPORT* fig. 4.2, at 41 (2008), http://www.trilateral.net/statistics/tsr/statisticsreport/fullreport.pdf.


\(^{73}\) JAFFE & LERNER, *supra* note 3, at 143.

\(^{74}\) Our statistical analysis is based, by necessity, on grants for the USPTO. Annually, applications are roughly twice grants.
1980s.75 Allison et al. report a 3.2% litigation rate.76 The litigation rate from these studies depend on a number of things including the stock of patents and time period evaluated.

For our purposes, we are constructing annual estimates of the cost of substandard patents. Federal statistics indicate that there are approximately 3000 patent cases filed annually. Allison et al. show that most litigated patents are younger, typically being three years or less.77 Thus, we construct a patent stock of relatively recent patents. Over the most recent five years for which there is data (2004-2008), the USPTO has granted nearly one million patents, so we assume the stock of patents is 1 million and construct an annual litigation rate using that stock.78 So, a reasonable proxy for the annual litigation rate is 0.3% (or 3 cases per 1000 patents) on the stock of patents (both valid and substandard).

The cost of litigation varies substantially across patents, but the average is typically claimed to be in the $1 million to $4 million range for the discovery phase (about half the cost of a full trial).79 Allison et al., citing the American Intellectual Property Law Association, state that a patent case can cost $1.5 million per side.80 Only about 5% of cases actually go to trial, with 95% being settled at some point in the process.81 In a recent economic simulation of patent litigation, Graham and Harhoff use a cost of litigation of $5 million based on estimates from the American Intellectual Property Law Association.82

In light of the evidence, as a benchmark we assume a litigation rate of 0.3% and a litigation cost of $2 million per case. The stock of patents is assumed to be one million (which approximates patents granted in the past five years) and we assume that half the patent stock is substandard.83 Thus, the approximate

76 Allison et al., supra note 39 at 477 fig. 1.
77 Id.
80 Allison & Lemley, supra note 71.
82 Graham & Harhoff, supra note 69.
deadweight loss from the litigation of substandard patents is $3 billion annually. While this is also a very large number, it again remains much smaller than the $21 billion annual cost of research deterrence.

C. Review of the Evidence

Our analysis shows that the cost of a “loose” patent system that is prone to grant substandard patents is very high. Much of the cost is attributable to the reduced innovation, but the administrative and litigation costs are non-trivial. We estimate that annually, the deadweight loss from reduced innovation is $21 billion, administrative costs $1.5 billion, and litigation costs $3 billion. The total of these deadweight losses that we calculate is $25.5 billion annually. We stress that these estimates are preliminary. As such, we have provided a range of probable values to demonstrate the change in estimates given alternative assumptions. Certainly more research is needed on this very important topic.

CONCLUSION

The purpose of patent policy is to balance the incentive to invent with the ability of the economy to utilize and incorporate new inventions and innovations. Because patent law grants de jure monopolies to patent holders and provides those holders with substantial rights to prevent infringement or sue for substantial damages, it is crucial that such patents be awarded only for truly original innovations. As Justice Kennedy recently wrote,

We build and create by bringing to the tangible and palpable reality around us new works based on instinct, simple logic, ordinary inferences, extraordinary ideas, and sometimes even genius. These advances, once part of our shared knowledge, define a new threshold from which innovation starts once more. And as progress beginning from higher levels of achievement is expected in the normal course, the results of ordinary innovation are not the subject of exclusive rights under the patent laws. Were it otherwise patents might stifle, rather than promote, the progress of useful arts.

increase the estimated litigation costs, so the estimates here might be considered conservative.

84 The calculation is $0.50*1000000*0.003*2000000 = $3 billion.
LEAVING ROOM FOR RESEARCH:
THE HISTORICAL TREATMENT OF THE COMMON LAW
RESEARCH EXEMPTION IN CONGRESS AND THE COURTS, AND ITS
RELATIONSHIP TO BIOTECH LAW AND POLICY

Maureen E. Boyle*

12 YALE J.L. & TECH. 269 (2010)

ABSTRACT
The recent suit over the validity of gene patents between the
American Civil Liberties Union and Myriad Genetics has
highlighted the troubling ways in which patents may be interfering
with the willingness of scientists and companies to engage in basic
biotechnology research on matters of vital importance to human
health and disease. Many scholars have argued for a legislative
research exemption to protect this sort of research. Theoretically,
the common law already contains an exemption to protect certain
uses of a patented product from being deemed patent infringement.
This Article evaluates the history of the common law research
exemption alongside the history of biotechnology policymaking
since the 1970s, identifying how confusion over the scope of the
judicial research exemption may have led to legislative stagnation
on the issue of protecting research. Even during the infancy of
biotechnology, members of Congress believed in the existence of a
robust research exemption when making policy decisions about
whether to create a legislative exemption. Now that the scope of
the research exemption has been narrowed significantly by recent
Federal Circuit decisions, at a time when the intellectual property
regime permits patents on human building blocks as basic as
genes, this Article highlights the need for a clear exemption. It also
overviews and comments on existing policy solutions scholars have
offered to counteract the chilling effect that the lack of a clear
exemption might be having on basic research, including research
in the biotechnology sector.

* J.D. candidate, Yale Law School (2011). Many thanks to Daniel J. Kevles for
his helpful assistance and comments on this manuscript.
### Table of Contents

**Introduction** ................................................................. 271

I. The History of the Common Law Research Exemption 274  
   A. Origins of the Exemption ......................................... 275  
   B. Subsequent Interpretations of the Exemption Prior to Biotech ........................................... 278

II. 1970s and 1980s: Biotechnology and the Bayh-Dole Act .................................................... 280

III. 1980s and 1990s: Congressional Understanding of the Research Exemption ................. 285  
   B. Patent Competitiveness and Technological Innovation Act of 1990 ........................................ 289

IV. 2000s: Judicial Evisceration of the Common Law Exemption ............................................ 293  
   A. Madey v. Duke .......................................................... 295  
   B. Integra LifeSciences v. Merck ................................. 298

V. The Future of the Research Exemption in Biotechnology ................................................... 301  
   A. Liability Rules ........................................................... 304  
   B. A “Fair Use” Exemption ............................................. 305  
   C. Compulsory Licensing, Non-exclusive Licensing, and Patent Pools ........................................ 307  
   D. Legislation .............................................................. 308

**Conclusion** ........................................................................... 309
INTRODUCTION

In March, 2010, the American Civil Liberties Union made news when its lawsuit against a prominent genetics company—Myriad Genetics—won its case on summary judgment in a New York district court, after surviving an earlier summary judgment battle over its standing to bring the suit in November, 2009. Myriad holds a patent on the BRCA1 and BRCA2 genes, the presence of which indicate a woman’s predisposition to certain types of cancer. With its patent, Myriad has a monopoly over the gene, including all diagnostic testing related to it. Women cannot seek a second opinion and there is no cheaper alternative test; scientists cannot look at the gene, let alone perform research on it without Myriad’s permission. The heart of the ACLU complaint alleges that Myriad’s monopoly over the BRCA genes interferes with women’s health and doctors’ practices. But the complaint also alleges that Myriad’s patent prohibits independent, non-commercial research on the genes from taking place in university and nonprofit labs. Indeed, the other plaintiffs in the ACLU suit are researchers who received cease and desist letters from Myriad after engaging in unsanctioned work, work which could have provided valuable information about the gene itself and technologies directed to it. In preparation for trial, the ACLU argued that:

[Gene patents interfere with the ability of physicians and researchers to investigate complex diseases. For example, BRCA1/2 may be associated with cancers other than breast and ovarian cancer, but so long as the patents on these genes remain, no one will be able to include these genes in tests for other disease predispositions.]

Although the district court ruled for the ACLU summarily on other grounds relating to the invalidity of Myriad’s patents, the
court did not rule out the possibility that were a full trial to occur, it could be proven that Myriad’s patents were indeed functioning to prevent basic, beneficial research from continuing.8

The clause of the Constitution dealing with patents—Article I, Section 8, Clause 8—optimistically describes the patent monopoly as meant to “promote the Progress of Science and useful Arts” by promoting disclosure of novel and useful methods and inventions.9 Although patents do encourage the disclosure of beneficial ideas, patent holders use their patents for a number of other reasons in modern society: to encourage investor confidence in a new product or market; to gain bargaining chips for cross-licenses, sales, mergers, and acquisitions; or defensively, to secure freedom to work on a new technology or product without fear of infringement.10 More detrimentally, a patent holder may engage in behavior like Myriad’s—rarely licensing the patented technology, but instead enforcing the patent strategically to stifle basic research, the development of competitive alternatives, and other non-sanctioned uses. This type of guarded behavior preserves the patentee’s dominance, but may ultimately harm the public by impeding beneficial research on or with the patented technology.

Long before the advent of biotechnology, the fundamental importance of experimentation was recognized by the judiciary, and some research activities were granted qualified immunity from patent infringement suits. This immunity is known as the “research exemption” or “experimental use exemption.”11 Although the scope of the exemption is and always has been murky,12 since the nineteenth century, judges around the country have recognized that common sense seems to dictate that certain not-for-profit experimentation should not constitute patent infringement under the patent statutes.13 In recent years, however, the Federal Circuit—the federal court with exclusive appellate jurisdiction over patent suits—has narrowed the common law exemption substantially, leaving it difficult to discern whether there is any room for non-commercial research using patented technologies in

8 Id. at *81 (“[T]here exists a sharp dispute concerning the impact of patents directed to isolated DNA on genetic research and consequently the health of society. . . . [T]he resolution of these disputes of fact and policy are not possible within the context of these motions.”).
9 U.S. CONST. art. I, § 8, cl. 8.
11 The exemption is also called the “research exception” in other literature. I use it to mean the judicially-created immunity for users of patented technology who engage in non-commercial research. See infra Part I.
12 See infra notes 28-31 and accompanying text.
13 See infra notes 28-31 and accompanying text.
universities and nonprofits after the court’s recent holdings. Although the specter of a possible research exemption may have at least discouraged patent holders from suing non-commercial experimenters, the Federal Circuit’s erosion of the exemption makes it likely that any non-commercial experimenter, whether individual or institutional, could risk being sued if her work involves patented technology.

This Article examines how, historically, the research exemption has been discussed and relied upon in patent policymaking, and how the demise of the common law research exemption relates to practices in the biotechnology industry. Did the common law research exemption ever really exist? Were fundamental policy choices made in reliance on it? What results from the evisceration of the common law exemption, given the state of current policy toward biotechnology? What should legislators do about it?

Biotechnology is a particularly vulnerable technology because of its deep relationship to our understanding of health and disease. Continued research is vital to confirm the accuracy of genetic tests, to discover potential flaws and fixes, and to allow researchers to find suitable alternatives or substitutes if possible. In an industry so intertwined with life and death, the threat of an anticommons is particularly worthy of concern. Without the space and freedom to research, patients, doctors, and society at large are at the patentee’s mercy. A person’s health may depend on the patentee granting licenses, choosing a reasonable price for products incorporating the monopolized technology, and doing further research that may improve or cheapen the technology. As Myriad’s behavior has demonstrated, a bad actor has little incentive to do any of these things. The pro-competitive goal of patent law is undermined by the anti-competitive effect of patents on genetic material: with a gene, there is no way to invent around the patented technology, so the patentee need not fear competition for the term of the patent. In a competitive environment, the patentee would be incentivized to do more research, to charge reasonable prices, and potentially to cross-license the technology. In an environment free of competition, profit-maximizing behavior and progress-maximizing behavior may be at odds.

14 See generally Michael A. Heller & Rebecca S. Eisenberg, Can Patents Deter Innovation? The Anticommons in Biomedical Research, 280 SCIENCE 6918 (1998) (discussing the seriousness of the patent thicket impeding continued research in the biomedical field).

15 See id. There is some specific evidence that biotechnology patents are being used to slow progress or impede competition. Lori Andrews has identified a case in which GlaxoSmithKline pursued a patent on a test which would examine the effectiveness of one of their drugs, not because they intended to develop the test, but rather so that no one could do further work on it. Lori B. Andrews, Genes
exemption might help alleviate at least some of these problems, and legislators and policymakers should consider ways in which the research exemption might be reinstated and clarified now that the common law exemption has been eviscerated.

Part I of this Article tells the story of the common law research exemption as it evolved prior to the advent of biotechnology. Part II discusses the beginnings of biotechnology and the passage of the Bayh-Dole Act, which essentially gave researchers (including academic and nonprofit researchers) a duty to commercialize and license their work, a subtle yet dangerous threat to the underpinnings of the research exemption. Part III examines the ways in which biotechnology policymakers, aware of the threats to public health posed by biotechnology patents, discussed and relied upon the research exemption in their decisions during the 1980s and 1990s. Part IV overviews the recent narrowing of the common law exemption and its ramifications, specifically for the most recent advancement in the modern biotechnology industry—genetic analysis and testing. Part V sets forth the solutions that have been advanced by academics and policymakers to address the current system’s chilling effect on basic, beneficial research, and concludes with some recommendations for future action.

I. THE HISTORY OF THE COMMON LAW RESEARCH EXEMPTION

Although many authors have discussed the origins of the research exemption, their interpretations of the exemption vary as widely as the interpretations advanced by various courts over the years. This section attempts to briefly overview the history of the experimental use or research exemption prior to the advent of biotechnology, highlighting its inconsistent application and meaning. While perhaps offering no clear answers to questions about the traditional meaning or scope of the exemption, the

and Patent Policy: Rethinking Intellectual Property Rights, 3 NATURE REVIEWS GENETICS 803, 804 (2002). Progress and profit may not always be in competition, though; a company that obtains a patent might work to cheapen the production of the patented biotechnology, or to develop technologies that enhance the value of the patented product, in cases where the ability to charge monopoly prices would allow the patent holder to reap additional profits. I thank Bret Hembd, Executive Editor of the Yale Journal of Law and Technology, for these suggestions.

history does demonstrate that there would be at least some basis to believe that certain applications of patented technology—particularly uses for the purposes of testing the accuracy of an invention or testing its proper enablement by the specification—are protected from infringement because of the absence of harm to the patentee.

A. Origins of the Exemption

The common law research exemption originated in an 1813 case from Massachusetts, *Whittemore v. Cutter.* The defendant, who was charged with infringement for constructing the plaintiff’s patented machine, challenged a jury instruction which stated that making a machine with “a design to use it for profit” constituted infringement. Justice Story, sitting in his appellate capacity on the Massachusetts federal circuit court, affirmed the instruction, noting that making a patented technology for profit was within the purview of the Patent Act of 1793; it was not-for-profit use of the patented technology that might not be covered. Justice Story stated that “it could never have been the intention of the legislature to punish a man, who constructed such a machine merely for philosophical experiments, or for the purpose of ascertaining the sufficiency of the machine to produce its described effects.”

Justice Story thus believed that Congress intended to punish persons deriving profit from their use of the invention, but not those who used the patent for certain other purposes.

Justice Story again discussed the issue of profit as a component of infringement just five months later in *Sawin v. Guild,* another Massachusetts circuit court case. The defendant, a deputy sheriff, seized and sold the plaintiff’s patented nail cutting machine as part of an execution of the plaintiff’s debts. In holding that this was not infringement, Justice Story referenced *Whittemore* in dicta while remarking that the Act of 1793 had already been construed. He stated that

[For] the making of a patented machine to be an offence within the purview of [the statute], [it] must

---

18 *Whittemore,* 29 F. Cas. at 1121.
19 Id.
20 21 F. Cas. 554 (C.C.D. Mass. 1813) (No. 12,391).
be the making with an intent to use for profit, and not for the mere purpose of philosophical experiment, or to ascertain the verity and exactness of the specification. In other words, that the making must be with an intent to infringe the patent-right, and deprive the owner of the lawful rewards of his discovery.\textsuperscript{21}

One commentator has interpreted these two 1813 cases to mean that Justice Story believed that the experimental use exemption consisted of two separate requirements: “(1) the activity must not be performed with the intent to gain profit and (2) the activity must be either (a) for philosophical experiments or (b) for ascertaining the verity and exactness of the specification.”\textsuperscript{22}

While it may be easy enough to look at the question of intent, and it is a matter of fact whether the use was for ascertaining “verity” or “exactness,” the problem lies in interpreting what Justice Story meant by “philosophical experiments.” One interpretation would be that Justice Story contemplated only a man “tinkering around” in his basement with another’s invention; however, it seems unlikely that Justice Story would have limited philosophical experiments to such an invisible, individual use.\textsuperscript{23} Another view states that “philosophical experiments” would include use of the invention in the course of developing new technologies, although this would seem to extend directly to future for-profit uses that Justice Story would likely consider infringement.\textsuperscript{24}

Rebecca Eisenberg has advocated an interpretation somewhere in the middle of these two extremes: “[t]he first prong of Justice Story’s experimental use privilege, permitting ‘philosophical experiments’ . . . seems to permit subsequent researchers to use the patented invention at least in traditional basic research with no commercial implications.”\textsuperscript{25} Eisenberg defines

\textsuperscript{21} \textit{Id.} at 555 (internal citation omitted).
\textsuperscript{22} Karp, supra note 17, at 2171.
\textsuperscript{23} Bee, supra note 16, at 367.
\textsuperscript{24} Chisum, supra note 16, at 1019 n.203. This view is probably the weakest. The nineteenth century case \textit{Poppenhusen v. Falke} held that use of patents to develop future technology is not protected, and similar fact patterns were also held not to be experimental uses by other courts. 19 F. Cas. 1048, 1049 (C.C.S.D.N.Y. 1861) (No. 11,279) (“[The defendants] are rivals of the complainant in the very business to which his patents relate . . . . The answer alleges that all the defendants have thus far done since the organization of said company, has been done by way of experiment, for the purpose of hereafter working under certain patents, grants, and licenses of their own . . . it can hardly be necessary for the respondents to experiment with the complainant's inventions in order to perfect their own . . . .”).
\textsuperscript{25} Eisenberg, supra note 16, at 224.
“basic research” as “‘pure’ research directed solely toward expanding human knowledge, as opposed to ‘applied’ research directed toward solving practical problems.” Eisenberg’s definition encompasses the basement inventor, but leaves out researchers who use the invention for eventually for-profit purposes. More importantly for this inquiry, Eisenberg’s interpretation exempts researchers who aim to test an invention or use it to add to human knowledge and understanding, a more liberal construction of “philosophical experiments” than one which would protect only the casual, curious experimenter in his basement. Eisenberg’s definition is also consistent with the more recent research done by Janice Mueller, who evaluated other nineteenth century uses of the word “philosophical” and suggested that “philosophy referred to natural philosophy, which in turn meant science generally.” Under this definition, “philosophical experiments” might thus cover scientific research on a patented invention to ascertain its workings and to either evaluate them or attempt to design around them.

In any case, by the close of the nineteenth century, it was almost unanimously agreed that a narrow exemption for experimental use existed at common law. One nineteenth century treatise on patents stated that “where [the invention] is made or used as an experiment, whether for the gratification of scientific tastes, or for curiosity, or for amusement, the interests of the patentee are not antagonized.” The experimental use exemption was narrow from the outset—even prior to 1900, courts typically found that various uses of patented inventions by commercial infringers were not experimental—but even in the cases where the courts found no experimental use, the courts acknowledged that some exemption did exist for not-for-profit uses. As early as

26 Id. at 178 n.1.
27 Mueller, supra note 17, at 929.
28 But see Clerk v. Tannage Patent Co., 84 F. 643 (3d Cir. 1898) (holding that contracts were required even to conduct experimental testing); Albright v. Celluloid Harness-Trimming Co., 1 F. Cas. 320 (C.C.N.J. 1877) (No. 147); Palmer v. United States, 20 Ct. Cl. 432 (1885), aff’d on other grounds, 128 U.S. 262 (1888). These latter two cases held that clearly experimental uses—one, testing the performance of patented molds in the process of manufacturing trimming, and the other, testing a knapsack for its wartime practicality—were indeed infringements. However, the majority of cases both before and after followed Story’s logic rather than these aberrant holdings.
29 3 WILLIAM C. ROBINSON, THE LAW OF PATENTS FOR USEFUL INVENTIONS § 898 (1890).
30 See U.S. Mitis Co. v. Carnegie Steel Co., 89 F. 343, 351 (C.C.W.D. Penn.) (holding that “use in the course of business and for profit” is not experimental), aff’d without opinion, 90 F. 829 (3d Cir. 1898); Cimiotti Unhairing Co. v. Derboklow, 87 F. 997, 999 (C.C.E.D.N.Y. 1898) (acknowledging a “legitimate use for experimental purposes only”); Bonsack Mach. Co. v. Underwood, 73 F.
1861, one court even called it "well settled" that an experimental use exemption existed at common law, but the conflicting interpretations later given in courts around the country demonstrate that the scope of that exemption and the nature of the activities that would fall under it were hardly clear.

