

INTRODUCTION TO THE STANFORD NPE LITIGATION DATASET

Shawn P. Miller

Ashwin Aravind, J.D. '18
Bethany Bengfort, J.D. '16
Clarisse De La Cerda, LL.M. '17
Matteo Dragoni, LL.M. '16
Kevin Gibson, J.D. '17
Amit Itai, J.S.D. '20
Charles Johnson, J.D. '17

Deepa Kannappan, J.D. '16
Emily Kehoe, B.S. '17
Hyosang Kim, J.D. '16
Katie Mladinich, J.D. '18
Roberto Pinho, LL.M. '17
John Polansky, J.D. '18
Brian Weissenberg, J.D. '16

ABSTRACT

Despite widespread interest in the impact of patent assertion entities (“PAEs”) on our patent system, there has been no publicly available dataset that categorizes more than a small fraction of lawsuits as involving practicing entities, non-practicing entities (“NPEs”), or PAEs. To address this knowledge gap, Stanford Law School student researchers in the Patent Litigation Practicum (the “Practicum”), led by Mark Lemley and Shawn Miller, created and continue to compile the first comprehensive patent litigation dataset to categorize the type of plaintiff involved in every U.S. patent lawsuit filed since 2000. The dataset captures changes in patent litigation patterns occurring during an era of heightened activity and policy reform. It reveals trends before and after passage of the America Invents Act, key Supreme Court patent cases including *eBay* and *Alice*, and various executive orders focused on increasing transparency and reducing costs in patent suits.

We created the Stanford NPE Litigation Dataset (“the Dataset”) in partnership with Lex Machina and notable scholars who have previously studied NPEs.¹ We are reviewing every patent lawsuit filed in U.S. district court since 2000, categorizing the owner of every patentasserter in each case as a practicing entity or as one of 11 types of NPEs. By the end of 2017, we will complete the 54,000 lawsuits filed between 2000 and 2015, and then make the full Dataset publicly available. Thereafter, we will continue to release subsequent, updated versions of the Dataset that include more recently filed lawsuits. In addition to eventually aiding Congress and the United States Patent and Trademark Office in developing effective laws and policy, the Dataset provides a valuable tool to help policy makers, judges, litigators and scholars to better understand the nature of the entities filing patent suits.

At present, we completed and made public a random sample of 20% of the lawsuits filed from 2000 through 2015 (roughly 10,800 lawsuits). As we report in our preliminary findings, this initial random sample is sufficient to reveal changes in the extent of patent litigation attributable to each type of patent owner and other litigation trends. We hope this data will be used to further policy discussions and therefore conclude this paper with suggestions for how the data can be used to reveal the impact of different types of patent owners on the patent system.

¹ We thank Colleen Chien, Christopher Cotropia, Jay Kesan, Mark Lemley, Shawn Miller, Michael Risch and David Schwartz for their contribution to this project of patent owner datasets they developed for prior analysis of various subsets of patent litigation.

I. Introduction

The term “patent troll” has become part of the American legal lexicon, though few can agree on a precise definition. The concept of the patent troll first appeared in the 1994 short film, “The Patents Video,” which featured a troll demanding fees from would-be bridge-crossers as a satirical illustration of the modern patent system.² The origin of the term in common usage, however, is most often credited to Peter Detkin, counsel to Intel in the late 1990s.³ At the time, Intel was hit by a swath of what it believed to be unmeritorious lawsuits by businesses that did not manufacture products. One lawsuit in particular infuriated Detkin, prompting him to designate the lawyer who filed it a “patent extortionist.”⁴ In retaliation, the lawyer filed suit for libel.⁵ Detkin realized that he would need to create a term that was similarly evocative yet appreciably less literal. He conducted an internal contest at Intel to create a new moniker and thus the term “patent troll” was born.⁶

To date, the interest groups and scholars that use the term disagree about which types of patent owners are and are not patent trolls. And many refuse to use the term at all. Among those who use it, almost all agree that a patent troll must be a non-practicing entity (“NPE”), or an entity that owns patents but does not create or sell products or services. Most definitions further specify that a patent troll must be an entity that exists to assert patents against other actors, birthing the term “patent assertion entity” (“PAE”). While the terms PAE and NPE are often used interchangeably, the difference in definition is important. For example, universities and technology development companies are NPEs, but most commentators would not consider these PAEs. We define PAEs as entities that employ patents primarily to obtain license fees, rather than to support the transfer or commercialization of technology.

Over the last decade, a growing debate among scholars, the bar, industry, and public agencies about the impact of PAEs has sparked government scrutiny. Critics have come to believe that the steady escalation of PAE enforcement activity, including litigation, is harming innovation and is acting as a tax on producers and consumers. While there are a variety of theoretical business models for PAEs, with different predicted effects, critics essentially argue that PAE enforcement activity imposes costs on business, consumers, and the courts that exceed the financial benefits to PAEs themselves and the incentives their gains might generate for inventors.

