Suffolk University
Digital Collections @ Suffolk

Suffolk University Law School Faculty Works | Suffolk University Law School

1-1-2020

Fair Use and Machine Learning

Stephen M. McJohn
Suffolk University, smcjohn@suffolk.edu

Follow this and additional works at: https://dc.suffolk.edu/suls-faculty

Part of the Intellectual Property Law Commons

Recommended Citation
12 N.E.U.L.R. 99 (2020)

This Article is brought to you for free and open access by the Suffolk University Law School at Digital Collections @ Suffolk. It has been accepted for inclusion in Suffolk University Law School Faculty Works by an authorized administrator of Digital Collections @ Suffolk. For more information, please contact dct@suffolk.edu.

Fair Use and Machine Learning

By Stephen McJohn and Ian McJohn***

* Professor of Law, Suffolk University Law School.
** Massachusetts Institute of Technology, Class of 2022.

Table of Contents

I. INTRODUCTION

Commentators have long discussed whether artificial intelligence could perform legal reasoning; some projects have attacked specific legal domains.[1] For decades, artificial intelligence in general was more important theoretically than practically. Machine learning, however, has in recent years shown tremendous practical progress.[2] That progress stems partly from new algorithms but also from the tremendous increase in computing resources and availability of large data sets, which have given life to decades-old theoretical work. At the same time, the domains proving most amenable to machine learning have been quite different than legal reasoning. Such areas as machine vision, internet searching, language translation, handwriting recognition,[3] credit scoring, email spam detection and viewing recommendations[4] are relatively focused compared to legal reasoning, which is free-ranging, semantic and conceptual. This paper explores whether the surprising success of machine learning might extend to using those techniques to apply copyright's fair use doctrine.

There would be a beaten path to the maker of software that could reliably state whether a use of a copyrighted work was protected as fair use. The question, "Is this fair use?" arises millions

1 *See* Stephen M. McJohn, *Review of Artificial Legal Intelligence*, 12 HARV. J.L. & TECH. 241, 244 (1998) ("Thus, there have been a number of projects that claim some progress toward automating legal reasoning. This naturally raises the question, to what extent do the programs actually model the task at issue, or, alternatively, succeed in producing results similar to human decisions?").

2 This paper relies heavily on PEDRO DOMINGOS, THE MASTER ALGORITHM (2015); IAN GOODFELLOW ET AL., DEEP LEARNING (2016); AURÉLIEN GÉRON, HANDS-ON MACHINE LEARNING WITH SCIKIT-LEARN & TENSORFLOW (Nicole Tache et al. eds., 1st ed. 2017).

3 *See* Yann LeCun et al., MACHINE LEARNING IN HANDWRITING, http://yann.lecun.com/exdb/mnist/ (last visited Sept. 24, 2019) (a training database for "teaching" computers how to understand handwriting); Niek Temme, *Using TensorFlow to Create Your Own Handwriting Recognition Engine*, BUSINESS ANALYTICS: HARNESSING THE VALUE OF INFORMATION (Feb. 21, 2016, 7:24PM), https://niektemme.com/2016/02/21/tensorflow-handwriting/; Shan Carter et al., *Four Experiments in Handwriting with a Neural Network*, DISTILL (Dec. 6, 2016), http://distill.pub/2016/handwriting/; Alexander Mordvintsev & Abid K., *OCR of Hand-written Data Using kNN*, OPENCV-PYTHON TUTORIALS (2013), http://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_ml/py_knn/py_knn_opencv/py_knn_opencv.html.

4 GOODFELLOW ET AL., *supra* note 2, at 445–82 (discussing applications of deep learning, such as computer vision, speech recognition, natural language processing and recommender systems).

of times a day. A student using block quotes in a paper, a playwright parodying a blockbuster movie, an activist passing on someone else's video of a campaign speech, a fan posting a song on YouTube and a documentary filmmaker using old news stories, all could benefit if software could give an accurate prediction if they were infringing or making fair use.[5] From the copyright holder's point of view, a novelist who sees her work copied into fan fiction (with advertising), a photographer whose work is picked up without permission by major sites, and a music company that sees its songs posted on YouTube might likewise wonder if fair use applied. In the most likely application, sites that host user content, like YouTube or Twitter, could use such a fair use daemon to help deal with the multitude of postings each day. In addition, Google Translate's most recent incarnation uses machine learning to produce accurate translations between languages.[6] An able fair use analyzer would be useful in many contexts.

Fair use determinations must consider four broad factors in light of a vast[7] amount of case law which has flowed in various directions over time. Today's software could not replicate the process that an experienced lawyer would use to assess a case.[8] But it is well worth exploring how one might try to use machine learning on fair use. First, the exercise of looking at how specific machine learning algorithms might be used in fair use analysis can show which sorts of algorithms might ultimately be best suited to the task.[9] Second,

5 Presently, for many potential fair uses, the default is that the site may block the work, even if it qualifies for fair use. *See generally* Natalie Marfo, *Playing Fair: Youtube, Nintendo, and the Lost Balance of Online Fair Use*, 13 BROOK J. CORP. FIN. & COM. L. 466 (2019).

6 Yongui Wu, et al., *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation* 20 (2016), https://arxiv.org/pdf/1609.08144v1.pdf.

7 Vast in the sense of case law to research, not in the sense of Big Data, as discussed below.

8 Niva Elkin-Koren, *Fair Use by Design*, 64 UCLA L. REV. 1082, 1095 (2017) (citing Mark A. Lemley, *Rationalizing Internet Safe Harbors*, 6 J. ON TELECOMM. & HIGH TECH. L. 101, 110-11 (2007)) ("Skeptics believe that fair use analysis cannot be automated. One concern is that it involves a high degree of complexity, which requires discretion while weighing each of the four factors in light of the purpose of copyright law.").

9 Considering how such algorithms might work can also bear on the consideration that fair use programs might replicate implicit or explicit biases in the design of the software. *See* Dan L. Burk, *Algorithmic Fair Use*, 86 U. CHI. L. REV. 283, 285 (2019) ("Consequently, it may seem desirable to incorporate fair use metrics into copyright policing algorithms, both to

examining how machine learning might fit fair use analysis can be a useful way of studying how fair use analysis itself works in actuality.

Third, software will likely be used anyway to, in effect, make fair use determinations. We can see that by comparison with a related issue: whether material potentially infringes copyright. As with fair use, today's software is not ready to make the initial, subtle determination of whether a copyright has been infringed. That determination requires considering whether original expressive material has been copied (or adapted, distributed, performed, or displayed). For example, to see if a song posted on YouTube potentially infringes (i.e. without even considering whether the post is fair use), one would have to identify original elements in the copyrighted work that were copied into the accused copy, and then filter out any non-protected elements that were copied, such as non-original elements copied from still other works, or non-protected ideas (ideas may be copied without infringing copyright). Notable recent music copyright cases such as the *Blurred Lines* case[10] show how difficult that assessment can be.

Assessment of potential infringement is a subtle analysis. Nevertheless, *in effect*, most copyright infringement analysis today is done automatically by software. Copyright holders use software to crawl the web, search for copies of their works, and generate take-down notices – thereby making implicit infringement determinations, but bluntly, without actually considering such questions as originality or the non-protection of ideas, let alone fair use. This brute force approach may be accurate in the majority of cases, but it is harsh in some. YouTube likewise relies on automated processes. YouTube allows copyright holders to submit works to YouTube's Content ID System. When a work is uploaded to YouTube, it is compared to the Content ID database, and if identified as infringing, the copyright

protect against automated over-deterrence, and to inform users of their compliance with copyright law. In this paper I examine the prospects for algorithmic mediation of copyright exceptions, warning that the design values embedded in algorithms will inevitably become embedded in public behavior and consciousness."). The potential use of AI for legal decision-making raises philosophical and jurisprudential issues. *See* Lawrence B. Solum, *Artificially Intelligent Law*, BIOLAW J. - RIVISTA DI BIODIRITTO 58-61 (April 14, 2019), http://dx.doi.org/10.2139/ssrn.3337696; Rebecca A. Williams, *Rethinking Deference for Algorithmic Decision-Making* 2-3 (Aug. 31, 2018) (unpublished research paper) (available at http://dx.doi.org/10.2139/ssrn.3242482).

10 *See* Williams v. Gaye, 885 F.3d 1150, 1172 (9th Cir. 2018) (upholding jury finding of infringement in controversial music copyright case).

holder is given the choice between taking the work down or allowing it to remain in place with advertising revenue going to the copyright holder. In many online contexts, then, whether copyright infringement has occurred is implicitly decided by software – not as a judicial matter, but as a practical one. Online services also have incentive to assess whether a posting is not infringement, but fair use. This may allow users to post more material, which may make the service more attractive to users. It may help the services respond with more nuance to take-down orders generated automatically by copyright holders.

This paper explores whether fair use determinations are amenable to machine learning. In short, fair use is too complicated and nuanced, and involves such a broad range of subject matter like music, video, literature, and far more, to simply write a program to handle it. Machine learning has successfully seen computers teach themselves to identify spam emails, recognize objects in images and video, translate languages, and play chess. This paper surveys the leading schools of machine learning and how each may have advantages and disadvantages in dealing with fair use. It may be that fair use machine learning in early generations may have some success by looking at a few statistically, if not necessarily legally, relevant factors. This paper further suggests that whether later generations can make finer discriminations may depend on the progress of machine learning, and related disciplines in computer science, in identifying concepts and representing knowledge. Ironically, some of that work may stem from knowledge engineering, the branch of artificial intelligence that has been somewhat eclipsed by the spectacular achievements of machine learning in recent years. This paper also touches on possible legal issues with using machine learning to apply fair use (e.g., whether automated analysis of fair use is inconsistent with the case-bound analysis required by courts, and possible application of the recent data protection regulations in Europe), including by private actors such as sites that host user-posted content.

II. ARTIFICIAL INTELLIGENCE: FROM KNOWLEDGE ENGINEERING TO MACHINE LEARNING

Discussions of artificial intelligence came considerably earlier than even the most rudimentary electronic computer. Alan Turing established the basic mathematical principles of computing, most notably with the concept of a Turing machine, a theoretical prototype of today's programmed devices, and his proofs on the

limits on computability.[11] Turing also explored theoretically whether machines could think like humans, formulating what Turing called the "Imitation Game," and what everyone else has called the "Turing Test," broadly understood as whether a machine could carry on a conversation sufficiently intelligently to be indistinguishable from a human. Once electronic computers were developed, programmers worked at, among many other things, using that modest computing power to do tasks the way that a human did them. "Artificial Intelligence" became one, ill-defined area of computer science. No single definition can capture the variety of projects. Some consider AI the attempt to do with computers what humans do with their brains. A useful way of defining AI is "that it consists of finding heuristic solutions to NP-complete problems."[12] In other words, some mathematical problems can be solved and checked. Those can be written in computer programs. Some mathematical problems are so complex that we cannot even check if a proposed solution is correct. In the middle are NP-complete problems, which are too complex to solve even with gigantic computer resources, but where a proposed solution may be checked.[13] Humans are good at approximating solutions to some such problems, like driving a car without exactly figuring out all the practically infinite physics problems. AI can be considered the endeavor to likewise formulate approximations to such problems with computers. Measuring the capabilities of computers against humans has become less important as it becomes clearer that computers accomplish tasks in much different ways, however comparison to humans is still relevant in many cases. Whether self-driving cars should be permitted might depend, in part, on how well they perform compared to human

11 On whether Turing's Halting Problem places theoretical limits on the ability to use artificial intelligence to analyze legal problems, see Jeffrey M. Lipshaw, *Halting, Intuition, Heuristics, and Action: Alan Turing and the Theoretical Constraints on AI-Lawyering*, 5 SAVANNAH L. REV. 133, 147 (2018) (footnotes omitted) ("In the abstract, then, any modern digital computer that runs on stored programs is a universal Turing machine. Kearse, if it were to exist in the foreseeable future, would be a Turing machine and quite capable. But it would also be subject to the mathematical constraints of a Turing machine, namely, the inability to determine for *every* program that it might run whether it would, on one hand, complete that program and generate an answer, or, on the other hand, get stuck in a loop.").

12 DOMINGOS, *supra* note 2, at 33.

13 Other NP-complete problems include "the shortest route to visit a set of cities, the best layout of components on a microchip, the best placement of sensors in an ecosystem . . . and (most important) your Tetris score." *Id.*

drivers. But the central questions now might be simply how and whether machines perform tasks, and what safeguards might be appropriate.

AI can usefully be divided into two approaches: machine learning and knowledge engineering.[14] Machine learning includes such areas as neural networks. The human brain learns with networks of neurons, although it remains disputed how it learns and the extent to which knowledge and abilities are "hard-wired" from birth.[15] This question prompted researchers to try to develop electronic networks that could learn. The Perceptron was a very notable early example, which "was simple, yet it could recognize printed letters and speech sounds just by being trained with examples."[16] However, others proved mathematically that the Perceptron would be unable to learn some things, most notably the XOR operation, a basic logical operation dear to the hearts of computer scientists.[17] With these apparent limitations on the capabilities of networks to learn, research in that area faltered.

For a period of time, AI research's center of gravity shifted to the other approach: knowledge engineering – the idea that programs would be written to accomplish tasks previously reserved to humans.[18] One set of such programs has been called Symbolic AI, including programs that could prove mathematical theorems or play a decent game of chess, which could be expanded to the extent that reasoning was based on logic and symbols. Another category was the expert system.[19] This approach was to carefully observe how a human expert in a particular domain accomplished a task, then write a program that followed that approach step-by-step. Early hopes for this approach, however, were not realized. The theorem-proving approach proved limited to circumscribed areas like mathematics, because real-world applications required the need for implicit knowledge about the world[20] and for robust

14 *See id.* at 102–04.
15 *See generally* STEVEN PINKER, THE BLANK SLATE (2003).
16 DOMINGOS, *supra* note 2, at 100.
17 *Id.* at 100–01 (discussing MARVIN L. MINSKY & SEYMOUR A. PAPERT, PERCEPTRONS (1969)).
18 *Id.* at 101.
19 On the efforts to build expert systems in law, see RICHARD E. SUSSKIND, EXPERT SYSTEMS IN LAW (1987).
20 Note that although machine learning has found many useful applications, these issues remain at the frontiers of artificial intelligence research. *See* GOODFELLOW ET AL., *supra* note 2, at 486 (emphasis removed) (first citing

reasoning which allowed for fuzzy concepts. Building expert systems similarly turned out to require much more than could be specified in a program.[21] Experts may follow a series of steps, but also employ a great deal of background knowledge and make judgements much more flexibly than the steps of an algorithm.[22] There were heroic efforts to compile the sort of background factual knowledge that humans use.[23] The Cyc project attempted to simply catalog as many facts as possible about the world.[24] After bloating with new facts for over thirty years, Cyc still has yet to achieve even commonsense reasoning.[25]

In the field of law, early projects in applying artificial intelligence techniques to legal reasoning followed a similar arc. Case-based reasoning systems sought to find weights of various factors identified in case law and use those to predict future cases or formulate arguments.[26] Expert system projects arose, such as

Roberto Navigli & Paola Velardi, *Structural Semantic Interconnection: A Knowledge-Based Approach to Word Sense Disambiguation*, 27 IEEE TRANSACTIONS ON PATTERN ANALYSIS & MACHINE INTELLIGENCE 1075, 1075–86 (2005); then citing Antoine Bordes et al., *Joint Learning of Words and Meaning Representations for Open-Text Semantic Parsing*, 22 PROC. OF MACHINE LEARNING RES. 127 (2012)) (discussing the challenge of knowledge representation and "word-sense disambiguation . . . which is the task of deciding which sense of a word is the appropriate one, in some context.").

21 DOMINGOS, *supra* note 2, at 89–90 ("In the 1970s, so-called knowledge-based systems scored some impressive successes, and in the 1980s they spread rapidly, but then they died out. The main reason they did was the infamous knowledge acquisition bottleneck: extracting knowledge from experts and encoding it as rules is just too difficult, labor-intensive, and failure-prone to be viable for most problems. Letting the computer automatically learn to, say, diagnose diseases by looking at databases of past patients' symptoms and the corresponding outcomes turned out to be much easier than endlessly interviewing doctors.").

22 *See id.* at 90.

23 *Id.* at 35.

24 *Id.*

25 *Id.* ("Thirty years later, Cyc continues to grow without end in sight, and commonsense reasoning still eludes it. Ironically, Lenat has belatedly embraced populating Cyc by mining the web, not because Cyc can read, but because there's no other way.").