**B. Subsequent Interpretations of the Exemption Prior to Biotech**

As is evident from the limited history thus far, the scope of the exemption was murky from its outset. Although most courts recognized that, according to common sense, some experimental use could not have been intended to be infringement by the legislature, they frequently conflicted in their interpretations of what exactly permissible experimentation was or would be. This pattern of inconsistent interpretation continued for the majority of the early twentieth century, and overwhelmingly, plaintiffs prevailed against a defendant's claim of experimental use.

However, one interesting pattern during this period is of note: although strictly commercial enterprises were almost never exempted on the grounds of experimental use, in those cases in which experimental use was found, the defendant was the U.S. government, a frequent government contractor, or a nonprofit educational institution.

The educational institution absolved from infringement was the Colorado School of Mines. The school and its faculty and students were released from liability in a 1935 decision, *Ruth v. Stearns-Roger Manufacturing Co.* The disputed technology was a certain type of patented flotation machine. Although the named defendant, a commercial enterprise, was found guilty of

---

206, 211 (C.C.E.D.N.C. 1896) ("It is true that, if an infringing machine is made or used as an experiment merely, it does not infringe former patents."); 31 Poppenhusen v. Falke, 19 F. Cas. 1048, 1049 (C.C.S.D.N.Y. 1861) (No. 11,279) ("It has been held, and no doubt is now well settled, that an experiment with a patented article for the sole purpose of gratifying a philosophical taste, or curiosity, or for mere amusement, is not an infringement of the rights of the patentee."). 32 See 5 Donald S. Chisum, Chisum on Patents § 16.03(1)(b) (2010); see also Steven P. Caltrider & Paula Davis, The Experimental Use Defense: Post-Madey v. Duke and Integra LifeSciences I, Ltd. v. Merck KGaA, 86 J. Pat. & Trademark Off. Soc'y 1011 (2004) (providing an overview of the parameters of the exemption in individual cases throughout this period). In his 1957 article, Richard Bee also has a very detailed (although overwhelmingly critical) case-by-case description of these continuously inconsistent interpretations of experimental use. Bee, supra note 16, at 370-75. 33 See Bee, supra note 16, at 377; Eisenberg, supra note 16, at 222. 34 13 F. Supp. 697 (D. Colo. 1935), rev'd on other grounds, 87 F.2d 35 (10th Cir. 1936).
infringement, the school (which bought parts from the company) was immune from liability because the school used the technology in “laboratory machines used for experimental purposes, and consequently did not contribute to an infringing use.” Although it is not completely clear what the type of experimentation was, the court seems to have overlooked the fact that even educational institutions are in a sense commercial, in that they are in the business of attracting students and endowment investors. The court seems only to have considered that the use of the technology was in the lab and was for the purpose of satisfying scientific inquiry, an educational and experimental activity which it held to be exempt.

In addition to covering educational use of patented technology, the exemption seems also to have covered some work for government research. Although not explicitly for government use, one wartime case, Dugan v. Lear Avia, involved a type of technology for a direction-finding and position-indicating system in airplanes, and since Lear was an essential government contractor during World War II, one might imagine that the suit had implications for national defense. Although the case was decided on other grounds—the invalidation of the plaintiff’s patents—the court stated that “defendant built [one of the allegedly infringing] device[s] only experimentally and that it has neither manufactured it for sale nor sold any.” The device was only constructed to understand how it worked—a form of reverse engineering and industrial research that the court stated would be free from liability under the experimental use exemption. The exemption covered more obvious, explicit government research in a later case which found the United States not guilty of infringement: Chesterfield v. United States. In dicta, the court referenced the experimental use exception, stating that the government’s use of an alloy as part of government experiments was not infringement; unfortunately, it is completely unclear how or for what purpose the technology was used. The court stated only that “a portion of the 422-19 alloy procured by the defendant was used only for testing and for

---

35 Id. at 703.
36 I contrast this to work for government use—for example, use of the technology in warfare or as part of national defense. This type of use is clearly not experimental or research-based, and the “experimental use” defense has failed for the government in these situations. See Pitcairn v. United States, 547 F.2d 1106 (Ct. Cl. 1977); Palmer v. United States, 20 Ct. Cl. 432 (1885).
38 Id. at 229.
40 Id. The patent was invalidated in this case, so the experimental use discussion is therefore dicta—the court need not have reached the question of experimental use.
experimental purposes, and there is no evidence that the remainder was used other than experimentally. Experimental use does not infringe." The issue in both Dugan and Chesterfield seems to be whether the invention was being used by the government or a contractor in a strictly non-commercial sense: testing the sufficiency of an item for its own sake, or reverse engineering an item to see how it works without the intention of producing a copy.

Although these decisions indicate that courts were perhaps more likely to find a nonprofit or governmental entity engaged in basic research to be protected by the experimental use exemption, the application and construction of the exemption remained far from clear prior to the 1970s. There appears to have been some recognition that functionally non-commercial enterprises—university research, and perhaps certain research by the government—should not give rise to liability for the use of patented technology in non-commercial ways. However, in ways unforeseen, the lines between commercial and non-commercial were about to be blurred. With an action as small in scale as the introduction of DNA into a host bacterium, the business of biotechnology was on its way.


Biotechnology is generally defined as “any technique that uses living organisms (or parts of organisms) to make or modify products, to improve plants or animals, or to develop microorganisms for specific uses.” Beginning in the mid-1970s, with advances in genetic technology, the contemporary biotechnology industry was born. Molecular biologists researching recombinant DNA—a method of splicing, cloning, and isolating genetic material—quickly realized its implications and possibilities for the future of scientific research, given that they now possessed the ability to single out DNA segments and analyze their structure and function. However, as recombinant DNA technology became widespread, many others, including patent attorneys for universities, speculative venture capitalists, and even enterprising scientists themselves, recognized the commercial possibilities of recombinant DNA technology. The rise of biotechnology and the

41 Id. at 845-46.
44 See id.
battles the fledgling industry faced have been extensively chronicled and analyzed. Although biotechnology faced a number of detractors who feared its capabilities and hazards, many viewed biotechnology as an industry with the ability to stimulate much-needed domestic economic growth. In 1980, when news broke that one of the earliest biotechnology companies, Genentech (a combination of the first syllables of “genetic engineering technology”), had produced synthetic insulin with recombinant DNA technology, its stock price more than doubled on the day it went public. Start-up companies sold promises of medical miracles to their investors, and established pharmaceutical and chemical companies began investing millions in biotechnology research and development. The fruits of the biotechnology industry include the creation of many synthetic hormones with profound implications for human health, and in the following decades, genetic testing used to indicate biological predisposition for certain diseases.

The term “industry” brings to mind the private sector and private development, but from its very beginnings, the public and nonprofit sectors were at the heart of the biotech industry. It was an academic lab at Stanford University that spawned recombinant DNA technology, not a private-sector team of inventors. Academic molecular biologists were increasingly courted by biotechnology corporations with promises of funding and profits. In addition, and perhaps most troubling, academic scientific research was largely being funded by the government. The National Institutes of Health, Department of Defense, Department of Agriculture, Department of Energy, National Science Foundation, and other federal groups spent billions of dollars on university research and development over the course of the 1970s and 1980s. Alerted to the conflicts of interest inherent in public money funding private enterprise, members of the media began to cover biotechnology with no shortage of skepticism and cynicism. The concerns largely fell into two categories: first,

---

46 Krimsky, supra note 45, at 25.
48 Krimsky, supra note 45, at 30-37.
49 Hughes, supra note 43, at 541-42.
50 Krimsky, supra note 45, at 60.
51 Id. at 66-68.
52 Id. at 70-71.
concerns over the “commingling of funds” and whether scientists were using publicly funded labs and materials for commercial work, and second, the concern that private companies were appropriating the profits and the fruits of publicly funded academic research, making the public “pay twice for its investment.”

Congress took notice of the controversies and the excitement surrounding biotechnology. Initially, Congress’s focus was on regulation and driven by safety concerns, however, as private firms found success with commercial applications of recombinant DNA technology, Congress recognized that biotech could provide a serious boost to the American economy, and thus began to focus on ways the government could support the industry and ensure American dominance. Long before the 1970s, both universities and the government had encouraged the patenting of publicly funded research results. However, in the 1970s, two factors were different: first, the amount of federal money in R&D had increased dramatically, and second, the profits to be gleaned from the exploitation of biotechnology research were absolutely enormous compared to the paltry amount universities received from controlling and licensing their pre-biotechnology patents.

As the biotechnology frenzy swept the U.S. economy, the government was not equipped to quickly commercialize the results of the research it funded; besides, the commercial infrastructure was set up already by private biotech companies and start-ups. The nexus between government and the private sector was nonprofit and university research, but with the amount of funding and profits at stake, clear guidelines for ownership and transfer of technology from the universities to the private sector were needed.

Hence, Congress took action, first, to enable universities to retain ownership in the results of their federally funded research, and second, to facilitate (and all but mandate) the transfer of that technology to the commercial private sector. In 1980, Congress passed two pieces of legislation—the Stevenson-Wydler

53 Id. at 71.
54 Hughes, supra note 43, at 566-68.
56 See Kevles, supra note 47.
57 See Rebecca S. Eisenberg, Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research, 82 VA. L. REV. 1663, 1677-84 (1996) (discussing initiatives from the 1960s and earlier to encourage the patenting of inventions created with government funding).
58 Kevles, supra note 47, at 298.
59 Id. at 298-99.
60 Eisenberg, supra note 57, at 1663-64.
Technology Innovation Act and the Bayh-Dole Act—in order to encourage the commercial development of university and government discoveries, promote the creation of new jobs, and thereby stimulate the U.S. economy. The Bayh-Dole Act, also known as the University and Small Business Patent Procedures Act, has had a lasting effect on the university’s role in the patent system; it grants universities—as opposed to government agencies—title in inventions made with government funding, provided that these universities satisfy a number of requirements, including obtaining patents in the technology and actively pursuing “practical application,” or the commercial development of the invention, through licensing if necessary. The Bayh-Dole Act was Congress’s response to what U.S. universities perceived as a lack of clarity about their rights in federally funded technology, and moreover, to a fear that beneficial research would languish in university labs that lacked the tools to commercialize it. In addition to giving universities clearer rights and duties, the Act also provided the government with “march-in rights” to grant licenses to other contractors regardless of the patentee university’s willingness to license, if deemed necessary to hasten commercialization, “meet requirements for public use,” or “alleviate health and safety needs.” By requiring universities to find commercial outlets for their patented research (or else face government intrusion), the provisions of the Bayh-Dole Act have been interpreted by universities as creating an “implied duty to commercialize” any inventions or technologies created with public money.

Although the congressional hearings contained discussions about whether patent rights would be allocated to the government or the universities, Congress does not seem to have discussed the dedication of the developed technologies to the public domain. From the outset, patent protection was viewed as the best means for facilitating technology transfer, as opposed to open sharing of

---

63 Eisenberg, supra note 57, at 1663-65.
65 Eisenberg, supra note 57, at 1663-64.
66 35 U.S.C. § 203. The march-in rights may be exercised against the university and against licensees, despite the provisions of any existing contracts.
university- or nonprofit-developed inventions and methods. There is only a hint that some senators may have been considering public dedication in the remarks of Representative Jack Brooks (D-TX), contained in the house report on the Bayh-Dole bill:

My concern is simply the role of the government and the rights of the people in the patent process. When a private company risks its own money to develop new products and procedures it deserves and receives the profits that may result. There should not be a different standard applied when it is the government that risks the taxpayers' money. The rewards of successful research and development conducted at government expense should go to all the people.\(^\text{68}\)

The final form of the bill ensured the opposite: universities were to hold patents that would be licensed to private firms and developers. By the early 1980s, many universities had already established deep ties to the commercial sector.\(^\text{69}\) In 1980, the Supreme Court decision *Diamond v. Chakrabarty* encouraged further ties and investment in university biotech research, by clarifying that living material was not per se unpatentable subject matter.\(^\text{70}\) *Chakrabarty* paved the way for universities to work toward patents on DNA material, microorganisms, and farther down the road, even higher life forms.\(^\text{71}\)

With the advantage of hindsight, it is now apparent that the creation of a “duty to commercialize” stands in direct conflict not only with certain academic norms,\(^\text{72}\) but also with the university’s function as a center of basic research.\(^\text{73}\) Before the 1970s and 1980s, the experimental use exception may have protected universities from being liable for their research work using patented technologies—at the very least, the exception was murky enough that patent holders might not have been willing to gamble time and money to sue universities and nonprofits. But the passage

---


\(^\text{69}\) Kevles, *supra* note 47, at 303.

\(^\text{70}\) 447 U.S. 303 (1980).

\(^\text{71}\) For a history of the patenting of animals, see Daniel J. Kevles, *The Advent of Animal Patents: Innovation and Controversy in the Engineering and Ownership of Life*, in *INTELLECTUAL PROPERTY RIGHTS IN ANIMAL BREEDING AND GENETICS* 17, 17-30 (Max Rothschild & Scott Newman eds., 2002).

\(^\text{72}\) For example, norms encouraging the sharing of research, or the independence and integrity of chosen research projects. See Eisenberg, *supra* note 16; Kevles, *supra* note 47.

\(^\text{73}\) Eisenberg, *supra* note 16, at 224.
of the Bayh-Dole Act blurred the line between basic research and applied research in this setting, leaving it difficult to determine whether certain types of research on patented technologies in university or nonprofit labs would constitute infringing uses. In biotechnology, the stakes were financially high, but additionally, in fields touching public health and disease, the progress of certain kinds of research are critical: verifying and testing health-related technologies and methods, or encouraging and developing new ideas to design around preexisting inventions. After Bayh-Dole, with universities becoming heavily invested in commercialization, the ability of the experimental use defense to cover basic nonprofit and university research was jeopardized. But policymakers believed that an exemption existed, and that it would protect valued types of research. Indeed, in considering regulation and guidance for the development of the biotechnology industry in the 1980s and 1990s, legislators seem to have relied on the existence of the common law research exemption to ensure that critical and beneficial basic research would continue.

III. 1980s and 1990s: Congressional Understanding of the Research Exemption

In the legislative history of the Bayh-Dole Act, there is a notable absence of concern about the protection of university and nonprofit research activities. However, Congress was confronted again with biotechnology policy questions (including questions about the “experimental use” protection for basic research) in the subsequent decade, most notably during the debates on the patenting of transgenic animals and attempts to pass policies which would clarify U.S. patent law and bring it into line with global practices. This Part will examine the legislative history surrounding two bills in particular—the Transgenic Animal Patent Reform Act of 198874 and the Patent Competitiveness and Technological Innovation Act of 199075—neither of which was ever enacted. Although they never became law, the legislative history of the bills preserves the ways in which members of Congress discussed the value of university and nonprofit research, perceived the research exemption, and made choices about the codification of the common law exemption in proposed legislation.

A. The Transgenic Animal Patent Reform Act of 1988

It was not too long before the advancement of biotechnology rendered scientists able to genetically modify higher

74 H.R. 4970, 100th Cong. (1988).
life forms which satisfied the criteria of patentability—cancer-susceptible mice, for example, or genetically modified pigs capable of producing more meat. The technology involved in *Diamond v. Chakrabarty* was a kind of bacteria, perhaps more easily viewed as a patentable man-made composition of matter than as a living, breathing animal. Moral and environmental opposition to the patenting of these higher life forms again drew the attention of Congress to the biotech industry. Congress thus began to consider whether a moratorium on the granting of animal patents would be appropriate, and moreover, whether and what guidelines were necessary to govern patentability and infringement questions with regard to animal patents specifically. Representative Robert Kastenmeier (D-WI), Chairman of the House Judiciary Subcommittee that handled patents, held hearings on the issue and began formulating a bill to cover the patenting of transgenic animals, called the Transgenic Animal Patent Reform Act.

Prior to the drafting of the bill, in the hearings held by the Committee on the Judiciary, a statutory “research exemption” came up in the testimony of three individuals: Robert Merges, a professor of law at Columbia, Reid Adler, a patent attorney at Finnegan Henderson in Washington, D.C., and Leo Walsh, dean of the College of Agriculture at the University of Wisconsin. Ostensibly, the research exemption was suggested because such an exemption would mirror the exemption Congress inserted in the Plant Variety Protection Act of 1970 (PVPA).

Unfortunately, there is very little legislative history clarifying why the research exemption appeared in the PVPA. It seems likely that legislators included the research exemption because it was mandatory if the United States wished to become a member of the International Union for the Protection of New

---

77 447 U.S. 303 (1980).
79 *Id.* at 24, 28.
81 *Id.* at 12 (“Both [a farmer’s exemption and a research exemption] are paralleled in legislation Congress passed under the Plant Variety Protection Act.”).
Section 114 of the bill—covering the “research exemption”—is explained in the report only by the statement that “[u]se and production for research is not to constitute infringement.” *Id.* at 5093. Section 111 of the bill—covering the “infringement of plant variety protection” clarifies that “[u]se of the protected variety as one source of germ plasm to breed a novel variety is permissible” under the research exemption, *id.*, seeming to indicate that Congress wished to protect the ability of experimenters to design around the patented variety to produce diverse, novel varieties.
Varieties of Plants (UPOV). UPOV is an intergovernmental organization which encourages intellectual property protection for plant breeders’ creations internationally. The organization sets forth uniform legal standards that member nations must comply with—one of which is a robust research exemption. The reason for the exemption may be as simple as this: in order for the United States to join UPOV, and gain the attendant benefits of membership, Congress passed the PVPA with the required research exemption. But post-hoc rationalization of the inclusion of the research exemption is also instructive for interpreting how later legislators understood the importance of the exemption. After the passage of the PVPA, legislators have stated that the exemption exists because (1) there was concern about granting private entities exclusive control over federally funded technology, and a research exemption alleviated this concern, and (2) they were trying to protect valuable germplasm from being locked up in patents, preventing experimenters from using patented germplasm as a source to develop novel and diverse varieties of plants.

The latter reason is strikingly evocative of the fair use doctrine in trademark law, which prevents the holder of a trademark from removing particular language from public discourse (or controlling use of the language) on First Amendment grounds. Similarly, experimental use in the PVPA seems to try to prevent a patent holder from removing important germplasm from the collection of germplasm available to plant breeders.

---

85 Crocker, supra note 83, at 81-83.
86 H.R. REP. NO. 101-960, at 32 (1990) (“This amendment [creating a research exemption in the PVPA] was made, in part, because of the involvement of publicly funded research on plants.”).
87 This is supported by congressional debate surrounding the Plant Variety Protection Act Amendments of 1993: “The research exemption [in the 1970 bill] was included to promote the free flow of germplasm—essential to the maintenance of genetic diversity.” 139 Cong. Rec. S10841-02, S10868 (daily ed. Aug. 7, 1993) (statement of Sen. Kerrey). It is also supported by the “design around” provisions. See supra note 82.
88 See Mattel, Inc. v. MCA Records, Inc., 296 F.3d 894, 900 (9th Cir. 2002) (discussing the fair use doctrine in relationship to the First Amendment). For a more thorough discussion of the relationship between trademark fair use and the research exemption in PVPA, see Mark D. Janis & Stephen Smith, Technological Change and the Design of Plant Variety Protection Regimes, 82 Chi.-Kent L. Rev. 1557, 1563-65 (2007).
Germplasm and genetically modified animals share basic similarities, in that they are composed of identifiable genetic material and thus tied to life and the environment; it seems deleterious to permit patents to remove basic building blocks of life from the research scientist’s tool kit, whether those building blocks are germplasm or genetic sequences. While this theory is completely speculative, perhaps this connection between plants and animals motivated Merges, Adler, and Walsh to suggest that a research exemption comparable to the one in PVPA be included in any legislation covering transgenic animal patenting.

