How Not to Invent a Patent Crisis

F. Scott Kieff and Henry E. Smith

A well-functioning patent system fosters innovation, creates jobs, and helps generate capital investment and overall economic growth. But a patent system can also be plagued by unnecessary and expensive court and administrative procedures and extreme uncertainty. Most patent reform proposals these days are designed to give officials and courts more power to weaken or eliminate “unworthy” patents, with so-called patent trolls as the bêtes noires du jour. No patent system is perfect, and our present system could be meaningfully improved. But in light of the rapid, and we would argue excessive, changes that have already occurred in the courts, what patent law needs is a tweaking of existing safety valves and processes—not opening the floodgates to more discretion and uncertainty.

Although a great deal of daylight may lie between the Bush and Obama administrations on a range of issues, when it comes to patent reform the bills brought before Congress during both administrations are remarkably similar, with the new bills introduced in early 2009, like those from 2007 and 2008, proposing significant changes to the patent statutes that have been in place since the 1952 Act.

During the past few years intense debates have brewed between those emphasizing the benefits of strong patents and those seeking
to decrease baseless litigation and administrative process. Both groups raise important concerns and both have lobbied Congress and the White House in a curious set of political alignments. Those in the pro-patent camp are unions joining forces with Republicans and small businesses siding with big Pharma. Among the patent skeptics are the big players in High Tech joining with Democrats. Similarly, both the New York Times and The Wall Street Journal have advocated a major overhaul of patents (Timiraos 2007; Editorial 2007; Lee 2007; Sewell 2007; Fitzgerald 2007). The crowded reform bandwagon has a great deal of momentum.

The emerging consensus for flexibility targets three problems in current patent law. First is the concern with junk patents and the flood of patent applications at the Patent Office. Overworked and necessarily unaware of all the prior art that has preceded the patent applicants, examiners have let through a number of weak patents—leading to tales of patents for peanut butter sandwiches. Second, and relatedly, some industries are said to be subject to “anticommons” in which multiple rights overlap on the same subject (think gene fragments) also sometimes seen as “patent thickets” in which multiple patent rights need to be assembled to bring a product to market. We believe these two problems are overblown and can easily be handled within the framework of the 1952 Act, especially as supplemented by current case law. Third is the problem of “patent trolls,” creatures who are notoriously hard to define or to spot in the wild. For some, a troll is any nonpracticing entity, a business that does not manufacture products covered by the patent. Such a definition threatens the specialization function of patents: Why should an inventor or even a company not be able to concentrate on research and development and leave manufacturing to others without running the risk of a being hit with a compulsory license? We recognize that a conceptually narrower problem does occur with patentees who induce reliance and then try to capture ex post the investments by the now-infringer. Sometimes these trolls are also called submarines, in that they surface and threaten to torpedo
unsuspecting commercial traffic. But as we discuss in a moment, existing equitable safety valves are more than adequate to this task.

**The Reform Impulse**

Proposed reformers tend to fall into two groups. First are those who would alter the process to tighten the standard for granting patents, typically by giving some official or judge the discretion to decide what was within the skill of those in the prior art without being tethered (as heretofore) to factual inquiries into laboratory notebooks, printed publications, sample products, and the like. Such discretion-based proposals go by various names, including enhanced examination, opposition, reexamination, and second-window review. Second among reformers are those who would alter patent remedies, with injunctions being disfavored and more-tailored damages as the “solution.”

Interestingly, recent U.S. Supreme Court decisions have already altered patent law greatly along the lines favored by the pro-flexibility reformers. The Supreme Court decision in the 2007 *KSR* case injected more discretion into the determination of obviousness, thereby making it easier to deny patents. The Court also has rejected, in the 2006 *eBay* case, the long-prevailing rule that a patentee with a patent adjudicated to be both valid and infringed should be able almost automatically to get an injunction; it now seems to require such a patentee to demonstrate its entitlement to an injunction on the more discretionary standard for obtaining injunctions from other areas of law. The Court has also, in the 2008 *Quanta* case, made it more difficult for patentees to license firms without at the same time licensing their customers.