26 *See* McJohn, *supra* note 1, at 242–43 ("Anne Von der Lieth Gardner's GP program attempted to use previous cases to distinguish easy from hard cases in the area of contract law . . . . The Norwegian Research Centre for Computers and Law developed SARA, an attempt to model the differing weights given to relevant factors in applying legal norms . . . . Kevin Ashley's HYPO system used a database of some thirty cases to compare a case to precedent cases, examining whether similarities existed with respect to given factors . . . ."

the use of ten key questions to determine the ownership of found property.[27] Other projects sought to develop formal representations of legal concepts in such areas of the law as tax and negotiable instruments.[28] As in AI generally at the time, the early projects were seen as stepping stones toward future, practical applications of artificial intelligence.[29] The greatest value of the projects may have been in identifying the aspects of legal reasoning which put it beyond the artificial intelligence techniques of the day: legal reasoning is ill-defined, complex and context-dependent, and deals with rules and cases that are vague and inconsistent.[30] As with case-based reasoning and expert systems in other areas, these projects did not lead to others with anything like human-level performance.[31] As the author of a book on expert systems in law later recognized, in the legal area software became useful not for legal reasoning, but to automate lower-level tasks, such as document searching and preparation, albeit with sophisticated technology.[32] That trend has continued until today, where technology in legal practice is geared toward areas like document identification in discovery and provision of forms (legal forms have long constituted the most useful knowledge-based

---

(footnote omitted)).

27 McJohn, *supra* note 1, at 243 ("Alan Tyree's FINDER program sought to automate the analysis of deciding whether a found piece of property belonged to its finder by asking ten key questions and attempting to determine the result of the case from the answers.").

28 *See* McJohn, *supra* note 1, at 243 (discussing L.T. McCarty's TAXMAN program and Carole Hafner's LIRS project).

29 Edwina L. Rissland, *Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*, 99 YALE L.J. 1957, 1980 (1990) ("Legal reasoning is complex. Our current AI models, albeit too simple, are but steps to more subtle and complete models, and at each step we understand more."). As the author of the landmark TAXMAN project put it: "In general, then, the detailed analysis of the TAXMAN system has tended to support what should surely be a lawyer's intuition: that the current TAXMAN paradigm fails to capture many of the significant facts about the structure of legal concepts and the process of legal reasoning . . . . Although we would ultimately come to the conclusion, not unlike the lawyer's intuition, that nothing as complex as legal reasoning could ever be represented in a computer program, I believe it would be possible to sketch out a formal computer model somewhat more realistic than the current version of TAXMAN." L. Thorne McCarty, Reflections on Taxman: An Experiment in Artificial Intelligence and Legal Reasoning, 90 HARV. L. REV. 837, 892–93 (1977).

30 *See* McJohn, *supra* note 1, at 250–51.

31 Rissland, *supra* note 29, at 1980.

32 *See* McJohn, *supra* note 1, at 253.

expert system for lawyers).[33] Technology-assisted review ("TAR") is now standard in litigation: "the case law has developed to the point that it is now black letter law that where the producing party wants to utilize TAR for document review, courts will permit it."[34]

Meanwhile, in artificial intelligence research generally, the artificial neural network approach found new life.[35] The technique of backpropagation overcame the limits of the Perceptron.[36] Rather than simply training the outer layer of a network, in backpropagation errors are communicated to the lower layers, so that inner nodes of the network adjust. A node that contributed to a correct classification may be reinforced so that it becomes more influential in later decisions. A node that contributed to an error may conversely get less weight. Networks were trained to accomplish numerous tasks, showing the potential of learning.[37] At the same time, these projects were more narrowly focused. Rather than attempting to create a broad-ranging intelligent machine, networks were designed to learn quite specific tasks.

Machine learning's successes were not in domains similar to legal reasoning, so the growth in machine learning has not been matched with a resurgence in applying those techniques to legal reasoning. Instead, machine learning has found broad applications

33 *See, e.g.*, Pamela Radford, *Harnessing Technology During Discovery*, 82 ADVOCATE *39, *40 (2018) ("A second product like iControl's Envise or BrainSpace 6 is the leap or 'predictive coding' piece. Those tools use 'active' machine learning technology, sometimes referred to as 'Continuous Active Learning' (CAL) which feeds the 'review' software what it learns. Using only the passive learning technology will get you to the same result in document review, but adding true 'active learning' gets you to the finish line faster and using less resources."); Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 87–88 (2014) ("It misses a class of legal tasks for which current AI technology can still have an impact even given the technological inability to match human-level reasoning . . . this Article will suggest that there may be a limited, but not insignificant, subset of legal tasks that are capable of being partially automated using current AI techniques despite their limitations relative to human cognition.").

34 Rio Tinto P.L.C. v. Vale S.A., 306 F.R.D. 125, 127 (S.D.N.Y. 2015).

35 GOODFELLOW ET AL., *supra* note 2, at 17 ("In the 1980's, the second wave of neural network research emerged in great part via a movement called connectionism or parallel distributed processing . . . . The connectionists began to study models of cognition that could actually be grounded in neural implementations . . . .")

36 *See id.* at 18.

37 *See id.* ("During this time, neural networks continued to obtain impressive performance on some tasks.").

in legal practice beyond abstract legal reasoning. In discovery, for example, machine learning can be used to detect patterns that are likely to produce documents responsive to requests.[38] Rather than simple key word searches, which depend on the party thinking of all the words that might relate, a machine learning system might produce a more robust set of responsive documents.[39] Machine learning is finding a place in empirical research about the legal system itself. One study had considerable accuracy in predicting dissents, using such factors as the length of the opinion, the number of citations and voting valence among judges.[40]

At present, machine learning techniques have the definite advantage over knowledge engineering, in purely engineering terms. Machine learning has shown greater capabilities and is more widely deployed, at least in areas that might be considered artificial intelligence. There may also be policy reasons to prefer machine learning. Knowledge engineering seeks to write programs that mimic the practices of experts. It may yet prove feasible to build an expert system for fair use, coding how a judge might apply the various rules (statutory and case law) comprising fair use. But, in an area as vague and flexible as fair use, that may hazard replicating the biases of the system designer.[41] To the extent machine learning extracts patterns from decided cases, that may be less of a hazard, although constructing a machine learning system requires human steps that will also introduce bias.

38 *See, e.g.*, Radford, *supra* note 33, at 40 ("A second product like iControl's Envise or BrainSpace 6 is the leap or 'predictive coding' piece.").

39 *See id.* ("Using only the passive learning technology will get you to the same result in document review, but adding true 'active learning' gets you to the finish line faster and using less resources.").

40 Shivam Verma et al., *The Genealogy of Ideology: Predicting Agreement and Persuasive Memes in the U.S. Courts of Appeals, in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 253 (2017), https://dl.acm.org/citation.cfm?id=3086512&picked=prox ("We employ machine learning techniques to identify common characteristics and features from cases in the US courts of appeals that contribute in determining dissent. Our models were able to predict vote alignment with an average F1 score of 73%, and our results show that the length of the opinion, the number of citations in the opinion, and voting valence, are all key factors in determining dissent.").

41 *See* Burk, *supra* note 9, at 283 (warning that to the extent it is possible to "incorporate fair use metrics into copyright policing algorithms," that "the design values embedded in algorithms will inevitably become embedded in public behavior and consciousness").

III. FAIR USE AND ITS COMPLEXITIES

The Copyright Act states a four-factor rule for fair use:

> 107. Limitations on exclusive rights: Fair use
>
> Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—
>
> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
>
> (2) the nature of the copyrighted work;
>
> (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
>
> (4) the effect of the use upon the potential market for or value of the copyrighted work.
>
> The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.[42]

At first blush, the rule might seem designed for machine learning: a binary classification (fair use or not fair use), based on four factors. But, like all legal rules, the dynamics of the provision are more complex. As applied by the courts, fair use has many complicating, sometimes confounding, aspects.

Each of the factors has been interpreted to bring in different, sometimes conflicting, policies. The first factor, purpose of the use,

42 17 U.S.C. § 107 (2012).

calls for the court to ask whether the use falls into a favored category (the preamble specifically identifies some: "criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research").[43] Beyond the specific activity, the underlying purpose must also be considered ("including whether such use is of a commercial nature or is for nonprofit educational purposes").[44] Many of the categories favored in the first question (criticism, comment, news reporting, teaching, scholarship, research) are commercial, and so fall into the disfavored category in the second question.

In addition to the categories stated in the statute, courts have classified the purpose of the use in various respects. A long-standing distinction was between reproductive uses (merely making a copy) and productive uses (using the copyrighted work as a basis for another work, or simply using the copyrighted work in a valuable manner, to some courts).[45] In recent years, courts have differentiated between "transformative" and non-transformative uses.[46] The Supreme Court adopted the approach in *Campbell*,[47] holding that fair use was favored because the rap parody version of the country song *Pretty Woman* transformed the work. "Transformative" before too long became almost a magic word, with courts applying the term to uses that did not change the form of (transform) the work, but rather used the work in a different way. The Second Circuit noted several uses deemed transformative which did not adapt the original into a

43 *Id.*

44 *Id.*

45 Stephen M. McJohn, *Fair Use and Privatization in Copyright*, 35 SAN DIEGO L. REV. 61, 93 (1998) ("Courts have also used fair use in a more subtle way to balance the incentives of copyright by distinguishing between 'productive' and 'reproductive' uses. A reproductive use simply makes copies that compete with the copies authorized by the copyright holder. Where a use is productive, however, defendant goes beyond copying to contribute some independent value.") (footnotes omitted).

46 Laura A. Heymann, *Everything is Transformative: Fair Use and Reader Response*, 31 COLUM. J.L. & ARTS 445, 447 (2008) (footnotes omitted) ("In *Campbell v. Acuff-Rose Music*, the Supreme Court, relying on a 1990 law review article by Judge Pierre Leval, suggested that an important factor to consider in whether a use was fair was whether the second use was 'transformative' – whether it 'adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message.' Although some uses are more appropriately considered with regard to whether they are 'transformative' than others, the term has since become as fundamental a part of any fair use analysis as the statutory language itself.").

47 Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 587 (1994).

new, creative form: "scanning books to create a full-text searchable database and public search function (in a manner that did not allow users to read the texts) . . . copying works into a database used to detect plagiarism . . . displaying tiny, low-resolution "thumbnail" reproductions of art works to provide links serving as Internet pathways to the appropriate websites containing the originals . . . and copying by one who has acquired the right to view the content of a telecast to enable a single, non-commercial home viewing at a more convenient time."[48]

Another dimension (quite literally, if we are to discuss applying machine learning to fair use) on "the purpose and character of the use" is whether an expressive use triggers considerations of First Amendment protections. The Supreme Court has effectively constitutionalized fair use. *Eldred* held that Congress did not run afoul of the First Amendment in retroactively extending the terms of existing copyrights in 1998,[49] although that would keep works out of the public domain longer, thereby restricting the ability of others to use those works expressively. One rationale was that the Copyright Law contains "built-in" First Amendment protections, namely fair use and the nonprotection of ideas.[50] Along those lines, courts consider First Amendment interests in fair use,[51] such as using protected works to criticize public officials.[52]

The second factor, the nature of the copyrighted work, likewise has prompted a variety of judicial approaches.[53] Courts regularly differentiate between works with thick and thin copyright protection.[54]

48 Capitol Records, LLC v. ReDigi Inc., 910 F.3d 649, 660–61 (2d Cir. 2018) (citations omitted).

49 Eldred v. Ashcroft, 537 U.S. 186, 221 (2003).

50 *Id.* at 219–20.

51 *See, e.g.*, Nat'l Rifle Ass'n of Am. v. Handgun Control Fed'n of Ohio, 15 F.3d 559, 562 (6th Cir. 1994) ("This contrast with commercial activity helps show that the purpose and character of HCF's use is far removed from that which the copyright law centrally protects and instead falls within the realm of the designated fair use purposes. The document was used primarily in exercising HCF's First Amendment speech rights to comment on public issues and to petition the government regarding legislation.").

52 See Kienitz v. Sconnie Nation LLC, 766 F.3d 756, 758 (7th Cir. 2014) (holding that use of a modified picture of mayor used on T-shirt was fair use).

53 Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions, 1978-2005*, 156 U. PA. L. REV. 549, 610 (2008) (footnote omitted) ("Factor two instructs courts to consider 'the nature of the copyrighted work.' The data with respect to factor two are seemingly as ambiguous and open to interpretation as the statutory language itself.").

54 *See, e.g.*, Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P., 756 F.3d 73, 87 (2d

Copyright protects only the original expressive elements of a work.[55] A work may be composed largely of non-protected elements.[56] Software, mainly functional, qualifies for copyright protection as a literary work,[57] perhaps to the surprise of literature professors. A database is composed of nonoriginal and so nonprotected facts, but may have the necessary creativity for copyright protection by virtue of a creative selection, arrangement or coordination of those facts.[58] Such thinly protected works are more subject to fair use on the theory that the use is largely exploiting unprotected aspects of the work. In a subset of those cases, copying may be done in order to extract unprotected elements from a work, such as where software is copied to reverse-engineer its functionality[59] or where a database is copied to copy its unprotected facts.[60] The nature of the work can be characterized in other ways. *Harper & Row*[61] put considerable weight on the fact that the work copied, the autobiography of former United States President Gerald Ford, was soon to be published, so the copier in effect appropriated the right of first publication (not a right that appears in the Copyright Act).

The third factor sounds simply quantitative, and so readily amenable to machine learning: "the amount and substantiality of

Cir. 2014) ("As relevant here, this factor requires us to consider the extent of Swatch's copyright in the recording—the 'thickness' or 'thinness' of Swatch's exclusive rights—as well as whether or not the recording had been published at the time of Bloomberg's use.").

55 *See* Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 344 (1991).

56 *See, e.g.*, *Swatch Grp.*, 756 F.3d at 89 ("[E]ven within the field of fact works, there are gradations as to the relative proportion of fact and fancy. One may move from sparsely embellished maps and directories to elegantly written biography. The extent to which one must permit expressive language to be copied, in order to assure dissemination of the underlying facts, will thus vary from case to case." (quoting Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 563 (1985))).

57 17 U.S.C. § 101 (2012).

58 *See Feist Publications*, 499 U.S. at 344–45.

59 *See, e.g.*, Sega Enters. Ltd. v. Accolade, Inc., 977 F.2d 1510, 1514 (9th Cir. 1992).

60 Assessment Techs. of Wis., LLC v. WIREdata, Inc., 350 F.3d 640, 645 (7th Cir. 2003) ("Similarly, if the only way WIREdata could obtain public-domain data about properties in southeastern Wisconsin would be by copying the data in the municipalities' databases as embedded in Market Drive, so that it would be copying the compilation and not just the compiled data only because the data and the format in which they were organized could not be disentangled, it would be privileged to make such a copy, and likewise the municipalities.").

61 *Harper & Row*, 471 U.S. at 569.

the portion used in relation to the copyrighted work as a whole."[62] In many cases, it is.[63] Where the defendant copies the entire work, fair use is less likely, and where the defendant copies only a small portion, fair use is more likely.[64] But the courts have introduced many qualitative elements into consideration of this factor. The Supreme Court cases contrast. In *Sony*, the defendants copied the entire television programs at issue, but fair use nevertheless applied.[65] In *Harper & Row*, the Nation magazine copied only a few pages of former President Gerald Ford's biography, but fair use did not apply: the Nation had copied "the heart of the book," the few pages where Ford described how he came to pardon his predecessor, Richard Nixon.[66] In addition, courts consider not just the amount copied, but how well that amount corresponds to the favored use.[67]

The last factor is "the effect of the use upon the potential market for or value of the copyrighted work."[68] This factor sounds quantitative, redolent of micro-economics and finance. But courts again bring in many qualitative considerations. On its face, the factor presents a conundrum. One reading of the factor could make fair use redundant. Fair use, by definition, applies to unauthorized uses. Without fair use, such users would have to seek a license from the copyright holder. That being the case, every such use could negatively affect the potential market for the work. So if the copyright holder and potential user were in a position to negotiate, the use without authorization must have denied the copyright holder potential revenue. There could be some cases where the copyright holder and potential user were not in a position to negotiate. In particular, if the

62 17 U.S.C. § 107(3).

63 *See, e.g.*, Authors Guild v. Google, Inc., 804 F.3d 202, 221 (2d Cir. 2015) ("A finding of fair use is more likely when small amounts, or less important passages, are copied than when the copying is extensive, or encompasses the most important parts of the original.").

64 *Swatch Grp.*, 756 F.3d at 89-90 (quoting Infinity Broadcast Corp. v. Kirkwood, 150 F.3d 104, 109 (2d Cir. 1998)) ("In general, 'the more of a copyrighted work that is taken, the less likely the use is to be fair.'").

65 Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 449–50 (1984) (holding that fair use could apply even though the entire works were copied).

66 *Harper & Row*, 471 U.S. at 564–65 ("The portions actually quoted were selected by Mr. Navasky as among the most powerful passages in those chapters.").

67 Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 587 (1994) (quoting *Harper & Row*, 471 U.S. at 564) (internal quotations omitted) ("[E]ven substantial quotations might qualify as fair use in a review of a published work or a news account of a speech but not in a scoop of a soon-to-be-published memoir.").