Adler and Walsh went into deeper detail than Merges on the scope of the proposed statutory research exemption. Walsh expressed fears that animal patents would concentrate valuable resources in the hands of a few patentees and licensees, and thus recommended the legislation include “a university research exemption, compulsory licensing of the patent, public research focusing efforts on helping the smaller firms stay competitive in the market place, [and] public institutions cooperating in establishing and maintaining a gene bank,” among other suggestions which would protect university and nonprofit research. Adler seems to have argued that although a common law exemption existed, a statutory exemption was necessary because “the boundary between permissible research uses and impermissible infringement [was] not totally clear” from the case law. He further expressed concerns that because of the ambiguous precedents, courts might not recognize basic research on transgenic animals as exempt, even when “no direct commercial benefit” was at stake for the research scientists. The record thus demonstrates that Congress was warned by a few prominent advocates that a research exemption would be necessary in order to keep valuable genetic information in the public domain for basic research purposes.

Yet prior to the bill’s passage by the House of Representatives, the House Committee on the Judiciary deleted a proposed statutory research exemption. The reason: “a statutory exception was unnecessary in light of the existing judicially fashioned doctrine.” It was not oversight or lack of consideration that kept the Act from including a research exemption: it was reliance on the existence of a common law “experimental use” exemption that would protect basic research activities from constituting infringement.

90 Id. at 13
91 Id.
92 Id. at 3 (emphasis added).

288
The Transgenic Animal Patent Reform Act died in the Senate after being passed in the House. However, the debate about the Act is instructive for viewing how contemporary legislators viewed the function and strength of the experimental use doctrine. Two years later, in the debates surrounding another bill advanced by Kastenmeier, it would become even clearer that legislators believed that a robust common law research exemption existed.

B. Patent Competitiveness and Technological Innovation Act of 1990

The Patent Competitiveness and Technological Innovation Act of 1990 was broadly intended to “improv[e the] country’s patent law.” Like the Transgenic Animal Patent Reform Act, the bill was introduced by Kastenmeier, and it contains sections regulating everything from inventions made in space to genetically engineered animals. For our purposes, the critical component of the bill is Title IV, which would have created a statutory research exemption for basic scientific research activities. Title IV of the Patent Competitiveness Act would have amended 35 U.S.C. § 271, a section of the patent law, by adding a subsection which would state that “[i]t shall not be an act of infringement to make or use a patented invention solely for research or experimentation purposes.”

To contextualize the drafting of Title IV, it is essential to realize that the legislators viewed the bill as “an attempt to codify and clarify current case law in the United States which currently excludes experimental use or research as an act of infringement,” and stated that it was a “central tenet of American patent law that there is a right to use scientific information to create new and better inventions in competition with the patented invention.” Legislators thus did not see the bill codifying the research exemption as a departure from current case law, but rather as the legislation of an already existing common law exemption.

The report by the House Committee of the Judiciary on the bill contains some clarification of which activities would constitute protected research and which would not.

[T]he making or using of a patented invention solely for research or experimentation shall not be

---

93 Kevles, supra note 71, at 28.
an act of patent infringement unless the patented invention has a primary purpose of research or experimentation. If the patented invention has a primary purpose of research or experimentation (such as a transgenic mouse used for cancer research or a laboratory implement such as a microscope), it shall not be an act of infringement to manufacture or use one of these inventions to study, evaluate, or characterize it or to create a product outside the scope of the patent covering the particular invention.  

The House Report identified six additional examples of "experimental use":

1. Testing an invention to determine its sufficiency or to compare it to prior art;
2. Tests to determine how the patented invention works;
3. Experimentation on a patented invention for the purpose of improving on it or developing a further patentable invention;
4. Experimentation for the purpose of "designing around" a patented invention;
5. Testing to determine whether the invention meets the tester's purposes in anticipation of requesting a license; and
6. Academic instructional experimentation with the invention.

These permissible uses fall broadly into two groups: (1) research on the technology, or in other words, evaluations and studies of the technology itself; and (2) use of the patented technology in an effort to design around the technology. Both seem to fit within at least some interpretations of Justice Story's original formulation, and moreover, both are important parts of biotechnology research. Indeed, the clarification of biotechnology policy was expressly mentioned as reason to support the statutory exemption. Citing the progress of university-industry partnerships following the Bayh-Dole Act, the House Report stated that allowing scientists and researchers to remain confused over

98 Id.
99 Id. at 35-36.
100 See supra notes 25-27 and accompanying text.
101 H.R. REP. NO. 101-960 at 34-35 ("The field of biotechnology would particularly [sic] benefit from a statutory research exception.").
which research activities were permissible and exempt would be “contrary to sound public policy.” In addition to alleviating confusion, legislators cited two other main reasons to support a statutory research exemption for biotechnology: first, the prevalence of public funding in the biotechnology industry, and second, the fear that basic testing activities would be sent to countries with robust research exemptions, such as Japan and the countries in Western Europe. To indicate the widespread support for a statutory research exemption in biotechnology, the House Report quotes professors, economists, and scientists, all in support of the proposition that without a clear exemption, “[u]nnecessary litigation occurs, excessive threats are levelled, transaction costs are raised, and experimentation and research are chilled.”

Though legislators emphasized that legislating an exemption would merely be codification of the case law, the House Report also identified a strong tradition within Congress of supporting statutory research exemptions, evidenced by the PVPA and the Drug Price Competition and Patent Term Restoration Act of 1984 (commonly called the Hatch-Waxman Act). The House Report states that the PVPA research exemption “was made, in part, because of the involvement of publicly funded research on plants,” and suggests that Title IV was thus appropriate because, as of 1990, “more than 50 per cent of all scientific research and experimentation is Federally funded.”

The argument in the House Report about the statutory exemption in the Hatch-Waxman Act is particularly interesting because in that legislation, Congress was responding to a Federal Circuit case from 1984 which confronted the experimental use exception: Roche Products v. Bolar Pharmaceutical. In Roche, the Federal Circuit rejected Bolar’s argument that its use of patented drugs in order to ensure FDA approval of generic drugs (meant to hit market immediately after the patent expired) was experimental use, because of its commercial purpose. Congress overturned this decision by including a narrow statutory research exemption in the Hatch-Waxman Act, which established that “the use of a patented invention in preparation for the submission of data to the Food and Drug Administration in connection with approval for marketing a drug was not an act of patent infringement,” thus shielding some

102 Id. (“It only stands to reason in this public-private partnership that government and university scientists should not be confused about the permissible parameters of their research and experimentation. Clarity about research will promote competitiveness and creativity.”).
103 Id.
104 Id. at 35.
105 Id. at 33.
106 733 F.2d 858 (Fed. Cir. 1984).
107 Id.
biomedical and pharmaceutical research from being considered infringement. Using the PVPA and Hatch-Waxman Act as examples, the House Report argued that both common law and congressional tradition supported a strong research exemption to protect basic research.

The report accompanying the Patent Competitiveness Act therefore provides clear guidance as to how legislators perceived the common law research exemption: the parameters of the exemption were murky, yes, but legislators considered the exemption itself to be well-established and completely necessary. Because the bill contains a research exemption fashioned after the common law exemption, the provisions contained in the bill provide some indication as to what legislators believed the parameters of the common law exemption were and should be. They certainly believed an exemption existed, and that it should cover experiments to “research on” and “design around” patented technology.

Indeed, even the main critic of the bill, Representative Carlos Moorhead (R-CA), recognized the existence of the common law exemption in his dissenting remarks (in fact, as a reason not to legislate an exemption):

I am aware that since 1813, the doctrine of “experimental purpose” has been recognized as an exemption to patent infringement. Throughout the years, U.S. courts have recognized that making or using a patented invention for the purpose of studying or analyzing how the invention works has not given rise to patent infringement liability, so long as this is done in a way which does not directly interfere with the commercial interests of the patentee. This long standing legal principle is sound and is a recognized feature of the patent system. I am not aware of any reason to believe that there is a need for Congress to codify this doctrine.

To Moorhead, Title IV was unnecessary not only because of the existing common law exemption, but also because it sought to protect university research which he could not perceive as endangered:

The stated purpose of this title is to protect university research activity. I fail to understand what universities are being protected from. There

---

109 Id. at 56 (emphasis added).
THE HISTORICAL TREATMENT OF THE COMMON LAW RESEARCH EXEMPTION AND ITS RELATIONSHIP TO BIOTECH LAW AND POLICY

has never been a case, to my knowledge, where a university has been sued for patent infringement for carrying on research on a patented invention. If the existing patent law is harming universities or interfering with their research, I believe they should come forward and explain the nature of the problem.¹¹⁰

At the time, it may have seemed unthinkable that a university would be sued for its basic research activities involving patented technology. And in any case, the bill evidently was not at the forefront of Congress’s agenda: the Patent Competitiveness Act, like its predecessor the Transgenic Animal Patent Reform Act, languished in Congress for several years without being passed,¹¹¹ probably due in part to the defeat of its main proponent, Kastenmeier, in the 1990 primary election.¹¹² In the coming years, as the biotechnology sector failed to live up to both positive and negative expectations, biotechnology policy fell off of the public agenda, and a statutory research exemption fell away with it.¹¹³ However, there were hints—particularly in Roche v. Bolar—that if confronted with an experimental use defense, the Federal Circuit would construe the research exemption strictly and narrowly. These hints foreshadowed future judicial decisions that would dramatically alter researchers’ understanding of the common law exception, spurred on by something that may have been unimaginable to Moorhead and his contemporaries: a university was sued for its research work.

IV. 2000s: JUDICIAL EVISCERATION OF THE COMMON LAW EXEMPTION

While the controversies surrounding biotechnology played out in the 1980s, changes in the federal court system were taking place—specifically, the Court of Appeals for the Federal Circuit was created in 1982.¹¹⁴ The Federal Circuit has subject matter jurisdiction over patent appeals from U.S. district courts. Its decisions in patent cases are crucial, because they are binding

¹¹⁰ Id. at 57.
¹¹¹ See H.R. REP. NO. 102-18, at 334 (1991) (“In the Second Session of the 101st Congress, the Subcommittee developed and the full Committee reported legislation (Title IV of H.R. 5598) to provide a research exemption to the patent laws of the United States. The bill was not considered in the House, and activity may resume on this matter in this Session.”).
¹¹² Kevles, supra note 71, at 28.
¹¹³ Id. at 28-29.
precedent in district courts throughout the United States. Indeed, shortly after coming into existence, the Federal Circuit had the opportunity to create binding precedent on the scope of the research exemption in *Roche v. Bolar*, although its decision to interpret the common law research exemption extremely narrowly was overturned quickly by Congress. In the early 2000s, the Federal Circuit had new opportunities to rule on the scope of the common law research exemption—and the court has clarified just how narrow it perceives the exemption to be.

After *Roche v. Bolar*, the next experimental use case to come up in the Federal Circuit was *Embrex, Inc. v. Service Engineering Corp.* *Embrex* had a patent on a method of inoculating chicks against diseases before they hatched; Service Engineering evaluated the patented method in an effort to design around it. Because Service Engineering planned to compete with Embrex, the Federal Circuit held that its use of the patented technology was impermissible commercial use that could not be protected by the research exemption.

This case could have come out either way: on the one hand, because Service Engineering intended to eventually profit from designing around Embrex’s technology, its experiments with the technology may not have been experimental use under Justice Story’s original formulation. But on the other hand, the patent bargain requires patentees to disclose their inventions so that others might invent new and better methods around the technology, not so that patentees can stifle attempts to design around it. The facts of *Embrex* might actually be a “paradigm case of exempted experimental use”: the researchers at Service Engineering were using the technology only to understand how to avoid infringement, and the intent to profit was only remotely related to the use. In any case, *Embrex* reaffirmed that the Federal Circuit would not permit an experimental use defense if the alleged infringer would receive commercial gain and eventual profit from experimenting with patented technology. But what about functionally non-commercial research by nonprofit entities? The Federal Circuit illustrated just how remote the commercial connection that barred the experimental use defense could be in

---

115 See supra note 106 and accompanying text.
116 216 F.3d 1343 (Fed. Cir. 2000). It is not completely clear why the exemption is brought up so infrequently as a defense, but speculation suggests that the exemption’s track record of failure in federal courts may explain why defendants do not raise it as an affirmative defense as frequently as, say, patent invalidity.
117 Id. at 1346-47.
118 Id. at 1349.
119 See supra note 22 and accompanying text.
120 Mueller, supra note 17, at 935.
two subsequent cases: Madey v. Duke University and Integra LifeSciences v. Merck.

A. Madey v. Duke

The facts and posture of Madey are worth discussing in some detail. John Madey had formerly worked for (and directed) Duke University’s Free Electron Laser lab, and invented and owned certain equipment used in the lab. Prior to the lawsuit, Madey and Duke had a particularly vicious falling out, Madey left the lab, and Duke University scientists continued using his patented equipment in non-commercial research. In addition to suing Duke on employment-related claims, Madey sued Duke for patent infringement stemming from the continued use of his equipment. The North Carolina district court dismissed the patent infringement claim on summary judgment, based in part on the experimental use defense presented by Duke and its lawyers: Duke’s use of the technology was exempt because it was in the course of non-commercial, not-for-profit research.

The Federal Circuit reviewed this judgment. Madey argued for an extremely narrow interpretation of experimental use, which would make any beneficial use of the patent infringing; Duke countered that the experimental use defense protected the university’s basic, non-commercial scientific research. Both Duke and the district court cited Ruth v. Stearns-Roger Manufacturing Co.—a 1935 case in which the research exemption protected the Colorado School of Mines regarding its experiments with patented technology—as evidence that basic university research was protected by longstanding precedent.

The Federal Circuit ultimately rejected Duke’s arguments, overturning Ruth in the process. Not only did the court reaffirm

121 307 F.3d 1351, 1352 (Fed. Cir. 2002).
122 331 F.3d 860 (Fed. Cir. 2003). This Article will only generally cover the holdings of Madey and Integra, in order to demonstrate how they conflict with legislators’ understanding of the common law exemption. There is already a wealth of scholarship on the ramifications of these cases for the common law exemption. See, e.g., Caltrider & Davis, supra note 32; Chester G. Moore, Comment, Killing the Bayh-Dole Act’s Golden Goose, 8 TUL. J. TECH. & INTELL. PROP. 151, 163-68 (2006); Mueller, supra note 17, at 936-61.
124 Id. at *6.
125 Madey, 307 F.3d at 1352.
126 Appellant’s Reply Brief at 7-13, Madey, 307 F.3d 1351 (No. 01-1567).
127 Brief of Defendant-Appellee at 15-22, Madey, 307 F.3d 1351 (No. 01-1567).
128 Id. at 22.
129 Madey, 307 F.3d at 1362.
130 See supra notes 34-35 and accompanying text.
prior holdings that no commercial use is protected by the exemption, it also established that even an extremely remote relationship between the use and the profit might prevent utilization of the experimental use defense:

Our precedent clearly does not immunize use that is in any way commercial in nature. Similarly, our precedent does not immunize any conduct that is in keeping with the alleged infringer’s legitimate business, regardless of commercial implications. For example, major research universities, such as Duke, often sanction and fund research projects with arguably no commercial application whatsoever. However, these projects unmistakably further the institution’s legitimate business objectives, including educating and enlightening students and faculty participating in these projects. These projects also serve, for example, to increase the status of the institution and lure lucrative research grants, students and faculty.

In short, regardless of whether a particular institution or entity is engaged in an endeavor for commercial gain, so long as the act is in furtherance of the alleged infringer’s legitimate business and is not solely for amusement, to satisfy idle curiosity, or for strictly philosophical inquiry, the act does not qualify for the very narrow and strictly limited experimental use defense. Moreover, the profit or nonprofit status of the user is not determinative.

In the present case, the district court attached too great a weight to the nonprofit, educational status of Duke, effectively suppressing the fact that Duke’s acts appear to be in accordance with any reasonable interpretation of Duke’s legitimate business objectives. On remand, the district court will have to significantly narrow and limit its conception of the experimental use defense. The correct focus should not be on the nonprofit status of Duke but on the legitimate business Duke is involved in and whether or not the use was solely for amusement, to
satisfy idle curiosity, or for strictly philosophical inquiry.131

The Federal Circuit thus held that experimentation or research in the university setting with “no commercial application whatsoever” may not be protected by the research exemption because of the university’s business of attracting students, faculty, and grants. At least when researchers could rely on the distinction between commercial and non-commercial work, there was some guidance as to which activities would be protected. After Madey, only the old, vague guidelines protecting experiments for “philosophical inquiry” and “idle curiosity” remained, creating more confusion for nonprofit researchers than there may have been before the holding. As one commentator has put it, under the strict test in Madey, “it appears that any use of patented tools by researchers and faculty engaged in the constant pursuit of funding, whether in the form of research grants or licensing arrangements for inventions developed at the institution, is unlikely to be experimental use.”132

Duke immediately petitioned the Supreme Court for certiorari, identifying a number of concerns: first, that all nonprofit research institutions, because they are in the business of seeking grants and attracting researchers, would no longer be eligible for the research exemption. Second, Duke argued that the unavailability of the defense would create high licensing demands and transactions costs for nonprofits facing a thicket of corporate patents in the way of their research.133 The Supreme Court took some interest in these arguments, and invited the Solicitor General to submit a brief on the issue of whether certiorari should be granted.134

The Solicitor General’s brief recommended that the petition for certiorari be denied, which it ultimately was.135 The brief reasoned that the Federal Circuit ruling was not directly antagonistic to prior experimental use precedent, nor was it an inaccurate ruling given the facts of Madey’s case.136 The arguments in the brief are also direct evidence that the model of

131 Madey, 307 F.3d at 1362-63.
136 Brief for the United States as Amicus Curiae, supra note 133, at 6-13.
university patenting promoted by the Bayh-Dole Act rendered the research exemption untenable, at least in the view of the Department of Justice. The brief states that after Bayh-Dole, the university’s role as a center of non-commercial research was no longer “clear-cut” given the rise of deep university-industry partnerships, and that universities and other research institutions deserved no blanket exemption as a result.\textsuperscript{137} Not-for-profit research institutions were no longer primarily considered centers for advancement of human knowledge, but rather became institutions with deep corporate ties and conflicts of interest. The advent of biotechnology, the Bayh-Dole duty to commercialize, and the lack of clearly defined exempt uses combined to create a perfect storm, jeopardizing the continuation of basic research in even the most independent settings.

Nevertheless, the Solicitor General stated that the experimental use defense might be ripe for legislative (as opposed to judicial) consideration. The brief identified the “weighty concerns” raised by Duke about the scope of permissible research and the feasibility of licensing, and identified Congress as the authority most capable of evaluating those concerns and creating a solution.\textsuperscript{138} The brief identified the Hatch-Waxman Act (Congress’s response to \textit{Roche v. Bolar}), the Transgenic Animal Patent Reform Act, and the Patent Competitiveness Act as evidence of Congress’s willingness and ability to address the experimental use exception if necessary.\textsuperscript{139} After certiorari was denied, universities were left questioning whether their activities were protected research, and unfortunately, the legislature took no immediate action to clarify.

\textbf{B. Integra LifeSciences v. Merck}

\textit{Integra} is less instructive for this study because the research exemption was ultimately determined to be a collateral issue by the majority of the Federal Circuit panel.\textsuperscript{140} Moreover, the Supreme Court ultimately overturned the Federal Circuit decision in favor of the defendants, but on grounds not involving the

\textsuperscript{137} \textit{Id.} at 12-13.

\textsuperscript{138} \textit{Id.} at 15-16. The Solicitor General discussed the judiciary’s ability to address these concerns: “Indeed, it seems improbable that a 190-year-old, judge-made defense with little rooting in any statutory text could anticipate the challenges of the modern academic and research environment and adequately accommodate the competing policy concerns raised by the parties in this case.”


\textsuperscript{140} Integra LifeSciences I, Ltd. v. Merck KGaA, 331 F.3d 860 (Fed. Cir. 2003).
experimental use debate. Nevertheless, in the Federal Circuit case, the dissenting judge, Judge Newman, offered an interesting view of the experimental use exemption that merits discussion, particularly in view of some recent developments in biotechnology and some of the recently proposed solutions for the research exemption problem.