All this amounts to a sea change in patent law, and all without a single revision to the patent statute. Although we certainly have bones to pick with some of these and other similar decisions, they do raise the question, Why don’t we give them a chance to work
(this was after all what advocates in those cases and of patent reform generally have told us those decisions were designed to do) before tinkering with the statute itself and its valuable, innovation-promoting architecture?

The best place to start considering the role—and drawbacks—of greater flexibility is in the central question of patent law: Does an invention satisfy the nonobviousness requirement for obtaining a patent? That is, in light of possibly far-flung and disparate pieces of prior art, is the invention something that a person having ordinary skill in the art would have already been just about to do—or not?

The 1952 Act makes prior art largely a question of fact, based on evidence such as documents and factual testimony, as compared with opinion testimony. Some see the recent *KSR* decision as standing for the proposition that government decision makers such as judges now have increased discretion to pronounce what the prior art teaches and would like to extend such discretion to patent examiners. Others think the case was narrowly decided on its facts and that the relevant inquiry remains an objective determination of precisely what was taught by the particular combination of relevant pieces of prior art (Epstein and Kieff 2007; Haber, Kieff, and Parades 2007).

Importantly, proposed statutory changes would implement the same flexible approach urged by one side of the *KSR* debate. We think that flexibility can be carried too far and that the flexibility approach on offer relies on two false premises about how the system actually works.

The first false premise is that beefing up the patent examiner’s resources would help her find the key prior art. Of course, our examining corps should have good access to Internet databases and ample time and training to peruse them. But no realistically available amount of time and training will help an examiner at his desk obtain the laboratory notebook of an individual researcher at some company or university or an obscure student thesis on the bookshelf of a foreign library, which is where the key prior art is often found.
The second false premise is that discretionary decision making, whether in court or the Patent Office, can be immune from political and other pressure. Asking a decision maker to use her legal or technical expertise as the primary basis for deciding what she thinks the state of the art was at a particular time in history gives her greater discretion than asking an ordinary jury whether a particular document or sample product existed at a particular time and what that document actually contains. By increasing the discretion of government bureaucrats, flexibility increases uncertainty and gives a built-in advantage to large companies with hefty lobbying and litigation budgets.

Lessons from History

The historical consensus about patents turns out to be the exact opposite of today’s. Concepts such as across-the-board flexibility, balance, discretion, and subjectivity are not new to our patent system; we’ve tried them before, in ways strikingly similar to those proposed today. They were the hallmarks of the patent systems of the 1940s and of the 1970s, and, although the product of well-intentioned efforts, the results in each setting were consistent and bad.

Like other proposals to deal with our current crises, dialing down the patent system gained steam in the New Deal. Created in 1938, President Roosevelt’s Temporary National Economic Commission specifically targeted patents under the misguided belief that they led to the “concentration of economic power” (Public Resolution). By a decade later, the entire patent system had become practically decimated by the courts.

Determinations about a patent’s validity in those days typically boiled down to a flexible but tautological standard: to be patentable, an invention had to constitute what a judge considered to be an “invention.” Some courts, including the Supreme Court in its
The 1950 A&P case, treated this as a “synergism” test under which patents would be valid only when the claimed invention combined existing elements to achieve a mystically synergistic effect. The test became so vague and difficult to satisfy that Justice Jackson remarked in the 1949 Jurgensen case that “the only patent that is valid is one which this court has not been able to get its hands on.”

At the same time, a patentee’s options for licensing or bringing infringement suits were severely curtailed throughout the 1930s and 1940s as courts virtually eliminated the patent law doctrines that hold a defendant accountable for causing infringement by third parties. By the late 1940s, the Supreme Court was seeing patent misuse and antitrust concerns everywhere.