The principal motivation for this project is the perceived rise of the PAE business model. However, it is important to emphasize that not all NPEs are PAEs as traditionally understood. Some NPEs are individual inventors hoping to collect on others’ use of their innovations and there is disagreement on whether inventors who assert their patents on their own inventions are PAEs. Others are universities seeking to monetize the innovations of their faculty and students. Still

² See Paula Natasha Chavez, *The Original Patent Troll*, (Jan. 28, 2007), <https://www.youtube.com/watch?v=1OGozFzHkhs>.

³ See Ira Glass, *When Patents Attack!*, THIS AMERICAN LIFE 4 (July 22, 2011), http://www.thisamericanlife.org/sites/default/files/TAL441_transcript.pdf.

⁴ *Id.*

⁵ *Id.*

⁶ *Id.*

others are startups that failed to commercialize a patented idea that may still be financially worthwhile. At a minimum, any further legislative reform should consider the differences in the reasons why different types of NPEs exist, as well as differences in their assertion behavior generally and their litigation activity in particular. Indeed, research including ours reveals a complex ecosystem of patent litigation, filled with various practicing entities and NPEs employing diverse litigation strategies.

Definitions aside, critics of PAEs believe they represent a real problem for producers, consumers, and innovation. Some argue abusive litigation tactics conducted by PAEs, coupled with excessive verdicts, cost the U.S. economy \$1.5 billion per week.⁷ PAEs have been described as taxes on innovation and colossal drains on manufacturing; critics argue that PAEs ostensibly assert patents of dubious validity and value in order to extract licensing fees from the entities that are actually making products available to the public.⁸ Similarly, many scholars and defense attorneys argue that PAEs are a corrupt product of existing law, which enables PAEs to operate under the veil of shell companies to hide the true nature of their litigation activities and rent empty office spaces exclusively to take advantage of favorable federal jurisdictions for filing suit.⁹ Finally, critics believe—and meager empirical evidence supports the idea—that in many cases only a tiny fraction of the value of the patent extracted by a PAE ever reaches the original inventor, thus distorting the incentive structure designed by the patent system to encourage innovation.¹⁰ This final criticism of PAEs is a response to the principal argument of PAE defenders—that PAEs serve as efficiency-enhancing business intermediaries between inventors and commercializers.¹¹

⁷ Larry Downes, *Everyone Hates Patent Trolls, But Here's the Root Problem with Our Broken System*, WASH. POST (May 4, 2015),

<https://www.washingtonpost.com/news/innovations/wp/2015/05/04/everyone-hates-patent-trolls-but-heres-the-root-problem-with-our-broken-system>.

⁸ See James Bessen, *The Evidence is in: Patent Trolls do Hurt Innovation*, Harvard Bus. R. (Nov. 2014), <https://hbr.org/2014/07/the-evidence-is-in-patent-trolls-do-hurt-innovation> (finding the more R&D a firm performs, the more likely it is to be sued for patent infringement).

⁹ See *Judge Takes On Patent Troll With 'Sham Employee'; Forces Troll To Defend Practice Before A Jury*, (Sept. 26, 2013), <https://www.techdirt.com/articles/20130926/02410524662/judge-takes-patent-troll-with-sham-employee-forces-troll-to-defend-practice-before-jury.shtml> (finding PAE created a shell company, rented empty office space in East Texas, and hired sham employees). However, the Supreme Court's recent *TC Heartland* ruling may limit patent-litigation forum shopping to some extent.

¹⁰ See James Bessen & Michael J. Meurer, *The Direct Costs from NPE Disputes*, 99 CORNELL L. REV. 387, 411 (2014) (finding that payments to individual inventors come to only 7% of PAE licensing revenues and that less than one-quarter of these PAEs' revenues flow to innovative activity, which Bessen and Meurer define as purchases of other patents or direct R&D expenses).

¹¹ See, e.g., Daniel F. Spulber, *Patent Licensing and Bargaining with Innovative Complements and Substitutes*, Northwestern Law & Econ Research Paper No. 16-12 (August 3, 2016); Michael Risch, *Licensing Acquired Patents*, 21 Geo. Mason L. Rev. 979 (2014); Ryan Holte, *Trolls or Great inventors: Case Studies of Patent Assertion Entities*, 59 St. Louis. U. L.J. 1 (2014); James F. McDonough III, *The Myth of the Patent Troll: An Alternative View of the Function of Patent Dealers in an Idea Economy*, 56 EMORY L.J. 189 (2006); Ashish Arora & Robert P. Merges, *Specialized Supply Firms, Property Rights, and Firm Boundaries*, 13 Indus. & Corp. Change 451 (2004).