68 17 U.S.C. § 107(4) (2012).

transactions costs of the negotiation were greater than the potential value to the user, then the transaction would not occur.[69] So, fair use could authorize such uses. But those uses are also so inconsequential that litigation to enforce the copyright would likewise be not worth it. Thus, fair use would be reduced to inconsequential cases that would be unlikely to make it to court. Fair use would also shrink as new licensing techniques developed.

Courts have neither followed that narrow approach nor resolved the underlying conundrum. Rather, courts follow a number of approaches in considering whether and to what extent there is a loss of market or value.[70] *Campbell* held that certain losses would not be cognizable.[71] If a parody or a review was such effective criticism that it led to fewer sales of the work, that would not be a cognizable loss of market.[72] Similarly, where copying was done to reverse engineer a product by copying its nonprotected functional aspects, the introduction of a competing product would not represent a cognizable loss to the copyright holder. *Campbell* also recognized that certain losses would not count, where the copyright holder would not have taken advantage of them.[73] If the copyright holder of *Pretty Woman* would not have authorized a rap parody version, then that illusory lost licensing opportunity would not weigh against fair use.[74]

In addition to the varied interpretation of each factor, the factors are applied interdependently. *Campbell* held that "the more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use."[75] Where the use is noncommercial, there may be a greater burden for a party to show loss to its market.[76] Where the

69 *See generally* Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600, 1601 (1982).

70 Beebe, *supra* note 53, at 621 ("Factor four provides the analytical space for this balancing test to occur, and the various doctrinal propositions under factor four are merely there to tilt the scales one way or the other.").

71 *Campbell*, 510 U.S. at 591–92.

72 *Id.*

73 *Id.* at 592.

74 *Id.* at 592–94.

75 *Id.* at 579.

76 Zahr K. Said, *Foreword: Fair Use in the Digital Age, and* Campbell v. Acuff-Rose *at 21*, 90 WASH. L. REV. 579, 582 (2015) ("[*Campbell*] reversed the momentum created by two key presumptions in prior case law, namely that commercial uses were presumptively unfair, and that plaintiffs were allowed to presume harm when defendants' uses were commercial.").

use is commercial, some loss may be presumed. The amount of the work copied may be excusable depending on the nature of the use, especially if the amount is geared toward the amount necessary to make such a use.[77] Which factors are most important is quite unsettled. The law review article that inspired *Campbell* to look to the "transformative" use referred to the first factor as "the soul of fair use,"[78] while courts sometimes refer to the fourth factor, effect on the market, as the most important factor.[79]

As a legal rule, Section 107 does not guide the decision maker. It simply states that fair use is not infringement, and that the court should consider four factors. The rule does not state how the factors should be weighed. Nor does it state whether additional factors may be considered, although the factors may be read broadly enough to accommodate almost any fact or consideration a court wishes to include. Discerning how courts have in fact applied those factors is the strength of machine learning.

Likewise, machine learning can discern patterns (although sometimes unreliably), which may be apt to address doctrinal uncertainty. Fair use has been considered unpredictable; in Judge Learned Hand's words, "the issue of fair use . . . is the most troublesome in the whole law of copyright."[80] Whether a proposed use qualifies for fair use is often an uncertain decision. As Professor Madison put it, "twenty-five years after the doctrine was codified in the Copyright Act of 1976, courts are no closer to a meaningful understanding of the doctrine than Congress appeared to be at the time of the law's enactment."[81] Other commentators have found

77 Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 449–50 (1984) ("Moreover, when one considers the nature of a televised copyrighted audiovisual work, see 17 U.S.C. § 107(2), and that timeshifting merely enables a viewer to see such a work which he had been invited to witness in its entirety free of charge, the fact that the entire work is reproduced, see id., at § 107(3), does not have its ordinary effect of militating against a finding of fair use.").

78 Said, *supra* note 76, at 581–582 ("The Court had relied on Judge Pierre N. Leval's seminal Harvard Law Review article, *Toward a Fair Use Standard*, to articulate a framework for assessing the reason for a defendant's use of plaintiff's work. . . . He has referred to the factor one analysis as 'the soul of fair use.'").

79 Beebe, *supra* note 53, at 617 ("Of the opinions following *Campbell*, 26.5% continued explicitly to state that factor four was the most important factor.").

80 Dellar v. Samuel Goldwyn, Inc., 104 F.2d 661, 662 (2d Cir. 1939) (per curiam opinion attributed to Judge Learned Hand).

81 Michael J. Madison, *A Pattern-Oriented Approach to Fair Use*, 45 WM. & MARY L.

more regularity in fair use case law. Professor Samuelson found fair use "more coherent and more predictable than many commentators have perceived once one recognizes that fair use cases tend to fall into common patterns, or what this Article will call policy-relevant clusters."[82]

Fair use cases often involve disagreement even between the judges. The first two fair uses cases to reach the Supreme Court in modern times did not yield an opinion because the Justices split 4-4.[83] Of the three key Supreme Court cases to follow, the court was divided: 5-4 in *Sony,*[84] 6-3 in *Harper & Row,*[85] and a unanimous opinion in *Campbell.*[86] Reversals are frequent in fair use cases.[87] To use examples of key recent fair use cases: in *Cambridge University Press*[88] (on whether course packs are fair use), the panel reversed the trial court and then reversed it again when the case returned after remand; in *Cariou*[89] (whether painting on photographs, termed by some "appropriation art," is fair use), the Second Circuit reversed with respect to most of the images at issue; in *Oracle v. Google* (whether it was fair use for Google to copy application programming interfaces of Java), a case with "huge stakes" for the software industry,[90] the Federal Circuit reversed a jury verdict of fair use.[91]

REV. 1525, 1549 (2004).

82 Pamela Samuelson, *Unbundling Fair Uses,* 77 FORDHAM L. REV. 2537, 2541–42 (2009) ("The policies underlying modern fair use law include promoting freedom of speech and of expression, the ongoing progress of authorship, learning, access to information, truth telling or truth seeking, competition, technological innovation, and privacy and autonomy interests of users.").

83 Pierre N. Leval, *Toward a Fair Use Standard,* 103 HARV. L. REV. 1105, 1106–07 (1990) (discussing Williams & Wilkins Co. v. United States, 420 U.S. 376 (1975) and Columbia Broad. Sys. v. Loew's, Inc., 356 U.S. 43 (1958)).

84 Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417.

85 Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539 (1985).

86 Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569 (1994).

87 An empirical study found, however, no disparity between reversal rates in fair use cases compared to other areas of the law. Beebe, *supra* note 53, at 554.

88 Cambridge Univ. Press v. Albert, 906 F.3d 1290, 1293–1302 (11th Cir. 2018).

89 Cariou v. Prince, 714 F.3d 694, 695, 699, 712 (2d Cir. 2013) ("Prince is a well-known appropriation artist. The Tate Gallery has defined appropriation art as 'the more or less direct taking over into a work of art a real object or even an existing work of art.'").

90 Timothy B. Lee, *Google asks Supreme Court to overrule disastrous ruling on API copyrights,* ARS TECHNICA (Jan. 25, 2019), https://arstechnica.com/tech-policy/2019/01/google-asks-supreme-court-to-overrule-disastrous-ruling-on-api-copyrights/.

91 Oracle Am., Inc. v. Google LLC, 886 F.3d 1179, 1186–87 (Fed. Cir. 2018) ("Because we conclude that Google's use of the Java API packages was

None of those cases ruled categorically on the practice at issue, meaning that fair use may, to some undetermined extent, apply to course packs and appropriation art. Notably, a thorough empirical study of fair use cases determined that "the lower courts repeatedly and systematically inverted Supreme Court dicta to favor the defendant, so that if the Court stated, for example, only that 'not *x*' favors the plaintiff, the primary lesson the lower courts would draw from this is that '*x*' favors the defendant."[92] There are differences in how the various federal circuits apply fair use. The Seventh Circuit rejected other circuit's broad reading of "transformative."[93] Whether fair use protects uses to exploit the functional aspects of software has yielded an apparent divide.[94]

The reported cases indicate that fair use is uncertain.[95] But the reported cases are not a random sample of fair use cases. Rather, they represent cases where the parties had incentive and resources to litigate the case to judgment and appeal. It may be that the reported cases are those where the application was sufficiently unpredictable such that the parties' predictions differed. Plaintiff thought fair use would not apply, defendant thought it would, and

---

not fair as a matter of law, we reverse the district court's decisions denying Oracle's motions for JMOL and remand for a trial on damages.").

92 Beebe, *supra* note 53, at 556.

93 Kienitz v. Sconnie Nation LLC, 766 F.3d 756, 758 (7th Cir. 2014) (citing Cariou, 714 F.3d at 706) ("We're skeptical of *Cariou's* approach, because asking exclusively whether something is 'transformative' not only replaces the list in § 107 but also could override 17 U.S.C. § 106(2), which protects derivative works. To say that a new use transforms the work is precisely to say that it is derivative and thus, one might suppose, protected under § 106(2). *Cariou* and its predecessors in the Second Circuit do not explain how every 'transformative use' can be 'fair use' without extinguishing the author's rights under § 106(2)."); *see also* Jiarui Liu, *An Empirical Study of Transformative Use in Copyright Law*, 22 STAN. TECH. L. REV. 163, 240 (2019) ("Nonetheless, it is difficult to say with confidence that transformative use is an improvement over its ancestors. While the new label has harmonized fair use rhetoric, it falls short of streamlining fair use practice or increasing its predictability.").

94 *Compare* Sony Comput. Entm't, Inc. v. Connectix Corp., 203 F.3d 596, 608 (9th Cir. 2000) (holding fair use authorized copying of code to reverse engineer functional aspects) *and Oracle*, 886 F.3d at 1200 (holding fair use did not authorize copying of application programming interfaces of Java done to facilitate interoperability).

95 The extent of that uncertainty is subject to disagreement. *See* Matthew Sag, *Predicting Fair Use*, 73 OHIO ST. L.J. 47, 85 (2012) ("The final, and perhaps most important contribution of this Article is that it offers considerable evidence against the oft-repeated assertion that fair use adjudication is blighted by unpredictability and doctrinal incoherence.").

both chose litigation (in the face of the other parties' resistance to settling the dispute), because of their different assessments. Cases where fair use clearly does not apply may be less likely to make it to court, because defendants will cave. Conversely, where fair use applies, plaintiffs will not go to court. But that game-theoretical explanation is slightly undercut by the fact that copyright cases are rarely single issue cases. Defendant is likely to argue that the work does not qualify for copyright, that plaintiff lacks standing to enforce the copyright, that there is no proof that defendant copied, that defendant copied only non-protected elements, that there were no damages from defendant's copying and that fair use protected defendant. Where fair use is often only one of many issues in litigation, the predictability of its application may only be a minor factor in whether the parties pursue the case or settle. It remains to be seen whether fair use case law does in fact have regular patterns that guide its application. Machine learning may be one way to find that out.

IV. RENDERING DATA FOR FAIR USE MACHINE LEARNING

Machine learning techniques, with some notable exceptions, rely on data.[96] Key questions in applying machine learning to address fair use issues are what data to employ and how to prepare it. A number of sources could supply data. There are hundreds of reported judicial opinions on fair use. But using those to train software to predict fair use raises a number of issues. Comparison to a prototypical[97] machine application can highlight the issues.

A textbook example of machine learning is image classification. A project to recognize pictures of cats is a classic.[98]

96 *See, e.g.*, Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 880 (2016) ("'Machine learning' is part of a nest of concepts in the artificial intelligence arena, including 'data mining,' 'knowledge discovery in databases,' and 'big data,' that are often used interchangeably and confusingly in academia, government, and popular media.").

97 Abdellatif Abdelfattah, *Image Classification using Deep Neural Networks*, MEDIUM (July 27, 2017), https://medium.com/@tifa2up/image-classification-using-deep-neural-networks-a-beginner-friendly-approach-using-tensorflow-94b0a090ccd4 ("We will build a deep neural network that can recognize images with an accuracy of 78.4% while explaining the techniques used throughout the process.").

98 *See, e.g.*, *id.* (teaching example of deep neural network to recognize pictures of cats).

The project might start with a database of, say, 60,000 images.[99] The images would be tagged in advance, labeled as containing cats or not containing cats. The images would then be preprocessed to yield a form more convenient for machine learning.[100] Two common ways would be to convert the image to greyscale (a numerical range from white to black) or RGB values (giving the combination of red, green, and blue).[101] The network would be trained on a subset of the images (or in sequence on smaller subsets, to reduce the computational demands) and then tested against a subset not used in training.[102] If the accuracy is not satisfactory, adjustments could be made and training repeated until a satisfactory level of accuracy is met. Presented with a new image, the network would classify it as a picture with or without a cat, perhaps with an assessment of probability.

Where machine learning algorithms typically run on many thousands, if not millions, of examples, the reported fair use cases may prove to be a rather sparse data set.[103] Examples for a data set of fair use cases could be compiled with some imagination. Google sometimes presents a CAPTCHA to a user, showing several images and requiring the user to check a box for each image that contains a certain image, like a stop sign.[104] That may be a security measure, but the classification may also be used in Google's image recognition[105] or self-driving car endeavors. For fair use, every take-down notice presents a possible example. Although the take-down sender has implicitly labeled it as not fair use, it could also be seen as an unlabeled example. The take-down procedure also

99 *Id.*

100 *See* GÉRON, *supra* note 2, at 59–68 (discussing preparing the data for machine learning).

101 Abdelfattah, *supra* note 97.

102 GÉRON, *supra* note 2, at 29, 49–50 ("When you estimate the generalization error using the test set, your estimate will be too optimistic and you will launch a system that will not perform as well as expected. This is called *data snooping bias*.").

103 *Cf.* Beebe, *supra* note 53, at 550 (discussing an empirical study that found 306 "reported federal opinions that made substantial use of the section 107 four-factor test for fair use through 2005").

104 *See* Dennis Goedegebuure, *You Are Helping Google AI Image Recognition*, MEDIUM (Nov. 29, 2016), https://medium.com/@thenextcorner/you-are-helping-google-ai-image-recognition-b24d89372b7e.

105 *Id.* The author heard a discussion of this at a talk by Jonathan Frankle, Machine Learning and Neural Networks for Lawyers, Talk at the Boston University School of Law and Hariri Institute for Computing (Oct. 31, 2018).

allows the poster of content to require that the material be put back up, pending resolution of the case. So a smaller number of take-down/put-back-up examples could be grist for the data base. YouTube now largely channels copyright holders to use YouTube's Content ID system as an alternative to statutory take-down notices. That system, and others like it, could likewise present a rich vein of examples. Any site that has human moderators, vetters, or other supervisors deciding whether to take down user content could be a source of decisions on fair use, although the quality associated with those samples would depend on such things as whether the humans in question had training in the law or fair use specifically. Getting those examples would not be easy, requiring cooperation from the sites, which in turn could depend on the input of many constituencies and consideration of legal issues, including whether fair use authorized all that copying, distribution and adaptation of copyrighted works. With some investment and preprocessing of data, services like Amazon's Mechanical Turk could be used to have nonlawyers classify cases as fair use or not fair use, or even to create the cases. Those would not be borderline cases if the classifiers did not have whatever fine legal reasoning skills might be required to apply fair use. But a trove of clear cases could be helpful in training a machine learning program.[106] In some respects, easy cases may be better than close cases for programs to "learn" from.

Additional examples may be generated by automated means, termed "dataset augmentation."[107] In image classification, for

106 Other sources of data for legal applications may be unearthed with creativity. The Learned Hands project of Suffolk University Law School's Legal Innovation and Technology (LIT) Lab and the Stanford Legal Design Lab, created a data set for classification machine learning on spotting legal issues from "75,000 legal questions posted on Reddit . . . dealing with family, consumer, criminal and other legal issues." Jason Tashea, *New Game Lets Players Train AI to Spot Legal Issues*, A.B.A. J. (Oct. 16, 2018), http://www.abajournal.com/news/article/new_game_lets_players_train_ai_and_close_the_ justice_gap.

107 Goodfellow et al., *supra* note 2, at 240, 259–60; *see also* E. Alpaydin & Fevzi Alimoglu, *Pen-Based Recognition of Handwritten Digits Data Set*, Univ. Cal. Irvine Mach. Learning Repository, https://archive.ics.uci.edu/ml/datasets/Pen-Based+Recognition+of+Handwritten+Digits (last visited Jun. 30, 2019) ("We create a digit database by collecting 250 samples from 44 writers. The samples written by 30 writers are used for training, cross-validation and writer dependent testing . . ."); Mike Szczys, *Machine Learning Lets Micro Decode Your Handwriting*, Hackaday (May 3, 2012), https://hackaday.com/2012/05/03/machine-learning-lets-micro-decode-your-handwriting/ ("This rig will take the letters you write on the touchpad using a stylus and turn them into digital characters. The system is very fast and displays near-

example, one image can be used to generate multiple examples by cropping the same image at different locations[108] or with "random translations, rotations, and in some cases, flips of the input."[109] Machine learning itself may generate additional examples.[110] If a system has learned to identify features (such as concepts or classes of objects), then a *generative model* can produce new, different examples sharing the same features[111] but in a different arrangement.[112] For

perfect recognition. This is all thanks to a large data set that was gathered through machine learning.").