In response, Congress passed the 1952 Patent Act, which aimed to reverse the Supreme Court’s expansive approach to patent misuse and set forth an objective test for patentability called “nonobviousness.” Avoiding the usual process of extensive interest-group lobbying, leading to “balance” in the sense of brute compromise, the 1952 Act represented the consensus views of legal technicians interested in developing a system that was balanced in the different sense of logical coherence. In 1948 the New York Patent Law Association asked its past president, Giles Rich, to draft a bill that would provide a more predictable framework for patent law. Rich collaborated with a Patent Office representative named Pat Federico to draft a bill, coordinate national discussion of the issues, and explain them to Congress. The result was the 1952 Act, which substantially remains the patent law today.

The 1952 Act was applauded for its predictability by the leading jurists and commentators of the time, such as Learned Hand and Jerome Frank of the New York-based U.S. Court of Appeals for the Second Circuit. Judge Hand had repeatedly called for courts to determine patent validity objectively. As he explained in the 1946 Safety Car Heating & Lighting case when criticizing the absurdity of the synergism test: “substantially all inventions are the combination of old elements; what counts is the selection, out of all their possible
permutations, of that new combination which will be serviceable.” Writing later, Giles Rich went further, explaining that a synergism test makes no sense because “[t]he laws of physics and chemistry in accordance with which all inventions perform do not permit of the judicially imagined magic according to which $2 + 2 = 5$” (Rich 1972). Judge Frank put the net impact of the more objective approach very simply in the 1942 Picard case when he explained how patents produced by such a system can be the vital slingshots smaller innovative “Davids” use to compete against large established “Goliaths.”

Although the progress made by the 1952 Patent Act had been seriously eroded by the end of the 1970s, leaders from both sides of the political aisle soon acted to reinvigorate patents. A key figure from the patent system of the 1970s, Pauline Newman, then head of chemical giant FMC Corporation’s patent operation and now a judge on the Federal Circuit, has been reminding policy makers and commentators for years that the push to re-strengthen the patent system that culminated in the 1982 act signed by President Reagan was the direct result of a serious effort launched by President Carter (Newman 2005). The 1982 act created the court on which judge Newman now sits: the Federal Circuit, which hears the appeals from most patent cases across the country. By 1978, when the economy had reached serious disarray, President Carter, through his Commerce Department, empanelled a group of experts to conduct a “Domestic Policy Review” to study domestic innovation. Its key findings focused on the destructive impact on commercializing innovation and economic growth caused by unpredictability in the patent system; its chief recommendations included strengthening the patent system through the increased predictability that could be implemented by a new court (Industry Subcommittee 1979).

We have now come full circle—again. After a consensus era of strong patent protection, the Supreme Court’s recent decisions, especially the current reform proposals, are back to the future. Why
not learn from the twentieth-century history of patents, instead of repeating it?

**Improving the Patent System**

Much of the impulse for root and branch change comes from an incomplete view of what the patent system does and how it does it. Proponents of flexibility measure the system’s success by how well it achieves the “correct” reward for inventions. The ideal benchmark would be an all-knowing planner who would hand out checks in the exact minimum amount to induce the invention in question—and in any event no greater than the invention’s social value. What all this overlooks is that the system does more than provide a reward for invention; it also provides an overall architecture for innovation.

Getting an invention made and bringing it to market require coordination among its many complementary users, including developers, managers, laborers, other technologists, financiers, manufacturers, marketers, and distributors. Patents help achieve this socially constructive coordination by allowing those various actors to interconnect with each other like modules of a larger system. The underlying mechanism depends in at least three fundamental ways on the expectation that patents will be enforced with strong property protection. First, the credible threat of exclusion associated with a published patent acts like a beacon in the dark, drawing to itself all those interested in the patented subject matter. The beacon effect motivates those diverse actors to interact with one another and with the patentee, starting conversations among the relevant parties. Second, the widespread expectation that the patent will be enforced motivates each of these parties to reach agreements with one another over the use and deployment of the technology. That bargaining effect falls apart if the parties are unsure that the patent will be enforced; if the patent is seen as not
being enforced, there is significantly less need to reach agreement ex ante. Thus the fear of weak enforcement creates a disincentive for the necessary parties to work together at the outset. Finally, patent protection allows patentees to appropriate the returns to (rival) inputs to developing and commercializing innovation—labor, lab space, and so forth—without the law having to trace the relative contributions of these multifarious inputs. Instead patents form a platform on which coordination and development can take place (Kieff, 2007; Smith 2007).

