

Artificial Intelligence In Search of Protection

Author, Francesco Corea

A Data Science Foundation White Paper

May 2018

www.datascience.foundation

Copyright 2016 - 2017 Data Science Foundation

Artificial Intelligence In Search of Protection:

Is a patent a straw, stick or brick house?

I. Setting the stage

Working alongside early-stage companies, venture funds, and corporations, I often wonder what the word “**defensibility**” means nowadays for an AI startup. And even if I find myself often thinking and advising people on why a patent on a machine learning product/algorithm makes sense or not, I recently realized that this problem is actually related to how I see an AI company be protected in the long-run.

I will blog later about the broader concept of *defensibility* and rather focus here only on *IP protection*. I am not a patent lawyer, so there might likely be many details I am either missing or misunderstanding (and you should definitely check with one if you are thinking about patenting something) but from a business perspective, I think I have now a pretty good picture on what is going on.

So let's start with the fundamentals.

II. Why startups patent inventions (and why is different for AI)

Patenting an invention is one of the four forms of IP protection (i.e., patents, trademarks, copyrights, and trade secrets) and it basically gives you the chance to use and exploit the economic benefits of a certain invention for a quite long period (usually 20 years) demanding in turn that you make your invention public for the sake of the scientific and technological progress.

For a startup, having a patent has historically been a great advantage over the competition, but is also a burdensome cost to cover especially at a very early stage (around 20k both in Europe and US over a 3 to 5 years period). And this cost is not even getting into account the fact that obtaining a patent is the same of tossing a coin (only 55% of the applications result into a granted patent, eventually—see Carley et al., 2014 for more details).

So why companies get themselves into this tricky and expensive process in the first place? Well, there are multiple reasons for doing it. You can indeed create a strong competitive advantage or shape a new stream of revenues by licensing your technology, or simply increase your confidence that what you are building is being recognized by the rest of the world as useful.

Indeed one of the companies I am working with called [Meeshkan ML](#) recently applied for a patent for their distributed machine learning algorithm because in the founder's words “*we feel that we can share it with our local community of engineers and build new algorithms on top of it without worrying about losing our business advantage. We therefore made a calculated risk in investing time into the R&D necessary for the patent instead of pushing a product out.*”

But most of all, a startup often heavily depends on external financing—and investors love(d) patents. It is a simple way for many of them to achieve three things at once: they get sure the tech is legit (i.e., **they outsource the technical due diligence** to patent lawyers and offices); they are more confident the **technology is feasible** (reducing product risks); they get more confident the **team can actually build** what they claimed and that is committed to it (reducing execution and team risks).

This is especially true for deep science ventures as well as emerging technologies where understanding those three aspects is incredibly cumbersome. AI is clearly one of those fields, but I am also claiming it stands apart from other technological inventions for a bunch of different reasons:

- **A strong open-source community:** many AI algorithms/libraries/packages are completely open-sourced, and you usually tend to build on top of those (which makes patenting extremely hard to be managed);
- **Confusion around the (real) invention:** it is still hard to identify which part of an AI algorithm is the real source of invention. Is it the source code? Or rather the data used? Or maybe the process or the human-machine blending interaction?
- **Continuous evolution:** feedback loops in machine learning push the code to keep evolving, which complicates the understanding of whether a new invention breaks the initial patent.
- **It stems from an academic community:** traditionally many inventions are industry-driven while AI is historically intrinsically related to academia, where the approach to scientific research and development is culturally very different;
- **It is still half science half art:** even though a patent requires a fairly decent degree of details around the invention in question, when it comes to AI the devil is into the smallest details (e.g., tuning). In other words, even with a full access to the process or to the algorithm, you might not be able to use it and implement it correctly.

III. The advantages of patenting your product

So, if you are an AI entrepreneur, you might now question the utility of a patent and above all interrogating yourself with the big question: **should I even try to get a patent in the first place** or invest my money in more meaningful activities?

Quick answer: **it depends.**

In addition to common sense reasoning (which is more or less what I mentioned above), there is empirical evidence that a patent may be beneficial to your business.

In fact, Mann and Sager (2007) proved that **having patents is positively correlated to startup performance indicators**, such as the number of financing rounds, total investment received, exit status, longevity and late-stage financing. This is not an isolated example since Conti et al. (2013) for instance also showed that **having a patent increases the likelihood of getting VC funding** (but no angel funding, because angels apparently value other things more than IP protection). Haeussler et al. (2014)

interpreted patents as a quality signals for VC financing, finding that **one patent application reduces time to venture funding by 76%** (and VCs are apparently very good at understanding the quality of a patent) and therefore increases the likelihood of getting funding (an interesting side evidence is that also patent oppositions are well seen and increase likelihood of funding, maybe because they are signs of market potential and validation). This probability is even higher when you complement patents with trademarks (Zhou et al., 2016).