The facts of Integra are somewhat complex, but essentially involve the experiments of a scientist, David Cheresh, at the (nonprofit) Scripps Research Institute. Integra had a patent directed toward recombinantly-produced peptides (RGD peptides) and certain uses for them, chiefly for healing wounds and adhering prosthetics, although Integra was never successful in commercializing its patents. Cheresh discovered a new use for certain forms of the RGD peptides: inhibiting blood vessel growth, which could have profound implications for inhibiting cancerous tumor growth. Recognizing the possibilities of this technology, Merck, a German pharmaceutical company, entered into an agreement with Scripps to develop it.

The majority did not discuss experimental use, but in her dissent, Judge Newman expressed the opinion that the experimental use exception would have properly protected some of Cheresh’s early work. Judge Newman expressed her concern that the “right to [use patented technology to] conduct research to achieve [basic] knowledge need not, and should not, await expiration of the patent,” and her frustration at the majority’s decision to further “disapprove[] and essentially eliminate[] the common law research exemption.” Judge Newman distinguished “research” from “development,” and stated that the exemption should protect the former:

[T]here is a generally recognized distinction between “research” and “development,” as a matter of scale, creativity, resource allocation, and often the level of scientific/engineering skill needed for

142 Integra, 331 F.3d at 862-63.
143 Mueller, supra note 17, at 949-50.
144 Id.
145 Id.
146 Integra, 331 F.3d at 864 n.2 (stating that the experimental use exemption was not before them in the case, but suggesting that even if it had been briefed or argued, “the Patent Act does not include the word ‘experimental,’ let alone an experimental use exemption from infringement”).
147 Id. at 874. (Newman, J., dissenting) (stating that “either the common law research exemption or the development associated with § 271(e)(1) immunity embraces all of [the allegedly infringing] activities”).
148 Id. at 873.
the project; this distinction may serve as a useful
divider, applicable in most situations. Like “fair
use” in copyright law, the great variety of possible
facts may occasionally raise dispute as to particular
cases. However, also like fair use, in most cases it
will be clear whether the exemption applies.149

Despite leaving the parameters of the exemption open, Judge
Newman did give some guidance as to the types of research
activity that should be protected:

The subject matter of patents may be studied in
order to understand it, or to improve upon it, or to
find a new use for it, or to modify or “design
around” it. Were such research subject to
prohibition by the patentee the advancement of
technology would stop, for the first patentee in the
field could bar not only patent-protected
competition, but all research that might lead to such
competition, as well as barring improvement or
challenge or avoidance of patented technology.
Today's accelerated technological advance is based
in large part on knowledge of the details of patented
inventions and how they are made and used.
Prohibition of research into such knowledge cannot
be squared with the framework of the patent
law.150

Judge Newman’s language is evocative of the “research
on”/“research with” dichotomy that has been advanced by many
scholars, including Rebecca S. Eisenberg and the National
Research Council,151 and is also evocative of the protected uses
outlined in the statutory exemption contained in the Patent
Competitiveness Act.152 Under Judge Newman’s formulation, pre-
commercial research on the technology as an end in itself—
intended to help researchers understand the invention or avoid
infringement—would be exempted, while research using the
technology as a tool or a means to another end would be
infringement.

149 Id. at 876 (footnote omitted).
150 Id. at 875.
151 See Mueller, supra note 17, at 957-59; Rebecca S. Eisenberg, Patents and the
Progress of Science: Exclusive Rights and Experimental Use, 56 U. CHI. L. REV.
1017, 1074-75 (1989); Report of the National Institutes of Health Working
news/researchtools/appendd.htm.
152 See supra notes 98-99 and accompanying text.
The majority opinion in *Integra* was ultimately vacated and remanded by the Supreme Court on other grounds, and in any case, Judge Newman’s dissent would have had no precedential force. However, the dissent reflects the desperate need for guidance in delineating the bounds of the research exemption: even a basic, vague line between “research” and “development” might aid researchers and courts in their application and assessment of permissible research activities. As the Federal Circuit was narrowing the experimental use exception in the legal sphere, new challenges for experimental use were arising in the scientific world. The patenting of genes and genetic sequences was in full swing. And as the 2000s continued, the tension between gene patents and the progress of basic research would further illustrate the need for clear guidelines to govern not-for-profit research on patented technology.

V. THE FUTURE OF THE RESEARCH EXEMPTION IN BIOTECHNOLOGY

The need for clarification of the research exemption has been heightened by the rise of gene patenting, as new questions arise about how researchers can use basic DNA strands to do beneficial research on health and disease. The history of gene patenting has been written elsewhere; suffice it to say that since around the year 2000, the Patent Office has granted patents on small strands of complementary DNA—not on methods of using them, but on the fragments themselves—allegedly because a human’s actions in isolating and purifying the fragments renders them patentable. Gene patenting has held great promise for the biotechnology industry, but has also generated objections from groups with moral and ethical concerns about patenting sequences found naturally in human and animal bodies.

Deep controversies have arisen surrounding gene patenting. Some of the most troubling questions implicate the

153 For more in-depth discussion of the *Integra* case and the research exemption, see Alcorn, supra note 132; Rebecca Lynn, Note, Merck KGaA v. Integra Lifesciences I, Ltd.: Judicial Expansion of 271(e)(1) Signals a Need for a Broad Statutory Experimental Use Exemption in Patent Law, 21 BERKELEY TECH. L.J. 79 (2006).


fate of basic research on these genes and sequences, many of which are correlated with predisposition to certain health problems or diseases. Most obviously, there is no way to “invent around” a particular genetic sequence. Researchers who wish to work on genetic therapy or diagnostic testing related to a patented gene fragment often cannot do so without infringing the patent.\textsuperscript{157} A patent on a gene sequence can hinder research on the technology and attempts at the development of tests and therapy around it; any progress that occurs must happen in the labs and by the employees of the patentee and its licensees.

As with every new advancement in biotechnology, members of Congress attempted to legislate a research exemption, this time one that would allow non-commercial research on patented genes. The Genomic Research and Diagnostic Accessibility Act of 2002\textsuperscript{158} would have protected the use of patented genetic information in the course of “systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge,” as well as use of the patented sequence in the creation of “any test, designed to detect disease, to predict the potential for a medical disorder, or to predict the effectiveness of therapeutics.”\textsuperscript{159} But like nearly every other attempt to legislate a statutory research exemption, the bill was never acted on after being referred to committee, and it expired at the end of that session of Congress.\textsuperscript{160} Another bill, the Genomic Research and Accessibility Act of 2007,\textsuperscript{161} sought to remedy the research problems by prohibiting gene patenting altogether,\textsuperscript{162} but it, too, died after being referred to committee.\textsuperscript{163}

Confusion about the scope of permissible research may deter researchers from doing work on patented genes: after a patent was granted on a gene pertaining to hemochromatosis, thirty percent of the U.S. laboratories surveyed ceased their work on it.\textsuperscript{164} Many other researchers also report stopping their work after learning that a patent has been granted, or after being contacted by the patent holder with threats of suit or offers of exorbitant

\begin{footnotes}
\item[158] H.R. 3967, 107th Cong. (2002).
\item[159] Id. §§ 2(E), 3(F).
\item[161] H.R. 977, 110th Cong. (2007).
\item[164] Andrews, supra note 15, at 805.
\end{footnotes}
licensing fees.\textsuperscript{165} In any event, the effects of confusion deter beneficial competition and development, perhaps not for the most brazen researchers, but for enough researchers that it is a serious problem.\textsuperscript{166} Restrictions on basic research threaten to stunt advances in technology that could have profound effects both for public health and for economic growth. The confusion about the permissibility of basic research on DNA fragments remains.

Gene patents are only one subset of biotechnology patents which seem to be affected by the murky research exemption. Although recent scholarship has indicated that patents and patent infringement suits may not be the most serious obstacle that scientific researchers face in their efforts to access and use patented technology,\textsuperscript{167} the extant confusion created by the unclear exemption is terrible policy for a host of reasons. Researchers and universities are somewhat less likely to engage in activity which might be infringing; excessive licensing breeds high prices and transactions costs that may be transferred to the government and the public funding the research;\textsuperscript{168} there is also the possibility that basic research that would otherwise take place in U.S. labs is being taken on by countries with robust and clear research exemptions.\textsuperscript{169} Even if many researchers are not afraid of being sued for their work with patented technology, the norms which might be keeping researchers from being sued could one day be violated (envision another Myriad, or worse, universities suing one another over their patent portfolios).

Up to this point, this Article has attempted to point out the tensions between congressional understanding of experimental use and judicial understanding of experimental use. The two branches seem to be talking past one another: in the past half-century, judges have stated that a broader exemption could only be created by the legislature, while various legislators have relied on the existence of a common law exemption in deciding whether to support or amend a statutory exemption. Adding to the mess, important advances in biotechnology—most recently, the technology involved in gene patents—have resulted in broad patents that may be having a chilling effect on the continuation of basic research, particularly

\textsuperscript{165} Id.
\textsuperscript{166} Although Ouellette’s research demonstrates that the patent problem may be overstated, she also notes that some studies indicate that “DNA patents are limiting both the availability of testing and the development of new genetic tests.” Ouellette, supra note 156, ¶ 56.
\textsuperscript{167} See id.
\textsuperscript{169} See John F. Duffy, Harmony and Diversity in Global Patent Law, 17 BERKELEY TECH. L.J. 685, 717-19 (2002) (discussing the possibility of outsourcing as a result of the lack of a research exemption in the United States).
since Madey and Integra, which rendered the parameters of exempted non-commercial research even more unclear. This Part will provide a brief overview of a few of the solutions and alternatives that scholars and policymakers have advanced for the research exemption problem, suggestions which would protect the continuation of basic research without damaging the value and the incentives that the patent system provides. Although I discuss the arguments made for different types of solutions generally—in other words, the solutions recommended below could cover research using all patented technologies—many of these solutions could in theory be narrowed to specifically address the biotechnology sector.

A. Liability Rules

One of the solutions that has been proposed by scholars such as Rebecca S. Eisenberg and Janice M. Mueller would be to change the rule protecting the patentee’s intellectual property from a property rule to a liability rule; a property rule protects an owner’s rights against any non-consensual use of the property, while a liability rule permits non-consensual use with the payment of damages after-the-fact. The liability rule would not permit a broad research exemption, but it would allow courts to examine uses of patented technology case-by-case, balancing the harm to the patent owner against the scope, nature, and necessity of use by the infringer, and adjusting damages accordingly. This approach has been advocated by Judge Rader of the Federal Circuit.

In its recent decision in eBay Inc. v. MercExchange, L.L.C., the Supreme Court held that injunctions are not an automatic remedy for patent infringement, and the Court suggested that a balance-of-harms approach and adjusted damages may be more appropriate in particular cases. The public interest in the progress of basic research and improvements in public health thus might make a liability rule a good substitute for the experimental use exception, by permitting non-commercial university and nonprofit researchers to use technology in de minimis, non-commercial ways that would likely not cause great damages. However, there are two obvious problems: first, while assessing

---

172 See Mueller, supra note 17, at 934-36 (discussing Judge Rader’s arguments for a similar rule in Embrex and Integra).
damages on past research may be possible, valuing the damages caused by prospective or ongoing experimentation may be very difficult; and second, the patent holder’s uncertain ability to wield exclusive control over use and licensing of the patent will make the patent far less valuable and will also damage the overall value of the incentives provided by the patent system.

B. A “Fair Use” Exemption

While Justice Story is famous for creating the experimental use exception that is the topic of this Article, he is also responsible for creating another intellectual property doctrine: copyright fair use. Copyright fair use allows certain users to reproduce copyrighted material without permission, under limited circumstances. Though the doctrine originated at common law, it was somewhat murky (as experimental use is today); hence, in the Copyright Act of 1976, Congress enacted a fair use exemption meant to codify the existing common law standard. The Copyright Act provides that a court considering whether a defendant’s use is “fair use” and thus not copyright infringement should consider four factors:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
2. the nature of the copyrighted work;
3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. the effect of the use upon the potential market for or value of the copyrighted work.

Many scholars have suggested that patent law might benefit from a “fair use” type exemption, in which courts could use a multi-factor test to determine whether a defendant’s use of patented technology should be protected or not. An exemption

174 See Mueller, supra note 17, at 979.
178 Copyright fair use is distinct from trademark fair use, discussed briefly in my discussion of the PVPA above. See supra note 88 and accompanying text.
could look very similar to the “four factors” test of copyright law: for example, courts could consider the purpose of the use of the patented technology, the amount of the use, and so on in determining whether a defendant’s activities would be permissible research or impermissible infringement. The test could be created through either legislative guidance or common law decisions. An unambiguous, “fair use” style research exemption in patent law would help ensure that creation and sharing would still be incentivized, and would also protect the valid interests of the patentee in the fruits of his or her hard work.

However, the suggestion to create a “fair use” affirmative defense is not without its problems. It is a broad solution requiring significant legislative or judicial action that risks unintended consequences and might result only in added complexity and confusion. It may also be expensive. Although some have alleged that creating a clear exemption would reduce litigation and the costs of case-by-case adjudication and modification of an unclear research exemption, one could envision another scenario in which defendants would be more encouraged to present the defense or bring declaratory judgment actions, and hence, litigation might increase as courts build a body of “experimental use” precedent. There are also problems with the complexity of the evidence to be considered. Although factors like the “amount of use” are comparatively easy to judge in copyrightable works, in the patent scenario, courts might be confronted with particularly detailed, scientific, and subjective documents and testimony on the amount and purpose of use, further increasing the cost of litigation. Courts would also have to face “difficult pricing decisions” to decide the scope, amount, and intent of the use, in order to perform the balancing test. Furthermore, the current process, which provides courts with flexibility in computing damages for infringement, may already provide a sort of multi-

---

*Doctrine of Fair Use in Patent Law*, 100 Colum. L. Rev. 1177, 1205 (2000) (“The preceding analysis identified five factors relevant to a fair use finding: (i) the nature of the advance represented by the infringement; (ii) the purpose of the infringing use; (iii) the nature and strength of the market failure that prevents a license from being concluded; (iv) the impact of the use on the patentee’s incentives and overall social welfare; and (v) the nature of the patented work. While this test resembles that of copyright fair use, it diverges to reflect the different incentive scheme of patent.”); see also Eisenberg, supra note 151, at 1018 n.6. (“The U.S. copyright laws also exempt some research uses of copyrighted works from infringement liability under the ‘fair use’ doctrine.”).

180 See O’Rourke, supra note 179, at 1230-31.
181 See id. at 1242.
182 Id.
183 Id. at 1246-47.
factor test which protects de minimis, nonprofit researchers from having to pay large damages to patent holders, rendering a “fair use”-type defense redundant given the significant trouble creating a new fair use doctrine might entail.

C. Compulsory Licensing, Non-exclusive Licensing, and Patent Pools

Patent pools and compulsory licenses are traditional tools of patent law that could be harnessed to ensure the continuation of valued research, and this approach has been advocated by Lori Andrews as well as the National Research Council Committee on Science, Technology, and Law. Similar consortia and collective licensing programs have been utilized with great success in Europe. Patent pools are agreements in which two or more patent holders agree to license their technology to one another (or to third parties for a set fee). These agreements prevent parties seeking to use the technologies from having to seek licenses from each individual patentee. Because of the existing norms of sharing and scholarship in academia, universities and nonprofits who invent technology might be predisposed to make commitments to join these voluntary associations.

As a similar alternative, the norms in the scientific community could be utilized to promote non-exclusive licensing of university- or nonprofit-developed technology. This approach might require universities and nonprofits to sacrifice the high payments that come with exclusive licenses, but non-exclusive licensing could promote more widespread research and development on and around patented work. Although it would be more difficult, some scholars have advocated that the non-exclusivity reforms go even further: using public domain projects such as the Human Genome Project and some components of pharmaceutical research as examples, they have suggested that basic research in universities and nonprofits that is funded by the public go directly into the public domain. This would require revision of the Bayh-Dole Act, but one could imagine a revision

187 See Cassier, supra note 168, at 94-95.
permitting and requiring universities to pursue patent protection for applications of basic research, but not for the results of basic research and new knowledge itself.\textsuperscript{189}

Compulsory licensing, in contrast, requires no volunteerism on the part of the patent holder; instead, the government can grant a license on a patent to serve the public interest.\textsuperscript{190} The march-in rights in the Bayh-Dole Act would clearly seem to give the government this ability, but to date, these rights have not been utilized.\textsuperscript{191} The obvious problem with this and all of the licensing/pooling approaches is that they require volunteer, collective, or administrative actions, all of which are susceptible to high transactions costs and inertia.

\textbf{D. Legislation}

With the common law research exemption in its current state, the best choice would be congressional action to legislate a research exemption. In the past, Congress has legislated exemptions for certain types of possibly infringing activity involving patented technology when matters of life and health are at stake: Congress has protected research use of germplasm in the PVPA, drug experimentation in the Hatch-Waxman Act, and most recently, it has exempted doctors using patented surgical procedures from being sued for infringement.\textsuperscript{192} A legislative research exemption has been advocated by nearly every single scholar who has considered the problems inherent in the current system.\textsuperscript{193} The exception could be narrowly drawn—for example, protecting only genetic testing or certain types of research on certain biotechnologies—or it could guide all basic research in the field of public health.

As this paper has shown, for the past century, nearly every scholar, legislator, and judge to consider what kinds of experimentation and research \textit{should} be protected has agreed upon two broad categories: (1) \textit{research on} the patented technology and (2) research to \textit{design around} the technology. The National

\begin{footnotesize}
\begin{itemize}
\item[\textsuperscript{189}] See id.
\item[\textsuperscript{190}] Andrews, \textit{supra} note 15, at 807.
\item[\textsuperscript{191}] See \textit{supra} note 66 and accompanying text; see also Nat'l Research Council, \textit{supra} note 186, at 96.
\item[\textsuperscript{193}] See, e.g., Nat'l Research Council, \textit{supra} note 186, at 14; Andrews, \textit{supra} note 15, at 806-07; Sherizzaan Minwalla, A Modest Proposal To Amend the Patent Code 35 U.S.C. \textsection 287(c) To Allow Health Care Providers To Examine Their Patients' DNA, 26 S. Ill. U. L.J. 471, 473 (2002); Mueller, \textit{supra} note 17, at 980.
\end{itemize}
\end{footnotesize}
Research Council Committee has drafted four useful guidelines that outline the work which it believes should be exempted:

[M]aking or using a patented invention should not be considered infringement if done to discern or to discover:

a) the validity of the patent and scope of afforded protection;
b) the features, properties, or inherent characteristics or advantages of the invention;
c) novel methods of making or using the patented invention; or
d) novel alternatives, improvements, or substitutes.

Such an exemption would seem to protect both the patentee’s interests and the user’s. Research with the patented technology for commercial development would be infringing, but research incidental to commercial business—for example, in the course of developing an alternative, or ascertaining the veracity of the patent’s specifications—would not.

While any proposed legislation would require extensive hearings and the input of judges, scholars, researchers, and private firms, the suggestions already contained in past judicial opinions and legislative history are a good start. A clear exemption would ultimately benefit not only public health and non-commercial biotechnology research, but would also help private, commercial industry. Such an exemption would offer patent holders better guidance as to when they should pursue costly and time-consuming enforcement of their patents, and when enforcement would be unsuccessful or inappropriate. Moreover, clearly defined exempt uses would allow patent holders to cut through the thicket of both commercial and non-commercial researchers that may be using their technology, allowing them to focus on strategically licensing their patents to those engaging in non-exempt work.

CONCLUSION

There is a complex web of law and policy surrounding biotechnology and the university’s role in basic biotech research: at best, this Article has strived to identify some of the incongruities

---

194 NAT’L RESEARCH COUNCIL, supra note 186, at 14.
195 Id.
196 The difficulties in outlining the scope of such an exemption are briefly discussed in Eisenberg & Nelson, supra note 188.
between Congress's historical understanding of the common law research exemption and the narrow judicial reality, and how these misunderstandings have ultimately resulted in confusion that is affecting researchers and universities alike. Legislation to protect basic biotechnology research is necessary.