But when it comes to evaluating the system, today’s proponents of flexibility tend to treat it as a bundle of features or levers that can be tweaked on an industry-by-industry or even a case-by-case basis. How high should the bar for patentability be? What should the size of the reward be for a contribution? Would an injunction lead to too much bargaining power? The list goes on and on.

As we have seen, those designing and implementing our present patent system realized that many of the benefits of patents could be achieved holistically. The benefits in terms of stability and coordination from the system are not results achieved by this or that part of the system, which could be better tailored by greater levels of official discretion, in the interest of wringing out all the errors from the system.

The current consensus for flexibility problematically—and ironically—deals in absolutes. If the system does not provide good notice all the time, it is “not doing its job” and thus any change is presumptively an improvement. Theoretical problems with multiple patents impeding commercialization were declared an impending disaster before any empirical investigation. Modest problems along these lines exist, but their very modesty suggests that altering the mix of certainty and flexibility is better than declaring the system a failure and opening the flood gates of discretion (Murray & Stern 2007; Walsh et al. 2004).

This is not to say that the system cannot be evaluated empirically
or that it cannot be improved. But it does suggest caution and creativity in evaluating the case against the system and for radical reform. Although we think the flexibility-based solutions offered by patent critics are imprudent, some of the underlying concerns they raise are important to address. Here are our proposals.

**Remove the presumption of patent validity.** The patents that drive most calls for reform are indeed pernicious because they allow patentees to threaten expensive but meritless litigation against competitors, although the extent to which the U.S. patent system is afflicted with a disproportionate number of “bad” patents is a topic of serious debate (Katznelson 2007). Under the present system, an issued patent is presumed valid, which requires a challenger to prove invalidity by a higher standard of proof (“clear and convincing evidence”) than usually prevails in civil cases (“preponderance of the evidence”).

The costs under the present system of knocking out even an obvious patent can be large. The infamous litigation over Amazon’s “one-click-shopping-patent” probably required Barnes and Noble to spend more than two million dollars in litigation costs alone to defeat the preliminary injunction that had wreaked havoc during the 1999 Christmas season rush. The threat of such expensive litigation over even a questionable patent is precisely what is said to terrorize potential defendants, large and small, about the present patent system. But this *in terrorem* problem can be greatly mitigated through more targeted measures than injecting large dollops of discretion.

Dialing down the present presumption of validity to something like the ordinary standard for civil cases would decrease the bad, *in terrorem*, effect. When litigation is needed, the carefully crafted Federal Rules of Civil Procedure govern the procedures for join-der, compulsory counterclaims, and against relitigating issues and claims decided in previous litigation, which are collectively designed to avoid abusive and repetitive process. The Federal
Rules also provide streamlined procedures such as summary judgment, which avoids long trials where there is no genuine issue of material fact.

In the end, a decrease in the presumption of validity would be particularly good for the “Davids,” because it directly protects them from the in terrorem effect of junk patents and the threat of expensive but baseless litigation to defend against patents whose “validity” rests entirely on the present presumption. It also indirectly helps them raise the funds needed to litigate against a baseless opponent, regardless of whether they are asserting patent infringement or invalidity.

