NEW YORK UNIVERSITY

JOURNAL OF INTELLECTUAL PROPERTY AND ENTERTAINMENT LAW

VOLUME 7 SPRING 2018 NUMBER 2

THINK BIG! THE NEED FOR PATENT RIGHTS IN THE ERA OF BIG DATA AND MACHINE LEARNING

HYUNJONG RYAN JIN

From personalized medical diagnostics to election prediction, recent advancements in machine learning enables unprecedented, powerful applications of big data. Machine learning users can extract insights hidden in massive amounts of data, gaining an indispensable advantage against the competition. Investment in the process of gathering and analyzing data has now become a necessity to maintain a successful enterprise. Yet the difficulty of obtaining software patents since the 2014 Alice decision raises the question whether the current intellectual property framework may adequately protect inventions related to machine learning. This Note explores how we may utilize IP protection to harness the societal benefits we hope to enjoy through the advances in machine learning. The Note discusses the current framework of patent law, copyright, and trade secret in the context of machine learning inventions, and argues that patent rights for computational inventions adequately balances the concern of patent monopoly and promoting innovation. The Note concludes by applying the Alice framework to the proposed computational inventions, and demonstrates that the current patent system may still protect machine learning innovations.

INTRODUCTION

With AlphaGo's triumph over the 9-dan Go professional Lee Sedol in March 2016, Google's DeepMind team conquered the last remaining milestone in board game artificial intelligence.[1] Just nineteen years after IBM Deep Blue's victory over the Russian chess grandmaster Garry Kasparov,[2] Google's success exceeded expert predictions by decades.[3]

AlphaGo demonstrated how machine learning algorithms could enable processing of vast amounts of data. Played out on a 19 by 19 grid, the number of possible configurations on a Go board is astronomical.[4] With near-infinite number of potential moves, conventional brute-force comparison of all possible outcomes is not feasible.[5] To compete with professional level human Go players, the gaming artificial intelligence requires a more sophisticated approach than the algorithms employed for chess—machine learning. The underlying science and implementation of machine learning was described in a Nature article two months prior to AlphaGo's match with Lee. In the article, the Google team described how a method called "deep neural networks" decides between the insurmountable number of possible moves in Go.[6] The AlphaGo model was built by reinforcement learning from a database consisting of over thirty million moves of world-class Go players.[7] This allowed the algorithm to optimize the search space of potential moves, therefore reducing the required calculations to determine the next move.[8] In other words, the algorithm mimics human intuition based on the "experience" it gained from the database "fed" into the algorithm, which drastically increases computational efficiency by eliminating moves not worth subsequent consideration. This allows the algorithm to devote computational resources towards the outcomes of "worthwhile" moves.

---

[1] Sang-Hun Choe & John Markoff, *Master of Go Board Game Is Walloped by Google Computer Program*, N.Y. TIMES (March 9, 2016), https://www.nytimes.com/2016/03/10/world/asia/google-alphago-lee-se-dol.html (reporting the shocking defeat of Go Master Lee Se-dol to Google DeepMind's AlphaGo).

[2] Laurence Zuckerman, *Chess Triumph Gives IBM a Shot in the Arm*, N.Y. TIMES (May 12, 1997), http://politics.nytimes.com/library/cyber/week/051297ibm.html (detailing IBM's highly publicized win through Deep Blue's victory over world chess champion Garry Kasparov).

[3] See Choe & Markoff, *supra* note 1.

[4] David Silver et al., *Mastering the game of Go with deep neural networks and tree search*, 529 NATURE 484, 484 (2016).

[5] *Id.*

[6] *Id.*

[7] *Id.* at 485.

[8] *Id.*

The advent of such powerful analytical tools, capable of mimicking human intuition alongside massive computation power, opens endless possibilities—early stage cancer detection[9], accurate weather forecasting,[10] prediction of corporate bankruptcies,[11] natural event detection,[12] and even prediction of elections.[13] For information technology ("IT") corporations, investment in such technology is no longer an option, but a necessity. The question that this Note addresses is whether the current state of intellectual property law is adequate to harness the societal benefits that we hope to enjoy through the advances in machine learning. In particular, are patents necessary in the age of big data? And if they are, how should we apply patent protection in the field of big data and machine learning?

Part I of this Note examines the need for intellectual property rights in machine learning and identifies the methods by which such protection may be achieved. The differences between trade secret, copyright and patent protection in software are discussed, followed by the scope of protection offered by each means. This background provides the basis to discuss the effectiveness of each method in the context of machine learning and big data innovations.

Part II discusses the basics of the underlying engineering principle of machine learning and demonstrates how the different types of intellectual property protection may apply. Innovators may protect their contributions in machine learning by defending three *areas*—(1) the vast amount of data required to train the machine learning algorithm, (2) innovations in the algorithms itself including advanced mathematical models and faster computational methods, and (3) the resulting machine learning model and the output data sets. Likewise, there are three distinct *methods* of protecting these intellectual properties: patents, copyright, and secrecy.[14] This Note discusses the effectiveness of each method of intellectual property protection with three principles of machine learning innovation in mind:

---

[9] *See* Andre Esteva et al., *Dermatologist-level classification of skin cancer with deep neural networks*, 542 NATURE 115 (2017).

[10] *See* Sue Ellen Haupt & Branko Kosovic, *Big Data and Machine Learning for Applied Weather Forecasts*, IEEE SYMPOSIUM SERIES ON COMPUTATIONAL INTELLIGENCE (2015).

[11] *See* Wei-Yang Lin et al., *Machine Learning in Financial Crisis Prediction: A Survey*, 42 IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS 421 (2012).

[12] *See* Farzindar Atefeh & Wael Khreich, *A Survey of Techniques for Event Detection in Twitter*, 31 COMPUTATIONAL INTELLIGENCE 132 (February 2015).

[13] *See* Corey Blumenthal, *ECE Illinois Students Accurately Predicted Trump's Victory*, ECE ILLINOIS (Nov. 18, 2016), https://www.ece.illinois.edu/newsroom/article/19754.

[14] For the purpose of this Note, secrecy refers to the use of trade secret and contract based non-disclosure agreements.

facilitating data sharing, avoiding barriers to entry from data network effects, and providing incentives to address the key technological challenges of machine learning. This Notes proposes that patents on computational methods adequately balance the concern of patent monopoly and promoting innovation, hence should be the primary means of intellectual property protection in machine learning.

Part III then visits the legal doctrine of patentable subject matter starting with the United States Supreme Court's *Alice* decision. While *Alice* imposed a high bar for software patents, the post-*Alice* Federal Circuit decisions such as *Enfish*, *Bascom*, and *McRO* suggest that certain types of software inventions are still patentable. Specifically, this section will discuss the modern framework pertinent to subject matter analysis: (1) inventions that are directed to improvements of computer functionality rather than an abstract idea, (2) inventions that contain an inventive concept, and (3) inventions that do not improperly preempt other solutions. The Note will apply this framework to innovations in machine learning.

The Note proposes that patents for computational methods balance the need for intellectual property protection while permitting data sharing, paving the pathway for promoting innovation in machine learning. The Note further argues that machine learning algorithms are within patentable subject matter under 35 U.S.C. §101.

I
NEED FOR INTELLECTUAL PROPERTY RIGHTS IN MACHINE LEARNING

> "*He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me.*"
>
> – Thomas Jefferson
>
> "*I'm going to destroy Android, because it's a stolen product. I'm willing to go thermonuclear war on this. They are scared to death, because they know they are guilty.*"
>
> – Steve Jobs

The two quotes above demonstrate the conflicting views on protecting intangible ideas with intellectual property law. Thomas Jefferson implied that the free circulation of inventive ideas and thoughts would not dampen the progress of innovation nor disadvantage innovators. On the other hand, Steve Jobs exhibited fury over the similarity between the iOS and the Android OS. Why? Was it

because his company was worse off due to the similarity between the two products? Would Apple have refrained from inventing the iPhone had it known others would enter the smartphone market?