Furthermore, Baum and Silverman (2004) also found a **positive association between patent applications and pre-IPO financing**, while Farre-Mensa et al. (2017) instead pushed this analysis one step forward. They did not simply show that the first patent you obtain **increases your likelihood of raising venture capital funding in the following three years by 2.3** percentage points (a 53% increase relative to the 4.3% unconditional probability of raising VC funding, although the effect is smaller during the first year and ramps up later) as well as the likelihood of securing a loan by pledging the patent as collateral, but also showed the impact of them on the daily operations.

In fact, startups that file a patent have, on average, **55% higher employment growth and 80% higher sales growth five years later**. They also pursue more **follow-on quality innovation** and have a higher probability of getting subsequent patents (up by 49%) of a higher quality (with the average number of citations per subsequent patent increasing by 26%).

Eventually, have a portfolio of patents will have a strong positive effect also in case of an IPO.

Hence, let's quickly recap what we learned: a patent may bring you more funding, more quick funding, attract talents and increase sales, put you on the right road to keep building great things and increase your exit value.

Well, truth be told, it sounds like a great benefits package to me. So why is not everyone rushing to patent every single invention developed? Here comes the catch.

First of all, the results that we mentioned above do not apply to everyone (Farre-Mensa et al., 2017). Generally speaking, they apply only to startups that **have raised little or no money before patenting** (very early-stage), lead by **inexperienced entrepreneurs** (serial entrepreneurs do not need patent to prove the company potential and tech feasibility), mostly **operating in IT**, and located in **hot technology hubs** (where getting investors' attention is otherwise difficult). Moreover, **additional patenting through funding rounds do not increase the investment amounts**.

Block et al. (2013) indeed show that an **inverted U-shaped relationship** exists between the number of trademarks and the company's financial valuations (i.e., it is positively related up to a certain threshold and then becomes negative), and only for early-stage startups. The same type of relationship also exists between the breadth of start-ups' trademark portfolios and the financial valuations, and it is also valid for the relationship between persified patenting activity and performance (Fernhaber and Patel, 2012; Li et al., 2012; Qian, 2002).

So, if you are an early-stage startup trying to find your place in the world and equipped with a brilliant technical innovation in your hand, I would at least encourage you to have this conversation with a patent lawyer. The first patent you filed is definitely worth its value.

IV. Reasons behind not looking for patent protection

Of course is not all peaches and dandelions here. If from one hand patents increase valuation and attract investors' attention, some studies seem to suggest otherwise (Smith and Cordina, 2014). Looking at the patents portfolio is a shortcut for less sophisticated investors, but for the best one is nothing more than an extra data point in their complex risk-return simulation modeling.

Many startups deliberately choose then to not patent any internally generated innovation, and most of the time they do so for one or more of the following reasons:

- **They are too early-stage:** the majority of companies working nowadays in AI are relatively young (according to CBInsights, 69% or more of AI deals since 2012 have gone to startups still in the early stages). This often means they have no means, either financial or in terms of resources and time, to file a patent;
- **They might convey too much information:** even though a single patent might not say much about your business, a portfolio of them can actually give many more strategic insights to a careful observer (and probably more than what you aim for);
- **The patent landscape is messy:** as we already mentioned above, the presence of widespread open-source libraries drastically increases the complexity of filing a patent;
- **Patents are becoming less meaningful:** it is extremely easy to violate a patent without even knowing it, and incredibly hard to enforce one. Add to this problem the corollary issue that patent litigation (when you are actually able to recognize a violation, which is not obvious) costs you around half a million (a trademark instead between 300k and 500k) and you will not want to get yourself into that problematic spot at all;
- **Cultural pergence** (i.e., they do not believe in it): I might spend many words here, but I rather prefer to quote [Tesla announcement of a couple of years ago](#):

“Tesla will not initiate patent lawsuits against anyone who, in good faith, wants to use our technology. [...] Technology leadership is not defined by patents, which history has repeatedly shown to be small protection indeed against a determined competitor, but rather by the ability of a company to attract and motivate the world’s most talented engineers. We believe that applying the open source philosophy to our patents will strengthen rather than diminish Tesla’s position in this regard” (Elon Musk).