Some may argue that increasing the reliance on opinions of counsel will make it harder for lawyers to give advice, which is where the old tension underlying attorney-client privilege comes into play. On the one hand, decision makers often need to verify whether a party acted with good advice of counsel. On the other hand, it will be hard for a lawyer and client to openly discuss the strengths and weaknesses of various approaches if they know that all their communications are likely to be subject to open review later in court. But this is a false dichotomy. One lesson our society has learned from corporate scandals such as Enron is that it can be important to decouple auditing from advising. An opinion of counsel about a patent can be an important auditing tool that should be kept separate from the important advising resources a client needs in the competitive market and in litigation. This distinction should motivate the Federal Circuit as it works to clarify the evolving case law relating to attorney-client privilege for patent opinions of counsel after the 2004 Knorr case. Also, to prevent opinions of counsel from becoming a box to check and a whitewash for misconduct, courts have shown little hesitation to sniff out bogus opinions of counsel or to specifically call out their authoring attorneys and law firms, as was done by both the trial and the appellate courts in their respective 1997 and 1998 decisions in Johns Hopkins v. CellPro.
**Institute symmetric fee shifting.** Imagine a patent system in which both patentees and potential infringers had good access to fee shifting when the other side’s case was baseless. Under today’s rules, the patentee wants to educate the alleged infringer about the strength of the infringement case early in the process because this increases the patentee’s chance of getting enhanced damages, such as attorney fees. For the same reason, the alleged infringer has a strong incentive to avoid notice by avoiding communication.

Symmetric fee shifting would allow alleged infringers to collect attorney fees from a patentee who brings an infringement case after having been warned, for example, about a particular item of invalidating prior art. This practice of fee shifting when a patentee makes baseless arguments in defense of the patent’s validity would match the present rules in Sections 284 and 285 of the statute that allow patentees to get fees and (potentially) treble damages from infringers who should have known about their own infringement and have thus mounted baseless arguments in their defense. Such symmetry in fee shifting would encourage parties to exchange information and resolve disputes before undertaking expensive litigation (Kieff 2003; 2009). Under our rule, the alleged infringer would similarly want to educate the patentee about any validity-destroying prior art. Symmetry in fee shifting helps align both parties’ incentives to communicate with each other.

Such a system would mean that the existing markets for audit-type opinions of counsel would grow. Under today’s rules, the alleged infringer often wants to get an opinion of counsel early in the process so as to bolster later arguments that it had a good-faith basis for believing it had not infringed valid patent rights, thereby decreasing the chances of paying enhanced damages or attorney fees if it loses the case. Using our rule, the patentee would also want to get an opinion of counsel early in the process to establish its good-faith basis for its position, so as to avoid having to pay the alleged infringer’s attorney fees.

Under our proposed practice, it will be easier for third parties to
essentially spread the costs across multiple customers by opening businesses that provide rating services like those seen in today’s capital markets that evaluate a particular company’s stock or bond offerings. The ability to get attorney fees in baseless cases also opens up the market for contingent and other flexible fee arrangements for those who are too liquidity constrained to fight on their own.

**Equity in remedies.** Injunctions have come under a lot of fire, and recent flexibility-based reform proposals have centered on substituting tailored damages for injunctions in patent cases. Most of the time injunctions provide the certainty needed to avoid disputes and, when disputes do arise, encourage the parties to resolve them early. With injunctions, infringers have a hard time gaming the system or using the possibility of undercompensatory damages to do an end run around negotiations. Infringement is the violation of a property right; an injunction forces the infringer to stop and enforces the delegation of valuation questions to patentees and their contractual partners, with a view toward markets for inputs and products, rather than officials, courts, and experts for hire (Kieff 2001; Smith 2009).

But as illustrated by the furor over trolls, injunctions have their limits. When an infringer in good faith makes large, irreversible investments that include reliance on a patent that would have been easy to design around ex ante but hard to do so ex post, the infringer does have a claim on our sympathies (Denicolò et al. 2008). But contrary to what critics sometimes claim, this problem does not distinguish patent law from tangible property, which faces the same problem. In building encroachments, the good-faith encroacher—for example one building close to the lot line in reliance on a faulty survey—can argue against an injunction and pay only damages. But this emphatically does not apply to bad-faith encroachers: those who know they are committing a wrong. Many in the patent reform camp claim that patent law is not like property because it fails to give notice of claims: patent law cannot capture
inventions in language, and patents are too hard to find (Bessen and Meurer 2008). We think the reforms we have discussed would, by providing individualized notice, help ensure that others infringe (if at all) in bad faith. Another possible rule is that for those cases of infringement that are based on the doctrine of equivalents, which extends the literal claims to cover a penumbra of hard-to-foresee additional invention space, in the remedy should be presumed (strongly) to include only damages, not an injunction (Smith 2007).