This section discusses the motives behind the grant of intellectual property rights and whether such protection should be extended to machine learning innovations. Basics of patent law, copyright law, and trade secret are introduced to provide the analytical tools for subsequent discussion on which type of intellectual property protection best promotes the socially-beneficial effects of machine learning.

A. Do We Need Intellectual Property Rights for Machine Learning?

The primary objectives of intellectual property rights are to encourage innovation and to provide the public with the benefits of those innovations.[15] In the context of machine learning, it is not clear whether we need any additional incentives to promote participation in this field. Machine learning is already a "hot field," with countless actors in industry and academia in active pursuit to keep pace.[16] Hence investment incentivizing may not be a valid justification for granting intellectual property rights in machine learning. Rather, such protection is crucial to promote competition and enhance public benefits.

The *quality* of inferences that may be drawn from a given data set increases exponentially as the aggregation diversifies, which is why cross-industry data aggregation will greatly enhance the societal impact of machine learning.[17] Companies will need to identify new data access points outside of their own fields to gain access to other data sets to further diversity their data. Yet the incentive structures of behemoth corporations may not be well-suited to identify and grow

---

[15] Mark A. Lemley, *The Surprising Virtues of Treating Trade Secrets As IP Rights*, 61 STAN. L. REV. 311, 332 (2008) ("Patent and copyright law do not exist solely to encourage invention, however. A second purpose — some argue the main one — is to ensure that the public receives the benefit of those inventions.").

[16] Andrew Ng et al., *How Artificial Intelligence Will Change Everything*, WALL STREET JOURNAL (March 7, 2017), https://www.wsj.com/articles/how-artificial-intelligence-will-change-everything-1488856320.

[17] Limor Peer, *Mind the Gap in Data Reuse: Sharing Data Is Necessary But Not Sufficient for Future Reuse*, LONDON SCH. ECON. & POLI. SCI. (Mar. 28, 2014) http://blogs.lse.ac.uk/impactofsocialsciences/2014/03/28/mind-the-gap-in-data-reuse ("The idea that the data will be used by unspecified people, in unspecified ways, at unspecified times . . . is thought to have broad benefits").

niche markets.[18] It would be up to the smaller, specialized entities to find the gaps that the larger corporations overlooked and provide specialized services addressing the needs of that market. Protective measures that assist newcomers to compete against resource-rich corporations may provide the essential tools for startups to enter such markets. Sufficient intellectual property protection may serve as leverage that startups may use to gain access to data sets in the hands of the Googles and Apples of the world, thus broadening the range of social benefits from machine learning.

B. The Basics of Patent Law

> "*To promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries*"
>
> – United States Constitution, Article I, § 8

The United States Constitution explicitly authorizes Congress to promote useful arts by granting inventors the exclusive rights of their discoveries. Such constitutional rights stems from two distinct bases—(1) a quid pro quo where the government issues a grant of monopoly in exchange for disclosure to society, and (2) property rights of the inventor. The purpose for such rights is explicitly stated in the Constitution—to promote new inventions. The goal is to prevent second arrivers who have not invested in the creation of the initial invention from producing competing products and services at a lower price, undercutting the innovator whose costs are higher for having invested to create the invention. As an incentive for innovators willing to invest in new, useful arts, the patent system provides the innovator rights to exclude others from practicing the invention. Another purpose of such rights is the concept of "mining rights." Akin to the grant of mining rights to the owner in efforts to suppress aggressive mining, the inventor should have the right to define and develop a given field by excluding other people from the frontiers of that knowledge. Considering the importance of industry standards in modern electronics, such a purpose acknowledges the importance of early stage decisions that may define the trajectory of new technological advances.

---

[18] *See* Saeed Ahmadiani & Shekoufeh Nikfar, *Challenges of Access to Medicine and The Responsibility of Pharmaceutical Companies: A Legal Perspective*, 24 DARU JOURNAL OF PHARMACEUTICAL SCIENCES 13 (2016) (discussing how "pharmaceutical companies find no incentive to invest on research and development of new medicine specified for a limited population . . .").

C. The Thin Protection on Software Under Copyright Law

The Copyright Act defines a "computer program" as "a set of statements or instructions to be used directly or indirectly in a computer to bring about a certain result."[19] Though it may be counterintuitive to grant copyright protection for "useful arts" covered by patents, Congress has explicitly mandated copyright protection for software.[20] However, as will be discussed below, copyright protection of software has been significantly limited due to case law.

Copyright protects against literal infringement of the text of the program. Source code, code lines that the programmers "author" via computer languages such as C++ and Python, is protected under copyright as literary work.[21] In *Apple v. Franklin Corp.*, the Third Circuit Court of Appeals held that object code, which is the product of compiling the source code, is also considered a literary work.[22] Given that compiled code is a "translation" of the source code, this ruling seems to be an obvious extension of copyright protection. Removing the copyright distinction between source code and object code better reflects the nature of computer languages such as Perl, where the source code is not translated into object code but rather is directly fed into the computer for execution. However, the scope of protection on either type of code is very narrow. The copyright system protects the author against literal copying of code lines. This leaves open the opportunity for competitors to avoid infringement by implementing the same algorithm using different text.

Fortunately, in addition to protection against literal copying of code, copyright law may provide some protection of the structure and logical flow of a program. Equivalent to protecting the "plot" of a novel, the Second Circuit Court of Appeals ruled that certain elements of programming structure are considered an expression (copyrightable) rather than idea (not copyrightable), extending copyright protection to non-literal copying.[23] The *Computer Associates International v. Altai* court applied a three-step test to determine whether a computer program infringes other programs—(1) map levels of abstraction of the program; (2) filter out protectable expression from non-protectable ideas; and (3)

---

[19] 17 U.S.C. §101 (2012).

[20] *Id.*

[21] 17 U.S.C. §102(a) (Copyright exists "in original works of authorship fixed in any tangible medium of expression . . .").

[22] Apple Comput., Inc. v. Franklin Comput. Corp., 714 F.2d 1240 (3d Cir. 1983).

[23] Comput. Assocs. Int'l v. Altai, 982 F.2d 693 (2d Cir. 1992).

compare which parts of the protected expression are also in the infringing program.[24]

The merger doctrine is applied to step two of the *Altai* test to limit what may be protected under copyright law. Under the merger doctrine, code implemented for efficiency reasons is considered as merged with the underlying idea, hence not copyrightable.[25] Since most algorithms are developed and implemented for efficiency concerns, the *Altai* framework may prevent significant aspects of software algorithms from receiving copyright protection. This means that for algorithms related to computational efficiency, patents may provide significantly more meaningful protection than copyright. The Federal Circuit, in the 2016 case *McRO Inc. v. Bandai Namco Games America Inc.*, ruled that patent claims with "focus on a specific means or method that improves the relevant technology" may still be patentable.[26] Although preemption concerns may impede patentability, exemption of patent right by preemption is narrow compared to that of copyright by the merger doctrine.

Scène à faire doctrine establishes yet another limitation on copyright for computer programs. Aspects of the programs that have been dictated by external concerns such as memory limits, industry standards and other requirements are deemed as non-protectable elements.[27] For mobile application software, it is difficult to imagine programs that are not restricted by form factors such as mobile AP computation power, battery concerns, screen size, and RAM limitations. As for machine learning software, the algorithms determine the "worthiness" of computation paths based on conserving computational resources. The external factors that define the very nature and purpose of such machine learning algorithms may exempt them from copyright protection.