108 GOODFELLOW ET AL., *supra* note 2, at 459–60.

109 *Id.*

110 Generative machine learning can be used to make examples more like reality. *Cloud Machine Learning Engine*, GOOGLE, https://cloud.google.com/ml-engine/ (last visited June 30, 2019) (quoting Mathias Ortner) ("Google Cloud Machine Learning Engine enabled us to improve the accuracy and speed at which we correct visual anomalies in the images captured from our satellites. It solved a problem that has existed for decades. It will allow Airbus Defence and Space to continue to provide unrivaled access to the most comprehensive range of commercial Earth observation data available today.").

111 Alex Graves, *Generating Sequences with Recurrent Neural Networks*, UNIV. TORONTO at 1–2 (June 5, 2014), https://arxiv.org/pdf/1308.0850.pdf; *see also Recurrent Neural Network Handwriting Generation Demo*, https://www.cs.toronto.edu/~graves/handwriting.html.

112 *See* GOODFELLOW ET AL., *supra* note 2, at 540–43. In general, the ability of machine learners to generate new works raises the issue, whether those works are copyrightable and who would own the copyright. *See* Nina I. Brown, *Artificial Authors: A Case for Copyright in Computer-Generated Works*, 20 COLUM. SCI. & TECH. L. REV. 1, 5–6 (2018) (presenting "an argument for recognizing copyrights in computer-generated works"); Lin Weeks, *Media Law and Copyright Implications of Automated Journalism*, 4 N.Y.U. J. INTELL. PROP. & ENT. L. 67, 93 (2014) ("Where a strict moral conception of copyright might assert that all Polaroid photographs owe dependency to Edwin H. Land,104 or that the creator of an algorithm has a claim in everything that algorithm output, our current pecuniary conception values the promulgation of technologies into more and newer works."); Jessica Fjeld & Mason Kortz, *A Legal Anatomy of AI-generated Art: Part I*, HARV. J. L. & TECH.: JOLT DIG. (Nov. 21, 2017), https://jolt.law.harvard.edu/digest/a-legal-anatomy-of-ai-generated-art-part-i; Daniel J. Gervais, *The Machine as Author*, IOWA L. REV. (forthcoming 2019) (manuscript at 283, 285) (on file with authors); Jane C. Ginsburg & Luke Ali Budiardio, *Authors and Machines* (Colum. Pub. L. Res., Paper No. 14-597, 2018). The technology also raises issues of false works. *See* Mark Gibbs, *Machine Learning and Forgery*, NETWORK WORLD (Aug. 19, 2016), http://www.networkworld.com/article/3109557/software/machine-learning-and-forgery.html ("A new algorithm from University College London researchers creates forged handwritten text that never was . . ."); Douglas Harris, *Deepfakes: False Pornography is Here and the Law Cannot Protect You*, 17 DUKE L. & TECH. REV. 99 (2019). *But see* Alex Hern, *Google Says Machine Learning is the Future. So I Tried it Myself: If Deep Learning Will Be as Big as the Internet, It's Time for Everyone*

lawyers, this reminds us of the Socratic approach of slightly changing the facts of hypotheticals, but also reminds us of the need to develop systems that can identify facts and legal concepts within source material.

If the data available for fair use is somewhat limited, that increases the importance of considering the various algorithms discussed below. Where there is a huge amount of data to train the algorithm, then the various types of machine learning algorithms tend to approach the same level of performance, at least in theory.[113] With smaller data sets, the type of algorithm becomes more important.[114]

In the abstract, fair use is a classification problem: case of fair use or not fair use? But the data are less squarely defined. An image classifier is trained on a set of images, then attempts to classify images, all presented in the same format. A fair use case could be in any media (images, songs, pantomimes, architectural blueprints, etc.). Fair use does not depend only on examining the work. Rather, it would involve comparing the relevant work to the copyrighted work (what elements were used and were those elements themselves protected by copyright), then considering a number of factors extrinsic to the work (the purpose of the use, the effect on the market). In addition, fair use depends on comparison to precedent, so the learner could be trained on the facts in the case law. In short, the potential data to be considered are much more variegated than the typical machine learning application.

One key determination might be at what level of abstraction we define the features for the program to operate on.[115] One could think of fair use determinations as strictly comparing the instant case to the factors as determined in other cases. In other words, a case presenting the same set of factors should be decided the same way as precedent cases presenting the same set of factors. At the

to Start Looking Closely at It, THE GUARDIAN (Jun. 28, 2016), https://www.theguardian.com/technology/2016/jun/28/google-says-machine-learning-is-the- future-so-i-tried-it-myself (discussing an unsuccessful attempt to use machine learning, trained with vast amounts of The Guardian, to synthesize possible headlines from a "parallel universe").

113 *See* GERÓN, *supra* note 2, at 23 (discussing Michele Banko & Eric Brill, *Scaling to Very Very Large Corpora for Natural Language Disambiguation* and Peter Norvig et al., *The Unreasonable Effectiveness of Data*).

114 *Id.*

115 *See generally* ALICE ZHENG & AMANDA CASARI, FEATURE ENGINEERING FOR MACHINE LEARNING: PRINCIPLES AND TECHNIQUES FOR DATA SCIENTISTS (Rachel Roumeliotis & Jeff Bleidel eds., 2018).

other end, viewing fair use as ultimately a determination on the facts of particular cases, one could disregard the factors and concentrate solely on the facts. If a case is most similar to a precedent case on similar *relevant* facts, then it should be decided in the same way. The second approach would likely require a larger set of features to accommodate the broad range of fact settings in fair use cases.

One reason to define the features broadly is to avoid the hazard of *overfitting*. Overfitting, in machine learning as in statistics, occurs when the learner fits the function to the data, only too well.[116] The concept of overfitting is quite similar to the hazard in legal reasoning of tying a rule too closely to the facts of a case. For example, if a learner were trained only on music cases, including *Campbell*, it could include in its function that one required element is that the work be a musical work. One guard against overfitting is to include more examples in training. Another is testing the function against other test examples.[117] The model would fail when tested against examples where fair use applied in non-music cases. Another way to guard against overfitting is simply to use fewer features. If all the example cases happened to be music cases but that element was not included as a feature (or inferable from other features), then overfitting to music would be much less likely.[118]

Another way to reduce the number of dimensions is to preprocess the data. Unsupervised machine learning may be used to find clusters, which can be used as the data for supervised machine learning, such as classification.[119] It may become possible to use machine learning to put the examples into categories, such as the elements of the relevant works, upon which other programs may learn to classify the examples.[120] Copyright law itself supplies the relevant principles, like Learned Hand's abstractions test, used to determine whether a second author copied unprotected ideas or protected expression:

> Upon any work, and especially upon a play, a great number of patterns of increasing generality will fit

116 GERÓN, *supra* note 2, at 26–28.

117 *See, e.g.*, *id.* at 29 ("The only way to know how well a model will generalize to new cases is to actually try it out on new cases.").

118 *Id.* at 302–10 (discussing techniques to avoid overfitting).

119 DOMINGOS, *supra* note 2, at 211 ("Machine learners call this process dimensionality reduction because it reduces a large number of visible dimensions (the pixels) to a few implicit ones (expression, facial features.").

120 *Id.* at 210.

> equally well, as more and more of the incident is left out. The last may perhaps be no more than the most general statement of what the play is about, and at times might consist only of its title; but there is a point in this series of abstractions where they are no longer protected, since otherwise the playwright could prevent the use of his 'ideas,' [sic] to which, apart from their expression, his property is never extended.[121]

Hand's characterization is echoed by the use of machine learning to group data into conceptually linked groups. Unsupervised learning can be used to find clusters within material, which in turn can be treated as objects which can be grouped into more abstract clusters.[122] The technique is already used in "topic extraction," to find the themes being discussed in a collection of text documents,[123] which would be a start toward characterizing the components of legal documents like judicial opinions.

Data may also be stored in more efficient structures, depending on its characteristics. A technique with resonance for legal reasoning is "distributed representation."[124] Rather than identifying each feature independently, features can be identified by a set of shared attributes. "Cat" and "dog" are quite different but can be associated by a number of shared attributes (fur, four legs, etc.).[125] Legal reasoning often depends on finding the similarities between fact patterns, and so would seem amenable to that sort of representation. In particular, fact patterns (or policy issues) may be similar in the sense of "family resemblance," where there is no common core that two items share, but rather overlapping characteristics.[126]

---

121 Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930) (citing Holmes v. Hurst, 174 U.S. 82, 86 (1899); Guthrie v. Curlett, 36 F.2d 694, (2d Cir. 1929)).

122 DOMINGOS, *supra* note 2, at 210.

123 *See* ANDREAS C. MÜLLER & SARAH GUIDO, INTRODUCTION TO MACHINE LEARNING WITH PYTHON: A GUIDE FOR DATA SCIENTISTS 131 (Dawn Schanafelt ed., 2017) ("Here, the task is to find the unknown topics that are talked about in each document, and to learn what topics appear in each document.").

124 *See, e.g.*, GOODFELLOW ET AL., *supra* note 2, at 548–54.

125 *Id.* at 550.

126 *See* Lawrence B. Solum, *The Unity of Interpretation*, 90 B.U. L. REV. 551, 551–78 (2010) (discussing legal philosopher Ronald Dworkin's use of "family

A number of projects, including in areas other than machine learning, have attempted to structure legal material in a way that would facilitate computer processing at a higher conceptual level. It might be possible to render data into useful forms for processing by using tagging, which could be done, in turn, by software.[127] Research projects have attempted such tagging at a conceptual level, which would both make the tag's features for learning and also reduce the number of features, by combining similar facts into a single concept.[128] Cases could be annotated, ideally by automated means, to provide a common data structure.[129] In the area of contracts,

---

resemblance," which comes from Ludwig Wittgenstein's, *Philosophical Investigations*).

127 Kyoko Sugisaki, *Supertagging for Domain Adaptation: An Approach with Law Texts*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, *supra* note 40, at 249 ("Abstract: In this paper, we present a German supertagger that analyses syntactic functions in linear order. We apply a statistical sequential model, conditional random fields (CRF), to Swiss law texts, in a real world scenario in which the training data of the domain is missing. We show that the small amount of in-domain training data that was informed by linguistic hard and soft constraints and domain constraints achieved a label accuracy of 90% in the domain data, thus outperforming state-of-the-art parsers.").

128 *See* Matthias Grabmair et al., *Introducing LUIMA: An Experiment in Legal Conceptual Retrieval of Vaccine Injury Decisions Using a UIMA Type System and Tools*, *in* FIFTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 69 (2015), https://dl.acm.org/citation.cfm?id=2746090 ("This paper presents first results from a proof of feasibility experiment in conceptual legal document retrieval in a particular domain (involving vaccine injury compensation). The conceptual markup of documents is done automatically using LUIMA, a law-specific semantic extraction toolbox based on the UIMA framework. The system consists of modules for automatic sub-sentence level annotation, machine learning based sentence annotation, basic retrieval using Apache Lucene and a machine learning based reranking of retrieved documents. In a leave-one-out experiment on a limited corpus, the resulting rankings scored higher for most tested queries than baseline rankings created using a commercial full-text legal information system."); *see also* Milagro Teruel et al., *A Low-Cost, High-Coverage Legal Named Entity Recognizer, Classifier and Linker*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, *supra* note 40, at 9–18 ("In this paper we try to improve Information Extraction in legal texts by creating a legal Named Entity Recognizer, Classifier and Linker. With this tool, we can identify relevant parts of texts and connect them to a structured knowledge representation, the LKIF ontology.").

129 *See, e.g.*, Vern R. Walker et al., *Semantic Types for Computational Legal Reasoning: Propositional Connectives and Sentence Roles in the Veterans' Claims Dataset*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, *supra* note 40, at 217–26 (2017) ("This paper

efforts have been made to identify common elements which would facilitate creation of software to handle such matters as drafting and interpretation,[130] although presently the wide variation in language makes the job of annotation best handled by humans, which is laborious.[131] Professor McCarty, long a leader in artificial intelligence and the law, has suggested that the building blocks of legal reasoning may be identified through machine learning, analogous to how machine learning has mimicked the way human vision identifies objects, such as faces.[132]

Another question would be whether the facts used to make the prediction of fair use should be limited to legally relevant facts. Political scientists have had considerable success in predicting the outcome of Supreme Court cases by considering such factors as ideological direction (liberal or conservative) of the lower court ruling and the general nature of the case – factors which the Court

announces the creation and public availability of a dataset of annotated decisions adjudicating claims by military veterans for disability compensation in the United States. This is intended to initiate a collaborative, transparent approach to semantic analysis for argument mining from legal documents.").

130 *See* Kathryn D. Betts & Kyle R. Jaep, *The Dawn of Fully Automated Contract Drafting: Machine Learning Breathes New Life into a Decades-Old Promise*, 15 DUKE L. & TECH. REV. 216, 227 (2017).

131 *See* Silviu Pitis. *Methods for Retrieving Alternative Contract Language Using a Prototype*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, *supra* note 40, at 277; Sugisaki, *supra* note 127, at 249 ("Under these circumstances, the best approach is the manual annotation of a large amount of new domain data in which a parser can be trained. However, this is also the most cost-intensive solution."); Matías García-Constantino et al., *CLIEL: Context-Based Information Extraction from Commercial Law Documents*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, *supra* note 40, at 79; Illias Chalkidis, Ion Androutsopoulos & Achilleas Michos, *Extracting Contract Elements*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, *supra* note 40, at 19.

132 *See* L. Thorne McCarty, *How to Ground A Language for Legal Discourse in A Prototypical Perceptual Semantics*, 2016 MICH. ST. L. REV. 511, 526 (2016) ("We can now address the title of this talk: *How to Ground a Language for Legal Discourse in a Prototypical Perceptual Semantics*. Specifically, we will see how to use the machinery of manifold learning and deep learning to define a *semantics* for a logical language. Why do I call this a 'prototypical perceptual semantics'? Well, it's a *prototypical* semantics because it is based on my model of prototypical clusters, as we will see. Why is it a prototypical *perceptual* semantics? Well, notice that our primary examples are drawn from the field of image processing, and therefore, if we can build a logic on these foundations, we will have a plausible account of how human cognition could be *grounded* in human perception.").

itself would of course deem irrelevant.[133] If the purpose is to predict how a court would decide a fair use case, then similarly one might include factors with no legal relevance but which have been shown to have predictive value. That, however, would make the potential facts include not just the relevant works and the fair use case law, but facts about the case law. Some statistical approaches would remain within the boundaries of judicial opinions, although relying on features that are not relevant to the legal analysis of fair use. For example, one project sought to use an even sparser statistical approach, examining whether the waxing and waning of "mimetic" phrases in judicial opinions might be used to predict likelihood of dissents.[134] Use of such "predictive analytics," rather than legal analysis, to predict cases may, however, undercut the value of the rule of law.[135]

The sources of data would include reported fair use cases. But that might be insufficient for some varieties of machine learning, reflecting a general need in machine learning. For many techniques, a machine learner can learn only with training examples. A cat-recognizing learner needs oodles of pictures, some of cats, some not, each labeled. Those thousands or millions of pictures need to be obtained. That may be done with web-crawling software[136] (although

133 *See* Theodore W. Ruger et al., Essay, *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking*, 104 COLUM. L. REV. 1150, 1150 (2004) ("For every argued case during the 2002 Term, we obtained predictions of the outcome prior to oral argument using two methods--one a statistical model that relies on general case characteristics, and the other a set of independent predictions by legal specialists. The basic result is that the statistical model did better than the legal experts in forecasting the outcomes of the Term's cases: The model predicted 75% of the Court's affirm/reverse results correctly, while the experts collectively got 59.1% right.").

134 Verma et al., *supra* note 40, at 253 ("In addition to the dissents, we analyze the notion of memetic phrases occurring in opinions - phrases that see a small spark of popularity but eventually die out in usage - and try to correlate them to dissent.").

135 Frank A. Pasquale & Glyn Cashwell, *Prediction, Persuasion, and the Jurisprudence of Behaviourism*, 68 U. TORONTO L.J. 63 (2018).