The ACLU's victory in its suit against Myriad at the district court level leaves the future of gene patents uncertain. Although the court decided the case on limited grounds at summary judgment, namely, the unpatentability of products of nature, it did not rule out the possibility that gene patents may be impeding beneficial, basic research, and that further findings of fact might reveal that to be the case. If the ACLU suit survives an appeal from Myriad, litigation might be enough to effect changes in gene patenting, but as history has demonstrated, with each new advance in biotechnology there have been new problems and challenges implicating the research exemption, and clear legislative policy going forward would help alleviate the problem. At the very least, even if the ACLU suit is ultimately unsuccessful, it is possible that it will draw public attention to the problems inherent in the obstacles to basic research and will force some legislative action.

While the confusion may not yet have led to a crisis implicating public health, it would be better to have Congress act prematurely than to act too late. Since Madey and Integra, given the importance of continued basic research in biotechnology fields such as human genetics, a clear, legislated exemption to guide researchers and the universities and nonprofits that employ them is badly needed. A clear exemption would free Progress, that lofty aim of the patent law, from the patents that currently may be stifling it.

198 Id. at *77.
COURTING CHAOS:
CONFLICTING GUIDANCE FROM COURTS HIGHLIGHTS THE NEED FOR CLEARER RULES TO GOVERN THE SEARCH AND SEIZURE OF DIGITAL EVIDENCE

Lily R. Robinton*

12 YALE J.L. & TECH. 311 (2010)

ABSTRACT

Today’s digital devices allow users to store an astounding amount of personal information and data of all types. People now favor hard drives and e-mails over file cabinets and letters. When conducting criminal investigations in today’s high-tech world, forensic analysts may compare digital fingerprints rather than physical ones. Investigators must obtain search warrants before examining any digital device for evidence of criminal activity, just as they would before searching a suspect’s car, home, or office. In the digital context, however, the warrant requirement goes awry. Traditional search and seizure rules fail to prevent general, exploratory searches, which threaten individual privacy rights. Courts recognizing this problem have adopted “special approaches” for conducting digital media searches. Although these approaches provide greater protection for privacy rights, they often severely hamper legitimate law-enforcement interests. In order to both preserve privacy rights and promote justice, legislatures must enact laws directed at the search and seizure of digital media. These laws should (1) require investigators to follow narrow search protocols, but allow expanded searches where necessary; (2) require investigators to obtain a second warrant before seizing out-of-scope evidence, with a narrow exception; and 3) require a taint team to review digital media containing privileged or third party files.

* UCLA School of Law, J.D., Managing Editor, UCLA Law Review, Senior Articles Editor, UCLA Journal of Law & Technology, 2008-2009, U.C. San Diego, B.S., Biology. I would like to thank the attorneys at the U.S. Attorney’s Office for the Northern District of California, San Jose, for inspiring me to write this article, with thanks in particular to Assistant U.S. Attorneys Hanley Chew, Jeffrey Schenk, Daniel Kaleba, Susan Knight, and Jeff Nedrow for taking the time to answer my questions and guide me to resources. I would also like to thank Chris Beeson, Director of the Regional District Forensics Lab for the Northern District of California, for setting aside time in his busy schedule to give me an overview of how forensics analysts proceed with the complicated task of searching for digital evidence. And of course, thanks to my friends and family for edits and encouragement.
TABLE OF CONTENTS

INTRODUCTION ........................................................................................................313
I. THE WARRANT REQUIREMENT, AND THE “I HAVE NOTHING TO HIDE” RETORT .................................................................316
   A. The History and Framework of the Warrant Requirement ..................................................316
   B. The “I Have Nothing To Hide” Retort .................................................................320
II. LET’S GET DIGITAL! DIGITAL! .................................................................................321
   A. Digital Context Complications .................................................................323
   B. Current Law Enforcement Methods for Conducting Digital Searches ...........................................324
III. DEVELOPING DIGITAL RULES ..............................................................................327
   A. Traditional Rules Allow Investigators To Scan All Digital Media Files—No Search Protocol Required ......328
   B. Problems with the Plain View Doctrine in the Context of Digital Searches ..............................331
   C. Plain View Problems for Privileged Data and Third Parties 334
   D. The Carey-Winick “Special Approach” to Digital Searches 339
IV. A STATUTORY SOLUTION ......................................................................................341
   A. With Privacy and Justice for All .................................................................342
      1. The Narrow Search Protocol Requirement..................343
      2. The Plain View Doctrine and the Second Warrant Requirement ..........................344
      3. Protecting Third Parties and Privileged Documents 345
CONCLUSION ..............................................................................................................346
INTRODUCTION

On August 26, 2009, searches and seizures of digital property within the Ninth Circuit ground to a screeching halt. The culprit? A landmark opinion authored on that date by the Chief Judge of the Ninth Circuit, Alex Kozinski, in a case already fraught with controversy: United States v. Comprehensive Drug Testing, Inc. (CDT). CDT stemmed from the federal government’s investigation into the illegal distribution and use of steroids in Major League Baseball (MLB), which implicated well-known players such as Barry Bonds, Alex Rodriguez, Sammy Sosa, and Manny Ramirez. Several years of investigation gave the government probable cause to believe that at least ten Major League Baseball players had received illegal steroids from Bay Area Labs Cooperative. Federal investigators obtained a warrant to search the computer records of a private company retained by the MLB Players’ Association to oversee its drug testing program. The warrant authorized seizure of drug test records pertaining to those ten named players, but prosecutors discovered and reviewed a directory containing hundreds of records relating to other sports’ drug testing programs. Prosecutors then sought additional warrants to seize records and specimens pertaining to approximately one hundred other players who had tested positive for steroids. This move led to a heated debate in several lower courts over whether the government acted properly in reviewing and seeking additional warrants for data that fell outside the scope of the initial search.

In CDT, a limited en banc panel comprised of eleven judges overturned an earlier Ninth Circuit opinion written by a three-judge panel in favor of the government. The CDT majority

---

1 See Brief for the United States in Support of Rehearing En Banc by the Full Court at 1, 6, United States v. Comprehensive Drug Testing, Inc. (CDT II), 579 F.3d 989 (9th Cir. 2009) (en banc) (Nos. 05-10067, 05-15006, 05-55354) (“The government is accordingly laboring under the direct effects of this new legal regime. Many United States Attorney’s Offices have been chilled from seeking any new warrants to search computers.”).

2 CDT II, 579 F.3d.


4 See United States v. Comprehensive Drug Testing (CDT I), 513 F.3d 1085, 1090-94 & n.5 (9th Cir. 2008), rev’d en banc and vacated, 579 F.3d 989.

5 See CDT II, 579 F.3d at 993-94.

6 See CDT I, 513 F.3d 1085 (9th Cir. 2008), rev’d en banc and vacated, 579 F.3d 989.
held that the government had willfully disregarded the limitations of the search warrant to obtain out-of-scope evidence illegally; however, it was not the majority’s disapproval of the government’s actions that catalyzed the ensuing squall. In his August 26, 2009, opinion, Chief Judge Kozinski wrote, “This case is about a federal investigation into steroid use by professional baseball players. More generally, however, it’s about the procedures and safeguards that federal courts must observe in issuing and administering search warrants and subpoenas for electronically stored information.” After determining that the government had acted improperly, the opinion set forth extremely restrictive guidelines to govern the search and seizure of digital property. The majority took a significant step in shifting its focus from the facts of the case to the general issue of how magistrates and federal agents should issue and execute search warrants for electronically stored information. Styled as “guidelines,” but viewed by magistrate judges as mandatory, the new rules set forth in CDT have wreaked havoc on government investigations in the Ninth Circuit, and have been criticized for departing from controlling precedent.9 The opinion has caused such a stir that Solicitor General Elena Kagan, along with every U.S. Attorney’s Office in the Ninth Circuit, and

7 579 F.3d at 993.
8 Chief Judge Kozinski summarized his rules as follows:
   1. Magistrates should insist that the government waive reliance upon the plain view doctrine in digital evidence cases.
   2. Segregation and redaction must be either done by specialized personnel or an independent third party. If the segregation is to be done by government computer personnel, it must agree in the warrant application that the computer personnel will not disclose to the investigators any information other than that which is the target of the warrant.
   3. Warrants and subpoenas must disclose the actual risks of destruction of information as well as prior efforts to seize that information in other judicial fora.
   4. The government’s search protocol must be designed to uncover only the information for which it has probable cause, and only that information may be examined by the case agents.
   5. The government must destroy or, if the recipient may lawfully possess it, return non-responsive data, keeping the issuing magistrate informed about when it has done so and what it has kept.

Id. at 1006 (internal citations omitted); see Hugh Kaplan & Christine Mumford, Attorneys, Academics Sort Through Landmark Case on Computer Searches, 85 CRIM. L. REP. 688 (2009).
9 See Brief for the United States in Support of Rehearing En Banc by the Full Court, supra note 1, at 1, 5, 8-14.
five top attorneys from Main Justice, have petitioned the full twenty-seven member court to reconsider the decision.\(^\text{10}\)

This is not the first controversy to arise over the search and seizure of digital property. Crooks and innocents alike store information relating to every facet of their lives on digital devices, making them attractive targets for criminal investigators.\(^\text{11}\) Courts across the nation have struggled to apply Fourth Amendment principles to digital searches to ensure the searches do not expand into exploratory hunts that threaten individual privacy. Their attempts have produced a tangle of conflicting authority, and as demonstrated by \textit{CDT}, a digital search resolution remains elusive. To add to the confusion, different government agencies may disagree about how to approach and execute a search and seizure of digital property.\(^\text{12}\)

Regardless of whether the Ninth Circuit accepts or declines the federal government’s petition to reconsider its decision in \textit{CDT}, or whether the Court overturns the decision, the storm of controversy created by \textit{CDT} has underscored the need for a uniform set of rules that successfully balances individual privacy concerns against legitimate law enforcement interests. As Chief Judge Kozinski stated, “Everyone’s interests are best served if there are clear rules to follow that strike a fair balance between the legitimate needs of law enforcement and the right of individuals and enterprises to the privacy that is at the heart of the Fourth Amendment.”\(^\text{13}\) His rules, however, along with the rules of other courts, have thus far fallen woefully short of achieving this balance. This Article addresses these conflicting interests and argues for a legislative solution that combines and harmonizes existing rules.

Part I of this Article begins by introducing the history and framework of the warrant requirement, which grew from the Fourth Amendment. Part I also addresses the significance of the threat to privacy posed by unlimited digital searches in response to those who claim that law-abiding citizens need not worry about privacy intrusions. Part II addresses the complications introduced by digital media, and Part III explains the conflicting ways in which courts have responded to these complications. Part IV argues that legislatures should create statutory schemes to address

\(^{10}\) See Elias, \textit{supra} note 3; Laura Ernde, \textit{Prosecutors: Steroid Ruling Hurting Other Investigations: Obama Asks 9th Circuit To Reconsider Steroid Ruling}, DAILY J., Nov. 27, 2009. Due to the large size of the Ninth Circuit, a limited en banc panel consisting of eleven judges usually convenes to hear appeals. See Elias, \textit{supra} note 3.

\(^{11}\) See \textit{infra} Part II.

\(^{12}\) See Kaplan & Mumford, \textit{supra} note 8.

\(^{13}\) \textit{CDT II}, 579 F.3d 989, 1006 (9th Cir. 2009).
the issue. Part IV goes on to propose rules governing search and seizure of digital property, which would tackle the unique privacy concerns raised by these searches without hampering government investigations.

I. THE WARRANT REQUIREMENT, AND THE "I HAVE NOTHING TO HIDE" RETORT

A. The History and Framework of the Warrant Requirement

In 1761, the citizens of Massachusetts lived without the protections of the Fourth Amendment. They lived in a world where "writs of assistance," a type of general warrant, authorized meticulous searches of their private homes and businesses, and allowed searching officials to pry open locks, cast aside bars, and seize offending articles on no more than bare suspicion. In February of 1761, a lawyer named James Otis gave a passionate, five-hour speech against the perils of the "writ of assistance." What Otis lacked in brevity, he made up for in emotion as he effectively conveyed the fear incited by the specter of the general warrant:

Every one with this writ may be a tyrant; if this commission be legal, a tyrant in a legal manner, also, may control, imprison, or murder any one within the realm. In the next place, it is perpetual; there is no return. A man is accountable to no person for his doings. Every man may reign secure in his petty tyranny, and spread terror and desolation around him, until the trump of the

16 See Ridpath, supra note 14, at 48; Otis, In Opposition to Writs of Assistance, supra note 15; see also Akhil Reed Amar, The Fourth Amendment, Boston, and the Writs of Assistance, 30 Suffolk U. L. Rev. 53, 53 (1996). James Otis represented a group of Boston Merchants opposing the writs before the Superior Court of Massachusetts. Id. at 76.
Archangel shall excite different emotions in his soul.\textsuperscript{17}

Otis' plea to banish the writ of assistance fell on deaf ears,\textsuperscript{18} but a few years later, in the cases of Wilkes v. Wood\textsuperscript{19} and Entick v. Carrington,\textsuperscript{20} the English court refused to allow the government to rely on general warrants lacking probable cause to justify the arrests of political activists and subsequent searches of their homes and belongings. These two cases have been called "the O.J. Simpson and Rodney King cases of their day,"\textsuperscript{21} and likely influenced the Framers of the Constitution as they drafted an amendment that would protect the American citizens against the terrors preached by James Otis.\textsuperscript{22}

The Fourth Amendment, straight from the quills of the Framers, ensures that

\[
\text{[t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.}
\]

A government-instigated search of anything a person deems private must be reasonable to pass muster under the Fourth Amendment.\textsuperscript{24} The Supreme Court has developed two procedural tools to ensure the protections of the Fourth Amendment. The first

\textsuperscript{17} See RIDPATH, supra note 14, at 53; Otis, In Opposition to Writs of Assistance, supra note 15.
\textsuperscript{18} Writs of Assistance, supra note 14.
\textsuperscript{19} Wilkes v. Wood, (1763) 98 Eng. Rep. 489 (C.P.); Lofft 1.
\textsuperscript{21} Amar, supra note 16, at 65.
\textsuperscript{22} See id. at 64-65; Trepel, supra note 15, at 123-24.
\textsuperscript{23} U.S. CONST. amend. IV.
\textsuperscript{24} See id.; Carroll v. United States, 267 U.S. 132, 155 (1925) (setting forth a reasonableness standard for probable cause); Investigations and Police Practices, 38 GEO. L.J. ANN. REV. CRIM. PROC. 3, 43-44 (2009) ("Under the Fourth Amendment, every search or seizure by a government agent must be reasonable."); Regensburger, supra note 3, at 1156. The Fourth Amendment protects individuals from government intrusions that invade privacy. Investigations and Police Practices, supra, at 5-8. A person must have a legitimate expectation of privacy to merit protection under the Fourth Amendment. Id. If a person has a subjective expectation of privacy, and society accepts that expectation as objectively reasonable, the Supreme Court will deem that expectation legitimate. Id.
is the warrant requirement. With some exceptions, warrantless
searches are per se unreasonable under the Fourth Amendment.25

Before investigators can obtain a search warrant, they must
have probable cause to believe they will discover evidence of the
alleged crime during the search. A neutral magistrate must
consider the facts and circumstances presented in a warrant
application, and may issue the warrant only after finding a
substantial basis that probable cause exists to search the named
area and seize any evidence.26 The warrant must describe with
particularity the places investigators plan to search and items they
hope to seize. The particularity requirement defines the scope of
the warrant, and protects the privacy interests in a person’s home
and possessions from broad, exploratory rummaging by ensuring
that each search is narrowly tailored to the justifications presented
to the magistrate.27

A warrant contains sufficient particularity when it leaves
nothing to the discretion of the executing officers and officers “can
with reasonable effort ascertain and identify the place intended.”28
However, an overbroad warrant, or a warrant containing mistaken
information may be “cured” if executing officers can rely on
personal knowledge to narrow and identify the place intended to be
searched.29 An affidavit incorporated by reference or attachment to

25 The warrant requirement applies to any place in which a person holds a
reasonable expectation of privacy. See Raphael Winick, Searches and Seizures
26 See id.; see also Illinois v. Gates, 462 U.S. 213, 238 (1983) (“The task of the
issuing magistrate is simply to make a practical, common-sense decision
whether, given all the circumstances set forth in the affidavit before him,
including the ‘veracity’ and ‘basis of knowledge’ of persons supplying hearsay
information, there is a fair probability that contraband or evidence of a crime
will be found in a particular place.”); Investigations and Police Practices, supra
27 See U.S. CONST. amend. IV (requiring that warrants shall “particularly
describe[e] the place to be searched, and the persons or things to be seized”);
Andresen v. Maryland, 427 U.S. 463, 480 (1976) (“[T]he problem [posed by the
general warrant] is ‘not that of intrusion Per se, but of a general, exploratory
rummaging in a person’s belongings.’ . . . [T]he Fourth Amendment addresses
the problem] by requiring a ‘particular description’ of the things to be seized.”
(alterations in original) (quoting Coolidge v. New Hampshire, 403 U.S. 443, 467
(1971)); Investigations and Police Practices, supra note 24, at 27-28; Winick,
supra note 25, at 86.
28 Steele v. United States, 267 U.S. 498, 503 (1925); see also Andresen, 427
29 See, e.g., United States v. Judd, 889 F.2d 1410, 1413 (5th Cir. 1989) (holding
that a warrant that failed to describe a company’s second office still sufficiently
particular because agents checked city business license records, bank records,
corporate filings, and the address on the company’s letterhead to determine the
location to be searched; the offices were only 25-30 feet apart; and the company
had only leased the second office three weeks prior to the search). But see
United States v. Ellis, 971 F.2d 701, 704 (11th Cir. 1992) (holding that officers’
the warrant that lists items not mentioned in the warrant itself may also cure an overbroad warrant.\(^{30}\)

When a defendant challenges the particularity of a warrant authorizing the search of his or her records, courts will deem the warrant particular if it has been narrowed as much as the information available to the agents will allow. For example, in *United States v. Gardiner*,\(^{31}\) the Sixth Circuit found a warrant listing a variety of personal and business records to be sufficiently particular because it sought items pertaining to the time frame of the crime, and all of the listed records would logically relate to the alleged financial crimes.\(^{32}\) In *United States v. Mathison*,\(^{33}\) the Eighth Circuit deemed particular a warrant seeking all records pertaining to the suspect’s seventeen distinct businesses, as well as information with respect to eleven individuals. The Eighth Circuit reasoned that the affidavit supporting the warrant contained considerable evidence of the suspect’s involvement in illegal activities and demonstrated that the records sought would substantiate the suspect’s involvement.\(^{34}\) In finding the warrant sufficiently particular, the court in *Mathison* declined to employ the exclusionary rule, which is the second procedural tool designed to ensure the protections of the Fourth Amendment.\(^{35}\) The exclusionary rule functions to suppress from the record any evidence obtained through an illegal search and seizure. This can include evidence obtained during execution of an overbroad warrant and evidence that falls outside the scope of the warrant.\(^{36}\)

The exclusionary rule does not proceed directly from the Fourth Amendment as a means of “curing the invasion of the defendant’s rights which he has already suffered,”\(^{37}\) but rather developed as a judicially created remedy designed to deter

\(^{30}\) See, e.g., *In re Search of Office of Tylman*, 245 F.3d 978, 980-81 (7th Cir. 2001); *United States v. Towne*, 997 F.2d 537, 548 (9th Cir. 1993).

\(^{31}\) 463 F.3d 445 (6th Cir. 2006).

\(^{32}\) Id. at 471.

\(^{33}\) 157 F.3d 541 (8th Cir. 1998).

\(^{34}\) Id. at 547-49.

\(^{35}\) Id. at 549.


constitutional violations.\(^{38}\) Thus, where the exclusion of evidence would not advance the purpose of the rule, the court will allow its introduction under the so-called good faith exception.\(^{39}\) For example, in *United States v. Leon*,\(^{40}\) the district court found that the magistrate who had granted the warrant had done so in error, as the evidence submitted with the warrant application failed to establish probable cause.\(^{41}\) Nevertheless, the Supreme Court declined to suppress evidence on the grounds that the exclusionary rule “cannot be expected, and should not be applied, to deter objectively reasonable law enforcement activity.”\(^{42}\) The good faith exception to the exclusionary rule thus highlights the importance of issuing warrants based on a proper articulation of places to be searched and items to be seized.