D. Comparing Trade Secret and Non-disclosures with Patents

The crucial distinction between trade secret and patent law is secrecy. While patent applicants are required to *disclose* novel ideas to the public in exchange for a government granted monopoly, trade secret requires owners to keep the information *secret*. Though trade secret protection prevents outsiders from acquiring the information by improper means, it does not protect the trade secret against independent development or even reverse engineering of the protected

---

[24] *Id.*
[25] *See id.* at 707-09.
[26] 837 F.3d 1299, 1314 (Fed. Cir. 2016).
[27] *Altai*, 982 F.2d at 698.

information. In trade secret doctrine, the existence of prior disclosed art is only relevant for discerning whether the know-how is generally known, a different and simpler analysis than the issue of novelty in patent law.[28]

The United States Supreme Court has specified in *Kewanee Oil* that all matters may be protected under trade secret law, regardless whether it may or may not be patented.[29] The *Kewanee Oil* court predicted that inventors would not resort to trade secret when offered a presumptively stronger protection by patent law:

> The possibility that an inventor who believes his invention meets the standards of patentability will sit back, rely on trade secret law, and after one year of use forfeit any right to patent protection, 35 U.S.C. § 102(b), is remote indeed.[30]

Trade secret is an adequate form of protection for innovators that are concerned with the limits of what may be patentable. The secrecy requirement of trade secret inherently provides protection that may potentially outlive any patent rights, provided a third party does not independently acquire the secret. This coincides with an interesting aspect of machine learning and big data—the need for massive amounts of data. Developers need data to "train" the algorithm, and increase the accuracy of the machine learning models. Companies that have already acquired massive amounts of data may opt to keep their data secret, treating the aggregated data as a trade secret.

In addition to the amount of amassed data, companies have all the more reason to keep their data secret if they have access to meaningful, normalized data. Even if a company amasses an enormous amount of data, the data sets may not be compatible with each other. Data gathered from one source may have different reference points or methodologies that are not immediately compatible with data from another source. This raises the concern of "cleaning" massive amounts of data.[31] Such concerns of data compatibility mean that parties with access to a single, homogenous source of high quality data enjoy a significant advantage over parties that need to pull data from multiple sources.

---

[28] *See* Dionne v. Se. Foam Converting & Packaging, Inc., 240 Va. 297 (1990).

[29] Kewanee Oil v. Bicron Corp., 416 U.S. 470 (1974).

[30] *Id.* at 490.

[31] Nikolay Golova & Lars Rönnbäck, *Big Data Normalization For Massively Parallel Processing Databases*, 54 COMPUTER STANDARDS & INTERFACES 86, 87 (2017).

However, data secrecy may not be a suitable strategy for companies that are aiming for cross-industry data aggregation. Institutions such as Global Alliance for Genomics and Health are promoting data sharing between research participants. The Chinese e-commerce giant Alibaba announced a data sharing alliance with companies such as Louis Vuitton and Samsung to fight off counterfeit goods.[32] To facilitate the development of technology and to mitigate risks, various companies and research institutions across diverse fields are engaging in joint development efforts and alliances. Seeking protection under trade secret runs against this trend of engaging in effective cross-industry collaboration. Yet there are countervailing arguments that trade secret promotes disclosure by providing legal remedies that can replace the protection of secrets.[33] Parties can sidestep the limitations of trade secrets by sharing proprietary information under the protection of contract law. While data sharing practices may void trade secret protection, the nature of continued accumulation of data and carefully drafted contractual provisions may provide sufficient protection for the data owners.

II
PLACING MACHINE LEARNING WITHIN INTELLECTUAL PROPERTY LAW

> "*Learning is any process by which a system improves performance from experience.*"
>
> – Herbert Simon, Nobel Prize in Economics 1978.

The concept of machine learning relates to computer programs that have the capability to improve performance based on experience, with limited intervention of the programmer.[34] Machine learning models have the capability to automatically adapt and customize for individual users, discover new patterns and correlations from large databases, and automate tasks that require some intelligence by mimicking human intuition.[35] This section dissects the mechanics of machine learning to identify the aspects of machine learning innovations that are at issue as intellectual property.

---

[32] Jon Russell, *Alibaba Teams Up with Samsung, Louis Vuitton and Other Brands to Fight Counterfeit Goods*, TECHCRUNCH (Jan. 16, 2017) https://techcrunch.com/2017/01/16/alibaba-big-data-anti-counterfeiting-alliance.

[33] Lemley, *supra* note 15, at 33

[34] *See* Lior Rokach, *Introduction to Machine Learning*, SLIDESHARE 3 (July 30, 2012), https://www.slideshare.net/liorrokach/introduction-to-machine-learning-13809045.

[35] *Id.* at 4.

A. Machine Learning Basics

Machine learning methods are divided into two different approaches—supervised machine learning and unsupervised machine learning. For supervised machine learning, models are typically established by applying "labeled" sets of data to a learning algorithm. Labeled data refers to data sets that have both relevant features and the target results that the programmer is interested in. For example, we may be interested in developing a machine learning model that classifies images with dogs in them. The data sets for supervised machine learning would indicate whether a given images has dogs or not. The learning process begins with the algorithm fitting trends found in the training data set into different types of models. The algorithm compares the prediction errors of the models by inputting the validation set data into each model, measuring their accuracy. This allows the algorithm to decide which of the various models is best suited as the resulting machine learning model. Finally, the machine learning model is then evaluated by assessing the accuracy of the predictive power of the model. The developed model is then applied to data without a correct answer to test the validity of the model. In unsupervised machine learning, the data sets are "unlabeled" data, which may not contain the result that the programmer is interested in. Returning to our dog image classification example, data sets for unsupervised machine learning will have pictures of various animals that are not labeled—the computer does not know which pictures are associated with dogs. The unsupervised machine learning algorithm develops a model that extracts common elements from the picture, teaching itself the set of features that makes the subject of the picture a dog. In essence, unsupervised machine learning uses data sets that do not have specific labels fed into the algorithm for the purpose of identifying common trends embedded in that data set.

The objective of developing such machine learning models varies. Sometimes the goal is to develop a prediction model that can forecast a variable from a data set. Classification, which assigns records to a predefined group, is also a key application of the algorithm. Clustering refers to splitting records into distinct groups based on the similarity within such group. Association learning identifies the relationship between features.

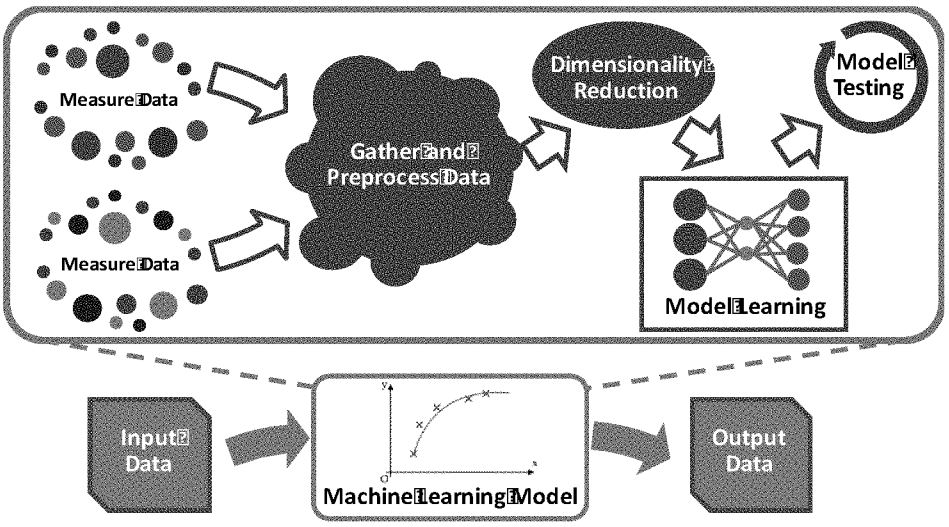

Figure 1. Overview of Machine Learning Model Development

Figure 1 illustrates the overall process of machine learning model development. The learning process of machine learning algorithms begins with aggregation of data. The data originates from an array of diverse sources ranging from user input, sensor measurement, or monitoring of user behavior.[36] The data sets are then preprocessed. The quality of data presents a challenge in improving machine learning models—any data that has been manually entered contains the possibility of error and bias.[37] Even if the data is collected through automatic means, such as health monitoring systems or direct tracking of user actions, the data sets require preprocessing to account for systematic errors associated with the recording device or method.[38] This includes data skews due to difference between individual sensors, errors in the recording or transmission of data, and incorrect metadata about the sensor.[39] Simply put, the data sets may have differing reference points, embedded biases, or differing formats. The "cleaning" process accommodates for the data skews.

---

[36] *Id.* at 10.