- **They have alternative protections:** you can easily supply the absence of patents with different traditional and non-traditional moats (by the way, if you haven't already, go and read [Gil Dibner's post](#) and [Jerry Chen's one](#) of this topic):

i) **Data moat:** most of the AI software need to be fed with millions of data points. Even if you have access

to the algorithm, without data you cannot basically use it. Furthermore, having access to different datasets (or multiple “*systems of records*”, as Jerry Chen called them) and use them together not only provides you with an exponentially higher value but also increases the likelihood you can deeply personalize your product and tailor it to your customers (which then increases their **switching costs** toward the competition);

ii) **Network effect** (also called **Metcalf’s law** or **AI flywheels**): the value of the network increases exponentially with the number of the nodes, which in other words means that new customers imply more data, which implies a better algorithm, which in turn implies new customers, and so on so forth. You need though to reach a [minimum algorithmic performance](#) (MAP) before the network can attract the first cohort of customers with a strong value proposition;

iii) **Talents**: AI as a field is dominated by academic-trained talents, which are incredibly expensive and hard to hire and often demands an open approach to research (i.e., open-source software and publications in scientific journals). There are very few strong AI researchers worldwide, and securing one major name and/or a good team can represent a strong barrier for a startup (although you really hit the jackpot if you hire someone with specific **domain expertise**, which is quickly becoming the real differentiation between succeeding and surviving);

iv) **Cost efficiency**: many of the most powerful AI-driven businesses today are able to scale while limiting costs related to teams and product development (and let’s not forget that until some time ago the median team size for startups acquired by big incumbents [was around 7](#)).

Again, patenting is not a solution that works for everyone, but if you have none of the moats above I would highly recommend you to speak to a patent lawyer and look for some additional IP protection.

V. The three Ws game: Who, What and Where

If you are an AI entrepreneur that wants to try to patent something, it is certainly useful to start from mapping the landscape of existing patents to see whether something similar has already been filed and whether an opportunity materializes for your innovation.

This is a very specific process related to what you want to build and goes beyond the scope of this post, but what it can be intriguing to present is the current AI spectrum of patents and players.

[CB Insights](#) analyzed over [1,150 AI companies in the last decade](#) (since 2009) finding that **21% applied** for a patent and **only 11% eventually got** at least one. In particular, in the last five years, the issue is becoming more tricky and the US IP law started drawing some clear lines (the most famous case up to date is *Alice Corp. v. CLS Bank*, where they rejected a patent application on computer software because too *abstract*). Nowadays, it seems therefore clear that training sets, proprietary information, a particular expression of source code, and many other steps in an AI value chain **cannot be patented**. On the other

side though, **trade secret protection** is suitable for a wide variety of circumstances (e.g., neural networks, training sets, AI-generated code, learning algorithm, etc.).

CBInsights further pided the [patents of the “Big 5” \(Apple, Google, Amazon, Facebook, and Microsoft\)](#) from the [patents obtained by startups](#). In this fashion, big tech giants patented several things, although many times because the **software is often attached to a hardware component** (Amazon for logistic robotics, Apple for the iPhone, and even Google for their smartphone products). Google leads the pack in applying for AI patents, while Microsoft is the most prolific in filing for patents overall. We are talking here about a relatively small magnitude (less than 40 AI patents in 2017, with a peak of 164 in 2015). On the startup side instead, Cortica and [Numenta](#) are dominating this space (with 38 and 37 patents, respectively) followed by Butterfly Network (27), SoundHound (26) and Smart Drive (24). Interestingly enough, in addition to traditionally strong-IP fields (e.g., healthcare), most of the patents have been filed on **horizontal AI type of applications** (in other words, platform, neuroscience approach to AI, GAI, core AI, etc.).

If we then widen the spectrum of what is being patented worldwide at any level and from any organization, it seems clear that most of the AI patents focus on **enabling intelligent robots** (e.g., self-driving cars, automated delivery drones, AI assistants, etc.), **deep learning, face recognition** and **AI hardware** (especially in China).

Finally, US and China are the two countries where most of this innovation (over 50% of the patents registered) has been happening over the last decade or so (which one of the two is leading is still controversial though), followed by Japan, South Korea, Germany, Canada, the U.K., Australia, India, and Russia. China seems to have an uncontested supremacy over **deep learning and machine vision**, where US is more keen to develop **NLP and other machine learning** technologies, and it is filing up patents at a much faster pace than its American counterparts (interestingly enough, Chinese researchers [more than doubled](#) the number of AI scientific papers published in 2017 compared with 2010, trend that is going downward in the US). Europe instead places itself in the middle, with a flattening growth in patents applications (OECD, 2017) and accounting for 10% of the US quantity. However, even if patenting is not so widespread in Europe, the scientific publication side is steadily growing (and of course, many [recent announcements](#) have been done to better position Europe in the AI race).