In response to PAE critics, public officials have begun to explore the nature and impact of these business entities. For example, in December 2012 the Federal Trade Commission (“FTC”) and the Department of Justice (“DOJ”) held a joint workshop on the behavior of PAEs, followed a month later by a Patent and Trademark Office (“PTO”) workshop on proposed sunshine rules that would provide more ownership transparency.¹² In October 2016, the FTC released an in-depth study of the behavior and organization of a select group of PAEs.¹³ Further, as part of the America Invents Act (“AIA”), Congress directed the Government Accountability Office (“GAO”) to conduct a study of the consequences of patent litigation brought by NPEs that do not make products or offer services.¹⁴ Additionally, President Obama took executive action aimed at curtailing PAE litigation, and issued an executive order requesting the PTO to initiate a rulemaking process requiring patent applicants and owners to regularly update ownership information.¹⁵ Congress also has considered—though not enacted—several substantive reform bills since the AIA. The Innovation Act, for example, would allow manufacturers to step in and defend end users in litigation and offer measures to reduce the often enormous costs of discovery.¹⁶

The goal of this paper is to contribute quantitative evidence to the ongoing policy debate surrounding PAEs by introducing and describing the first comprehensive public dataset to classify the type of entities asserting patents in every lawsuit from 2000 to present. With the release of the Stanford NPE Litigation Dataset (“the Dataset”), researchers will have access to data that can help to inform patent policy related to PAEs and other types of entities. We have made sufficient progress in amassing the data that we are now able to contribute early analysis in this paper as well as a full description of the Dataset and our underlying methodology.

In addition, this policy research paper invites comment from the PTO, FTC, GAO, and other public agencies, policymakers, scholars and practitioners, on the methodology used to create the underlying dataset, our initial findings on the scope of PAE and NPE litigation, and the direction that future research should take to help inform national legislation and PTO policy and regulation.

Key Findings

Preliminary results utilizing our 20% random sample reveal:

¹² See *List of Public Workshops*, U.S. DEP’T OF JUSTICE, ANTITRUST DIV., <http://www.justice.gov/atr/public/workshops/pae/index.html> (last visited May 12, 2016); *Roundtable on Real Party in Interest Information*, U.S. PATENT & TRADEMARK OFFICE, <http://www.uspto.gov/ip/officechiefecon/roundtable-RPI-agenda.1.pdf> (last visited May 12, 2016).

¹³ See *Patent Assertion Entities (PAE) Study*, U.S. FED. TRADE COMM’N, <https://www.ftc.gov/policy/studies/patent-assertion-entities-pae-study> (last visited May 12, 2016).

¹⁴ See *Intellectual Property: Assessing Factors that Affect Patent Infringement Litigation Could Help Improve Patent Quality*, GOV’T ACCOUNTABILITY OFFICE 17 (August 22, 2013), http://www.uspto.gov/sites/default/files/aia_implementation/GAO-12-465_Final_Report_on_Patent_Litigation.pdf.

¹⁵ *Attributable Ownership*, U.S. PATENT & TRADEMARK OFFICE, <http://www.uspto.gov/patent/initiatives/attributable-ownership> (last visited May 12, 2016).

¹⁶ See generally *Innovation Act*, H.R. 9, 114th Congress (2015-2016).

- **Litigation Growth:** Using the total number of defendants rather than lawsuits (hereinafter “defendant-lawsuit pairs”) to account for the AIA joinder rule, we find that the number of distinct patent disputes nearly doubled between 2000 and 2015.
- **Rise of PAEs:** Again using defendant-lawsuit pairs as a better count of distinct patent disputes than litigation filings, the share of litigation attributable to patent licensing firms that acquired their patents from third parties (our Category 1) increased from less than 5% in 2000 to over 30% on average since 2010.
- **Decrease in Practicing Entity Litigation:** Since about 2006, there has been a marked decrease in the share of all patent litigation attributable to practicing entities (our Category 8). Prior to 2006, about 70% of all lawsuits and 60% of all defendant-lawsuit pairs asserted practicing-entity patents; since then, those two percentages have dropped to around 45%.
- **Inventor and Inventor-Owned Licensing Firm Litigation:** There has been a slight decrease in individual inventor lawsuits (our Category 9), coupled with a small increase in cases where individual inventor-started patent licensing firms (our Category 5) assert patents—perhaps inventors have become more sophisticated or professional in increasingly utilizing the benefits of incorporation.
- **Software Patent Assertion:** In PAE lawsuits (our Categories 1, 4 and 5) filed in 2014, about half asserted a patent that meets a conservative definition of covering software. In contrast, only 11% of Category 8 practicing entity lawsuits asserted a patent that covered software.
- **Time to Termination:** Across our entire 20% random sample, suits involving only Category 1 patent asserters terminated most rapidly (average of 327 days). In that same sample, suits involving only practicing entities took significantly longer to end (average of 443 days).
- **Settlement:** Among lawsuits filed in 2014 that did not end in consolidation or transfer, 87% of PAE and 62% of non-PAE suits have already settled. This difference is statistically significant.
- **Merit Wins and Losses:** Across our entire 20% random sample, Category 1 patent asserters won a mere 13% of judgments on the merits while this number was 35% for practicing entities.
- **Patent Trial and Appeal Board (“PTAB”) Inter Partes Review:** Among lawsuits filed in 2014, a similar percentage of PAE and non-PAE suits asserted at least one patent that was the subject of an inter partes review (“IPR”) (35% and 34%, respectively). Among those suits in 2014 with at least one asserted patent “IPR’ed”, a similar percentage of PAE and non-PAE suits had at least one proceeding instituted (63% and 60%, respectively).