[37] *See* Lars Marius Garshol, *Introduction to Machine Learning,* SLIDESHARE 26 (May 15, 2012) https://www.slideshare.net/larsga/introduction-to-big-datamachine-learning.

[38] *Id.*

[39] *Id.*

The objective of machine learning models is to identify and quantify "features" from a given data set. The term "feature" refers to individually measurable property of an observed variable.[40] From the outset, there may be an extensive list of features that are present in a set of data. It would be computationally expensive to define and quantify each feature, and then to identify the inter-feature relationships, from massive amounts of data. Due to the high demand for the computational power required for processing massive amounts of data, dedication of computational resources to features that are outside the scope of the designer's interest would be a waste of such limited computational capacity.[41] The machine learning algorithm reduces waste of computational resources by applying dimensionality reduction to the pre-processed data sets.[42] The algorithm can identify an optimal subset of features by reducing the dimension and the noise of the data sets.[43] Dimensionality reduction allows the machine learning model to achieve higher level of predictive accuracy, increased speed of learning, and improves the simplicity and comprehensibility of the results.[44] However, the reduction process has limitations—reducing dimensionality inevitably imposes a limit on the amount of insights and information that may be extracted from the data sets. If the machine learning algorithm discerns a certain feature, the model would not be able to draw inferences related to said feature.

Following dimensionality reduction, the machine learning algorithm attempts to fit the data sets into preset models. Typically, three different types of data are fed into the machine learning model—training set, validation set, and test set.[45] The machine learning algorithm "trains" the model by fitting the training set data into various models to evaluate the accuracy of each selection. Then the

---

[40] *See* Lei Yu et al., *Dimensionality Reduction for Data Mining – Techniques, Applications and Trends*, BINGHAMTON UNIVERSITY COMPUTER SCIENCE 11, http://www.cs.binghamton.edu/~lyu/SDM07/DR-SDM07.pdf (last visited Feb. 23, 2018).

[41] *Id.*

[42] *See* Rokach, *supra* note 34, at 10.

[43] Yu et al., *supra* note 40.

[44] Laurens van der Maaten et al., *Dimensionality Reduction: A Comparative Review*, TILBURG CENTRE FOR CREATIVE COMPUTING, TiCC TR 2009-005, Oct. 26, 2009, at 1 ("In order to handle such real-world data adequately, its dimensionality needs to be reduced. Dimensionality reduction is the transformation of high-dimensional data into a meaningful representation of reduced dimensionality. Ideally, the reduced representation should have a dimensionality that corresponds to the intrinsic dimensionality of the data. The intrinsic dimensionality of data is the minimum number of parameters needed to account for the observed properties of the data").

[45] Andrew Ng, *Nuts and Bolts of Applying Deep Learning (Andrew Ng)*, YOUTUBE (Sept. 27, 2016), https://www.youtube.com/watch?v=F1ka6a13S9I.

validation set is used to estimate error rates of each model when applied to data outside the training set that was used to develop each model. Through this process, the machine learning algorithm selects the model that best describes the characteristics and trends of the target features from the test and validation sets.[46] The test set is then used to calculate the generalized prediction error, which is reported to the end user for proper assessment of the predictive power of the model.[47] Simply put, the training test and validation set is used to develop and select a model that reflects the trends of the given data set, and the test set is used to generate a report on the accuracy of the selected model.

The crucial elements in developing a machine learning model are (1) training data, (2) inventions related to the machine learning algorithm such as the method of preprocessing the training data, the method of dimensional reduction, feature extraction, and the method of model learning/testing, and (3) the machine learning model and output data.[48] An ancillary element associated with the three elements above is the human talent that is required to implement such innovation.[49] Innovators in the field of machine learning may protect their investments by protecting one or more of the elements listed above.

The difference between training data and output data, as well as the difference between the machine learning algorithm and the machine learning model, are best illustrated with an example. Let us assume a credit card company wants to use machine learning to determine whether the company should grant a premium credit card to a customer. Let us further assume that the company would prefer to grant this card to customers that would be profitable to the company while filtering out applicants that are likely to file for bankruptcy. Data sets about prior applicant information would correspond to *training data*. The company would apply a mathematical method of extracting insight about the correlation between features and the criteria that the company wants to evaluate (e.g., profitable for the firm or likely to file bankruptcy). The mathematical methods are referred as *machine learning algorithms*. The resulting mechanism, such as a scoring system, that determines the eligibility of card membership is the *machine*

---

[46] Andrew Ng, *Model Selection and Train/Validation/Test Sets*, MACHINE LEARNING, https://www.coursera.org/learn/machine-learning/lecture/QGKbr/model-selection-and-train-validation-test-sets (last visited Feb. 23, 2018).

[47] *Id.*

[48] *See* Rokach, *supra* note 34, at 10.

[49] Alex Rampell & Vijay Pande, *a16z Podcast: Data Network Effects*, ANDREESEN HOROWITZ (Mar. 8, 2016), http://a16z.com/2016/03/08/data-network-effects/.

learning model. The credit card applicant's personal data would be the *input data* for the machine learning model, and the *output data* would include information such as expected profitability of this applicant and likelihood of bankruptcy for this applicant.

B. Industry Trends in Machine Learning

Discussing incentive structures and trends behind the machine learning industry is essential in identifying adequate methods of intellectual property rights. The current trends in the world of machine learning will predict what intellectual property regime is most useful to companies to protect their work.

The United States has chronically struggled to maintain adequate supply of talent in the high-tech industry, a deficit of talent that continues in the field of machine learning.[50] From a report by the McKinsey Global Institute, the United States' demand for talent in deep learning "could be 50 to 60 percent greater than its projected supply by 2018."[51] Coupled with the dearth of machine learning specialists, the short employment tenure of software companies further complicates the search for talent. Software engineers from companies such as Amazon and Google have reported an average employment tenure of one year.[52] While some parts of the high attrition rate may be attributed to cultural aspects of the so-called "Gen Y" employees, the "hot" demand for programming talent has significant impact on the short employee tenure.[53] Job mobility within the software industry is likely to increase as the "talent war" for data scientist intensifies. Employee mobility and California's prohibition against "covenants not to compete" have been accredited as a key factor behind the success of Silicon Valley.[54] Another trend in the field is the rapid advances in machine learning methods. Due to the

---

[50] James Manyika et. al., *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, MCKINSEY GLOBAL INST., May 2011, at 11, *available at* https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_exec_summary.ashx.

[51] *Id.*

[52] Leonid Bershidsky, *Why Are Google Employees So Disloyal?*, BLOOMBERG (July 13, 2013, 11:41 AM), https://www.bloomberg.com/view/articles/2013-07-29/why-are-google-employees-so-disloyal-.

[53] *Id.*

[54] Rob Valletta, *On the Move: California Employment Law and High-Tech Development*, FEDERAL RESERVE BANK OF S.F. (Aug. 16, 2002), http://www.frbsf.org/economic-research/publications/economic-letter/2002/august/on-the-move-california-employment-law-and-high-tech-development/#subhead1.

fast-paced development of the field, data scientists and practitioners have every reason to work with companies that would allow them to work at the cutting edge of machine learning, using the best data sets. This may influence the attrition rates and recruiting practices of the software industry mentioned above.[55] Eagerness of employees to publish scientific articles and contribute to the general machine learning committee may be another factor of concern.

To accelerate innovation by repurposing big data for uses different from the original purpose, and to form common standards for machine learning, more industries are joining alliances and collaborations.[56] Cross-industry collaborations may enable endless possibilities. Imagine the inferences that may be drawn by applying machine learning methods to dietary data from home appliances, biometric data, and data on the weather patterns around the user. Putting privacy nightmares aside, machine learning with diverse data sets may unlock applications that were not previously possible. More companies are attempting to capitalize on commercial possibilities that data sharing may unlock.[57]

C. Machine Learning Innovators—Protect the Data or Inventions?

Though it may seem intuitive that patent protection may be the best option, innovations in machine learning may not *need* patent protection. Trade secret protection on the data sets may be sufficient to protect the interests of practicing entities while avoiding disclosure of their inventions during the patent prosecution process. Furthermore, numerous software patents have been challenged as unpatentable abstract subject matter under 35 U.S.C. §101 since the *Alice* decision in 2014.[58] Though subsequent decisions provided guidelines for types of software patents that would survive the *Alice* decision, it is not clear how the judiciary will view future machine learning patents. Such issues raise the question about the patentability of machine learning—should we, and can we, resort to patents to protect machine learning inventions?