136 *See* Elkin-Koren, *supra* note 8, at 1095–96 ("Another concern is that algorithms that analyze fair use will fail to process information that is external to the content itself. For instance, determining the nature of use may require external information and additional analysis of facts. Yet, algorithms could be programmed to extract and analyze data from external sources. For instance, educational use might be determined based on tagging the nature of the user. A program could detect the type of user (e.g., educational institution, governmental agency) based on the domain name (e.g., .edu, .gov) or by

that raises the issue of whether fair use protects making copies of images to use in machine learning).[137] The pictures also have to be labeled. That cannot be done automatically if no cat-recognizing learner has been built yet. The common alternative is simply to have humans view the pictures and label them, a labor-intensive process. Machine learning is often cited for its potential to displace humans from repetitious jobs, but it is also creating many repetitious jobs with its need for labeled data.[138]

The ultimate question may be whether fair use via machine learning is a project more like Cyc or like Google Translate. Cyc attempted to support knowledge engineering by building a database with as many facts as possible about the world, only to find the store of facts inexhaustible.[139] Google Translate took on machine translation, which might seem to require an equally daunting task: taking on all the possible sentences which might talk about facts (and much more).[140] It turned out that dealing with the meaning of sentences was unnecessary.[141] Google Translate, trained on a vast corpus of translations conveniently compiled by the United Nations in the course of its affairs, handles translation an entire sentence at a time, having learned on previous sentence translation.[142] A legal

checking registration in external databases.").

137 *See* Amanda Levendowski, *How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem*, 93 WASH. L. REV. 579, 622–25 (2018) (arguing that fair use should apply because "[u]sing copyrighted works as training data for AI systems is highly transformative"); Benjamin L. W. Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 COLUM. J.L. & ARTS 45, 97 (2017) (discussing policy arguments on both sides of the issue of whether fair use should protect the use of copies of work during machine learning).

138 Matthew Hutson, *The Future of AI Depends on a Huge Workforce of Human Teachers*, BLOOMBERG L. (Sept. 7, 2017), https://www.bloomberg.com/news/articles/2017-09-07/the-future-of-ai-depends-on-a-huge-workforce-of-human-teachers ("For an autonomous car to recognize pedestrians and stop signs, it's typically fed thousands or millions of photos, all hand-labeled. To nail a conversation, a digital assistant needs to be told over and over when it's failed. And so Rubin spends 10 to 30 hours a week on her phone or computer evaluating search results and chat retorts through a site called Clickworker. . . . All together, more than 1 million people around the world are chipping in, one click at a time.").

139 DOMINGOS, *supra* note 2, at 35 ("Thirty years later, Cyc continues to grow without end in sight, and commonsense reasoning still eludes it. Ironically, Lenat has belatedly embraced populating Cyc by mining the web, not because Cyc can read, but because there's no other way.").

140 *See* Wu, *supra* note 6, at 2, 20.

141 *Id.* at 13, 17–18.

142 *See* Gideon Lewis-Kraus, *The Great A.I. Awakening: How Google Used Artificial*

case is not as modular as a task in translation because the sentences are interdependent. The question will be whether the necessary features for machine learning can be extracted from legal texts (or from examples of fair use in other media, like images).[143] The domain of law faces an issue recognized to be one of the key areas of machine learning research, which is finding a way to represent semantic knowledge: in technical terms, a "research frontier is to develop embeddings for phrases and for relations between words and facts."[144] The succeeding sections rest on the hopeful assumption that some progress will be made in providing the data necessary for using machine learning in the context of legal reasoning.[145]

V. SUPERVISED LEARNING WITH NEURAL NETWORKS: LEARNING TO RECOGNIZE FAIR USE

Following the classification by Professor Domingos,[146] machine learning can be divided into several schools.[147] The school presently making the most notable progress, artificial neural networks

Intelligence to Transform Google Translate, One of Its More Popular Services – and How Machine Learning is Poised to Reinvent Computing Itself, N.Y. TIMES (Dec. 14, 2016), https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html.

143 As an indication of the pervasive issue of extracting conceptual information from legal texts, see *Call for Participation: Competition on Legal Information Extraction/Entailment (COLIEE)*, INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L., https://nms.kcl.ac.uk/icail2017/cfcoliee.php (last visited Oct. 2, 2019) ("There are two tasks in the competition. One is to extract articles from Japanese civil codes which contribute to solving a bar exam yes/no question; the second task is to check entailment of a question from given civil code article(s). We also provide various NLP tools' outputs for training data to help your information retrieval and textual entailment.").

144 GOODFELLOW ET AL., *supra* note 2, at 484.

145 Note that judicial reform could improve the availability of data generally in the legal system. *See* David Colarusso & Erika J. Rickard, *Speaking the Same Language: Data Standards and Disruptive Technologies in the Administration of Justice*, 50 SUFFOLK U.L. REV. 387, 388 (2017) ("[E]stablishing data standards for electronically sharing information across the justice system would propel existing technological innovations to greater prominence and effectiveness.").

146 *See generally* DOMINGOS, *supra* note 2.

147 One could divide machine learning more broadly, such as between instance-based and versus model-based learning. GÉRON, *supra* note 2, at 17. The former "learns the examples by heart, then generalizes to new cases using a similarity measure." *Id.* The latter builds "a model of these examples, then use[s] that model to make predictions." *Id.* at 18. One can see the analogy to legal reasoning, which can be viewed as case-based or as applying general rules which may be derived from cases or set out in statutes.

(including deep learning), is an approach inspired by the networks of neurons comprising the brain, although now they are "generally not designed to be realistic models of biological function."[148] A connected network of virtual nodes is adjusted to respond to experience: "[b]y adding more layers and more units within a layer, a deep network can represent functions of increasing complexity."[149] Such deep learning has found successful application in a number of areas, such as machine vision,[150] speech recognition,[151] handwriting recognition[152] and natural language processing.[153] Deep learning networks are in broad use in a number of other fields, potentially anywhere there is ample data.[154]

The model receives a set of input values and gives a set of output values.[155] The weights of the various nodes are adjusted until the desired output values for each input are received.[156] Just how to adjust those weights, of course, is the key to whether the network can learn as intended.[157] The first wave of network learners, in the 1950s and 1960s, including the Perceptron, went under the name of "cybernetics."[158] The leading algorithm employed was "stochastic gradient descent."[159] Gradient descent has been compared to

148 GOODFELLOW ET AL., *supra* note 2, at 13.
149 *Id.* at 166.
150 *Id.* at 452–57.
151 *Id.* at 458–60.
152 Stephen Wu & Justin Churchill, *Touchpad Figure Recognition*, CORNELL U. ELECTRICAL & COMPUTER ENGINEERING (Spring 2012), http://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/s2012/shw46_jec324/shw46_jec324/index.html ("Our project implements a touchpad input system which takes user input and converts it to a printed character. Currently, the device only recognizes the 26 letters of the alphabet, but our training system could be easily generalized to include any figure of completely arbitrary shape, including alphanumerics, punctuation, and other symbols.").
153 GOODFELLOW ET AL., *supra* note 2, at 461–477.
154 One commentator listed inspirational deep learning applications: "1. Colorization of black and white images. 2. Adding sounds to silent movies. 3. Automatic machine translation. 4. Object classification in photographs. 5. Automatic handwriting generation. 6. Character text generation. 7. Image caption generation. 8. Automatic game playing." Jason Brownlee, *8 Inspirational Applications of Deep Learning*, MACHINE LEARNING MASTERY (July 14, 2016), https://machinelearningmastery.com/inspirational-applications-deep-learning/.
155 DOMINGOS, *supra* note 2, at 108.
156 *Id.*
157 *Id.* at 109–11.
158 GOODFELLOW ET AL., *supra* note 2, at 13–15.
159 *Id.* at 15.

descending from a mountain.[160] If the hiker takes the steepest path down from any given point, she will continue to descend the mountain to the bottom, unless stuck in local minimum, a spot from where all paths go up.[161] The second wave, "connectionism," overcame limitations of the first wave, linear models.[162] Back-propagation techniques adjusted for errors throughout the network, allowing networks to learn more complex functions.[163] Distributed representation permitted more efficient representation of concepts within networks and gave greater power to compare multiple attributes of concepts.[164] The third wave, deep learning, uses some new software techniques, such as "greedy layer-wise pretraining," to create deeper networks with greater learning power.[165] But the new deep learning networks continue to use decades-old techniques such as gradient descent and back-propagation.[166] Much of the increased power of machine learning in this era of deep learning comes from the general development of computer network capability and deployment, along with availability of much larger sets of data.[167] Datasets used to train networks are exponentially larger than in past decades. The Street View House Numbers dataset, for example, includes over 600,000 images, and other datasets include up to "tens of millions of examples."[168] The computing resources devoted to deep learning are correspondingly extensive, both with respect to hardware[169] and software to run training software over a large

160 DOMINGOS, *supra* note 2, at 291–92.

161 *Id.* at 110.

162 GOODFELLOW ET AL., *supra* note 2, at 14–15, 18.

163 *Id.* at 15–18 (noting limitation of linear models and development of backpropagation).

164 *Id.* at 150, 541.

165 *Id.* at 18–19.

166 *Id.* at 14–15, 18–19.

167 *See, e.g.*, *id.* at 461 (discussing how increases in computing resources enables improvements in using neural nets in speech recognition: "Later, with *much larger and deeper models* and much larger datasets, recognition accuracy was dramatically improved by using neural networks to replace [Gaussian Mixture Models] for the task of associating acoustic features to phonemes.").

168 *Id.* at 19–22.

169 *See id.* at 445–47 ("Because the size of neural networks is of paramount importance, deep learning requires high performance hardware and software infrastructure."). Goodfellow describes progress from using fast central processing unit implementation to use of graphics processing units (GPUs). *Id.* ("Together, this results in graphics cards having been designed to have a high degree of parallelism and high memory bandwidth," which fit well with the needs for deep learning.).

network of computers.[170]

A network responds to training examples and responds with its output to the classification question (such as "cat," or "not cat"). The network then receives feedback on whether the response was correct and adjusts its internal weights accordingly. When the network is performing with the training examples to a satisfactory level, it can then be tested on a different set of examples. This reconfiguration of the network can be seen several different ways. In one sense, the network has adjusted in response to feedback on its performance, hence it has "learned." Put in terms of mathematics (which can be seen as the art of detecting patterns), the network now has determined responses to various types of examples, so it has configured a function. Put another way, the network has reified an algorithm to perform that function, so the network has written a computer program. This characterization captures one great advantage of machine learning: rather than having a person write a program to recognize pictures of cats (which may be too detailed a task for humans), the network itself writes the program, in a sense.

The program will not mirror the human thought process – that is well-illustrated by the concept of "generative adversarial networks."[171] Suppose a network has attained a high success rate on identifying pictures of cats. Another machine learning program could run on the output of the cat classifier, identify which features it deems salient, and produce "adversarial examples."[172] It might take a picture of a cat and change only a few key pixels so that a human would still readily identify it as a picture of a cat. The cat classifier would be fooled by the small change to classify it as not a picture of a cat. Generative adversarial networks remind us that machine learning may accomplish certain tasks with high proficiency, but that does not mean it has formed robust human concepts like "cat" applicable in many contexts.[173]

---

170 *See id.* at 448–49 ("In many cases, the computational resources available on a single machine are insufficient. We therefore want to distribute the workload of training and inference across many machines.").

171 *Id.* at 546–47 ("In this approach, a generative model is trained to fool a feedforward classifier.").

172 *Id.* at 268–69.

173 Generative adversarial networks pose practical security and reliability issues for the use of machine learning, and software generally, in a number of contexts. *See* Cory Doctorow, *Generative adversarial network produces a "universal fingerprint" that will unlock many smartphones*, BOING BOING (Nov. 15, 2018) (citing PHILIP BONTRAGER ET AL., DEEPMASTERPRINTS: GENERATING

Neural networks in the brain inspired, by analogy, artificial neural networks.[174] There is another analogy between the idea of a network learning through exposure to examples, with feedback, and the process of the common law. Edward Levi's classic *An Introduction to Legal Reasoning* described the course of common law rule development as courts formulating rules to explain the outcomes of cases such that a rule could be altered when it could no longer yield a just result, and so would be changed accordingly.[175] That resonates with the process of artificial neural networks, when the feedback from the outcome causes the nodes in the network to change their weights until gradually the outcome may be different from a similar case. Ronald Dworkin has likewise described the concept of the common law gradually adjusting as it is challenged by case after case.[176] As some legal theorists view jurisprudence, drawing on the work of John Rawls, an ideal judge considers all relevant matter until the judge reaches a state of reflective equilibrium.[177] An artificial neural network can be trained until it reaches a state of equilibrium with the training set of cases. In fair use in particular, the broad four-factor rule requires cases to give it content.

One issue would be acquiring sufficient cases to train the network. Deep learning has achieved impressive results in part due to the use of big data sets: "[a]s of 2016, a rough rule of thumb is that a supervised deep learning algorithm will generally achieve acceptable performance with around 5,000 labeled examples per

MASTERPRINTS FOR DICTIONARY ATTACKS VIA LATENT VARIABLE EVOLUTION (2018)), https://boingboing.net/2018/11/15/masterprints.html.

174 *See, e.g.*, GOODFELLOW ET AL., *supra* note 2, at 13 (internal citation omitted) ("Some of the earliest learning algorithms we recognize today were intended to be computational models of biological learning, i.e. models of how learning happens or could happen in the brain . . . While the kinds of neural networks used for machine learning have sometimes been used to understand brain function, they are generally not designed to be realistic models of biological function.").

175 Edward H. Levi, *An Introduction to Legal Reasoning*, 15 U. CHI. L. REV. 501, 501–03 (1948).

176 *See* Michael Gentithes, *Precedent, Humility, and Justice*, 18 TEX. WESLEYAN L. REV. 835, 891 (2012) (citing RONALD DWORKIN, LAW'S EMPIRE 400 (1986)) ("This claim is, of course, largely similar to the argument that the rules of society 'work themselves pure' through common law judicial decision making.").

177 Lawrence B. Solum, *Legal Theory Lexicon 069: Reflective Equilibrium*, LEGAL THEORY BLOG (May 15, 2011) https://lsolum.typepad.com/legal_theory_lexicon/2011/05/legal-theory-lexicon-069-reflective-equilibrium.html.

category, and will match or exceed human performance when trained with a dataset containing at least 10 million labeled examples."[178] The lower threshold of 5,000 examples might be met simply by using reported fair use cases augmented as discussed above, but would fall short of the 10 million required to match human performance. For some purposes, "acceptable performance" could suffice where the stakes are relatively low. In addition, human performance might not be as high a standard in classifying fair use cases as in, say, recognizing pictures of cats or celebrities. Humans appear to have a natural facility for recognizing faces. With fair use, as noted above, humans often disagree.

Assigning weight to cases used to train the network presents another issue. *Sony*, *Harper & Row*, and *Campbell* are just three of thousands of reported Supreme Court cases, but they are the governing precedent on how to interpret fair use. Other cases from the lower courts, in legal theory, simply serve to interpret the big three – or if they precede the big three, have diminished precedential value, because a subsequent Supreme Court case would have more precedential value than an earlier lower court opinion. That is viewing things through the lens of jurisprudence. One could also view the reported cases as simply a sample of how judges view fair use, as a means to predict the outcome of future cases. Through that lens, the big three might be accorded more weight but would not dominate. Indeed, perhaps some lower court case better predicts how fair use will be applied in the future.

Along the same lines, one could argue about whether certain cases should be excluded from the sets, either as training examples or test examples. The lower court opinions reversed by the Supreme Court would seem inapt examples to train our fair use learner. Less clear would be the many district court opinions that have been reversed by courts of appeals. Some of those very likely might have been affirmed by a different set of judges. Along the same lines, one could select a good number of fair use cases, especially predating *Sony* that might be decided differently today.

As noted above, a challenge for using machine learning in fair use (or any legal domain including cases) is likely to be the curse of dimensionality – the considerable number of variables that could enter into legal analysis. Networks do have some techniques to

178 GOODFELLOW ET AL., *supra* note 2, at 19–20; *see also* Frank Fagan, *Big Data Legal Scholarship: Towards a Research Agenda and a Practitioner's Guide*, 20 VA. J. L. & TECH. 1 (2016) (discussing applications of big data in the law).

reduce dimensionality. One of the most powerful is the autoencoder. An autoencoder takes a set of inputs and puts out the same set of inputs – which sounds pretty mundane.