As these initial results suggest, the Dataset has the potential to aid in dramatically improving the understanding of the role that different types of patent owners play in the patent system.

Accordingly, we hope that researchers will utilize the Dataset in independent analyses to determine if reform aimed at decreasing PAE patent enforcement is necessary and, if so, what policy action would be both proportionate and effective. We similarly encourage Congress, the PTO, FTC, GAO, and other decision-makers to utilize the Dataset to enhance patent policy. In that spirit, the Stanford PAE Symposium (May 10-11, 2017) brought together private sector stakeholders, policy makers, judges, lawyers, scholars and researchers to engage with the data, share findings, and think dynamically about the future of patent policy. Additionally, at the Symposium, sixteen teams of academic researchers presented their initial findings using our Dataset. Their work and future projects by other researchers will be linked at our project website as they are finalized and published.¹⁷

Roadmap

Part II describes the Dataset and methodology utilized to produce it. Various methods of quality control were implemented to ensure the most accurate input of the data. Thirteen different classifications were created to describe the different types of entities asserting one or more patents during litigation. These classifications include companies currently manufacturing products or selling services (called operating companies), PAEs started by individual inventors, and PAEs that acquired patents from others. To date, we have completed our review of over 40,000 of the 54,000 lawsuits filed between 2000 and 2015, including a random sample of 20% of all lawsuits during this time period. The random sample is available for download on the project website.¹⁸

In a soon-to-be-complete Part III of the paper, we will present summary statistics and basic time trends found in the Dataset. It continues with preliminary demonstration of how the Dataset can be combined with other data sources to analyze patent litigation. Our general findings indicate that litigation growth in the past fifteen years is due largely to the rise of PAE-initiated lawsuits. These lawsuits tend to involve software patents and resolve much more quickly than other lawsuits, lending credence to the theory that PAEs often assert with the goal of extracting quick money from defendants. When PAEs assert their patents through judgment on the merits, they fare much worse than practicing entities asserting patents.

Finally, in Part IV, we present our suggestions on how we hope the Dataset will be used. At its core, our research fills a knowledge gap by categorizing patent plaintiffs. As a result, our data may itself lead to a change in litigation strategy for patent practitioners; perhaps licensing firm PAEs will curtail or mask their behavior to avoid the attention of lawmakers, academics, and the public. We encourage Congress and other governing agencies to use our findings and future research findings that emerge from the Dataset as a basis for targeted patent reform. The Dataset furthermore offers federal district courts and the Federal Circuit clearer understanding of the activities of the various types of patent owners, creating an opportunity to craft nuanced legal precedent.

¹⁷ See Stanford NPE Litigation Dataset, <https://law.stanford.edu/projects/stanford-npe-litigation-dataset/#slnav-brief-dataset-methodology> (last visited October 12, 2017).

¹⁸ *Id.*

II. Creating the Stanford NPE Litigation Dataset

In partnership with Lex Machina and prominent scholars who have conducted research on NPE litigation,¹⁹ our project objectives are to: (1) review every patent infringement lawsuit, including declaratory judgments, filed in U.S. district court since 2000 and categorize the party (or parties) asserting the patent(s) in each case (hereinafter called “patent asserters”) as a practicing entity or as one of 11 types of NPEs (see Table 1); and (2) to conduct a preliminary analysis of the data to determine whether litigation trends differ by patentasserter type and whether there is variation in the characteristics of litigation across patentasserter type. Our categories are adapted from the taxonomy of Allison, Lemley, and Walker (2009)²⁰ with the names of the categories listed in Table 1 below.

Part II.1 discusses the patentasserter taxonomy used in the Stanford NPE Litigation Patent Dataset. Part II.2 describes the methodology we use to categorize patent asserters.