Following the discussion on the building blocks of machine learning and recent emerging trends in the field, this section discusses the mode and scope of

---

[55] *Id.*

[56] *See* Quentin Hardy, *IBM, G.E. and Others Create Big Data Alliance*, N.Y. TIMES (Feb. 15, 2015), https://bits.blogs.nytimes.com/2015/02/17/ibm-g-e-and-others-create-big-data-alliance.

[57] *See, e.g.*, *Finicity and Wells Fargo Ink Data Exchange Deal*, WELLS FARGO (Apr. 4, 2017), https://newsroom.wf.com/press-release/innovation-and-technology/finicity-and-wells-fargo-ink-data-exchange-deal.

[58] Alice Corp. Pty. Ltd. v. CLS Bank Int'l, 134 S. Ct. 2347 (2014).

protection that current legal system provides for each element pertinent to innovation in machine learning. The possible options for protecting innovations are (1) non-disclosure agreements and trade secret law, (2) patent law, and (3) copyright. The three options for protection may be applied to the three primary areas of innovation—(1) training data, (2) inventions related to computation, data processing, and machine learning algorithms, and (3) machine learning models and output data. This discussion will provide context about the methods of protection for innovations in machine learning by examining the costs and benefits of the various approaches.

1. Protecting the Training Data—Secrecy Works Best

Access to massive amounts of training data is a prime asset for companies in the realm of machine learning. The big data phenomenon, which triggered the surge of interest in machine learning, is predicated on the need for practices to analyze large data resources and the potential advantages from such analysis.[59] Lack of access to a critical mass of training data prevents innovators from making effective use of machine learning algorithms.

Previous studies suggest that companies resent sharing data with each other. [60] Michael Mattioli discusses the hurdles against sharing data and considerations involved with reuse of data in his article *Disclosing Big Data.*[61] Indeed, there may be practical issues that prevent *recipients* of data from engaging in data sharing. Technical challenges in comparing data from different sources, or inherent biases embedded in data sets may be reasons that complicate receiving outside data.[62] Mattioli also questions the adequacy of the current patent and copyright system to promote data sharing and data reuse—information *providers*

---

[59] Karen E.C. Levy, *Relational Big Data*, 66 STAN. L. REV. ONLINE 73, 73 n.3 (2013), https://review.law.stanford.edu/wp-content/uploads/sites/3/2013/09/66_StanLRevOnline_73_Levy.pdf (explaining that the big data phenomenon is due to the need of practices to analyze data resources).

[60] Christine L. Borgman, *The Conundrum of Sharing Research Data*, 63 J. AM. SOC'Y FOR INFO. SCI. & TECH. 1059, 1059-60 (2012) (discussing the lack of data sharing across various industries).

[61] *See* Michael Mattioli, *Disclosing Big Data*, 99 MINN. L. REV. 535 (2014).

[62] *See id.* at 545-46 (discussing the technical challenges in merging data from different sources, and issue of subjective judgments that may be infused in the data sets).

may prefer not to disclose any parts of their data due to the rather thin legal protection for databases.[63]

Perhaps this is why secrecy seems to be the primary method of protecting data.[64] The difficulty of reverse engineering to uncover the underlying data sets promotes the reliance on non-disclosure.[65] Compared to the affirmative steps required to maintain trade secret protection if the data is disclosed, complete non-disclosure may be a cost effective method of protecting data.[66] Companies that must share data with external entities may exhibit higher reliance on contract law rather than trade secret law. In absence of contract provisions, it would be a challenge to prove that the trade secret has been acquired by misappropriation of the recipient party.

The "talent war" for data scientists may also motivate companies to keep the training data sets secret. With a shortage of talent to implement machine learning practices and rapid developments in the field, retaining talent is another motivation for protecting against unrestricted access to massive amounts of data. Companies may prefer exclusivity to the data sets that programmers can work with—top talents in machine learning are lured to companies with promises of exclusive opportunities to work with massive amounts of data.[67] The rapid pace of development in this field encourages practitioners to seek opportunities that provide the best resources to develop their skill sets. This approach is effective since a key limitation against exploring new techniques in this field is the lack of access to high quality big data. Overall, secrecy over training data fits well with corporate recruiting strategies to retain the best talents in machine learning.

Non-disclosure and trade secret protection seems to be the best mode of protection. First, despite the additional legal requirements necessary to qualify as trade secrets, trade secret protection fits very well with non-disclosure strategy. On

---

[63] *See id.* at 552 (discussing how institutions with industrial secrets may rely on secrecy to protect the big data they have accumulated).

[64] *See id.* at 570 ("[T]he fact that these practices are not self-disclosing (i.e., they cannot be easily reverse-engineered) lends them well to trade secret status, or to mere nondisclosure").

[65] *Id.*

[66] *Id.* at 552.

[67] Patrick Clark, *The World's Top Economists Want to Work for Amazon and Facebook*, BLOOMBERG (June 13, 2016, 10:47 AM), https://www.bloombergquint.com/technology/2016/06/09/the-world-s-top-economists-want-to-work-for-amazon-and-facebook ("If you want to be aware of what interesting questions are out there, you almost have to go and work for one of these companies").

the other hand, patent law is at odds with the principle of non-disclosure. While trade secret law provides companies protection without disclosing information, patent law requires disclosure in exchange for monopolistic rights. Furthermore, neither patent nor copyright provide adequate protection for underlying data. Patent law rewards creative concepts and inventions, not compiled facts themselves. Copyright may protect labeling or distinct ways of compiling information, but does not protect underlying facts. Also, as a practical matter, the difficulty of reverse engineering of machine learning models does not lend well to detecting infringement. Analysis of whether two parties used identical training data would not only be time consuming and costly, but may be fundamentally impossible.

If companies were to seek protection of training data, it would be best to opt for secrecy by non-disclosure. This would mean companies would opt out of the cross-industry collaborations that were illustrated above. This may be less of a concern for innovation, as companies may still exchange output data as means of facilitating cross-industry collaboration.

2. Protecting the Inventions—Patent Rights Prevail

Adequate protection over inventive approaches in processing data is becoming increasingly important as various industries begin to adopt a collaborative alliance approach in machine learning. Cross-industry collaboration requires implementation of methods such as preprocessing diverse data sets for compatibility. As the sheer amount of data increases, more processing power is required. The machine learning algorithm needs to maintain a high degree of dimensionality to accurately identify the correlations between a high number of relevant features. The need for more innovative ideas to address such technological roadblocks will only intensify as we seek more complex applications for machine learning.

The three primary areas where novel ideas would facilitate innovations in machine learning are pre-training data processing, dimensional reduction, and the machine learning algorithm.

Access to massive amounts of data alone is not sufficient to sustain innovation in machine learning. The raw data sets may not be compatible with each other, requiring additional "cleaning" of data prior to machine learning

training.[68] The data provided to the machine learning algorithm dictates the result of the machine learning model, hence innovations in methods to merge data with diverse formats is essential to enhancing the accuracy of the models. As cross-industry data analysis becomes more prominent, methods of merging data will have more significant impact on advancing the field of machine learning than mere collection of large data sets. Cross-industry data sharing would be useless unless such data sets are merged in a comparable manner.[69]

Companies can opt to protect their inventive methods by resorting to trade secret law. The difficulty of reverse engineering machine learning inventions, coupled with the difficulty of patenting software methods provides incentives for innovators to keep such inventions secret from the public. However, two factors would render reliance on non-disclosure and trade secret ineffective—frequent turnover of software engineers and rapid speed of development in the field.