> The key is to make the hidden layer much smaller than the input and output layers, so the network can't just learn to copy the input to the hidden layer and the hidden layer to the output, in which case we may as well throw the whole thing out. But if the hidden layer is small, something interesting happens: the network is forced to encode the input in fewer bits, so it can be represented in the hidden layer, and then decode those bits back to full size.[179]

Autoencoders thus find more efficient ways to represent data. They find many uses in machine learning and have independent applications in other areas, such as data compression.[180] In natural language processing, an area which will be required to use judicial opinions as data, networks have used autoencoders in such applications in as word embeddings, machine translation, document clustering, sentiment analysis and paraphrase detection.[181]

Autoencoders can be stacked on top of each other, so that each one uses the output of the previous autoencoder. The first autoencoder can encode features at one level of abstraction, meaning the next autoencoder can encode features slightly more abstract. In image recognition, one level could be edges, one could be edges that are part of facial features, the next could be entire features like noses or mouths, and the top layer could identify entire faces. Indeed, the Google Brain network used stacked autoencoders and other network components to recognize cats.[182] By analogy (while recognizing the gulf between these domains), autoencoders could be used to

179 DOMINGOS, *supra* note 2, at 116.

180 *Cf.* Cory Doctorow, *Compression Could be Machine Learning's "Killer App,"* BOING BOING (Oct. 18, 2018) https://boingboing.net/2018/10/18/narrative-sensors.html.

181 *See* Venkata Krishna Jonnalagadda, *Sparse, Stacked and Variational Autoencoder*, MEDIUM (Dec. 5, 2018) https://medium.com/@venkatakrishna.jonnalagadda/sparse-stacked-and-variational-autoencoder-efe5bfe73b64; *see also* Müller & Guido, *supra* note 123, at 131–32 ("*Clustering algorithms*, on the other hand, partition data into distinct groups of similar items.").

182 DOMINGOS, *supra* note 2, at 117; Gideon Lewis-Kraus, *The Great A.I. Awakening*, N.Y. TIMES (Dec. 14, 2016) https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html.

recognize the features of fact patterns.

Another technique that could deal with abstraction is the convolutional neural network. Convolutional neural networks can abstract local features from examples.[183] The best-known application of convolutional neural networks may be in recognizing items in images.[184] Fair use analysis, as noted above, necessarily requires taking bare facts and reducing them with abstractions in order to apply the factors.

Evaluating the performance of the software would be both vital and vexing. Whether a use is fair use or not is often subject to disagreement.[185] In law, the assessment goes beyond the result to evaluating the reasoning supporting the result. One of the greatest policy issues with machine learning is that many (not all, as discussed below) algorithms are black boxes.[186] A network has been reconfigured, effectively writing an algorithm to classify examples, but the algorithm is far too complicated to simply read and evaluate. Some applications apply several machine learning algorithms to the same problem,[187] and then seek the best answer by comparing the

183 *See* SEAN GERRISH, HOW SMART MACHINES THINK 135–39 (2018).

184 *See, e.g.*, *id.* at 135 ("This pattern – convolutional layers followed by fully connected layers – turns out to be very common in networks used for image recognition.").

185 Some have called in general for more rigorous evaluation of AI and the law projects. *See* Jack G. Conrad & John Zeleznikow, *The Role of Evaluation in AI and Law*, 15 INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. PROC. 181, 181–86 (2015); *compare* Richard Tromans, *France Bans Judge Analytics*, ARTIFICIAL LAW. (June 4, 2019) https://www.artificiallawyer.com/2019/06/04/france-bans-judge-analytics-5-years-in-prison-for-rule-breakers/ ("In a startling intervention that seeks to limit the emerging litigation analytics and prediction sector, the French Government has banned the publication of statistical information about judges' decisions – with a five year prison sentence set as the maximum punishment for anyone who breaks the new law."), *with* Daniel L. Chen, *Machine Learning and the Rule of Law*, 2019 SANTA FE INST. PRESS (forthcoming) https://ssrn.com/abstract=3302507 ("Predictive judicial analytics holds the promise of increasing the fairness of law.").

186 McJohn, *Artificial Legal Intelligence*, *supra* note 1, at 244 ("Applied to the legal domain, a neural network would give a result without the reasons for it – a 'black-box' approach that fits poorly with the need for justifications in the legal world."); *see also* Roger A. Ford & W. Nicholson Price II, *Privacy and Accountability in Black-Box Medicine*, 23 MICH. TELECOM. & TECH. L. REV. 1, 18–21 (2016).

187 *See* GÉRON, *supra* note 2, at 181–83 (discussing Ensemble Learning); GOODFELLOW ET AL., *supra* note 2, at 452 (discussing "mixture of experts" approach).

various answers from the several algorithms, which would make it even more opaque as to what the cause of the result was.[188] The potential of undisclosed bias is pervasive in machine learning (indeed, machine learning necessarily requires some bias, in a statistical sense, in the data)[189] and other software. It has been raised in areas from credit-scoring to criminal sentencing. Bias can enter in through the operation of the algorithm (such as incorporating biases in the data) but also through the human-guided process of bringing machine learning to bear on a program. The software is not simply set loose. Rather, the software may be trained on a set of data, then the performance will be evaluated[190] and improvements sought.[191] The evaluation by humans can allow for some biases to creep in.

This area is one of intense interest, and means to make algorithms more transparent may be in the making. But until then, deployment of a fair use algorithm in some contexts (such as where it prevented material from being publicly disseminated) without grounds would raise considerable questions (especially if the government were involved).

One related network learning technique deserves mentioning. Perhaps the greatest difficulty in applying machine learning to the domain of law, as discussed above, is the need for sufficient data and for that data to come in a form amenable to the various techniques of machine learning. Neural networks in particular function better the greater the number of training examples are available.[192] A recent

188 *See* DOMINGOS, *supra* note 2, at 237 ("As it turns out, it's not hard to combine many different learners into one, using what is known as metalearning. Netflix, Watson, Kinect, and countless others use it, and it's one of the most powerful arrows in the machine learner's quiver.").

189 *See id.* at 64 ("Tom Mitchell, a leading symbolist, calls it 'the futility of bias-free learning.' In ordinary life, *bias* is a pejorative word: preconceived notions are bad. But in machine learning, preconceived notions are indispensable; you can't learn without them.").

190 *See* GÉRON, *supra* note 2, at 71–74 (giving practical example of evaluation of a machine learning model's performance).

191 *See, e.g.*, *id.* at 253–306 (chapter on Model Evaluation and Improvement). As the comic xkcd put it:

> "This your machine learning system?
> Yup! You pour the data into this big pile of linear algebra, then collect the answers on the other side.
> What if the answers are wrong?
> Just stir the pile until they start looking right."

Randall Adams, *Machine Learning*, XKCD, https://xkcd.com/1838/.

192 *See, e.g.*, GERRISH, *supra* note 183, at 143 ("As a result, they ended up with 2,000 times the amount of training data they started with, or about 2 billion

project, published in the journal Science in December 2018, provided an important exception to that rule.[193] Computer programs have been superior to humans in playing chess for many years now.[194] In general, those programs – such as the IBM's Deep Blue that defeated chess champion Gary Kasparov in 1997 – relied on learning using many thousands of human-played games as examples, along with heuristics programmed in.[195] AlphaZero, by contrast, was given no more domain knowledge than the rules of chess.[196] It generated its own example games. As Deep Mind put it:

> To learn each game, an untrained neural network plays millions of games against itself via a process of trial and error called reinforcement learning. At first, it plays completely randomly, but over time the system learns from wins, losses, and draws to adjust the parameters of the neural network, making it more likely to choose advantageous moves in the future.[197]

After about nine hours of training, AlphaZero was superior to any other computer chess program. "In one game, AlphaZero made a bold bishop sacrifice, sometimes used to gain positional advantage, followed by a queen sacrifice, which seemed like a colossal blunder until it led to a check mate many moves later that neither Stockfish nor humans saw coming."[198]

---

images with which to train their network. If they hadn't augmented their training data like this, they would have needed to use a much smaller—and less expressive—network.").

193 *See generally,* David Silver et al., *A General Reinforcement Learning Algorithm that Masters Chess, Shogi, and Go Through Self-Play*, 362 Science 1140 (Dec. 7, 2018).

194 *See generally,* Monty Newborn, Kasparov versus Deep Blue: Computer Chess Comes of Age (1997).

195 *Id.*

196 David Silver et al., *Shedding New Light on the Grand Games of Chess, Shogi and Go*, DeepMind (Dec. 6, 2018), https://deepmind.com/blog/alphazero-shedding-new-light-grand-games-chess-shogi-and-go/.

197 Terrence J. Sejnowski, The Deep Learning Revolution 20 (2018).

198 *Id.* (the author commented wryly, "[t]he aliens have landed and the earth will never be the same again"); *see also Blame it on the Robot?*, Mich. Tech. L. Rev., https://mttlr.org/2017/10/blame-it-on-the-robot/ (discussing earlier work on AlphaGo); Cade Metz, *DeepMind Can Now Beat Us at Multiplayer Games, Too*, N.Y. Times (May 30, 2019) https://www.nytimes.com/2019/05/30/science/deep-mind-artificial-intelligence.html.

So in some domains, the sheer power of deep learning may suffice, without the need for training examples.[199] It is difficult, however, to see how that would translate to the domain of fair use analysis. In chess, the rules of chess are given as the sole knowledge. That is sufficient for the network to play games according to those rules, to determine which side wins the game according to those rules, and to give itself feedback. If the only input were the rules of fair use, Section 107 of the Copyright Statute states the four-factor rule, but not in a manner that determines how to apply it. Nor could a program use the section to generate training examples. The rules of cases are not built-in to Section 107 nor do they flow from it automatically. But such pure reinforcement learning is still in early days. Perhaps it could operate with relatively small data sets, which would address one of the issues in adapting machine learning to legal reasoning.

VI. GOOD NEIGHBORS MAKE GOOD ANALOGIES

Neural networks have great power in classifying, but present considerable issues with respect to using them in fair use. Another set of machine learning algorithms rely on the concept of analogy, a basic reasoning process used in legal analysis as well.[200] The nearest neighbor approach, a standby of the analogy approach, simply attempts to identify, from the set of previous cases, which one is most similar to the present case.[201] More broadly, the approach may identify the most similar cases (k-nearest neighbors, meaning the k most similar cases) and use a statistical median of them. Netflix, for example, might attempt to predict a viewer's preferences by finding another viewer with the most similar viewing history and then offering a film the second viewer liked. To match a face in a database of images (such as Facebook might compile), it simply looks for the closest match.[202] One great advantage of the nearest neighbor approach is that it is "lazy."[203] Rather than training a network by

199 Note that Deep Mind, although relying on deep learning, could be viewed as unsupervised machine learning, where the algorithm is not given a specific goal, as opposed to the supervised learning algorithms that, for example, classify examples into predefined categories.

200 *See* DOMINGOS, *supra* note 2, at 178–202 (discussing analogical reasoning in machine learning); Cass R. Sunstein, *On Analogical Reasoning Commentary*, 106 HARV. L. REV 741 (1993) (discussing analogical reasoning in legal reasoning).

201 DOMINGOS, *supra* note 2, at 179–86.

202 *Id.* at 179–80.

203 *Id.* at 179–82.

exposing it repeatedly to examples, a nearest neighbor algorithm simply searches a database of examples to find the nearest match. The technique also does not require the enormous datasets on which the most sophisticated neural networks rely.

Such an algorithm could be used in a number of ways to apply fair use. The most straightforward approach would be to predict whether a case was fair use by simply relying on whether the most similar case had been held to be fair use. Nearest neighbor here gives two related advantages. There is no time spent training, although an algorithm must be crafted that can rapidly search through the database.[204] By the same token, there is no need to implicitly form a general model[205] about what is or is not fair use, rather simply to rely on the closest match. Different weights can also be assigned to the cases in running the program: "examples closest to the test example should count for more."[206] Fair use has been notoriously resistant to general characterization. But finding a close case may often resolve an issue in satisfactory fashion. Other areas of law with flexible rules might likewise be subject to this approach.

The success of this approach would depend critically on the features chosen to match. Not all facts in a case are equally relevant. An advantage of nearest neighbor is the savings in training a network (or using it where training a network would be impracticable). In balance, the feature engineering cannot rely on the network determining which features are determinative or how they should be weighed. Moreover, among the various machine learning algorithms, nearest neighbor is especially likely to be slowed by numerous features in the data, because it must search the entire database and compare with each of those features.[207] So selecting a spare number of features to compare would be necessary, by eliminating features of little relevance or finding a way to combine related features.[208] Preprocessing the data would be key to determining the predictive power of the program.

In some applications, that might not be as imposing a task as would appear at first. To use nearest neighbor to reach a legal

204 *Id.* at 179–80.

205 *Id.* at 179 ("If you want to learn to recognize faces and have a vast database of images labeled face/not face, just let it sit there. Don't worry, be happy. Without knowing it, those images already implicitly form a model of what a face is.").

206 *Id.* at 183.

207 *Id.* at 186.

208 *Id.* at 186–89.

conclusion as to sets of facts, would require first determining the characteristics of facts that would be considered, and processing both the sample cases and the case to be decided into those parameters. As noted above, that is not something any existing application can do. But, per the Netflix example, if all that is required is a prediction of the case (as opposed to legal analysis), then a simpler approach might have acceptable accuracy. To figure out what films might appeal to someone compared to other viewers, we might consider an unmanageable set of facts: their personality, their taste in entertainment, their sense of humor. But an acceptable predictor might consider no facts about the person, rather simply their viewing history, and look for another person with a similar viewing history. It might be that in predicting whether a user is posting material protected by fair use, comparing their posting history to a user (or set of k users) with a similar history might yield an acceptable prediction.

The lazy learning nature of nearest neighbor also might enable its use at early stages of fair use analysis, in finding relevant cases. Presented with a fact pattern, a lawyer or judge will wonder whether there were other fair use cases involving the similar subject matter (be it course web pages, marching band music, fan fiction, etc.) or similar questions about the interplay between factors (where the disfavored use may have somehow increased the market for the work). Nearest neighbor could be used in such a way that relevant features were chosen at run time and so used as a research tool or otherwise to find related precedent. That would depend on the data being compiled in a manner amenable to such flexible searching.

A more complex method of analogy in machine learning is the use of support vector machines (SVM). SVMs work, in essence, by drawing lines between sets of cases.[209] A SVM can separate cases in a two-dimensional graph by drawing a line, in three dimensions by drawing a plane, in four dimensions by drawing a hyperplane, and so on.[210] Legal reasoning often turns on the similar concept of drawing a line between cases (and other geometric analogies like not letting open the floodgates or letting things down a slippery slope). Like legal categorization, SVM's can handle outlier or mislabeled cases, simply by tolerating some error – drawing the line in a way that leaves some examples sitting in the wrong class.[211] Likewise,

209 *See* GÉRON, *supra* note 2, at 145–67.
210 *See id.* at 147–67.
211 *See id.* at 147.

fair use case law contains cases that were wrongly decided, or at least contrary to the present law of fair use.

The question whether machine learning algorithms in the analogy camp are most useful in legal applications may depend on a basic inquiry that remains as yet unsettled in jurisprudence (and psychology, and biology, and philosophy, etc.). To what extent does human reasoning depend on reasoning by analogy? There is a great deal of literature on reasoning by analogy, but no basic agreement even on such matters as what reasoning by analogy is and whether it limits or enables reasoning.[212] A useful way to think about reasoning is the framework of deduction, induction and abduction, from philosopher and logician Charles Sander Peirce.[213] To adapt one of Peirce's examples: if all the beans in a bag are blue, and a bean is from the bag, *deductive* reasoning (logical reasoning to certain conclusions) tells us the bean is necessarily blue. If all the beans we have drawn from another bag are green, *inductive* reasoning (generalizing from examples) suggests that the next bean will probably be green. If all the beans in another bag are pink, and there is a pink bean lying near the bag, *abductive* reasoning ("interference to the best explanation")[214] suggests that the bean came from the bag.[215] The modes of reasoning vary as to reliability and productivity. Deductive reasoning is completely reliable, but has lower productivity, because it cannot give new knowledge beyond what can be logically inferred from given premises. Inductive reasoning may be fairly reliable, but even with many examples may yield an untrue result. The sun comes up in the east every day, but some day will not. A turkey is fed by the farmer every day – until Thanksgiving comes.[216] Induction

212 *See, e.g.*, STEVEN PINKER, THE STUFF OF THOUGHT 253–59 (2007) (discussing various camps on whether analogy constrains or enables human reasoning).

213 *See* Stephen M. McJohn, *On Uberty: Legal Reasoning by Analogy and Peirce's Theory of Abduction*, 29 WILLAMETTE L. REV. 191, 192–93 (1993) [hereinafter McJohn, *Peirce*].

214 Lawrence Solum, *Legal Theory Lexicon: Inference to the Best Explanation (Abduction)*, LEGAL THEORY BLOG (Feb. 17, 2019), https://lsolum.typepad.com/legaltheory/2019/02/legal-theory-lexicon-inference-to-the-best-explanation-abduction.html.

215 McJohn, *Peirce*, *supra* note 213, at 197–98 ("Thus, where induction simply infers that characteristics of a sample will apply to the whole (inferring a general law from particulars), abduction infers an explaining hypothesis from a body of data (what Peirce termed inferring cause from effect): Induction classifies, abduction explains.").