II.1. Taxonomy

Table 1

PATENT ASSERTER CATEGORIES IN STANFORD NPE LITIGATION DATASET

| Category | Description |
|-----------------|-------------------------------------|
| 1 | Acquired patents |
| 2 | University heritage or tie |
| 3 | Failed startup |
| 4 | Corporate heritage |
| 5 | Individual-inventor-started company |
| 6 | University/Government/Non-profit |
| 7 | Startup, pre-product |
| 8 | Product company |
| 9 | Individual |
| 10 | Undetermined |
| 11 | Industry consortium |
| 12 | IP subsidiary of product company |
| 13 | Corporate-inventor-started company |

¹⁹ Scholars who shared how they categorized entities in prior work include: Colleen Chien, Associate Professor of Law, Santa Clara Law School; Chris Cotropio, Professor of Law, University of Richmond School of Law; Jay Kesan, Professor of Law, University of Illinois College of Law; Mark Lemley, William H. Neukom Professor of Law, Stanford Law School; Shawn Miller, Lecturer in Law, Stanford Law School; Michael Risch, Professor of Law, Charles Widger School of Law at Villanova University; and David Schwartz, Professor of Law, Pritzker School of Law at Northwestern University.

²⁰ John R. Allison, Mark A. Lemley & Joshua Walker, *Extreme Value or Trolls on Top? The Characteristics of the Most-Litigated Patents*, 158 U.Pa. L.Rev. 1, 10-11 (2009).

Of the thirteen categories, only Category 8 (product company) practicing entities actually make products or offer services for sale. The remaining categories are different types of NPEs.²¹ We believe those who use the term PAE or “patent troll” are generally referring to entities that fall within Category 1 (acquired patents), Category 4 (corporate heritage), or Category 5 (individual-inventor started company).

Category 1 includes any NPE primarily in the business of asserting patents that it has acquired from other entities. We include in this category large patent aggregators such as Acacia and Intellectual Ventures.²² By contrast, Category 5 is comprised of firms primarily in the business of asserting patents, but where the original inventor of the patents is the founder and/or owner of the NPE. Most Category 5 entities are limited liability companies owned by the inventor(s) of the asserted patents and exist solely to hold and enforce those patents. For example, we coded Ronald A. Katz Technology Licensing, LLC, as a Category 5 entity.

Category 6 includes any patent asserter that is a university, government entity, or a not-for-profit institution. Relatedly, Category 2 entities are not themselves universities but are rather “IP subsidiaries” of universities or separate licensing firms known to primarily assert patent rights on behalf of universities. Examples include the Wisconsin Alumni Research Foundation, the University of Colorado Foundation, and Competitive Technologies Inc.

Category 8 and Category 12 have a similar relationship as Category 6 and Category 2. Category 8 product companies manufacture products, sell products, or deliver services (unrelated to patent enforcement). We do not analyze whether a particular company is making use of the patent it is asserting in its products or services, but simply whether the company sells goods or services generally. Category 12 consists of the IP enforcement subsidiaries of practicing entities. For example, we coded AT&T Intellectual Property I, L.P., as a “12.”

Category 9 includes any individual litigant asserting patents. Usually these are the original inventors suing in their own name rather than through an LLC or other company they own. We also include in Category 9 inventors’ family trusts.

Category 13 “corporate-inventor-started” companies generally consist of what Cotropia, Kesan, and Schwartz (2014) describe as technology development companies.²³ These entities engage in substantial research and development in-house and obtain patents from that research. These entities

²¹ Where possible, we have attempted to avoid classifying entities as Undetermined (Category 10).

²² Note that Cotropia, Kesan & Schwartz, *Unpacking Patent Assertion Entities (PAEs)*, 99 *Minnesota Law Review* 649 (2014), and others use a taxonomy with a separate category for large patent aggregators. Note that most entities Cotropia, Kesan & Schwartz categorize as “patent holding companies” will also fall in our Category 1. They describe patent holding companies as follows: “[U]sually limited liability companies, that appear to have been formed solely to hold and enforce a patent or small portfolio of patents. As far as we can tell, these companies are not owned by the original inventor. Frequently, these companies were formed shortly before litigation commenced.” Cotropia, Kesan & Schwartz at 670.

²³ Cotropia, Kesan & Schwartz, *Unpacking Patent Assertion Entities (PAEs)*, 99 *Minnesota Law Review* 649, 670 (2014).

are generally much more likely to negotiate exclusive licenses with producers and have reoccurring relationships with the same producers. Valinge is an example of a corporate-inventor-started company that develops technology through R&D and licenses its IP to producing companies.²⁴

Failed startups in Category 3 are entities that originally intended to commercialize their inventions but were unable to bring their product to market before ceasing operations. In defining Category 3, we included entities that achieved only *de minimis* sales or production of prototypes before shutting down. One example is CardSoft.²⁵ Category 7 “pre-product” startups are companies that were still in operation at the time of the lawsuit but did not yet possess any relevant products or make significant sales. One example is T5 Labs.²⁶ For a firm to be a “startup” and fit into either Category 3 or Category 7, the entity had to be described as a startup, possess venture funding, or possess other indicators of being a startup as conventionally understood.