Rapid dissemination of information from employment mobility may endanger intellectual property protection based on secrecy. Furthermore, while the law will not protect former employees that reveal trade secrets to their new employers, the aforementioned fluid job market coupled with general dissemination of information make it difficult to distinguish between trade secrets from former employment and general knowledge learned through practice. The difficulties of reverse engineering machine learning models work against the trade secret owner as well in identifying trade secret misappropriation—how do you know others are using your secret invention? The desire for software communities to discuss and share recent developments in the field does not align well with the use of secrecy against innovations in machine learning. Secrecy practices disincentivize young data scientists from joining due to the limits against gaining recognition.[70]

The rapid development of machine learning technology also presents challenges against reliance on trade secret law. Secret methods may be

---

[68] BILL FRANKS, TAMING THE BIG DATA TIDAL WAVE 20 (2012) (discussing that the biggest challenge in big data may not be developing tools for data analysis, but rather the processes involved with preparing the data for the analysis).

[69] *See* Borgman, *supra* note 60, at 1070 ("Indeed, the greatest advantages of data sharing may be in the combination of data from multiple sources, compared or "mashed up' in innovative ways." (citing Declan Butler, *Mashups Mix Data Into Global Service*, 439 NATURE 6 (2006))).

[70] Jack Clark, *Apple's Deep Learning Curve*, BLOOMBERG BUSINESSWEEK, (Oct 29, 2015) https://www.bloomberg.com/news/articles/2015-10-29/apple-s-secrecy-hurts-its-ai-software-development.

independently developed by other parties. Neither trade secret law nor non-disclosure agreements protect against independent development of the same underlying invention.[71] Unlike training data, machine learning models, or the output data, there are no practical limitations that impedes competitors from independently inventing new computational methods of machine learning algorithms.

With such a fluid employment market, high degree of dissemination of expertise, and rapid pace of development, patent protection may provide the assurance of intellectual property protection for companies developing inventive methods in machine learning. Discussions on overcoming the barriers of patenting software will be presented in later sections.[72]

3. Protecting the Machine Learning Models and Results—Secrecy Again

The two primary products from applying the machine learning algorithms to the training data are the machine learning model and the accumulation of results produced by inputting data into the machine learning model. The "input data" in this context may refer to individual data that is analyzed by the insights gained from the machine learning model.

In a recent article, Brenda Simon and Ted Sichelman discuss the concerns of granting patent protection for "data-generating patents," which refers to inventions that generate valuable information in their operation or use.[73] Exclusivity based on patent protection may be extended further by trade secret protection over the data that has been generated by the patented invention.[74] Simon and Sichelman argue that the extended monopoly over data may potentially overcompensate inventors since the "additional protection was not contemplated by the patent system[.]"[75] Such expansive rights will cause excessive negative impact on downstream innovation and impose exorbitant deadweight losses.[76] The added protection over the resulting data derails the policy rationale behind the quid pro quo exchange

---

[71] Kewanee Oil v. Bicron Corp., 416 U.S. 470, 490 (1974).

[72] *See infra* Section III-B.

[73] Brenda Simon & Ted Sichelman, *Data-Generating Patents*, 111 NW. U.L. REV. 377 (2017).

[74] *Id.* at 379.

[75] *Id.* at 414.

[76] *Id.* at 415 ("[B]roader rights have substantial downsides, including hindering potential downstream invention and consumer deadweight losses . . .").

between the patent holder and the public by excluding the patented information from public domain beyond the patent expiration date.[77]

The concerns addressed in data-generating patents also apply to machine learning models and output data. Corporations may obtain patent protection over the machine learning models. Akin to a preference for secrecy for training data, non-disclosure would be the preferred mode of protection for the output data. The combined effect of the two may lead to data network effects where users have strong incentives to continue the use of a given service.[78] The companies that have exclusive rights over the machine learning model and output data gather more training data, increasing the accuracy of their machine learning products. The reinforcement by monopoly over the means of generating data allows few companies to have disproportionately strong dominance over their competitors.[79]

Market dominance by data-generating patents becomes particularly disturbing when the patent on a machine learning model preempts other methods in the application of interest. Trade secret law does not provide protection against independent development. However, if there is only one specific method to obtain the best output data, no other party would be able to create the output data independently. The exclusive rights over the only methods of producing data provides means for the patent holder to monopolize both the patent and the output data.[80] From a policy perspective, the excessive protection does seem troubling. Yet such draconian combinations are less feasible after the recent rulings on patentable subject matter of software, which will be discussed below. [81] Mathematical equations or concepts are likely directed to an "abstract concept," thus will be deemed directed to a patent ineligible subject matter.[82] Furthermore, though recent cases in the Federal Circuit have found precedents where software patents passed the patentable subject matter requirement, those cases expressed limitations against granting patents that would improperly preempt all solutions to a particular problem. [83] The rapid pace of innovation in the field of machine

---

[77] *Id.* at 417.

[78] Rampell & Pande, *supra* note 49.

[79] Lina Kahn, *Amazon's Antitrust Paradox*, 126 YALE L.J. 710, 785 (2017) ("Amazon's user reviews, for example, serve as a form of network effect: the more users that have purchased and reviewed items on the platform, the more useful information other users can glean from the site").

[80] Simon & Sichelman, *supra* note 73, at 410.

[81] *See infra* Section III-A.

[82] *Id.*

[83] *See infra* Section III-B.

learning compared to the rather lengthy period required to obtain patents may also dissuade companies from seeking patents. Overall, companies have compelling incentives to rely on non-disclosure and trade secrets to protect their machine learning models instead of seeking patents.

The secrecy concerns regarding training data applies to machine learning models and the output data as well. Non-disclosure would be the preferred route of obtaining protection over the two categories. However, use of non-disclosure or trade secrets to protect machine learning models and output data presents challenges that are not present in the protection of training data. The use of secrecy to protect machine learning models or output data conflicts with recruiting strategies to hire and retain top talent in the machine learning field. The non-disclosure agreements limit the employee's opportunity to gain recognition in the greater machine learning community. In a rapidly developing field where companies are having difficulty hiring talent, potential employees would not look fondly on corporate practices that limit avenues of building a reputation within the industry.[84]

Companies have additional incentives to employ a rather lenient secrecy policy for machine learning models and the output data. They have incentives to try to build coalitions with other companies to monetize on the results. Such cross-industry collaboration may be additional source of income for those companies. The data and know-how that Twitter has about fraudulent accounts within their network may aid financial institutions such as Chase with novel means of preventing wire fraud. The reuse of insights harvested from the large amount of raw training data can become a core product the companies would want to commercialize. Data reuse may have an incredible impact even for applications ancillary to the primary business of the company.

Interesting aspects of disclosing machine learning models and output data are the difficulty of reverse engineering and consistent updates. If the company already has sufficient protection over the training data and/or the computational innovations, competitors will not be able to reverse engineer the machine learning model from the output data. Even with the machine learning model, competitors will not be able to provide updates or refinements to the model without the computational techniques and the sufficient data for training the machine learning

---

[84] Jack Clark, *Apple's Deep Learning Curve*, BLOOMBERG BUSINESSWEEK (Oct 29, 2015), https://www.bloomberg.com/news/articles/2015-10-29/apple-s-secrecy-hurts-its-ai-software-development.

algorithm. In certain cases, the result data becomes training data for different applications, which raises concerns of competitors using the result data to compete with the innovator. Yet the output data would contain less features and insights compared to the raw training data that the innovator possesses, and therefore would inherently be at a disadvantage when competing in fields that the innovator has already amassed sufficient training data.

Grant of patents on machine learning models may incentivize companies to build an excessive data network while preempting competitors from entering competition. This may not be feasible in the future, as technological preemption is becoming a factor of consideration in the patentable subject matter doctrine. Companies may use secrecy as an alternative, yet may have less incentives to keep secrecy compared to the protection of training data.

D. Need of Patent Rights for Machine Learning Inventions in the Era of Big Data

The current system, on its surface, does not provide adequate encouragement for data sharing. If anything, companies have strong incentives to avoid disclosure of their training data, machine learning model, and output data.