But introducing more discretion into determinations of patent validity is exactly the wrong way to go. For the inevitable uncertainty and difficulty in giving notice that remain (in real property law too, more than patent skeptics acknowledge), traditional equitable principles are more than up to the job.

The 2006 Supreme Court *eBay* decision held that the standard for an injunction is based on a four-factor test applied in other settings, without much further guidance. Taking their cue from the concurrences in the case, some see this as an invitation to inject more discretion to withhold injunctions in the broadly defined public interest. Others see the case as largely confirming a strong tendency toward injunctions, once validity and infringement have been decided in court.

We think that the best way to implement *eBay* is to take this equitable approach seriously and apply it in the traditional (and sensible) fashion. Crucially, the equitable approach is a safety valve for those situations in which someone who is otherwise a good candidate for getting an injunction—such as a patentee whose patent has been infringed—should not get one because of some glaring injustice. The equitable approach is flexible but not boundlessly so, in contrast to currently proposed reforms that elevate discretion to new heights. Moreover this safety valve is probably all we would need.

First, the order of the factors in the test—irreparable harm, inadequacy of damages, balance of hardships, and public interest—is no
accident. In the old days when courts of equity were separate from courts of law, the inadequacy of the legal remedy (here damages) because of irreparable harm was jurisdictional. There is a functional reason for this as well: equity is a safety valve, not an ocean of free-floating discretion. It is designed to be effective only within a limited domain where it’s most needed and will not upset general expectations.

Turning to the balance of hardships and public interest, these two factors sound a lot broader than the safety valves they are meant to be. The balancing called for in the traditional equitable test is not a full-blown cost-benefit analysis or an even a weighing of hardship of both sides. Instead, traditional equitable analysis asked whether someone otherwise entitled to an injunction should not get one, in the judge’s discretion, in light of a grossly disproportionate hardship on the defendant (Epstein 1997; Schwartz 1964; American Jurisprudence 2d Injunctions 2005). The point here is not to get the exact correct reward but to avoid egregious errors in an otherwise robust system of injunctive relief.

Likewise, the public interest standard in equity is not an invitation to maximize official discretion. Rather, it is another safety valve to prevent major harm to third parties. Judges are simply not able to measure the public interest in some more specific fashion, such as a cost-benefit analysis or the judges’ values or who knows what.

Finally, the equitable approach embraces more than this four-factor test. Injunctions can be tailored to the harm and delayed to give an innocent infringer time to redesign (Lemley and Shapiro 2007). And a delay in suing or lurking in wait while good-faith reliance occurs can sometimes be grounds for denying injunctions.

Conclusion

The approach we propose will decrease slightly the average value of all patents because patentees will now have to fight harder on the
issue of validity when they assert their patents in court. But this is not necessarily bad. The costs of arguing to the Patent Office to get patent rights in the first instance will be less than in a system under which the examiners have largely unfettered discretion to reject applications.

Most important, the approach we propose directly addresses the fears of those held hostage under the current system by the threat of litigation costs surrounding patents that are merely presumed to be valid. Under a decreased presumption of validity, such a terrorizing effect largely evaporates. With fee shifting, meritless suits against infringers will be discouraged, and the full traditional but limited use of equitable discretion will provide all the safety valves we need for good-faith infringers and those facing true patent trolls. These approaches should be given time to work. The prudent course for the country is to embrace a strong patent system based on predictability and facts, which will benefit all players, large and small, in their contributions to American innovation and economic growth.

References


Public Resolution No. 113, 75th Cong. (1938).