The final two categories are Category 4 and Category 11, and occur rarely. Category 4 “Corporate heritage” entities are firms that were successful producers for a sustained period of years but then transitioned to a PAE business model. Examples include Encyclopedia Britannica, Inc., IMX, Inc. and PDL Biopharma, Inc. Finally, Category 11 “Industry consortiums” are industry groups that sue on behalf of their members. One example is MPEG L.A.

By the end of 2017, we anticipate completing our coding of the nearly 55,000 patent lawsuits filed between 2000 and 2015. Upon completion, we will make the full Dataset publicly available. Thereafter we will continue to release subsequent versions of the Dataset that include more recently filed lawsuits and corrections to previous coding as revealed to us through public usage and comment.

To date, we have completed our review of over 40,000 lawsuits, including a 20% random sample of lawsuits filed between 2000 and 2015. We made this completed random sample publicly available in May 2017. As we describe in the preliminary results section below, the random sample is sufficient to report observations on the extent of patent litigation trends attributable to each type of patent owner over the last 15 years, as well as significant differences in the characteristics of litigation involving these different types of patent asserters. Before delving into those results, however, we detail the methodology that we utilized to categorize patent infringement lawsuits.

²⁴ For a description of Valinge and its R&D, see <https://www.valinge.se/about>.

²⁵ See Aaron S. Lukas, *Will the Inventor of the Handheld Cash Register Ever Cash-In?*, *InventorsDigest.com* (Dec. 2014), https://www.cozen.com/Templates/media/files/Inventors%20Digest%20-%20December_pdf.pdf.

²⁶ See Dean Takahashi, *Who invented cloud gaming? T5 Labs tangles with OnLive (exclusive)*, *VentureBeat* (Feb. 15, 2011), <http://venturebeat.com/2011/02/15/t5labs-patent-onlive/>.

II.2. Coding Methodology

Every patentasserter involved in a lawsuit corresponds to one of the 13 categories, including “Undetermined” for entities where the category was not clear after our review (see Table 1). Each data point in the Dataset is a lawsuit; some lawsuits contain multiple patent asserters and some of these multiple-asserter cases contain patent asserters that fit into multiple categories.²⁷ Those cases are assigned to multiple patentasserter categories. For example, if an infringement action is filed by an individual inventor and a product company that has an exclusive license to use the technology, then that lawsuit fits in both Category 9 and Category 8 and is categorized both as an “8” and a “9.” Accordingly, as we review each lawsuit, we assign each entity in the case that asserted a patent a category and code it into a user interface maintained for the Dataset by Lex Machina.²⁸

To determine how to categorize each entity, student researchers consult several different sources. First, the coders review Lex Machina’s webpage for each lawsuit and read the initial complaint and other pleadings. The complaint in particular often contains useful identifying information, such as a company touting the success of their products in the marketplace (suggesting a Category 8), or the prowess of the individual inventor in starting the company (suggesting a Category 5). In addition, reviewing the complaint also enables students to distinguish declaratory judgments, where the patentasserter is the defendant in the action.

Second, student researchers utilize Lex Machina’s links to the patents asserted in a lawsuit to gather relevant information. For example, Lex Machina indicates whether the same patent was asserted in other cases, enabling students to review the pleadings in those cases to learn additional facts about the entity. Additionally, the Lex Machina interface allows students to determine whether the patent was asserted by the inventor, or if the patent has been assigned and asserted by non-inventors. Lex Machina’s patent page also contains links to Google Patents and the PTO’s online patent databases; student researchers use these links to gather further information about patent assignments and the geographic location of the inventors.

Third, student researchers conduct a comprehensive web search of each unknown patentasserter. These searches usually reveal useful sources like an entity’s website, which may offer background or applicable product information. Additionally, web searches may reveal media and industry news articles discussing the entity or its litigation, public corporate information such as SEC filings identifying the entity owners and sources of income, or other legal websites (like RPX) that would identify certain cases as those involving producers or non-practicing entities. Table 2 summarizes examples of the evidence used to code each particular entity.

²⁷ As we will report in our results section below, about 6 percent of lawsuits reviewed so far include multiple patent asserters that also fit into multiple entity categories.

²⁸ Note that the Dataset is now publicly available at <https://law.stanford.edu/projects/stanford-npe-litigation-dataset/#/slnav-brief-dataset-methodology>. Contact Shawn Miller, Research Lead, Stanford NPE Litigation Dataset (smiller@law.stanford.edu) with any questions.