Despite these concerns, data reuse may enable social impacts and advances that would not be otherwise possible. Previous studies have pointed out that one of the major barriers preventing advances in machine learning is the lack of data sharing between institutions and industries.[85] Data scientists have demonstrated that they were able to predict flu trends with data extracted from Twitter.[86] Foursquare's location database provides Uber with the requisite data to pinpoint the location of users based on venue names instead of addresses.[87] Information about fraudulent Twitter accounts may enable early detection of financial frauds.[88] The possibilities that cross-industry data sharing may bring are endless.

---

[85] Peer, *supra* note 17 ("The idea that the data will be used by unspecified people, in unspecified ways, at unspecified times . . . is thought to have broad benefits").

[86] *See* Harshavardhan Achrekar et al., *Predicting Flu Trends using Twitter data*, IEEE CONFERENCE ON COMPUT. COMMC'NS. WORKSHOPS 713 (2011), http://cse.unl.edu/~byrav/INFOCOM2011/workshops/papers/p713-achrekar.pdf.

[87] Jordan Crook, *Uber Taps Foursquare's Places Data So You Never Have to Type an Address Again*, TECHCRUNCH, (May 25, 2016) https://techcrunch.com/2016/05/25/uber-taps-foursquares-places-data-so-you-never-have-to-type-an-address-again/.

[88] *See* Rampell & Pande, *supra* note 49.

To encourage free sharing of data, companies should have a reliable method of protecting their investments in machine learning. At the same time, protection based on non-disclosure of data would defeat of purpose of promoting data sharing. Hence protection over computation methods involved with machine learning maintains the delicate balance between promoting data sharing and protecting innovation.

Protection over inventions in the machine learning algorithm provides one additional merit other than allowing data sharing and avoiding the sort of excessive protection that leads to a competitor-free road and data network effects. It incentivizes innovators to focus on the core technological blocks to the advancement of technology, and encourages disclosure of such know-how to the machine learning community.

Then what are the key obstacles in obtaining patents in machine learning inventions? While there are arguments that the definiteness requirement of patent law is the primary hurdle against patent protection of machine learning models due to reliance on subjective judgment, there is no evidence that the underlying *inventions* driving big data faces the same challenge.[89] Definiteness may be overcome by providing reasonable certainty for those skilled in the art of defining what the scope of the invention is at the time of filing.[90] There is no inherent reason why specific solutions for data cleaning, enhancement of computation efficiency, and similar inventions would be deemed indefinite by nature.

Since the United States Supreme Court invalidated a patent on computer implemented financial transaction methods in the 2014 *Alice* decision, the validity of numerous software and business method patents were challenged under 35 U.S.C. §101.[91] As of June 8th, 2016, federal district courts invalidated 163 of the 247 patents that were considered under patentable subject matter—striking down 66% of challenged patents.[92] The U.S. Court of Appeals for the Federal Circuit invalidated 38 of the 40 cases it heard.[93]

---

[89] *See* Mattioli, *supra* note 61, at 554 ("A final limitation on patentability possibly relevant to big data is patent law's requirement of definiteness").

[90] *See* Nautilus, Inc. v. Biosig Instruments, Inc., 134 S. Ct. 2120 (2014).

[91] *See* Alice Corp. Pty. Ltd. v. CLS Bank Int'l, 134 S. Ct. 2347 (2014).

[92] Robert R. Sachs, *Two Years After Alice: A Survey of the Impact of a "Minor Case" (Part 1)*, BILSKI BLOG (June 16, 2016), http://www.bilskiblog.com/blog/2016/06/two-years-after-alice-a-survey-of-the-impact-of-a-minor-case.html.

[93] *Id.*

Arguably, the public benefits more from such high rates of post-issuance invalidity. The public still has access to the disclosures from the patents and patent applications. In reliance on granted patents, companies may have already invested in growing related businesses, catering to the need of consumers. At the same time, the patent holder's monopolistic rights have been shortened as the result of litigation. Effectively, the price that the public pays to inventors in exchange for the benefits of disclosure is reduced.

Yet the high degree of invalidity raises several concerns for the software industry. Smaller entities, lacking market influence and capital, have difficulty competing against established corporations without the monopolistic rights granted through the patent system. Investors become hesitant to infuse capital into startups for fear that invalidity decreases the worth of patents. Reliance on trade secret has its own limitations due to the disclosure dilemma—the inventor needs to disclose the secret to lure inventors, but risks losing secrecy in the process. Copyright law does not provide appropriate protection. The restrictions imposed by the merger doctrine and scènes à faire doctrine constrain copyright protection of software. Though copyright provides an alternative method of protecting literal copying of code, it does little to protect the underlying software algorithms and innovation.

Ultimately, the increase of alliances and collaboration provides incentives for parties to obtain patent rights. Reliance on trade secret or copyright are not suitable methods of protecting their intellectual property. Furthermore, market power or network effects alone cannot sufficiently mitigate the risks involved with operating a business. Patents become even more important for startups since patents provide investors with assurance that in the worst case, the patents may still serve as potential collateral.

III
PATENTABILITY OF MACHINE LEARNING INNOVATIONS IN THE ERA OF BIG DATA

Patentable subject matter continues to be a barrier for patenting innovations in software. Additional doctrines such as enablement, written description, and obviousness are also serious obstacles against obtaining patents, yet such requirements are specific to each claimed invention and the draftsmanship of claims. Subject matter is considered a broader, categorical exclusion of patent rights. This section explores the current landscape of the patentable subject matter doctrine in the software context.

A. Alice: The Legal Framework of Patentable Subject Matter in Software

The complexity involved with software, coupled with the relatively broad scope of software patents, has presented challenges in identifying the boundaries of the claims. [94] Many members of the software community detest imposing restrictions on open source material and attest that many key innovations in algorithms are rather abstract.[95] Such hostility against patenting software has raised the question of whether patent rights should be the proper method of protecting innovations in software.

Alice was a case that embodied such opposition to the grant of software patents. The case involved patents on computerized methods for financial trading systems that reduce "settlement risk" when only one party to financial exchange agreement satisfies its obligation.[96] The method proposed the use of a computer system as a third-party intermediary to facilitate the financial obligations between parties.[97] The United States Supreme Court ruled that the two-step test established from *Mayo* governed all patentable subject matter questions.[98] In particular, for the abstract idea context, the Supreme Court established the following two-step framework for patentable subject matter of software inventions:

1. Step one: "[D]etermine whether the claims at issue are *directed to* a patent-ineligible concept. If so, the Court then asks whether the claim's [additional] elements, considered both individually and 'as an ordered combination,' 'transform the nature of the claim' into a patent-eligible *application*."[99]

2. Step two: "[E]xamine the elements of the claim to determine whether it contains an '*inventive concept*' sufficient to 'transform' the claimed abstract idea into a patent-eligible application. A claim that recites an abstract idea must include 'additional features' to ensure that the [claim] is more than a drafting effort designed to monopolized the

---

[94] Stephanie E. Toyos, *Alice in Wonderland: Are Patent Trolls Mortally Wounded by Section 101 Uncertainty*, 17 LOY. J. PUB. INT. L. 97,100 (2015).

[95] *Id.*

[96] Alice Corp. Pty. Ltd. v. CLS Bank Int'l, 134 S. Ct. 2347, 2349 (2014).

[97] *Id.*

[98] *Id.*

[99] *Id.* at 2350 (emphasis added) (citation omitted).

> [abstract idea]" which requires "more than simply stat[ing] the [abstract idea] while adding the words 'apply it.'"[100]

The *Alice* Court found that the patent on financial transaction was "directed to a patent-ineligible concept: the abstract idea of intermediated settlement," and therefore failed step one.[101] Furthermore, the Court ruled that the claims did "no more than simply instruct the practitioner to implement the abstract idea of intermediated settlement on a generic computer" and did not provide an inventive concept that was sufficient to pass step two.[102]

B. The post-Alice cases from the Federal Circuit

The *Alice* framework was considered as a huge setback for the application of patentable subject matter doctrine to software. It was a broad, categorical exclusion of certain inventions that were deemed "directed to" an abstract idea, natural phenomenon, or law of nature. The biggest misfortune was the lack of guidance in the *Alice* decision on the threshold for such categorical exclusion—we were left without any suggestions on the type of software patents that would be deemed as patentable subject matter.