216 DOMINGOS, *supra* note 2, at 61 (discussing Bertrand Russell's view on the

is productive in a limited way, generalizing from new examples. Abduction is the most productive form of reasoning, generating hypotheses that may constitute new knowledge.[217] But abduction is also the least reliable. Those pink beans may come from somewhere else. Many scientific hypotheses must be discarded, as must many daily explanations.

Peirce's framework is useful, but not fully established. We use induction widely, but philosophers, most notably Hume and Popper, have shown that induction lacks the sort of solid demonstration of deductive logic.[218] Abduction has received even less definition or justification. Abduction seems to capture a basic human reasoning ability. Even cognitive tasks that may be unconscious, such as visual perception, involve making small abductive inferences[219] but how we do it or what its logical structure is are yet to be elucidated. The overall framework is nevertheless a useful way to think about reasoning.

Analogical reasoning can be seen as a mixture of induction and abduction.[220] Analogies between closely similar examples (such as two factually similar cases) seem like induction. Analogies between rather different examples (cases with similarities at a more conceptual level) seem more like hypotheses. Whether algorithms that depend on similarity, like nearest neighbor, are useful in fair use will depend on whether that sort of similarity does play a reliable role in fair use. It could be (and courts insist on the fact-bound nature of fair use and its resistance to "mathematical" formulation) that abduction's leap of insight must always be available. It may be that machine learning algorithms can go beyond the sort of mechanical analogy making of nearest neighbor, into the creative hypothesis-formulation of abduction. But that would raise the issue of reliability.

---

problem of induction).

217 *See* McJohn, *Peirce*, *supra* note 213, at 200, 202 ("Abduction is synthetic, seeking to form general laws and bring order to apparently disconnected ideas. An abductive inference provides a way to see order and unity in what was previously confusing.").

218 Leah Henderson, *The Problem of Induction*, STAN. ENCYCLOPEDIA PHIL., https://plato.stanford.edu/entries/induction-problem/ (last updated Sept. 21, 2019).

219 *See* McJohn, *Peirce*, *supra* note 213, at 203 ("Under Peirce's view, we use abduction in all types of cognition, from the commonplace to the most abstract. He regarded the extreme case of abduction to be the interpretation inherent in the perceptive judgment.").

220 *See id.* at 209.

People make hypotheses with ease, but many of those turn out to be inaccurate. Some sort of check would have to be built in to software that approached that same terrain.

VII. DECISION TREES – TRANSPARENT LEARNERS

A decision tree provides a model of decision-making analogous to a flow chart. To give an unusually clear-cut example, a decision tree could classify animals with such questions as: does it have a backbone? Does it have fur? Does it give birth to live young? The decision tree classifies the example according to its features and returns a marsupial, or in a different path, a grass. The decision tree can be built by classifying the data to create the branches of the trees. The algorithm chooses a likely feature, uses it to divide the set, then uses another feature to divide the set, then continues the process until a maximum chosen depth is reached or there remains no feature that can reliably divide the set.[221] A set of cases may not divide perfectly into nodes at the ends of the decision tree branches. The tree can still be used to estimate the probability that an example belongs in a particular category by using the ratio of classes that end up in a particular node.[222] For example, Microsoft's Kinect video game controller uses a decision tree to "figure out where various parts of your body are from the output of its depth camera; it can then use their motions to control the Xbox game console."[223]

A decision tree could conceivably track the factors identified by the Copyright Act and big three fair use judicial decisions: Is the use commercial? Is the use transformative? Is there a negative effect on the market? Was the work unpublished? Different features might be used on various branches. The tree would be reliable if it produced leaf nodes[224] with reliable classification. That would mean that fair use cases could be broken into sets, such as that a noncommercial, transformative, no-negative-market-effect use was usually fair, even

221 GÉRON, *supra* note 2, at 171 ("Once it has successfully split the training set in two, it splits the subsets using the same logic, then the sub-subsets and so on, recursively. It stops recursing once it reaches the maximum depth (defined by the max_depth hyperparameter), or if it cannot find a split that will reduce impurity.").

222 *Id.* ("A Decision Tree can also estimate the probability that an instance belongs to a particular class *k*: first it traverses the tree to find the leaf node for this instance, and then it returns the ratio of training instances of class *k* in this node.").

223 DOMINGOS, *supra* note 2, at 88.

224 A leaf node is the end of a decision branch. *Id.* at 86.

if the copied work was unpublished. Alternatively, features could be defined by the facts of cases: Is the work software? Was the use reverse-engineering? Did defendant obtain the work lawfully? These categories might yield predictive classes. Features beyond the record of the case could have predictive value. A Decision tree built from statistics drawn from "six observable characteristics," such as circuit of origin and general legal area of a case, had a higher success rate in predicting affirmance or reversal in Supreme Court cases than a panel of legal experts (although the experts' predictions of the votes of individual judges were slightly better).[225]

Decision trees have a great advantage for use in the legal domain. A difficult policy issue with machine learning and much other software is the black box problem: programs make decisions or predictions without stating the underlying reasoning.[226] A complex neural net may classify examples from a set of data but will not state the reasons why. Decision trees, by their very nature, provide the structure of the decision process.

> For example, if a neural network says that a particular person appears on a picture, it is hard to know what actually contributed to this prediction: did the model recognize that person's eyes? Her mouth? Her nose? Her shoes? Or even the couch that she was sitting on? Conversely, Decision Trees provide nice and simple classification rules that can even be applied manually if need be[.][227]

If machine learning could produce a reliable decision tree for fair use, then, it would have perfect transparency.[228]

225 Andrew D. Martin et al., *Competing Approaches to Predicting Supreme Court Decision Making*, 2 PERSP. ON POL. 761 (2004). The characteristics used were "(1) the circuit of origin for the case; (2) the issue area of the case, coded from the petitioner's brief using Spaeth's protocol; (3) the type of petitioner (e.g., the United States, an injured person, an employer); (4) the type of respondent; (5) the ideological direction of the lower court ruling, also coded from the petitioner's brief using Spaeth's protocol; and (6) whether or not the petitioner argued the constitutionality of a law or practice." *Id.* at 762.

226 *See, e.g.*, Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 889, 907 (2018) ("To be sure, we may be able to tell what the AI's overarching goal was, but black-box AI may do things in ways the creators of the AI may not understand or be able to predict.").

227 GÉRON, *supra* note 2, at 170.

228 Note that multiple decision trees may be used, which improves the overall predictive power, in the Random Forest technique. The use of many trees,

"If", because the decision tree, compared to other algorithms, is more demanding in the form of data on which it can operate. For example, "the main issue with Decision Trees is that they are very sensitive to small variations in the training data."[229] Whatever the features are for legal data, they may be subject to variation. A case could be decided differently. A use might be considered transformative or not depending on which circuit's law applies. The reliance on reasoning by analogy between the facts of cases introduces flexibility, but also complicates categorization. The characterization of features in the training data – a key issue for machine learning and the law generally – is especially acute given the somewhat brittle nature of decision trees, compared to neural networks, with their gradual adjustment to the data.

Having said that, the method compensates by offering the ability to use probabilities. If the cases can be sorted into different nodes, then the decision process would predict whether the case is fair use but qualify that prediction with a probability by comparison with other cases generally sharing the same set of attributes.

Decision trees, then, have the disadvantage of being especially dependent on data with features amenable to classification, and sensitivity to small variation. That might render them unfit for handling fair use generally until efforts to preprocess legal data are more successful (whether by tagging or finding attributes by clustering or other approaches). But decision trees have the advantage of transparency, along with the option of predicting by probability. That could mean that they could be deployed in narrowly defined areas even before the battles with data are won. An automated fair use check for music uploads, for example, might have good predictive power with some definable features drawn from the case law (is the work copyrighted, how much is uploaded, does the uploader have a history of infringement, what is the difference between the works, if any). The algorithm would no doubt make errors, but the transparency of decision trees would make them more easily subject to dispute, such as in a counter-takedown process.[230] If a post was

however, greatly reduces the ability to trace the decision-making process. *See id.* at 181, 189–90.

229 *Id.* at 177.

230 Section 512 of the DMCA provides internet service providers immunity from copyright infringement by their users, provided the internet service provider, among other things, has a system in place to respond to takedown notices from copyright owners and counter takedown notices from the user that posted the content. *See, e.g.*, Kathleen O'Donnell, *Lenz v. Universal Music Corp.*

taken down, the poster could get notification along with the specific reasons why the post was taken down. Although lawyers might recoil, the notice could even contain a probability determination ("It is our experience that 85% of similar posts are infringing."). The poster could then have an opportunity to respond. Decision trees may prove to have practical value in lower stakes, repeat player settings.

VIII. UNSUPERVISED LEARNING: PATTERN DETECTION

Supervised machine learning is used to classify examples into pre-determined categories, such as potentially classifying examples as fair use/not fair use. The algorithm "learns" to detect patterns correlated to the specified classification. Unsupervised machine learning, by contrast, leaves it to the algorithm to find patterns of any sort, to see if useful information can be gained. Data mining is the classic example, where unsupervised learning is used on a large body of data to find patterns. Unsupervised machine learning might likewise reveal some useful patterns within fair use. As noted above,[231] unsupervised machine learning can find clusters in data and even operate on those clusters to find more abstract clusters.[232] Unsupervised machine learning may detect patterns: for example, "principal component analysis," used for determining which aspects of data contribute most to its distribution, can identify the factors that carry the greatest weight in functions.[233] A similar algorithm "has a surprising ability to zero in on the most important dimensions of complex data." [234] It might show which factors have the biggest say in fair use – and whether those factors are the ones that the courts expressly consider.

A common view is that fair use is replete with fact-specific case law rather than common principles.[235] Under this view, some

---

and the Potential Effect of Fair Use Analysis Under the Takedown Procedures of § 512 of the DMCA, 2009 DUKE L. & TECH. REV., 1, ¶10, ¶15 (2009).

231 *See* discussion of using unsupervised machine learning to preprocess data *supra* notes 20–21.

232 DOMINGOS, *supra* note 2, at 210.

233 *Id.* at 213–214 (using the example of examining the locations of businesses to find the underlying pattern, that they were arranged in proximity to the principal street in a town).

234 *Id.* at 217 (discussing Isomap, "[o]ne of the most popular algorithms for nonlinear dimensionality reduction").

235 *See, e.g.*, Michael W. Carroll, *Fixing Fair Use*, 85 N.C. L. REV. 1087, 1090 (2007) ("While the doctrine's attention to context has many salutary attributes, it is

cases might be clearly fair use (quoting a few lines in a book review) or clearly not fair use (selling unauthorized copies of movies for sheer profit) but the cases in the middle are not reliably predictable. A refined version is that fair use, as legal doctrine, may be incoherent,[236] but may be more predictable when analyzed in light of social and cultural patterns.[237] Others argue that fair use is no less predictable than other areas of case-bound law.[238]

There are several types of patterns which might be disclosed. It could be that the four factors, as elucidated by the courts, are applied in a consistent fashion, even if the language of the various judicial opinions masks that somewhat. As noted above, fair use cases may divide not so much by the four statutory factors but by "policy-relevant clusters."[239] It could also be the case that legal doctrine is vague enough to lack consistent predictive power, but that underlying social and cultural patterns drive the application of fair use in a consistent manner.[240] Other patterns might also be found. Reading fair use cases (as with many intellectual property cases) raises questions about such issues as home-town advantage.

---

so case-specific that it offers precious little guidance about its scope to artists, educators, journalists, Internet users, and others who require use of another's copyrighted expression in order to communicate effectively.").

236 Madison, *supra* note 81, at 1586–87 ("[T]he Supreme Court's formal jurisprudence has encouraged the courts of appeals, and presumably the district courts following their lead, to abstract the fair use inquiry to the point of incoherence").

237 *Id.* at 1622 ("I review what social and cultural patterns are and describe how an emphasis on those phenomena can be reconciled with the language of the current fair use statute. The second section below compares the generalized patterns model developed in the first section and applies it more broadly to fair use case law since enactment of the current copyright statute.").

238 Samuelson, *supra* note 82, at 2541–42 (footnote omitted) ("[F]air use law is both more coherent and more predictable than many commentators have perceived once one recognizes that fair use cases tend to fall into common patterns, or what this Article will call policy-relevant clusters. The policies underlying modern fair use law include promoting freedom of speech and of expression, the ongoing progress of authorship, learning, access to information, truth telling or truth seeking, competition, technological innovation, and privacy and autonomy interests of users.").

239 *Id.*

240 Madison, *supra* note 81, at 1622 ("I review what social and cultural patterns are and describe how an emphasis on those phenomena can be reconciled with the language of the current fair use statute. The second section below compares the generalized patterns model developed in the first section and applies it more broadly to fair use case law since enactment of the current copyright statute.").

An adaptation of a Manhattan-centric cover of *New Yorker* magazine, used as an advertisement for the film *Moscow on the Hudson*, failed to find support in fair use – when the case was decided in Manhattan, in the Southern District of New York.[241]

For fair use – and legal scholarship generally – the prospect of detecting patterns is of great interest. As noted above, legal reasoning relies on analogy, and yet the nature of analogical reasoning is elusive. Identifying the patterns that correspond to what judges and lawyers deem to be strong analogies could throw a lot on analogical reasoning. It could also bring yet another round of legal realism by showing what is going on underneath the hood of the legal engines of reasoning. That said, we must recall that one of the great challenges of machine learning is detecting patterns that do not correspond to any real phenomenon, such as overfitting the data to develop descriptions that are actually too complex. If patterns are detected, whether indicating legal, social, or cultural patterns, then appropriate skepticism should apply. The finding of patterns is a suggestion, not a determination: "It's even been said that *data mining* means 'torturing the data until it confesses.'"[242] With fair use in particular, the perhaps limited number of cases available to train the network may raise the risk of error.[243] As with other forms of machine learning, a detected pattern can be tested against sets of data other than the training set to see if the same pattern is found.[244] Unsupervised learning is often used as an exploratory approach, that may be a fruitful source of hypotheses, but those hypotheses must be further tested.[245]

IX. MECHANICAL FAIR USE PROHIBITED?

Legal issues may arise with the implementation of a fair use algorithm, especially depending on its specific deployment (who uses it and for what purpose, whether it simply is a research tool or effectively makes a decision, and who is affected by the use).

241 Steinberg v. Columbia Pictures Indus., 663 F. Supp. 706, 714 (S.D.N.Y. 1987).

242 DOMINGOS, *supra* note 2, at 72–73.

243 *See* GÉRON, *supra* note 2, at 22–23.

244 DOMINGOS, *supra* note 2, at 75 ("So how do you decide whether to believe what the learner tells you? Simple: you don't believe anything until you've verified it on data that *the learner didn't see*.").

245 Müller & Guido, *supra* note 123, at 132 ("As a consequence, unsupervised algorithms are used often in an exploratory setting, when a data scientist wants to understand the data better, rather than as part of a larger automatic system.").

Two issues are particularly worth noting. In addition to technical considerations of using machine learning to assess fair use, it is at least possible that legal considerations could limit the practice. A recent Eleventh Circuit case, *Cambridge University Press v. Patton*, squarely held that it was error for a court to rely on a "mathematical formula" to assessing the factors in fair use.[246] The court emphasized, as have others, that "fair use is not a mechanical determination."[247] Read for everything it's worth, that would bar machine learning assessment of fair use, at least by judges.

Machine learning is the implementation of statistics on a large scale. It may remain a moot point until judicial decisions are handed over to computers. More realistic is whether such a bar could extend to private parties. The Digital Millennium Copyright Act gives internet service providers immunity for infringement by their users, provided, among a number of conditions, that the service providers set up a process to respond when copyright owners send take-down notices.[248] A copyright claimant may send a take-down notice only in good faith.[249] The *Lenz* case notably held that a take-down could

246 Cambridge Univ. Press v. Albert, 906 F.3d 1290, 1300–01 (11th Cir. 2018) ("As the Supreme Court has explained and as we reiterated in Cambridge II, 'the four statutory factors may not be treated in isolation, one from another. All are to be explored, and the results weighed together, in light of the purposes of copyright.' We emphasized that 'fair use is not a mechanical determination,' and that a court must 'weigh[] ... the four factors in light of the facts of a given case,'. To be sure, the district court described its arithmetic weights as 'initial' and 'approximate,' and it stated that it would 'adjust[]' them when it found a 'noteworthy strength or weakness' among the factors. But the district court made such adjustments only four times, each time to bolster the importance of the third factor's weighing against fair use. And, on those four occasions, the district court did nothing to adjust the other factors in the overall fair-use calculus. We conclude that the district court's quantitative rubric was an improper substitute for a qualitative consideration of each instance of copying in the light of its particular facts.") (quoting Cambridge Univ. Press v. Patton, 769 F.3d 1232, 1260 (11th Cir. 2014)).