Table 2

STANFORD NPE LITIGATION DATASET PATENT ASSERTER TAXONOMY

| Category | Evidence |
|--------------------------------------|--|
| 1. Acquired patents | <ul style="list-style-type: none"> • The entity states in the pleadings or on its website that its purpose is to license patents or generate revenue from licensing patents • There is no evidence in the pleadings or on the entity’s website that it made, sold, or offered a product or service • The entity has been identified in news reports or legal websites as an NPE • The entity has never been sued for patent infringement • The entity has filed many lawsuits and/or has sued many different alleged infringers • The entity has filed suit in the Eastern District of Texas or the District of Delaware • The entity’s address was within the Eastern District of Texas • The entity has a small number of employees • None of the entity’s employees or founders match the name of the individual inventor on the patent • A PTO patent assignment search shows that the original patent owner (e.g., individual or producing company) transferred the patent to the entity that appears unrelated • The patent had been assigned to the entity shortly before a large increase in litigation • The entity generates largely variant amounts of income |
| 2. University Heritage or Tie | <ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the company has ties to a university or exists to license university intellectual property |

| | |
|---|--|
| | <ul style="list-style-type: none"> • There is no evidence suggesting the entity is actually the university itself or an alias for the university |
| 3. Failed startup | <ul style="list-style-type: none"> • The pleadings, company website, news reports or other filings indicate the entity’s former intention to make products or offer services, such as the creation of prototypes or attempts at marketing • The complaint, company website, news reports, or other filings indicate that the entity had received venture funding • The complaint, company website, news reports, or other filings indicate that the entity is a “startup” • There is evidence that the entity was a 1990s “dot.com” that failed when the tech bubble burst |
| 4. Corporate Heritage | <ul style="list-style-type: none"> • Websites, news, or other reports indicate that the entity formerly made products or offered services • Websites, news, or other reports indicate that the entity shifted its business to IP monetization • Directories indicate that the entity formerly operated retail establishments |
| 5. Individual-Inventor-Started Company | <ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the entity is owned, operated, or related to the inventor of the held patent(s) • There is no evidence available to suggest that the entity itself offers products or services for sale • A PTO PAIR search shows the inventor(s) resided in the same geographic location as the entity • Public company information indicates that the inventor(s) listed on the asserted patent owns, operates, founded or is somehow related to the entity (e.g., familial relative to the company owner) |

| | |
|--|---|
| <p>6. University/Government/NGO</p> | <ul style="list-style-type: none"> • The pleadings or website of the entity states that it is an institution for higher education, a government body, or a non-profit organization |
| <p>7. Startup, pre-product</p> | <ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the entity is a “startup” but has not yet developed a product or service • The complaint, company website, news reports, or other filings indicate that the entity has received venture funding • There is no evidence available to suggest that the entity actually offers products or services |
| <p>8. Product Company</p> | <ul style="list-style-type: none"> • The pleadings, company website, court filings, or SEC filings indicate that the plaintiff made a product, sold a product, or offered a service at the time the lawsuit was filed • The entity has a many employees • The entity operates brick and mortar retail establishments or a functioning website offering actual products or services • News or other reports describe the entities products or services |
| <p>9. Individual</p> | <ul style="list-style-type: none"> • The entity identifies itself in the complaint as an individual or trust • The entity has the name of a natural person • The entity is a family trust (typically for the inventor(s) family) • Usually, the name of the entity matches the name of the inventor on the patent |
| <p>10. Undetermined</p> | <ul style="list-style-type: none"> • There is no evidence available to support the classification of the entity asserting the patent |
| <p>11. Industry Consortium</p> | <ul style="list-style-type: none"> • The pleadings, entity website, news reports, or other filings indicate the entity is a standards setting or other type of industry organization operating on behalf of independent companies |

| | |
|---|---|
| | <ul style="list-style-type: none"> • The evidence suggests that the entity does not itself produce products or services beyond IP holding or assertion on behalf of its members |
| 12. IP Subsidiary of Product Company | <ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the entity is a holding company for another company’s intellectual property • There is no evidence available to suggest that the entity itself offers products or services |
| 13. Corporate-Inventor-Started Company | <ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the entity performs in-house research and development • The pleadings, company website, news reports, or other filings indicate that the entity has made significant expenditures on research and development • The pleadings, company website, news reports, or other filings indicate that the entity exclusively licenses its technology to manufacturers for commercialization • The entity has a larger number of employees than needed merely to enforce IP |

After determining the proper categories for the patent asserters in a lawsuit, student researchers assign the numbers for those categories to the patent asserters in the user interface. Before moving to the next unknown entity, the coders review every other lawsuit involving the entities and assign them to their appropriate category in every lawsuit where they appear. Occasionally, a student finds that another coder has previously categorized an entity differently in other lawsuits. When this occurs, the students resolve the inconsistency by reviewing the evidence and discussing a final judgment as a group.

When a student is unable to categorize an entity after reviewing all the sources of evidence, the student notes the entity and enlists the aid of the group through quality control measures described below.