**B. The “I Have Nothing To Hide” Retort**

A person will likely feel that the government has violated his right to privacy if agents begin rummaging through the medicine cabinet to search for stolen fifty-inch flat screens, which clearly cannot fit next to the Aspirin. The same person, however, may fail to perceive an examination of every file on his computer as intrusive. The obscure nature of the digital search, and the lack of any spatial correlation between the evidence sought and the files examined, can mask potential privacy violations. It is easy to discount the danger of the general digital search and argue complacently, “So what if the government looks through every file on my computer? I have nothing to hide. I’d rather sacrifice a little privacy for the sake of bringing criminals to justice.” On this view, justice should trump privacy: government intrusion into digital data would be a threat to criminals, but not to law abiding citizens. Most people do not sympathize with the white-collar criminal who gets caught with child pornography during a search of his computer for evidence of investment fraud. Most proponents of the “nothing to hide” viewpoint would argue that only criminals need fear a general search of digital media.\(^{43}\)

\(^{38}\) See *Leon*, 468 U.S. at 906; *Investigations and Police Practices*, supra note 24, at 201.

\(^{39}\) See *Leon*, 468 U.S. at 919-20; *Investigations and Police Practices*, supra note 24, at 204-06.


\(^{41}\) Id. at 900-03.

\(^{42}\) Id. at 919.

\(^{43}\) See Daniel J. Solove, “I’ve Got Nothing To Hide” and Other Misunderstandings of Privacy, 44 SAN DIEGO L. REV. 745, 746-47 (2007) (“The argument that no privacy problem exists if a person has nothing to hide is frequently made in connection with many privacy issues. . . . The nothing to
Scholar Daniel Solove addresses this common retort by pointing out that “[p]rivacy . . . is not the trumpeting of the individual against society’s interests, but the protection of the individual based on society’s own norms and values. . . . [P]rivacy has a social value.” Solove argues persuasively that a society without the cushion of privacy would be unlivable; life in a free society necessitates rules that may unintentionally shield criminal behavior.

Furthermore, even those who think they have nothing to hide might find themselves unpleasantly surprised by what might turn up in a probing search of their digital media. Solove captures this possibility with a colorful quotation from Friedrich Durrenmatt’s novella Traps, in which a man who believes himself innocent inquires as to his crime: “‘An altogether minor matter,’ the prosecutor replies[.] . . . ‘A crime can always be found.’” In a world where computers facilitate and store oceans of data about every aspect of our lives, it seems certain that some type of crime can always be found among the bits and bytes of the average hard drive.

II. LET’S GET DIGITAL! DIGITAL!

Fourth Amendment jurisprudence has evolved to suit searches conducted in a physical world with the human senses. “In the time before the atom, what we could see with our eyes was all there was. Similarly, when the country was young and the universe of searchable data was limited to ‘papers, and effects,’ law enforcement agents were able to literally see everything covered by Fourth Amendment protections.” This is no longer true. Technology now allows us to conduct much of our social and professional lives in cyberspace, while storing hoards of information of all types in digital format. People today use computers to store images, movies, documents, personal records, and correspondence. Computers double as “photo albums,
stereos, telephones, desktops, file cabinets, waste paper baskets, and televisions,”50 “postal services, playgrounds, jukeboxes, dating services, movie theaters, daily planners, shopping malls, personal secretaries, virtual diaries, and more.”51 The storage capacity of computers today is astonishing. As of April of 2009, the highest capacity commercial hard drives were capable of storing two terabytes of data. A terabyte can hold approximately 1000 hours of video, 250,000 four-minute songs, 1,000,000 thick books of about 500 pages each, or as much information as can be printed on the paper from 50,000 trees. A desktop hard drive might store between 120 gigabytes and two terabytes.52 Even a measly 80-gigabyte desktop drive stores the equivalent of 40 million pages of text.53

When investigators decide to search a suspect’s computer, they face vast quantities of information. If stored in written form, that data might fill an entire library. As technology advances to allow a user to squeeze larger quantities of data into tinier spaces, the amount of information that can be contained in digital format will continue to grow.54 Furthermore, in addition to the wealth of information stored on files purposely saved by an individual, investigators mine the hard drive for deleted files and glean information from metadata.55


50 Trepel, supra note 15, at 128.

51 United States v. Andrus, 483 F.3d 711, 718 (10th Cir. 2007) (quoting Orin S. Kerr, Searches and Seizures in a Digital World, 119 HARV. L. REV. 531, 569 (2005)).


53 Trepel, supra note 15, at 128-29.

54 “North Carolina State University engineers have created a new material that would allow a fingernail-size computer chip to store the equivalent of 20 high-definition DVDs or 250 million pages of text, far exceeding the storage capacities of today’s computer memory systems.” Researchers Develop Material That Could Boost Data Storage, Save Energy, PHYSORG.COM, Oct. 20, 2009, http://www.physorg.com/news175252581.html.

55 Metadata consists of information that characterizes the digitally stored data and answers the “who, what, when, where, why, and how about every facet of the data that are being documented” on a digital storage device. USGS, Frequently-Asked Questions on FGDC Metadata, http://geology.usgs.gov/tools/metadata/tools/doc/faq.html (last visited Oct. 14, 2009); see also Trepel, supra note 15, at 129.
A. Digital Context Complications

The advanced storage capabilities of today’s digital media complicate the scope of digital searches. For one, the particularity requirement can malfunction in the digital context because searching for evidence of a crime on a computer is akin to searching for a needle in a haystack. Investigators usually cannot predict where, or in what format, they might find the relevant information, and thus cannot “particularly describ[e]” the “place to be searched” or the “things to be seized.” As Professor Orin Kerr points out, “[i]n the physical world, different spatial regions are used for different purposes. This allows the police to make educated guesses as to where evidence may or may not be found . . . .” In the physical world, one might look for an incriminating letter in a file folder or a desk drawer. Drugs or guns might be stored in shoeboxes or bedside tables. The money might be under the mattress. In the computer context, however, the location of evidence does not necessarily depend on the character of the evidence itself. Information stored on a computer is represented by “zeros and ones of electricity,” making the format and location of any stored information flexible, and difficult to predict. Investigators searching a suspect’s house for stolen stereo equipment can logically rule out the medicine cabinet as a possible location, but anticipating the location of electronic evidence is inherently more difficult because “electronic evidence can be located anywhere. . . . [T]he investigator can never rule out a particular part of the hard drive ex ante.”

Some courts resolve the issue by allowing warrants to describe the media to be searched in general terms, without requiring investigators to pinpoint the particular files they plan to search. This rule recognizes the concern that investigators might not be able to predict whether evidence will be located on a suspect’s computer, an external hard drive, a CD, a flash drive, or some other external storage device. Other courts have taken a

---

56 U.S. Const. amend. IV; see Kerr, supra note 47, at 303.
57 Kerr, supra note 47, at 303.
58 Id. at 284.
59 Id. at 304.
60 See United States v. Lacy, 119 F.3d 742, 746 (9th Cir. 1997) (“[T]his type of generic classification is acceptable ‘when a more precise description is not possible.’” (alteration in original) (quoting United States v. Cardwell, 680 F.2d 75, 78 (9th Cir. 1982))).
61 Regensburger, supra note 3, at 1157; Stanley, supra note 48, at 217; see, e.g., United States v. Lacy, 119 F.3d 742, 746-47 (9th Cir. 1997) (“The government knew Lacy had downloaded computerized visual depictions of child pornography, but did not know whether the images were stored on the hard drive or on one or more of his many computer disks . . . there was no way to specify
more restrictive approach by requiring investigators to name at least the type of evidence sought. Most courts, however, allow broad particularity designations when investigators cannot predict precisely which files contain evidence. These courts recognize that investigators may need to seize information that appears innocuous, but that may later prove incriminating in conjunction with other evidence.

B. Current Law Enforcement Methods for Conducting Digital Searches

So how might investigators wade through this quagmire of data? The simplest option would be for the officer to sit down at the suspect’s computer and examine the data manually. The officer would simply turn the computer on, and begin opening files one-at-a-time in search of something incriminating.

Practical drawbacks, however, preclude the use of this method. For one, the investigator would have to sift through a forest’s worth of documents, making the search extremely inefficient. Such a search would fail to locate incriminating files deleted by the suspect, and the officer would risk destroying evidence during the search process. Simply opening a file or turning on a computer can overwrite deleted data, and may alter time stamps on the data, which investigators might need to show what hardware and software had to be seized to retrieve the images accurately.”).

62 See, e.g., United States v. Riccardi, 405 F.3d 852, 862 (10th Cir. 2005).
63 See United States v. Wuagneux, 683 F.2d 1343, 1349 (11th Cir. 1982) ("[C]rimes may require the assembly of a ‘paper puzzle’ from a large number of seemingly innocuous pieces of individual evidence: ‘The complexity of an illegal scheme may not be used as a shield to avoid detection when the State has demonstrated probable cause to believe that a crime has been committed and probable cause to believe that evidence of this crime is in the suspect's possession.’ It is universally recognized that the particularity requirement must be applied with a practical margin of flexibility, depending on the type of property to be seized, and that a description of property will be acceptable if it is as specific as the circumstances and nature of activity under investigation permit.” (quoting Andresen v. Maryland, 427 U.S. 463, 481 n.10 (1976))); Regensburger, supra note 3, at 1156-57; see also United States v. Jacob, 657 F.2d 49, 52 (4th Cir. 1981) (recognizing the complexity of the alleged crime as a factor in determining whether the warrant met with particularity requirements); United States v. Abrams, 615 F.2d 541, 548 (1st Cir. 1980) (Campbell, J., concurring) ("The investigators usually do not, and often cannot, know in advance precisely what they will find when they search through files pursuant to a warrant. They, therefore, may find it difficult to describe what they are seeking, other than to say that they expect to find, and will seize, documents constituting evidence of the particular fraud.").

the time the suspect created or last accessed a file. Not only would this method erase possibly relevant information, it would also defeat investigators’ attempts to authenticate the evidence and disprove tampering. An investigator might as well “walk[] into a murder scene with muddy boots, remov[e], bare-handed, a knife from the victim, drop[] it in his coat pocket and return[] to the office.”

It is nearly impossible to search a hard drive without the assistance of some kind of software program. To search digital media properly, investigators elicit the help of digital forensics specialists, who use a number of tools and forensics techniques. Courts have consistently recognized that requiring police to search digital media at the suspect’s home or office could create an extreme burden on both the individual’s privacy, as well as on police resources. Investigators might need to camp out for days to conduct a thorough search, which would severely disrupt the suspect’s life or business. Forensic analysts would need to cart their own computers, equipped with forensic tools and special programs, to the scene of every search. This would create an enormous hassle and burden investigative resources. To alleviate these concerns, courts generally permit removal of digital media to an off-site location for examination by experts, although some courts urge investigators to return equipment as soon as possible to minimize disruption of an individual’s activities.

---

65 Id. at 1094.
66 United States v. Long, 425 F.3d 482, 487 (7th Cir. 2005) (“[W]e observe that it is impossible to search computer hardware or software without using some type of software.”); Interview with Chris Beeson, Director of the Regional District Forensics Lab for the Northern District of California (Dec. 2, 2009). The Regional District Forensics Lab for the Northern District of California is one of fourteen regional computer forensics laboratories across the country. Regional computer forensic labs conduct digital forensic examinations for all law enforcement agencies, including the Federal Bureau of Investigation, within each region. Interview with Chris Beeson, supra.
68 See Thomas K. Clancy, The Fourth Amendment Aspects of Computer Searches and Seizure: A Perspective and a Primer, 75 MISS. L.J. 193, 267 (2006); Orin Kerr, Searches and Seizures in a Digital World, 119 HARV. L. REV. 531, 541 (2005); McLain, supra note 64, at 1093-94; see also, e.g., United States v. Walser, 275 F.3d 981, 985 (10th Cir. 2001); United States v. Hay, 231 F.3d 630, 637 (9th Cir. 2000); United States v. Upham, 168 F.3d 532, 535-36 (1st Cir. 1999) (upholding seizure and subsequent off-site search of computer for “needles in the computer haystack”); United States v. Hill, 322 F. Supp. 2d 1081, 1089 (C.D. Cal. 2004) (“To be certain that the medium in question does...
To avoid contaminating or damaging any digital evidence, forensic investigators first make a “bitstream” copy of the media they plan to search. The bitstream image captures every piece of information on the target drive, including files accessible by the normal user, deleted files, metadata, and empty space. Investigators save the bitstream copy in “read only” format to ensure they do not accidentally alter the evidence during analysis.  

Forensic investigators then narrow the set of data to be searched using “known fingerprints” or “hash value” programs, and forensic tools such as EnCase or Forensic Tool Kit (FTK).  

Before analyzing investigators’ search methods under the Fourth Amendment, one must have a basic understanding of how hash value programs operate. A “hash value” is an identifier that characterizes a data set. The relationship between a hash value and its data set compares roughly to the relationship between an organism and its DNA sequence; analysis of two separate data sets will rarely return the same hash value. Just as forensic analysts use DNA to determine the identities of criminal suspects or victims, digital forensic investigators use hash values to identify data—programs, images, files, etc.—on a computer.  

A hash value program converts each data set on the target drive into its corresponding identifier and matches the resulting identifiers with known identifiers. For example, investigators might suspect an individual of using a specific hacking program, or of downloading a particular image of child pornography. By comparing hash values from the suspect’s computer with known values for the hacking program or the image, investigators can

not contain any seizable material, the officers would have to examine every one of what may be thousands of files on a disk—a process that could take many hours and perhaps days. Taking that much time to conduct the search would not only impose a significant and unjustified burden on police resources, it would also make the search more intrusive.” (internal citation omitted)); cf. United States v. Leveto, 343 F. Supp. 2d 434, 441 (W.D. Pa. 2004) (approving of investigators’ steps to minimize upheaval of defendant’s business, including downloading and copying files at the scene rather than removing them for off-site review); United States v. Hunter, 13 F. Supp. 2d 574, 583 (D. Vt. 1998) (“At the very least, the government should copy and return the equipment as soon as possible.”).

69 See Clancy, supra note 68, at 265-67; Interview with Chris Beeson, supra note 66.

70 McLain, supra note 64, at 1094; Interview with Chris Beeson, supra note 66; Interview with Hanley Chew, Assistant U.S. Att’y, U.S. Attorney’s Office, San Jose Branch (Nov. 16, 2009).

71 Interview with Chris Beeson, supra note 66; Interview with Hanley Chew, supra note 70; see also Kerr, supra note 68, at 541 (“A hash is a complicated mathematical operation, performed by a computer on a string of data, that can be used to determine whether two files are identical.”).
determine if either exists on the suspect’s hard drive.\textsuperscript{72} Hash programs can also recognize “normal” files, such as Microsoft Windows or Word Perfect, which commonly turn up on computers. Forensic investigators negatively screen out these common operating system files and applications to reduce the size of the data set they will search.\textsuperscript{73}

Forensic tools such as EnCase and FTK allow investigators to access deleted files, eliminate common operating system files, preview image files, flag encrypted files, and search the entire hard drive or active files by keyword or phrase. These tools also allow investigators to identify mismatched file extensions. For instance, if a suspect attempts to hide incriminating evidence and mislead investigators by changing the .jpg extensions on images of child pornography to .doc extensions, the forensic program will alert the investigator to the altered files.\textsuperscript{74} However, criminals may devise other strategies for disguising incriminating evidence, which forensic tools will not detect. For example, the suspect might embed the image of child pornography within a word document, as opposed to changing the file extension. A forensic investigator searching for images in .jpg files may overlook the embedded evidence in the .doc file. The forensic program will not flag such files as altered or suspicious.\textsuperscript{75}

\textbf{III. Developing Digital Rules}

As digital media searches have become more frequent, courts face the challenge of applying Fourth Amendment principles, which were designed for discrete physical-world searches, to vast and amorphous digital spaces. Courts and scholars remain divided on the issue, and have roughly separated into two camps.\textsuperscript{76} Adherents of one viewpoint advocate for application of existing rules to digital searches, and argue that computers are nothing more than glorified containers holding files that represent physical documents.\textsuperscript{77} Followers of the other viewpoint argue that the “container analogy” is inadequate, and search of digital media

\begin{itemize}
  \item Kerr, \textit{supra} note 68, at 541; Interview with Hanley Chew, \textit{supra} note 70. Forensic analysts also use hashes to ensure the bitstream copy accurately matches the original drive. Kerr, \textit{supra} note 68, at 546.
  \item Interview with Chris Beeson, \textit{supra} note 66.
  \item McLain, \textit{supra} note 64, at 1094-95.
  \item Interview with Chris Beeson, \textit{supra} note 66.
  \item See Clancy, \textit{supra} note 68, at 196.
  \item See, e.g., \textit{id.} at 271 (arguing that “there is nothing ‘special’ in the nature of computer searches that differentiate [sic] them in any principled way from other document and container searches.”).
\end{itemize}
requires a “special approach” with new rules and procedures. The prevailing concern in both camps remains the same: whether existing principles suffice in the digital arena to prevent every digital search from becoming the kind of general, exploratory search prohibited by the Fourth Amendment.

A. Traditional Rules Allow Investigators To Scan All Digital Media Files—No Search Protocol Required

Courts willing to compare computers to file cabinets recognize that the versatility and massive storage capacity of computers complicate the task of parsing through intermingled files. Investigators examining computers face a plethora of intermingled data and cannot avoid combing through oceans of material not specified in the warrant. Courts have resolved this issue with respect to physical documents by allowing investigators to scan all documents in order to ascertain their relevancy. In Andresen v. Maryland, the court noted:

We recognize that there are grave dangers inherent in executing a warrant authorizing a search and seizure of a person’s papers that are not necessarily present in executing a warrant to search for physical objects whose relevance is more easily ascertainable. In searches for papers, it is certain that some innocuous documents will be examined, at least cursorily, in order to determine whether they

---

78 See, e.g., United States v. Carey, 172 F.3d 1268, 1275 & n.7 (10th Cir. 1999) (discrediting the comparison of computer searches to searches of file cabinets, and advocating for a “special approach” with respect to search and seizure of digital evidence); Winick, supra note 25, at 110 (“An analogy between a computer and a container oversimplifies a complex area of Fourth Amendment doctrine and ignores the realities of massive modern computer storage.”).

79 See United States v. Hunter, 13 F. Supp. 2d 574, 583 (D. Vt. 1998) (“Computer searches present the same problem as document searches—the intermingling of relevant and irrelevant material—but to a heightened degree.”).