VI. Conclusions

When I started thinking about this IP issue, my main question was “how is it possible to patent something when the market is driven by open-source forces?”, cause I was naively assuming the two elements being incompatible. The reality is that they are not because they might serve different purposes and appear in different stages of a company life.

Bottom line, keep your mind open about the chance of protecting your IP when you are very early-stage or to send a signal out there about what you are building, but do not waste too much money and time on it if you are already building a good traction quickly. Do not also assume that a patent will vouch for

everything you say and do because it won't, and it is not a substitute to a great co-founding team or a strong market validation, and by all means not having a patent does not imply your product is inferior or your company is worse than others.

To say that with terms that we all know, **“patents are vitamins, not painkillers”**.

I am always interested in speaking to, learning from or simply connecting with interesting founders working in highly impactful fields like life sciences, energy, and others. If you are one of them, [feel free to reach out here!](#)

References

Baum, J. A., Silverman, B. S. (2004). “Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology start-ups”. *Journal of Business Venturing* 19: 411-436.

Block, J. H., De Vries, G., Schumann, J. H., Sandner, P. (2013). “Trademarks and venture capital valuation”. *Journal of Business Venturing* 29 (4): 525-542.

Carley, M., Hegde, D., Marco, A. (2014). “What is the Probability of Receiving a U.S. Patent?”. *Yale Journal of Law and Technology* 17 (1): 204-224.

Conti, A., Thursby, M., Rothaermel, F.T. (2013). “Show me the right stuff: Signals for high-tech startups”. *Journal of Economics & Management Strategy* 22 (2): 341-364.

Farre-Mensa, J., Hedge, D., Ljungqvist, A. (2017). “What is a Patent Worth? Evidence from the US Patent “Lottery”. NBER Working Paper, 23268: 1-57.

Fernhaber, S.A., Patel, P. (2012). “How do young firms manage product portfolio complexity? The role of absorptive capacity and ambidexterity”. *Strategic Management Journal* 33 (13): 1516-1539.

Haeussler C., Harhoff D., Mueller E., (2014). “How Patenting informs VC investors—The Case of Biotechnology”. *Research Policy* 43 (8): 1286-1298.

Li, L., Qian, G., Qian, Z. (2012). “The performance of small and medium-sized technology-based enterprises: do product persity and international persity matter?”. *International Business Review* 21 (5):

Data Science Foundation

941-956.

Mann, R. J., Sager, T. W. (2007). "Patents, venture capital, and software start-ups". *Research Policy* 36: 193-208.

OECD. (2017). "Science, innovation and the digital revolution", in *OECD Science, Technology and Industry Scoreboard 2017: The digital transformation*, OECD Publishing, Paris.

Qian, G., (2002). "Multinationality, product persification, and profitability of emerging US small- and medium-sized enterprises". *Journal of Business Venturing* 17 (6): 611-633.

Smith, J. A., Cordina, R. (2014). "Patenting and the early-stage high-technology investor: evidence from the field". *R&D Management* 45 (5): 589-605.

Zhou, H., Sandner, P. G., Martinelli, S. L., Block, J. H. (2016). "Patents, Trademarks, and Their Complementarity in Venture Capital Funding". *Technovation* 47: 14-22.

About the Data Science Foundation

The Data Science Foundation is a professional body representing the interests of the Data Science Industry. Its membership consists of suppliers who offer a range of big data analytical and technical services and companies and individuals with an interest in the commercial advantages that can be gained from big data. The organisation aims to raise the profile of this developing industry, to educate people about the benefits of knowledge based decision making and to encourage firms to start using big data techniques.

Contact Data Science Foundation

Email: contact@datascience.foundation
Telephone: 0161 926 3641
Atlantic Business Centre
Atlantic Street
Altrincham
WA14 5NQ
web: www.datascience.foundation

Data Science Foundation

Data Science Foundation, Atlantic Business Centre, Atlantic Street, Altrincham, WA14 5NQ
Tel: 0161 926 3641 Email: contact@datascience.foundation Web: www.datascience.foundation
Registered in England and Wales 4th June 2015, Registered Number 9624670