Quality Control

The Practicum instituted a procedure for quality control to ensure reliable coding. The quality control (“QC”) procedure has several levels of review. First, researchers are instructed and encouraged to bring ambiguous or challenging classifications to the entire research group for discussion at weekly meetings. For example, if a researcher suspects—but is not certain—that an entity is a licensing firm that acquires patents (Category 1) as opposed to an individual-inventor-started licensing firm (Category 5), that researcher would email the entire research group for their thoughts and suggestions. Only after the group achieved consensus, with input from Professor Shawn Miller, would the researcher then code the entity in the Dataset.

Additionally, Professor Miller regularly reviews random samples of each researcher’s work product and, if necessary, directs that team member to revise any errors and revisit any categories that were consistently miscoded. Since the second year of the research project, more experienced law students have also performed QC for those newer to the project. The senior member reviews roughly 20 lawsuits out of an assigned batch of 200 (10% random sample). Inter-coder reliability between researchers ranges from 85 to 100% (i.e., researchers disagree on 0 to 3 of the 20 cases they reviewed). When agreement is less than 90%, or it appears that mistakes are due to a misunderstanding of our categorization taxonomy, Professor Miller and/or senior team members instruct the researcher to review their batch again and correct mistakes. Further, if the reviewing senior members were unsure of an entity coded by a junior member, the members would discuss the issue with the entire research team. Importantly, researchers were not rewarded for the number of cases they ultimately coded. Instead, the research Practicum emphasized the primacy of accuracy over speed.

We believe that the methodology and quality control measures we have employed thus far have resulted in a reliable Dataset. However, the best mechanism for ensuring reliability is its public accessibility. We encourage all who download the Dataset to contact us with any categorization disagreements and we will re-review the evidence and re-categorize entities as necessary.

III. Description of the Data: Summary Statistics, Time Trends, and Preliminary Findings

[To be added in Next DRAFT]

IV. Conclusions and Suggestions for Future Research

Motivated by the intense debate surrounding the impact of PAEs on the patent system, innovation, and the macro economy, we began our work creating the Stanford NPE Litigation Dataset nearly four years ago. Since then, we have carefully categorized patent asserters in over 40,000 cases to date including a random sample of 20% of all lawsuits filed between 2000 and 2015 (over 10,800 cases). Our preliminary data release consists of these cases, and will be updated in the coming months to include all 54,000 cases filed between 2000 and 2015.

Preliminary findings from our Dataset offer surprises. For one, licensing firms that acquired their patents from other entities were the type most responsible for the doubling of total defendant-lawsuit pairs since 2000. In 2000, these licensing firms asserted their patents in less than 5% of distinct defendant-plaintiff disputes, but in more recent years in over 30% of all disputes. Other findings reinforce what other researchers have already revealed. For example, PAE suits tend to involve software patents, terminate more quickly, and end in settlement more frequently than practicing entity cases. These facts demonstrate that PAEs have been responsible for a growing share of patent enforcement, and also lend support to the view that PAEs may assert patents of questionable validity (or more uncertain scope) and thus pursue a strategy of seeking quick nuisance value settlements.

The Dataset provides immediate value to scholars who wish to replicate past empirical studies using the full population of lawsuits over a significant span of years. Scholars who have previously conducted studies on PAE litigation can confirm past research conclusions utilizing our comprehensive Dataset. The data might also reveal that some findings derived from small non-random samples are not supported in analysis of the population or other subsets. For example, scholars can now study how PAEs of varying sizes operate, rather than just how the largest and most easily recognizable PAEs operate.

Dataset research will allow scholars to develop a deeper understanding of the realities of litigation in the patent system. Scholars can examine the near doubling of distinct defendant-lawsuit pairs revealed over the 16-year period of the random-sample analysis. A more in depth analysis of the entire Dataset might reveal the catalyst of increased patent litigation. Scholars may also be able to draw more robust conclusions about PAE settlement rates and the amount of time respective entities take to litigate.

The Dataset also offers scholars the ability to focus on industry specific players or patents. For example, in PAE lawsuits filed in 2014, approximately half of all PAE plaintiffs asserted a software patent. In contrast, only a small minority (11%) of Category 8 practicing entities asserted a software patent in 2014. Researchers can hone in on particular industries and their impact on the overall patent system. Another potential research topic is whether the patents found in certain industries are more likely to be subject to inter partes review.

These preliminary findings and suggestions for further research underscore the potential value of further utilization our Dataset. We encourage scholars, policymakers, and other decision-makers to use our Dataset to test new theories, craft new policy, and engage the PAE debate in a data-driven approach. Our shared goal is to bring transparency to patent litigation by reshaping that landscape with new tools and understandings that help to promote the public interest through innovation, entrepreneurship, and competition across sectors and industries.