80 See, e.g., Andresen v. Maryland, 427 U.S. 463 (1976); United States v. Schandl, 947 F.2d 462, 465 (11th Cir. 1991) (“It was inevitable that some irrelevant materials would be seized as agents searched through numerous documents for evidence of tax evasion and failure to file, crimes that are generally only detected through the careful analysis and synthesis of a large number of documents.”); United States v. Slocum, 708 F.2d 587, 604 (11th Cir. 1983) (“[A]n officer acting pursuant to such a warrant is entitled to examine any document he discovers, but that ‘the perusal must cease at the point of which the warrant’s inapplicability to each document is clear.’” (quoting United States v. Heldt, 668 F.2d 1238, 1267 (D.C. Cir. 1981))); United States v. Abbell, 963 F. Supp. 1178, 1198 (S.D. Fla. 1997) (“When executing a search warrant for documents, searching agents are entitled to at least cursorily examine each document at the specified search location.”).
COURTING CHAOS: THE NEED FOR CLEARER RULES TO GOVERN THE SEARCH AND SEIZURE OF DIGITAL EVIDENCE

are, in fact, among the papers authorized to be seized.81

By allowing investigators to peruse all documents in a suspect’s possession, as opposed to only those stored in folders with relevant labels, courts “recognize[] the reality that few people keep documents of their criminal transactions in a folder marked ‘drug records.’”82 Accordingly, most courts that accept the comparison between computers and file cabinets allow investigators to open and scan all digital files to ascertain the responsiveness of the data.83 In United States v. Gray,84 the court noted that “[c]omputer records are extremely susceptible to tampering, hiding, or destruction” and concluded that the searching agent “was not required to accept as accurate any file name or suffix and limit his search accordingly.”85 The courts in United States v. Hunter and United States v. Hill also declined to limit investigators’ search methods on the grounds that criminals can easily mask incriminating evidence so it will not be discovered using rigid, predictable protocols.86 In United States v. Fumo,87 the court stated that regardless of the search protocols or keywords used by the government, the government may open and briefly examine each computer file to determine whether it is in the description recited in the warrant . . . ‘no tenet of the Fourth Amendment prohibits a search merely because it cannot be performed with surgical precision.’88

Some scholars argue that limiting the ability of investigators to scour digital media might encourage criminals to hide evidence outside the range of the search. They contend that

81 Andresen, 427 U.S. at 482 n.11.
82 United States v. Riley, 906 F.2d 841, 845 (2d Cir. 1990).
83 See Clancy, supra note 68, at 198. For example, United States v. Gray compared digital evidence to paper records and documents, which “unlike illegal drugs or other contraband, may not appear incriminating on their face. As a result, in any search for records or documents, ‘innocuous records must be examined to determine whether they fall into the category of those papers covered by the search warrant.’” 78 F. Supp. 2d 524, 528 (E.D. Va. 1999) (quoting United States v. Kufrovich, 997 F. Supp. 246, 264 (D. Conn. 1997)).
84 78 F. Supp. 2d 524.
85 Id. at 529.
88 Id. at *16-17.
because criminals can disguise files, investigators have probable cause to view every file and should not be forced to employ restrictive search methods.\(^8\) Supporting the stance against search protocols, the Supreme Court has ruled outside the digital search context that warrants need not outline the methods investigators plan to employ in conducting a search:

Often in executing a warrant the police may find it necessary to interfere with privacy rights not explicitly considered by the judge who issued the warrant. . . . It would extend the Warrant Clause to the extreme to require that, whenever it is reasonably likely that Fourth Amendment rights may be affected in more than one way, the court must set forth precisely the procedures to be followed by the executing officers.\(^9\)

This reasoning has been applied in the digital search scenario.\(^9\) In fact, DOJ guidelines expressly direct prosecutors to oppose restrictions imposed by magistrates that require the government to specify how it will examine digital media to find evidence responsive to the warrant.\(^9\) Nevertheless, allowing agents to search every digital media file creates very real concerns. Many courts and commentators have reacted to this broad authorization by claiming that it operates with the plain view doctrine to transform every digital search into the type of general search prohibited by the Fourth Amendment.\(^9\)

\(^8\) See, e.g., Chang, supra note 49, at 48-50; Clancy, supra note 68; Regensburger, supra note 3, at 1196-97.
\(^9\) Dalia v. United States, 441 U.S. 238, 257-58 (1979) (holding that government agents need not specify the means by which they would execute installation of a wiretap authorized by warrant).
\(^9\) See, e.g., United States v. Vilar, No. S3 05-CR-621, 2007 U.S. Dist. LEXIS 26993, at *121-25 (S.D.N.Y. Apr. 4, 2007) (citing Dalia for the proposition that officers should not be required to specify ahead of time how they planned to search defendants computers, and stating that "[t]he majority view rejecting a protocol requirement makes good sense as there is no principle in the law that requires law enforcement officers to limit their investigative techniques ex ante, before conducting any kind of search.").
\(^9\) DOJ GUIDELINES, supra note 67, at 80.
\(^9\) See Chang, supra note 49, at 43-44; Kerr, supra note 47, at 304-05; Trepel, supra note 15, at 137-38; Winick, supra note 25, at 107-09.
B. Problems with the Plain View Doctrine in the Context of Digital Searches

The plain view doctrine operates as one of several exceptions to the warrant requirement. Under the plain view doctrine, investigators may seize incriminating evidence without a warrant if they encounter the evidence in plain view during lawful observation of the area. Horton v. California established three requirements that investigators must meet before lawfully seizing evidence in plain view. First, the investigator must have lawful authority to be in the position from which he had occasion to observe the evidence. Second, the evidence must be in plain view. Third, the incriminating character of the evidence must be “immediately apparent”—the plain view doctrine does not authorize further investigation to determine the evidentiary value of the evidence. A number of courts have expanded the plain view doctrine to encompass “plain touch,” “plain smell,” and “plain hearing” corollaries.

The plain view doctrine applies easily to items that appear incriminating at first glance, such as drugs or guns, and some documents such as fake ID’s, gambling records, and documents

94 See Investigations and Police Practices, supra note 24, at 44 (“[T]he kinds of searches and seizures are valid as exceptions to the probable cause and warrant requirements, including investigatory stops, investigatory detentions of property, warrantless arrests, searches incident to a valid arrest, seizures of items in plain view, searches and seizures justified by exigent circumstances, consent searches, searches of vehicles, searches of containers, inventory searches, border searches, searches at sea, administrative searches, and searches in which the special needs of law enforcement make the probable cause and warrant requirements impracticable.”).
95 See id. at 74-75.
97 See id. at 136-37; Arizona v. Hicks, 480 U.S. 321, 326-28 (1987) (noting that allowing further investigation beyond a cursory examination would “especially erode the plurality’s warning in Coolidge that the ‘plain view’ doctrine may not be used to extend a general exploratory search from one object to another until something incriminating at last emerges” (quoting Coolidge v. New Hampshire, 403 U.S. 443, 466 (1971))).
98 See Minnesota v. Dickerson, 508 U.S. 366, 375 (1993) (“We think that [the plain view] doctrine has an obvious application by analogy to cases in which an officer discovers contraband through the sense of touch during an otherwise lawful search.”).
99 See, e.g., United States v. Staula, 80 F.3d 596, 602 (1st Cir. 1996) (“[O]fficer evidence furnishes the officer with probable cause to conduct a search of the confined area.”).
100 See United States v. Ceballos, 385 F.3d 1120, 1124 (7th Cir. 2004) (citing a “plain hearing” exception to the search warrant requirement); United States v. Fisch, 474 F.2d 1071, 1077 (9th Cir. 1973) (likening the plain view doctrine to a situation in which officers eavesdropped on suspects from their rented hotel room).
linking co-defendants.\textsuperscript{101} It is less clear that incriminating evidence is in “plain view” when discovered as investigators examine intermingled documents in an attempt to separate responsive from non-responsive items. Difficult cases are easy to conceive. For example, imagine investigators have obtained a warrant to search a suspect’s belongings for evidence of bank fraud. The warrant allows the investigators to scan each letter in a stack of letters to determine which, if any, contain evidence of bank fraud. An investigator reads the first paragraph of one letter, but cannot determine whether the letter is relevant, so he continues reading. In the middle of the second paragraph, he reads the statement, “I hid the cocaine in the cookie jar. Just ask Jim for the ‘fresh baked goods’ and leave the money with him.” Assuming the remainder of the letter contains nothing relevant to bank fraud, its contents clearly fall outside the scope of the investigators’ warrant. The investigators would like to use the evidence against the suspect in a subsequent drug trial. Given that the investigator had to read the first two paragraphs of the letter before discovering any evidence of drug trafficking, was the evidence really “immediately apparent” as required by the plain view doctrine?

Courts answer this question with respect to physical evidence by allowing a “brief perusal of documents in plain view in order to determine whether probable cause exists for their seizure under the warrant. If in the course of that perusal, their otherwise incriminating character becomes obvious, they may be seized.”\textsuperscript{102} Courts have generally permitted investigators to seize documents discovered in “plain view,” accepting without much discussion that the investigators must conduct some degree of perusal to ascertain the relevance of the documents.\textsuperscript{103} Investigators need not be “absolutely certain” that documents or other items discovered in plain view constitute evidence of the crime at hand,\textsuperscript{104} and may “test their belief by proceeding with a limited inspection of the ‘incriminating object.’” However, “perusal must cease at the point at which the warrant’s inapplicability to each document is clear.”\textsuperscript{105}

\textsuperscript{101} See Regensburger, supra note 3, at 1197.
\textsuperscript{102} United States v. Heldt, 668 F.2d 1238, 1267 (D.C. Cir. 1981).
\textsuperscript{103} See United States v. Ochs, 595 F.2d 1247, 1258 n.8 (2d Cir. 1979).
\textsuperscript{104} See, e.g., id. at 1258; United States v. Duckett, 583 F.2d 1309, 1314 (5th Cir. 1978) (“There is no rule of law which requires an officer to know with absolute certainty that all elements of a putative crime have been completed when he seizes an article which reasonably appears to be incriminating evidence.”); United States v. Pugh, 566 F.2d 626, 627 (8th Cir. 1977); Mapp v. Warden, 531 F.2d 1167 (2d Cir. 1976); United States v. Smollar, 357 F. Supp. 628, 632 (S.D.N.Y. 1972) (“The plain view exception would be worthless if officers had to be ‘absolutely certain’ that what they saw was seizable”).
\textsuperscript{105} Heldt, 668 F.2d at 1267.
Courts have begun to apply this reasoning to digital data. Most cases involve application of the plain view doctrine to digital evidence that is incriminating on its face. The visual nature of child pornography makes it the most common type of evidence seized under the plain view doctrine during searches of digital property. However, courts have suggested the plain view doctrine would apply to admit evidence that is not graphic in nature if found after a brief scan of data on a computer screen.

This practice arguably transforms all digital search warrants into general warrants and raises the question: what if investigators are always in a position from which they can view everything on the computer screen? If courts authorize investigators to scan every piece of data on a hard drive to determine its relevancy, then investigators will always be in a lawful position from which to view evidence of unrelated crimes. The warrant’s scope would lose all relevance because any evidence not covered by the warrant could be seized under the plain view doctrine. Digital searches would become fishing expeditions.

The majority joining Chief Judge Kozinski’s ruling in CDT must have recognized this danger. The majority took the drastic view that to “avoid this illogical result, the government should, in future warrant applications, forswear reliance on the plain view doctrine or any similar doctrine that would allow it to retain data to which it has gained access only because it was required to segregate seizable from non-seizable data.”

The plain view doctrine, however, serves an important function in both physical and digital contexts. If police come upon out-of-scope evidence during the course of an otherwise lawful search it could be “dangerous—to the evidence or to the police

---


107 See, e.g., United States v. Gray, 78 F. Supp. 2d 524, 531 n.11 (E.D. Va. 1999) (“Agent Ehuan could have continued his systematic search of defendant’s computer files pursuant to the first search warrant, and, as long as he was searching for the items listed in the warrant, any child pornography discovered in the course of that search could have been seized under the ‘plain view’ doctrine.”); State v. Mays, 829 N.E.2d 773, 779 (Ohio Ct. App. 2005) (holding that an officer’s observation of the words “he will die today” on defendant’s computer screen while lawfully present in defendant’s home fell within the ambit of the plain view doctrine); Commonwealth v. Hinds, 768 N.E.2d 1067 (Mass. 2002) (holding that the officer “was not obligated to disregard files listed in plain view on the ‘Chuck’ directory whose titles suggested contents that were contraband”).

108 See Chang, supra note 49 (arguing for the abolition of the plain view doctrine in order to prevent general searches of digital property).

109 CDT II, 579 F.3d 989, 998 (9th Cir. 2009).
themselves—to require them to ignore it."¹¹⁰ A search for evidence of tax evasion might yield evidence that the suspects plans to kill his wife for insurance money. Police may need to seize out-of-scope evidence to prevent a heinous crime.

Requiring investigators to forswear reliance on the plain view doctrine in the Ninth Circuit has already had deleterious effects on law enforcement efforts to uphold the law and protect vulnerable individuals. In the Western District of Washington, federal agents received information from their counterparts in San Diego that two individuals had filmed themselves raping a four-year-old girl and traded the images via the internet. The agents did not obtain a warrant to search the suspects’ computers, however, because of concerns that any evidence discovered about other potential victims could not be disclosed by the filter team.¹¹¹

Fortunately, federal agents could refer the case to state authorities, who are not bound by the restrictions outlined in CDT. This example stresses that a blanket elimination of the plain view doctrine could create more threats to society than it prevents.

Eliminating the plain view doctrine may have less dramatic, but equally serious effects. If investigators could not seize out-of-scope evidence in plain view, investigators might “result in the loss of highly probative evidence about the very crime under investigation.”¹¹² This could occur, for example, “when a warrant contains a date restriction but the resulting search turns up evidence that the crime began or continued after officers previously had reason to believe.”¹¹³ Investigators have a legitimate interest in pursuing out-of-scope evidence to uphold society’s laws and thwart criminal activity, when possible.

C. Plain View Problems for Privileged Data and Third Parties

The view-all-use-all practices that result from the direct application of traditional rules to digital searches also raise hackles where privileged documents and third parties are concerned.¹¹⁴

¹¹¹ Brief for the United States in Support of Rehearing En Banc by the Full Court, supra note 1, at 6-7.
¹¹² Id. at 14.
¹¹³ Id.
¹¹⁴ See United States v. Abbell, 963 F. Supp. 1178, 1198-99 (S.D. Fla. 1997) (allowing perusal of all files, and approving of the government’s use of a taint team to protect privileged materials); State v. Viatical Servs., Inc., 741 So. 2d 334
Investigators may peruse all intermingled data to ascertain its relevancy, and in doing so may examine privileged documents or third party information regardless of whether it falls within the scope of the warrant. Incriminating information in “plain view” may be seized and used to charge third parties previously considered innocent.\(^\text{115}\)

In order to protect third party privacy, the Attorney General has issued rules requiring federal officers to pursue relevant evidence in the hands of disinterested third parties by issuing subpoenas rather than warrants.\(^\text{116}\) Pursuant to these rules, federal officers may only seek a warrant to obtain materials from a disinterested third party when it appears that “the use of a subpoena, summons, request, or other less intrusive alternative means of obtaining the materials would substantially jeopardize the availability or usefulness of the materials sought.”\(^\text{117}\) This policy sounds promising, but third parties receive little protection once the government decides to pursue evidence with a warrant because “failure to comply with this policy ‘may not be litigated, and a court may not entertain such an issue as the basis for the suppression or exclusion of evidence.’”\(^\text{118}\)

560, 563 (Fla. Dist. Ct. App. 1999) (“[T]he court must fashion a remedy to protect the privacy rights of innocent third parties while still allowing the state to proceed with its criminal investigation.”); Chang, \textit{supra} note 49, at 58 (discussing privilege as a possible limitation on the plain view doctrine); Regensburger, \textit{supra} note 3, at 1153, 1170-72 (expressing concern that a third party will get “caught up in the government’s dragnet,” and analyzing the use of taint teams to prevent prosecutors from accessing privileged data).

\(^\text{115}\) See, \textit{e.g.}, CDT II, 579 F.3d 989, 1005 (9th Cir. 2009) (The directory searched by the government “contained a huge number of drug testing records, not only of the ten players for whom the government had probable cause but hundreds of other professional baseball players, thirteen other sports organizations, three unrelated sporting competitions, and a non-sports business entity—thousands of files in all, reflecting the test results of an unknown number of people, most having no relationship to professional baseball except that they had the bad luck of having their test results stored on the same computer as the baseball players.”).

\(^\text{116}\) See \textit{DOJ GUIDELINES, supra} note 67, at 111 (citing 28 C.F.R. § 59.4(a)(1) (2009)).

\(^\text{117}\) 28 C.F.R. § 59.4(a)(1).

\(^\text{118}\) See \textit{DOJ GUIDELINES, supra} note 67, at 111 (quoting 28 C.F.R. § 59.5(b)). Congress has enacted statutory schemes to offer a higher degree of protection to third party internet service providers, and third party publishers, journalists, authors, or other individuals where a search of his or her possessions may implicate First Amendment concerns. See \textit{id.} at 101-09, 112-33. The Stored Communications Act (SCA), 18 U.S.C. §§ 2701-2712 (2006), regulates government access to electronic records stored by third-party service providers. See \textit{DOJ GUIDELINES, supra} note 67, at 112. The Privacy Protection Act (PPA), 42 U.S.C. § 2000aa, governs federal computer searches when agents have reason to believe they will encounter materials relating to freedom of expression. See \textit{DOJ GUIDELINES, supra} note 67, at 101. However, both acts
Where the government has obtained a warrant to search digital media containing privileged information, the DOJ Guidelines offer strategies for reviewing privileged computer files:

First, the court itself may review the files in camera. Second, the presiding judge may appoint a neutral third party known as a “special master” to the task of reviewing the files. Third, a team of prosecutors or agents who are not working on the case may form a “filter team” or “taint team” to help execute the search and review the files afterwards. The filter team sets up a so-called “ethical wall” between the evidence and the prosecution team, permitting only unprivileged files to pass over the wall.119

To protect privileged information, the Department of Justice prefers to segregate data using taint teams composed of attorneys or agents who are not members of the prosecution team. However, the use of taint teams is not mandatory, nor do all jurisdictions condone the use of taint teams.120 Computers belonging to medical and legal professionals often contain a spectrum of confidential or privileged material such as client or patient communications, medical records, or attorney work product.121 “[T]he use of computerized equipment for the storage and exchange of sensitive confidential information has become commonplace.”122 If a comprehensive search requires investigators to review every file, it seems that privileged documents must suffer some type of scrutiny. The fact that the scrutinizing eyes do not belong directly to the

---

119 See DOJ GUIDELINES, supra note 67, at 110. The “ethical wall” is also referred to as a “Chinese wall.” See In re Search Warrant for Law Offices, 153 F.R.D. 55, 57 (S.D.N.Y. 1994).

120 See Chang, supra note 49, at 58.


Courting Chaos: The Need for Clearer Rules to Govern the Search and Seizure of Digital Evidence

A prosecution team may offer little comfort. Courts have vacillated on whether in-camera review, review by special master, or review by a taint team most effectively protects privileged digital materials without unduly hampering the government’s investigation.

The court in United States v. SDI Future Health, Inc. noted, “[f]ederal courts have taken a skeptical view of the Government’s use of ‘taint teams’ as an appropriate method for determining whether seized or subpoenaed records are protected by the attorney-client privilege.” The court in In re Grand Jury Subpoenas cautioned that taint teams might leak confidential information. The government has an interest in uncovering every scrap of evidence to further the investigation, and while some leaks occur through human error, human nature may lead taint-team attorneys to violate their ethical obligations. “[T]he government’s fox is left in charge of the appellants’ henhouse, and may err by neglect or malice, as well as by honest differences of opinion.”

Courts have also noted the drawbacks to using a special master or neutral magistrate to separate privileged materials. In Black v. United States, the court pointed out that using special masters can incur high costs and fees, and they may take a prohibitive amount of time to review the contested materials. In one instance, appointment of a special master delayed a criminal trial for two and a half years. Such a delay could “effectively deprive the Government of any access to any of the seized information.” A magistrate or special master might have many duties that conflict with the task of reviewing millions of computer

123 See id. at 516 (“The Plaintiffs have a serious concern that disclosure to taint team prosecutors would not protect the confidentiality and privacy rights they here assert.”).
125 Id. at 1037.
126 In re Grand Jury Subpoenas, 454 F. 3d 511 (6th Cir. 2006).
127 Id. at 523.
128 Id.; see also United States v. Abbell, 914 F. Supp. 519 (S.D. Fla. 1995) (deciding that privileged materials should be reviewed by a special master despite the government’s appointment of a taint team); In re Search Warrant for Law Offices, 153 F.R.D. 55, 59 (S.D.N.Y. 1994) (“[T]his Court notes that reliance on the implementation of a Chinese Wall, especially in the context of a criminal prosecution, is highly questionable, and should be discouraged.”); cf. United States v. Neill, 952 F. Supp. 834, 841 n.14 (D.D.C. 1997) (criticizing the use of taint teams, but noting that “[h]owever unwise this policy decision may be, absent a showing of harm, it does not offend the Constitution”).
130 Id. at 514 n.4, 516 (citing Abbell, 914 F. Supp. 519).
131 Id. at 516.
files, whereas a taint team specifically designated to segregate privileged materials could do so in a timely manner.\textsuperscript{132}

The court in the case of \textit{In re 5444 Westheimer Rd. Suite 1570}\textsuperscript{133} determined that the lengthy amount of time a special master or magistrate judge would require to review privileged materials outweighed the protection this method might afford. Instead, the court allowed the government to proceed with taint team review.\textsuperscript{134} The court noted that the government’s taint team procedure did not prejudice defendants because the use of a taint team gave defendants the opportunity to challenge the taint team’s privilege determinations in front of the court. The court would then resolve any privilege disputes before the taint team could disclose materials classified as unprivileged to the prosecution team.\textsuperscript{135} In contrast, a neutral magistrate or special master may not offer the defendants the same opportunity to challenge privilege determinations.\textsuperscript{136}

The controversy incited by the government’s seizure of third party drug records in \textit{CDT} demonstrates how computers’ massive storage capacities have magnified the problem. The fact that the targeted computers contained vast quantities of third party data allowed the government to seize “thousands of medical records and test results involving every single Major League Baseball player,” and “thousands of other medical records for individuals in thirteen other major sports organizations, three unaffiliated business entities, and three sports competitions,” despite the fact that the government only had a search warrant for a small handful of MLB players, and none of the other individuals were the subject of any criminal inquiry.\textsuperscript{137} This result highlights the distinction between physical and digital searches: “people now have personal data that are stored with that of innumerable strangers. Seizure of, for example, Google’s email servers to look

\textsuperscript{132} See DOJ GUIDELINES, supra note 67, at 110; \textit{Black}, 172 F.R.D. at 516 n.8.
\textsuperscript{134} \textit{Id}. at *9.
\textsuperscript{135} \textit{Id}. at *9, *11 n.5.
\textsuperscript{136} See \textit{United States v. Grant}, which approved the use of a taint team. No. 04 CR 207, 2004 U.S. Dist. LEXIS 9462 (S.D.N.Y. May 25, 2004). The court noted that “after the initial privilege determination is made by the special master or judicial officer, the Government would not have the opportunity to brief or argue the ruling aided by the contents of the documents. Without the benefit of such a review, the privilege team would likely be unable to argue, for example, that no attorney-client privilege attached to the communication because of the crime-fraud exception, or that a document should be available for use at trial, regardless of work-product contents, because of necessity and unavailability by other means.” \textit{Id}. at *4-5.
\textsuperscript{137} \textit{CDT I}, 513 F.3d 1085, 1116-17 (9th Cir. 2008) (Thomas, J., dissenting).
COURTING CHAOS: THE NEED FOR CLEARER RULES TO GOVERN THE SEARCH AND SEIZURE OF DIGITAL EVIDENCE

for a few incriminating messages could jeopardize the privacy of millions.138

Investigators searching a suspect’s home or office could rarely net such an abundance of evidence. The court in Black v. United States might have been speaking to privileged documents when it called for “a re-thinking of some of the traditional approaches Courts have made in years gone by,”139 but it was not alone in seeking a new approach to regulating search and seizure of digital media.