The recent line of cases in the Federal Circuit provides the software industry with the much-needed clarification on the standards that govern patentability of software inventions. *Enfish v. Microsoft*, decided on March 2016, involved a "model of data for a computer database explaining how the various elements of information are related to one another" for computer databases.[103] In June 2016, the Federal Circuit decided another case on the abstract idea category for patentable subject matter. *Bascom Global v. AT&T Mobility* is on a patent disclosing an internet content filtering system located on a remote internet service provider (ISP) server.[104] Shortly after *Bascom*, the Federal Circuit decided *McRO v. Bandai Namco Games* in September 2016.[105] The case ruled that an automated 3D

---

[100] *Id.* at 2357 (emphasis added) (alteration in original) (citation omitted).

[101] *Id.* at 2350.

[102] *Id.* at 2351.

[103] Enfish, LLC v. Microsoft Corp., 822 F.3d 1327, 1330 (Fed. Cir. 2016).

[104] Bascom Glob. Internet Servs. v. AT&T Mobility LLC, 827 F.3d 1341, 1349 (Fed. Cir. 2016).

[105] McRO, Inc. v. Bandai Namco Games Am. Inc., 837 F.3d 1299, 1308 (Fed. Cir. 2016).

animation algorithm that renders graphics in between two target facial expressions is patentable subject matter.[106]

The rulings from the Federal Circuit on the aforementioned three cases provide guidelines along the two-step *Alice* test of patentable subject matter. The software patents in *Enfish* and *McRO* were deemed "directed to" a patent eligible subject matter, informing the public of what may pass the first set of the *Alice* test. *Bascom* failed the first step.[107] Yet the court ruled that those patents had inventive concepts sufficient to transform a patent ineligible subject matter into a patent eligible application. Combined together, the three cases give more certainty in what may pass the 35 U.S.C. §101 patentable subject matter inquiry.

Reiterating the *Alice* test, whether an invention is a patentable subject matter is determined by a two-step process—(1) is the invention *directed to*, rather than an application of, an abstract idea, natural phenomenon, or law of nature, and even if so, (2) do the elements of the claim, both individually and combined, contain an *inventive concept* that transforms this invention into a patent-eligible application? The Federal Circuit fills in the gaps that were left unexplained from the *Alice* ruling.

1. The Federal Circuit's Standard for Alice Step One

The *Enfish* court discussed what constitutes an abstract idea at the first step of the *Alice* inquiry. Judge Hughes instructs us to look at whether the claims are directed to a specific improvement rather than an abstract idea. In this case, the patent provides the public with a solution to an existing problem by a specific, non-generic improvement to computer functionality. The *Enfish* court ruled that such invention is patent eligible subject matter.[108]

McRO also ruled that the facial graphic rendering for 3D animation was not an abstract concept. Here, the Federal Circuit again emphasized that a patent may pass step one of the *Alice* test if the claims of the patent "focus on a specific means or method that improves the relevant technology."[109] The *McRO* court also noted that preemption concerns may be an important factor for the 35 U.S.C. §101 subject matter inquiry—that improper monopolization of "the basic tools of

---

[106] *Id.*
[107] *Bascom*, 827 F.3d at 1349.
[108] *Enfish*, 822 F.3d at 1330.
[109] *McRO, Inc.*, 837 F.3d at 1314.

scientific and technological work" is a reason why such categorical carve outs against granting patents on abstract ideas exist.[110]

Bascom provides the standards on what would fail step one of the *Alice* patentable subject matter inquiry. If the patent covers a conventional, well-known method in the field of interest, then the invention would be considered abstract. This is akin to the inventive concept considerations conducted at the second phase of the 35 U.S.C. §101 subject matter inquiry.

The main takeaway from *Enfish* and *McRO* is that in the first step of the *Alice* test, a patent application is not directed to an abstract idea if (1) the invention addresses an existing problem by specific improvements rather than by conventional, well-known methods and (2) the claims do not raise preemption concerns. This encourages practitioners to define the problem as broadly as possible, while defining the scope of improvement in definite terms.

2. The Federal Circuit's Standard for Alice Step Two, and the Overlap with Step One

The second step of the *Alice* test is an inquiry of whether the patent application, which is directed to a patent ineligible subject, still contains a patent-worthy inventive concept. *Bascom* ruled in favor of granting the patent following the second step of the *Alice* test.[111] While the patent at hand was considered directed to patent ineligible subject matter, the *Bascom* court found that the content filter system invention still had an inventive concept worthy of a patent.[112] Even if elements of a claim are separately known in prior art, an inventive concept can be found in the non-conventional and non-generic arrangement of known, conventional pieces. This inquiry seems like a lenient standard compared to the 35 U.S.C. §103 obviousness inquiry; hence, it is not clear if this step has an independent utility for invalidating or rejecting a patent. Nonetheless, the court found that merely showing that all elements of a claim were already disclosed in prior art was not sufficient reason to make an invention patent ineligible.[113]

While it is possible to infer sufficient reasons of ruling out inventive concepts from the *Bascom* case, it is still unclear what would warrant an invention to pass the second step of the *Alice* test. Cases such as *DDR Holdings v.*

---

[110] *Id.*
[111] *Bascom*, 827 F.3d at 1349.
[112] *Id.*
[113] *Id.*

Hotels.com have suggested that the second step of *Alice* is satisfied since it involved a solution to a specific technological problem that "is necessarily rooted in computer technology in order to overcome a problem specifically arising in the realm of computer networks."[114]

This interpretation of inventive concept becomes perplexing when comparing the two steps of *Alice*—both steps look to whether the proposed solution addresses problems that are specific to a given field of interest. While we would need additional cases to gain insight on whether the two steps have truly distinct functions, at the very least the Federal Circuit provided essential guidelines on what may be deemed as patentable software.

C. Applying Patentable Subject Matter to Machine Learning Inventions

As the *Bascom* court has taught, the first step in the *Alice* inquiry is to ask whether an invention (1) provides a solution to an existing problem by (2) a specific, non-generic improvement that (3) does not preempt other methods of solving the existing problem. Applying this test to inventions in machine learning, *mathematical* improvements and *computational* improvements would be treated differently.

As mentioned before, a key aspect of machine learning is the "noise" associated with the data sets.[115] Another concern is the fitting of a given algorithm to a certain model. Methods that facilitate the computations of the training process may be deemed as a specific improvement. However, machine learning algorithms themselves, including the base models that the algorithm fits the training into would not be pertinent to just a specific improvement. Hence, generic mathematical methods applicable to various problems are directed to an abstract idea. For example, an invention that addresses the issue of normalizing data from different sources would be a computational issue and hence would pass the *Alice* test given that it did not preempt other solutions to the problem of data normalization. On the other hand, a specific mathematical equation that serves as a starting model for the machine learning algorithm would be mathematical and hence directed to an abstract idea. Even if the mathematical starting model is only good for a specific application, the model is not a specific improvement pertinent to that application. Although the model may not necessarily be a good starting

---

[114] *See* Toyos, *supra* note 94, at 121; *DDR Holdings, LLC v. Hotels.com*, 773 F.3d 1245, 1257 (Fed. Cir. 2014).

[115] *See supra* Section II-A.

model for other applications, it is nonetheless a generic solution that applies to other applications as well.

CONCLUSION

While highly restrictive, the guidelines from the Federal Circuit still allow the grant of patent rights for the computational aspects of machine learning algorithms. The guidelines also would prevent highly preemptive mathematical innovations, including data-generating patents such as machine learning models.

The narrow range of patentability makes a patent regime appealing for computational methods. The recent emphasis on preemption concerns acts in favor of preventing data network effects based on data-generating patents. While not discussed in this paper, other patentability requirements such as obviousness or definiteness would further constraint the grant of overly broad data-generating patents.

Such an approach strikes the appropriate balance between promoting innovation and encouraging data reuse for societal benefits. Compared to other approaches of providing protection over innovations in machine learning, the narrowly tailored approach for patent rights for computational inventions fits best with the policy goal of promoting innovation through data reuse. The industry trends in collaboration and recruiting also matches the proposed focus on patent law protection.