D. The Carey-Winick “Special Approach” to Digital Searches

Both scholars and courts have referenced the enormous storage capabilities of digital media to justify the viewpoint that digital searches cannot and should not be compared to physical-world searches.140 In his influential article, Ralph Winick emphasized that the “very quantity and variety of information” on a computer “increases the likelihood that highly personal information, irrelevant to the subject of the lawful investigation, will also be searched or seized.”141 Winick recognized the threat created by allowing investigators to examine intermingled documents, and argued that application of traditional rules to searches of digital media “allows officers to gain a window into all aspects of a suspect’s life.”142 Instead, Winick advocated for applying the “intermingled-document” approach outlined in United States v. Tamura143 to digital media searches. In his proposal for a new approach, Winick discredited the theory that comprehensive computer searches require investigators to peruse every file on the hard drive. Instead of conducting a full review of digital files, he proposed investigators limit their search of the data by file type, or use key word searches to locate relevant files. He opined that government agents should seal any intermingled files, and submit specific search protocol to a neutral magistrate for approval before proceeding with review of those files.144

138 CDT II, 579 F.3d 989, 1005 (9th Cir. 2009).
140 See, e.g., United States v. Carey, 172 F.3d 1268, 1275 (10th Cir. 1999); Chang, supra note 49, at 35-36; Kerr, supra note 47, at 301-03.
141 Winick, supra note 25, at 105.
142 Id. at 111.
143 694 F.2d 591, 595-96 (9th Cir. 1982) (“In the comparatively rare instances where documents are so intermingled that they cannot feasibly be sorted on site, we suggest that the Government and law enforcement officials generally can avoid violating fourth amendment rights by sealing and holding the documents pending approval by a magistrate of a further search . . . .”).
144 Winick, supra note 25, at 107-09.
The Tenth Circuit has been the strongest proponent of using a “special approach” for digital media searches.\textsuperscript{145} The court in United States v. Carey\textsuperscript{146} became the first to adopt Winick’s “special approach” explicitly.\textsuperscript{147} Carey expanded upon Winick’s “special approach” by suggesting that investigators should obtain a second warrant before seizing out-of-scope evidence discovered in plain view.\textsuperscript{148}

While some have approved of the “Carey-Winick” approach,\textsuperscript{149} others have been quick to point out its flaws. Courts rejecting the approach have continued to allow investigators to peruse all intermingled documents because requiring search methods would be too restrictive.\textsuperscript{150} Scholars also questioned the wisdom of limiting the extent to which investigators could open and view files, citing the argument that criminals may disguise evidence in ways investigators may not be able to predict.\textsuperscript{151}

\textsuperscript{145} Regensburger, supra note 3, at 1157.
\textsuperscript{146} United States v. Carey, 172 F.3d 1268 (10th Cir. 1999).
\textsuperscript{147} See id. at 1275-76; Trepel, supra note 15, at 130.
\textsuperscript{148} Carey, 172 F.3d at 1271, 1276. DOJ guidelines note that when agents discover evidence in plain view that is not identified by the warrant, it would be a “safe practice” to obtain a second warrant. However, this practice is not mandatory. See DOJ GUIDELINES, supra note 67, at 90.
\textsuperscript{149} The term “Carey-Winick” was coined by David Ziff in a 2005 piece criticizing the approach’s limitations. See David J.S. Ziff, Note, Fourth Amendment Limitations on the Execution of Computer Searches Conducted Pursuant to a Warrant, 105 COLUM. L. REV. 841, 842 & n.4 (2005). For examples of courts approving of the Carey-Winick limitations, see United States v. Riccardi, 405 F.3d 852, 862 (10th Cir. 2005); United States v. Walser, 275 F.3d 981, 986-87 (10th Cir. 2001); Trulock v. Freeh, 275 F.3d 391, 411 n.2 (4th Cir. 2001); and People v. Carratu, 755 N.Y.S.2d 800 (N.Y. Sup. Ct. 2003).
\textsuperscript{150} See United States v. Adjani, 452 F.3d 1140, 1149-50 (9th Cir. 2006); United States v. Hill, 322 F. Supp. 2d 1081, 1090-91 (C.D. Cal. 2004) (“Forcing police to limit their searches to files that the suspect has labeled in a particular way would be much like saying police may not seize a plastic bag containing a powdery white substance if it is labeled ‘flour’ or ‘talcum powder.’ There is no way to know what is in a file without examining its contents, just as there is no sure way of separating talcum from cocaine except by testing it.”); United States v. Gray, 78 F. Supp. 2d 524, 529 (E.D. Va. 1999); Commonwealth v. McDermott, 864 N.E.2d 471, 488-89 (Mass. 2007) (specifically declining to adopt the special approach outlined in Carey).
\textsuperscript{151} See Clancy, supra note 68, at 207-08 (“[T]here are significant reasons to reject [Carey’s] position that a search be restricted by file names or file types. Professional investigators recognize that computer users attempt to conceal criminal evidence by storing it ‘in random order with deceptive file names,’ thus requiring a search of all the stored data to determine whether it is included in the warrant.” (citation omitted)); Trepel, supra note 15, at 134 (“According to Kerr, the process required by the Carey-Tamura approach is flawed for the very practical reason that ‘computer forensics is contingent, fact-bound, and quite unpredictable.’ An investigator will not know beforehand which operating system is on the device to be searched, which software is on it, or whether the suspect attempted to hide or disguise any incriminating files.”) (footnote

340

Published by Yale Law School Legal Scholarship Repository, 2010
COURTING CHAOS: THE NEED FOR CLEARER RULES TO GOVERN THE SEARCH AND SEIZURE OF DIGITAL EVIDENCE

Others have criticized the suggestion that a neutral magistrate should determine which files the government should access. When Winick first outlined his proposal, the idea of magistrate oversight was plausible: computers at that time only held 100 megabytes of data, the equivalent of 100,000 typed pages. Investigators could reasonably print the file directory for magistrate review. Modern computers, however, store considerably more information, and it would be prohibitively time consuming for lawyers to quibble in front of a magistrate over the contents of a huge volume of files.\footnote{See Ziff, \textit{supra} note 149, at 860-61.}

Finally, some have found fault with the use of a second warrant to pursue evidence discovered outside the scope of the original warrant. A second warrant may fail to protect privacy concerns implicated by the plain view doctrine because investigators will have already discovered the out-of-scope evidence without previous probable cause. Assume for the sake of argument that investigators should forswear reliance on the plain view doctrine when conducting digital searches. Under this rule, if investigators searching for evidence of bank fraud come across child pornography, they cannot seize it as evidence “in plain view,” but they can apply for a second warrant to expand their search based on the image they have just discovered. The application will likely be granted since investigators now have concrete evidence of a possible second crime. For the purposes of the second warrant, it matters not whether the evidence fell within the scope of the first. Therefore, the second warrant procedure creates a loophole to the ban on seizing evidence in plain view by authorizing investigators to seize the same out-of-scope evidence without relying on the plain view doctrine.\footnote{See Chang, \textit{supra} note 49, at 48, 50.} According to critics, this “easy-to-meet” procedure is “functionally equivalent to the plain view doctrine.”

There has been no consensus amongst different factions as to how investigators should execute search warrants for digital media without violating the Fourth Amendment. This dissonance has led to conflicting rules and results, highlighting the dire need for lawmakers to issue guidance or regulation in this area.

IV. A STATUTORY SOLUTION

To effect a solution, federal and state legislatures should adopt a set of rules that augment the Fourth Amendment with

\footnote{Id. at 50.}
respect to search and seizure of digital media. For example, Rule 41 of the Federal Rules of Criminal Procedure currently governs search and seizure pursuant to a warrant for federal investigations. This rule and corresponding state rules could be amended to incorporate specific sections pertaining to warrants for search and seizure of digital property. Legislative action offers several advantages over the solutions implemented by courts. Legislatures are not limited by stare decisis, and thus have more flexibility to design new rules. While courts may stray from precedent when changed conditions and increased knowledge render existing rules unworkable, a court’s influence may extend only as far as its jurisdiction. Legislatures can promulgate rules that span jurisdictions, thus facilitating consistent practices.

An opinion issued by the Supreme Court could settle controversy across jurisdictions, but the Supreme Court may only address issues presented in the case before it. The Supreme Court will never hear a case presenting every nuanced issue raised by a digital media search. Legislatures, on the other hand, do not need to wait until a problem presents itself. Legislatures can also effect changes much more quickly than many courts across many jurisdictions.

Additionally, legislatures may be more competent than courts to address the problem of digital media searches, because “it is difficult for judges to fashion lasting guidance when technologies are new and rapidly changing.” While legislatures receive information from a wide range of sources, including legislative hearings, advocacy groups, constituent and public input, national media, and special caucuses, judges receive information funneled through the briefs and oral arguments of two parties. Legislative branches are better situated to gather information about the developing technologies that both complicate and facilitate digital searches.

A. With Privacy and Justice for All

To walk the line between privacy and justice, legislatures need to adopt rules for digital property searches that offer more protection than the traditional approach to search and seizure, but fewer restrictions than the “special approach” described by the

155 FED. R. CRIM. P. 41.
156 See Kerr, supra note 47, at 308.
157 See Trepel, supra note 15, at 142.
159 Kerr, supra note 158, at 858.
160 Id. at 875-76.
Courting Chaos: The Need for Clearer Rules to Govern the Search and Seizure of Digital Evidence

Carey-Winick doctrine. The following three subsections each describes a rule that legislatures should adopt to strike this balance. Legislatures should: 1) require narrow search protocols, but allow file-by-file searches where necessary; 2) require a second warrant for the seizure of ambiguous out-of-scope evidence, but allow investigators to seize contraband under the plain view doctrine; and 3) require a taint team to review privileged or third party files. The taint team should afford the defense the opportunity to challenge privileged determinations. The third subsection also recommends a procedure for the taint team to follow if it encounters evidence that incriminates a third party.

1. The Narrow Search Protocol Requirement

As explained in Section IV.A, the traditional approach to search and seizure in theory allows investigators to scan every file in search of evidence. Investigators need not follow a pre-specified protocol to examine digital records. In reality, however, investigators have neither the time nor the resources to scan every piece of data on a hard drive. Forensic investigators will use whatever methods they can to narrow the subset of data they must search in order to discover evidence that responds to the warrant. Creating a rule that requires investigators to use reasonable methods to narrow their searches would serve the interests of both the Fourth Amendment and government investigators. Where less intrusive, more effective search methods exist, it would be unreasonable not to require investigators to employ those methods.

Legislatures should therefore require forensic investigators to begin searching digital media with available forensic tools such as hashing programs, EnCase, FTK, or other tools on the market. Due to the rapid pace of technological development, it would not be wise for legislatures to require the use of specific tools. Such a rule would tie investigators to outdated programs upon the invention of new search technology.

A new rule requiring the use of these programs will assuage fears that investigators will examine every piece of data, because investigators will always conduct hashes and key word searches as an initial step. If these steps returned the sought-after evidence, the search should cease. This rule would not hinder searches because most investigators already use these programs to narrow the set of data to be searched.

If the warrant authorizes investigators to search for images, and hashes or key word searches cannot locate the images sought,

161 Interview with Chris Beeson, supra note 66.
162 See McLain, supra note 65, at 1097-98.
investigators may scan all digital images, but should not extend this file-by-file search to word documents or other file types. However, the rule should recognize that criminals might disguise evidence. If an investigator discovers a file with an altered extension, the investigator would have reason to believe the suspect attempted to conceal evidence. In this case, the investigator may open the altered file even if that file type would not normally contain evidence associated with the alleged crime.

Forensic tools will not flag misnamed files in all cases, therefore, legislatures should allow investigators to begin opening and scanning all files if, and only if, narrow search methods fail to produce results. Investigators should not be required to obtain permission from the court before expanding their search. Some might argue that such a permissive rule would encourage unscrupulous investigators to ignore the narrow search requirement. However, codification of the narrow search requirement will keep law enforcement in check; if a defendant challenges the legality of the search, the law should place the burden on investigators to prove to the court that they conducted a narrow search before proceeding with a more intrusive one. The above-described rules align with many restrictions described by the Carey-Winick doctrine, but also square with traditional doctrine by authorizing investigators to conduct a more comprehensive search where required.

2. **The Plain View Doctrine and the Second Warrant Requirement**

Requiring investigators to waive the plain view doctrine, as described in *CDT*, is a drastic and dangerous step. “A search of a computer for evidence of fraud, for example, could reveal evidence of a planned terrorist attack or a search aimed at drug trafficking could reveal evidence of an ongoing violent crime or sexual abuse.”

Abolishing the plain view doctrine with respect to digital searches may create risks to society that outweigh those created by governmental intrusion into individual privacy.

Furthermore, if legislatures pass laws that limit digital searches as described above in Subsection V.A.1, the plain view doctrine becomes less problematic. Investigators will no longer have the authority to search every file in all cases, which diffuses the threat of the general search. If investigators discover out-of-scope evidence during the course of their search under the new rule, investigators may pursue evidence under the plain view

---

163 See Brief for the United States in Support of Rehearing En Banc by the Full Court, *supra* note 1, at 14.
164 See *supra* Part III.B.
COURTING CHAOS: THE NEED FOR CLEARER RULES TO GOVERN THE SEARCH AND SEIZURE OF DIGITAL EVIDENCE

doctrine if that evidence clearly demonstrates criminal behavior. For example, a person cannot legally possess child pornography. Thus, if child pornography exists on a person’s hard drive, a crime has occurred and investigators may seize this contraband under the plain view doctrine.

However, investigators should obtain a second warrant before seizing out-of-scope evidence if questions arise as to whether the evidence meets the “immediately incriminating” requirement of the plain view doctrine. In this scenario, the second warrant requirement will ensure that investigators indeed have probable cause to seize the potentially incriminating evidence. For example, suppose investigators searching for images of child pornography stumble upon an image of a letter that reads, “How much would it cost to hire a hit man to kill Joe and his family on Thursday night?” Either this statement could be evidence of murder for hire, or it could simply be exaggerated venting, or the exercise of a person’s right to free speech under the First Amendment. Requiring investigators to submit a second warrant application to a neutral magistrate will ensure that the government can legally admit this evidence at trial, and shields government agents from claims of misconduct.

In addition, a second warrant may authorize investigators to continue searching for evidence related to the second crime. Investigators could gather the evidence needed to bring charges quickly, which could prevent a dangerous person from committing an act of violence.

3. Protecting Third Parties and Privileged Documents

Searches of digital media containing privileged data, or data pertaining to third parties require special considerations. Where investigators must issue a warrant to search computers containing either type of data, the case agents and forensic investigators should first narrow the set of data to be searched by using forensic tools and other reasonable limitations. If the in-scope evidence exists amongst privileged data or data pertaining to third parties, the most practical and protective measure would be to allow a taint team to review those materials. The taint team should be composed of disinterested forensic investigators, agents, and

---

165 Interview with Chris Beeson, supra note 66 (explaining an example of out-of-scope evidence that does not necessarily point to illegal activity, and emphasizing THAT this evidence must be treated with “kid gloves,” unlike evidence of child pornography, which is illegal in-and-of itself to possess).

166 See supra Part III.C (discussing the threats posed by the plain view doctrine to third parties and privileged documents, and different methods courts endorse to mitigate these threats).
attorneys who would review privileged and third party materials to cull non-responsive items.

For cases involving privileged documents, the government should provide the defense with the opportunity to review any data categorized as unprivileged, and allow the defense to challenge that categorization in front of the court. In re 5444 Westheimer Rd. Suite 1570\textsuperscript{167} approved of this procedure and noted that it did not prejudice the defendant because privilege disputes would be resolved by the court before the taint team could disclose materials to the prosecution team.\textsuperscript{168} A neutral magistrate or special master would not offer the government the same opportunity to challenge privilege determinations. Additionally, use of a special master or neutral magistrate would strain government and judicial resources and slow the investigation, as described in Section IV.C.

Where an investigation involves examination of digital media containing the data of unrelated third parties, members of the prosecution team should not view the intermingled data. A disinterested individual or taint team should segregate the data and provide the prosecution team with information pertaining to the suspect. If the taint team discovers evidence that incriminates a third party, the taint team may disclose that information to the prosecution team if doing so will prevent harm to another person or entity. The taint team should consult a disinterested attorney to determine whether the team has an ethical obligation to disclose the evidence to prevent harm. It should be noted that the taint team would not be authorized to scan every privileged or third party file. Unless the narrower search failed, the taint team would be required to conduct a narrow search using forensic tools as described in Subsection IV.A.1. Additionally, if the taint team refers a piece of evidence that incriminates an unrelated third party to the prosecution team, the prosecution team must obtain a second warrant before pursuing that evidence. The second warrant requirement would be waived if the evidence is contraband or clearly demonstrates that a crime has occurred. These safeguards deter unscrupulous conduct, and ensure that the prosecution team will not conduct an exploratory search of third party data in an effort to discover evidence with which to charge new crimes.

**CONCLUSION**

Digital media has become an integral part of the lives of many Americans, and advancements in technology will continue to blur the line between physical and digital worlds. As we import,
upload, or download more of our personal lives onto digital media, privacy stakes rise. Without clarity from legislatures, courts will continue to grapple over the application of the Fourth Amendment to digital media searches. Courts applying traditional Fourth Amendment principles risk eroding the relevancy of search warrants by allowing every warrant authorize an exploratory search. Courts crafting new guidelines risk tying the hands of investigators and hopelessly frustrating the legitimate purposes of law enforcement. In order to formulate sound rules for governing the search and seizure of digital property, legislatures must strike a balance between these competing factions. To strike this balance, legislatures should:

1) Require narrow search protocols, but allow file-by-file searches where necessary;
2) Require a second warrant for the seizure of out-of-scope evidence that does not immediately demonstrate that a crime has occurred. Investigators may seize contraband without a second warrant under the plain view doctrine; and
3) Require a taint team to review privileged or third party files. The taint team should afford the defense the opportunity to challenged privileged determinations. If the taint team encounters evidence incriminating a third party it should consult a disinterested attorney to determine whether the team should disclose the evidence.

Legislatures adopting this approach will provide courts with the clarity needed to enforce privacy protections while preserving the legitimate goals of law enforcement.