247 *Id.* The decision can also be seen as addressing the question, whether fair use analysis can be structured with specific rules applicable to common fact settings. *See* Niva Elkin-Koren & Orit Fischman-Afori, *Rulifying Fair Use*, 59 ARIZ. L. REV. 161, 163–65 (2017) ("Can fair use be rulified in this manner? A recent Eleventh Circuit decision brought this issue to the forefront of legal discourse. In Cambridge University Press v. Patton, the court repudiated the attempt made by the lower court to offer a rule-like elaboration of fair use in the context of an educational e-reserve system.") (citation omitted).

248 *See* 17 U.S.C. § 512(c) (2012).

249 *Id.* On automated copyright enforcement generally, see Maayan Perel & Niva Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, 19 STAN.

be sent in bad faith if there was a failure to consider whether the material was posted as fair use.[250]

One might argue then, that *Cambridge University Press*'s disapprobation of a mathematical approach to fair use might mean that using machine learning to vet take-down notices before sending would be bad faith. Take-down senders, attempting to monitor and protect thousands of copyrighted works, are hardly to be held to the same standard as federal judges deciding individual cases. But there would be a requirement that the software meet some standard of assessing fair use, as opposed to a mere formality.[251] At a deeper level, a network trained on many cases to discern and classify examples of fair use is different than the approach disavowed in *Cambridge University Press*. There (at least in the appellate court's view), the trial court set mechanical numbers to weigh the various factors and applied them to the instances before it, without consideration of the interaction of the factors and the goals of fair use.

A network trained on cases to classify examples would necessarily have discerned in case patterns how those factors are interrelated. As opposed to mathematically applying preset weights, ignoring the policies of fair use, a reliable classifier would mathematically have adjusted the weights within the network to continue applying those policies as previous cases have (for better or worse). Nor is using machine learning on a problem a mechanical process. As with other software, the performance and modelling of a machine learner is evaluated and then the software adjusted to attempt to improve performance.[252] It would seem that automated fair use need not be mechanical and so could be consistent with fair use jurisprudence, notwithstanding the protean nature of fair use.

Another possible issue is the use of personal data. The European Union's General Data Protection Regulation ("GDPR") raised substantially the regulations of data online, prompting websites to modify their processes such that websites now tell us they use cookies and often ask for our consent, as many of us have by now experienced.[253] The GDPR restricts automated decisions

TECH. L. REV. 473 (2016).

250 *See* Lenz v. Universal Music Corp., 801 F.3d 1126 (9th Cir. 2015).

251 An analogous issue is whether "machine opinion evidence," the output of neural networks, should be admissible as evidence. *See generally* Curtis E.A. Karnow, *The Opinion of Machines*, 19 COLUM. SCI. & TECH. L. REV. 136 (2017).

252 *See, e.g.*, GÉRON, *supra* note 2, at 253–311 (chapter on Model Evaluation and Improvement).

253 *See* Sean O'Brien, *GDPR: Don't Forget to Bring a Towel*, BOING BOING (May

concerning individuals based on data,[254] albeit with numerous expansive exceptions.[255] The regulation is in early days, so its scope and exceptions remain to be fleshed out.[256] It may well be that the GDPR does not come into play, where decisions are made automatically and may involve some personal data but are not in the nature of profiling, where the data is used to make decisions about the individual[257] (automated decision-making about whether a post by an internet user was fair use would likely focus on the use itself, and not the post, thereby avoiding the sort of user profiling that likely triggers the GDPR).[258] But the copyright holders have been successful in persuading the European Union to tighten restrictions

25, 2018), https://boingboing.net/2018/05/25/gdpr-dont-forget-to-bring-a.html.

254 Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC, 2016 O.J. (L 119) 1, 46 (EU) ("The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.").

255 Maja Brkan, *AI-Supported Decision-Making Under the General Data Protection Regulation*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, *supra* note 40, at 3–4 (2017) ("The GDPR, in its Article 22, prohibits automated individual decision making, including profiling. On the first impression, it seems that this provision strongly protects individuals and potentially even hampers the future development of AI in decision-making. However, it can be argued that this prohibition, containing numerous limitations and exceptions, looks like a Swiss cheese with giant holes in it.").

256 *See, e.g.*, Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You are Looking For*, 16 DUKE L. & TECH. REV. 18, 20–21 (2019) ("There has been a flurry of interest in a so-called 'right to an explanation' that has been claimed to have been introduced in the General Data Protection Regulation (GDPR).").

257 On legal responsibility for AI decisions more broadly, see Ignacio N. Cofone, *Servers and Waiters: What Matters in the Law of A.I.*, 21 STAN. TECH. L. REV. 167, 170 (2018) ("Who should be responsible for accidents caused by self-driving cars? Who owns the copyright over work created by an algorithm? Do algorithms have free speech? Can an A.I. agent be responsible for a crime? Can it be an accessory to a crime?").

258 *Cf.* Maja Brkan, *AI-Supported Decision-Making Under the General Data Protection Regulation*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW, *supra* note 40, at 6 ("These examples demonstrate that the provision of the Data Protection Directive seemed to focus mostly on instances of profiling based on automated processing of data, not including other types of automated decision-making involving processing of personal data.").

on even linking to copyrighted content, to the extent that Google may shut down Google News in some European countries.[259] Future copyright legislation in Europe and elsewhere may take on the question of automated assessment of fair use.[260] In some contexts, safeguards should be in place to protect against the particular vagaries of machine learning.

X. INTELLECTUAL PROPERTY DOMAINS AMENABLE TO MACHINE LEARNING

Thinking about how various machine learning approaches would apply to fair use raises a question about whether other areas of intellectual property might be more or less suitable domains. Two possible applications seem to have more bounded data structures, which might make using machine learning more realistic.

Patent law, in general, could seem even iffier for machine learning, with its preference for strictly defined, concrete inventions.[261] Machine learning, however, can learn to classify by using examples.[262] Patent law, in a sense, defies categorization. A claimed invention must be deemed patentable or not; this is a classification problem. The standard for patentability (unlike copyright, trademark, or any other type of intellectual property) requires novelty and nonobviousness.[263] In short, every patentable invention, by the standard of patentability, should be something unique. Classification programs should be able to deal with unprecedented examples. But a patentable invention should be not

259 Cory Doctorow, *We lost the fight for balance in the EU's Copyright Directive, but here's what we won*, BOING BOING (Apr. 18, 2019), https://boingboing.net/2019/04/18/we-fight-on.html; *see also* Paris Martineau, *The UK's Tech Backlash Could Change the Internet*, WIRED (Apr. 9, 2019), https://www.wired.com/story/uk-tech-backlash-could-change-internet/.

260 As Danny O'Brien, international director of the Electronic Frontier Foundation (a digital rights nonprofit group that opposed the bill) and opponent of the bill, said about proposed (as yet unsuccessfully) strict online copyright rules in Europe, "[t]here's no way that those algorithmic filters are going to be able to decide that something is fair use, parody, a meme or a mash-up." Adam Satariano, *Tech Giants Win a Battle Over Copyright Rules in Europe*, N.Y. TIMES (July 5, 2018), https://www.nytimes.com/2018/07/05/business/eu-parliament-copyright.html.

261 *See* Alice Corp. v. CLS Bank Int'l, 573 U.S. 1090, 222–25 (2014) (limiting patents on abstract inventions, particularly applicable to software inventions).

262 *See, e.g.*, GÉRON, *supra* note 2, at 79–102 (chapter on classification using machine learning).

263 35 U.S.C. § 103 (2012).

just unprecedented (novel) but also significantly different from any previous invention (nonobvious).

In another respect, however, patent law channels inventors toward structuring data in a way that copyright does not. An author gets copyright automatically upon creating a work,[264] and may choose to register the work by filling out a form and depositing copies of the work.[265] An inventor, however, is entitled to a patent only if she (or, more commonly, her patent lawyer) drafts patent claims which distinctly claim the invention and differentiate it from previous technology.[266] That previous technology may be described in publications, in products or services available to the public, or – in the most commonly applied category – the claims of previous patents in relevant technology.[267] When patent examiners consider whether a claimed invention is new and nonobvious, the examiner looks primarily, often exclusively, at the claims of previous relevant patents.[268] A patent claim is one sentence long, albeit often one horrifically complex, opaque sentence.[269]

The determination of patentability, then, often consists of comparing the inventor's one-sentence-long claim to other patent claims. That is more bounded than fair use, which involves

264 17 U.S.C. § 102(4)(a) (2012) ("Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.").

265 *Id.* § 408 ("At any time during the subsistence of the first term of copyright in any published or unpublished work in which the copyright was secured before January 1, 1978, and during the subsistence of any copyright secured on or after that date, the owner of copyright or of any exclusive right in the work may obtain registration of the copyright claim by delivering to the Copyright Office the deposit specified by this section, together with the application and fee specified by sections 409 and 708. Such registration is not a condition of copyright protection.").

266 *See* 35 U.S.C. § 102(b) (2012).

267 *See id.* § 102(a).

268 *See, e.g.*, Jie Qi et al., *Prior* Art, PATENT PANDAS, https://patentpandas.org/resources/prior-art (last visited Oct. 14, 2019) ("Even though prior art officially comes in many forms, patent examiners pretty much only look at existing patent and patent applications for prior art.").

269 *See, e.g.*, Stephen Schott, *An Appeal to the New Patent Office Director: Repeal the Single Sentence Rule*, PATENTLY-O (Sept. 18, 2009), https://patentlyo.com/media/docs/2009/09/schott.sentence.patentlyo.pdf ("But after all this careful work, the drafter may be left with an impenetrable sentence of a length not seen since Matthew begatted Jesus's lineage back to Abraham.").

comparing two works to each other. Copyrighted works can be thousands of sentences, or musical notes, or pixels. Fair use can require comparing one type of work to another, such as the unauthorized use of unpublished letters to a biography about their author, or a play to a movie. So patentability will at least work with more structured data. The claim will likely be a sentence never before used, though linguistics tells us this of perhaps all sentences.[270] Perhaps surprisingly, most sentences in this paper may never have been written before. Yet Google Translate, which now operates on the sentence as the basic unit to translate,[271] could render a fairly good Chinese translation. Whether a sentence claims a patentable invention is quite a different task, but one that would seem to be in the class of engineering problems, as opposed to mere speculation.

The legal determinations of patentability in general are no less complex, raising such abstract issues as patentable subject matter, which in turn requires considering the boundaries of such concepts as abstract ideas, laws of nature, and natural phenomena. These can be skirted, however. One could limit the question to novelty and nonobviousness, thereby leaving out questions of patentable subject matter. Or one could simply not try to define those abstract categories, rather considering only whether the claim at issue was more similar to claims deemed patentable or not.

Not surprisingly, there have been some reported attempts to use machine learning in patent drafting.[272] After looking at the related issues in fair use, it seems likely that machine learning will be most useful where the focus is on patent claims and where the claims are relatively similar, such as where an improvement is made in a well-defined area of technology. Machine learning may be less useful where the differences in claims are more abstract. The flip side

270 *See* STEVEN PINKER, THE LANGUAGE INSTINCT: HOW THE MIND CREATES LANGUAGE 22 (1994) ("[Noam] Chomsky called attention to two fundamental facts about language. First, virtually every sentence that a person utters or understands is a brand-new combination of words, appearing for the first time in the history of the universe.").

271 *See* Wu, *supra* note 6.

272 *See* David Hricik, *Machine Aided Patent Drafting: A Second Look*, PATENTLY-O (Aug. 25, 2017), https://patentlyo.com/hricik/2017/08/machine-patent-drafting.html; Ben Hattenbach & Joshua Glucoft, *Patents in an Era of Infinite Monkeys and Artificial Intelligence*, 19 STAN. TECH. L. REV. 32, 35 (2015) ("Cloem is attempting (*not* satirically, it appears) to use brute-force computing to mechanically compose text for thousands of patent claims covering potentially novel inventions and also to generate defensive publications to prevent others from obtaining patent protection in the same field.").

of patent protection relying on verbal claims is that quite different sectors of technology may fall within a broadly worded patent claim. A patent on an invention to report falls by elderly people was infringed by the Wii video game controller, because both relied on using an accelerometer to detect rapid downward movement.[273] It is unlikely that a claim drafted for the Wii would have been verbally similar to the fall-detector claim. In the same vein, a basic problem with software patents is that broad attempts to claim software inventions yielded claims that subsequently could be read to cover technology well beyond the real scope of the first inventor's invention.

Trademark registration also offers an interesting possibility. One of the two most common reasons a trademark application is rejected by the United States Patent and Trademark Office (USPTO) is that the applicant's mark is too similar to a previously registered mark: in trademark parlance, the applicant's mark is likely to cause confusion with the registered mark.[274] If there is a likelihood of confusion, the mark will be denied registration.[275] Like fair use, likelihood of confusion is a multi-factor test: whether two marks are likely to be confused depends on such matters as how similar the two marks are and how distinctive the marks are, but also how close in the market the marks are, how well known they are, and how sophisticated relevant consumers are.[276] In some respects, however, it is more bounded. First, unlike fair use, likelihood of confusion has a standard, if a fuzzy one: whether the proffered mark is likely to cause confusion. Second, the consideration by the USPTO is more limited than the consideration of similar issue in trademark litigation. If a mark holder sues for trademark infringement, they must prove likelihood of confusion in the marketplace context, considering the actual uses of the two marks.[277] The USPTO limits

273 *See* Joe Mullin, *Jury Finds Nintendo Wii Infringes Dallas Inventor's Patent, Awards $10M*, Ars Technica (Sept. 1, 2017), https://arstechnica.com/tech-policy/2017/09/jury-finds-nintendo-wii-infringes-dallas-inventors-patent-awards-10m/.

274 *See* U.S. Patent & Trademark Office, *Possible Grounds for Refusal of a Mark* (July 11, 2016), https://www.uspto.gov/trademark/additional-guidance-and-resources/possible-grounds-refusal-mark ("The USPTO may be required to refuse registration of your mark on numerous grounds. The most common are: Likelihood of Confusion . . . .").

275 *Id.*

276 *See id.*

277 *See, e.g*, Viacom Int'l v. IJR Capital Invs., L.L.C., 891 F.3d 178, 198 (5th Cir. 2018) (analyzing likelihood of confusion between Krusty Krab restaurant and

its consideration to the categories of use defined in the trademark application (although other factors may permit consideration of real-world facts).[278] So rather than facts about the markets in which the marks operate, the analysis considers the numerical category of registration. The overall analysis is even more circumscribed by the examples considered. Where fair use compares two works, and patent law compares two complex claims, likelihood of confusion in trademark compares two marks. Of the many factors, the similarity of the two marks has been described as the most important factor.[279]

A likelihood of confusion program could be quite useful. Trademark searching is somewhat an art, rather than a science, and such a program would be a useful tool. The imperfections of machine learning might also be quite bearable in the search context. A prospective mark user would get some guidance as to whether their proposed mark would pass muster but need not place great reliance on the program's output. In particular, even if a false positive meant that the applicant needlessly abandoned their plans to use the mark and sought another, that may not be a great loss, if used at the stage before marketing and other trademark-related investment. Other marks are available and a rose by any other name would smell as sweet.[280]

However, a program based simply on the trademark register would only address one part of considering use of a mark. Unregistered marks may also have validity and priority against subsequent marks. Amazon.com learned this when it had to contend with the unregistered mark for Amazon Books, which might have had priority, limited to its area of actual use in

fictional Krusty Krab restaurant of SpongeBob SquarePants cartoon show).

278 TMEP § 1207.01(a)(iii) (Oct. 2018), https://tmep.uspto.gov/RDMS/TMEP/current#/current/TMEP-1200d1e1.html ("Reliance on Identification of Goods/Services in Registration and Application: The nature and scope of a party's goods or services must be determined on the basis of the goods or services recited in the application or registration.").

279 *Cf.* Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CALIF. L. REV. 1581 (2006) (finding that similarity of marks is the greatest factor in infringement likelihood of confusion cases).

280 The supply of potential trademarks is not unlimited. *See* Barton Beebe & Jeanne C. Fromer, *Are We Running Out of Trademarks? An Empirical Study of Trademark Depletion and Congestion*, 131 HARV. L. REV. 945, 951 (2018) ("Specifically, the data present compelling evidence of substantial word-mark depletion, particularly with respect to the sets of potential marks that businesses prefer most: standard English words, short neologisms that are pronounceable by English speakers, and common American surnames.").

Minneapolis.[281] Amazon settled that suit, and other companies have had to adjust after adopting marks only to learn later that a confusingly similar mark had priority.[282] A learner trained only on the trademark register would not take them into consideration.

XI. CONCLUSION

Applying machine learning to fair use faces considerable hurdles. Fair use has generated hundreds of reported cases, but machine learning works best with examples in greater numbers. More examples may be available, from mining the decision-making of websites, having humans judge fair use examples just as they label images to teach self-driving cars, and using machine learning itself to generate examples. Beyond the number of examples, the form of the data is more abstract than the concrete examples on which machine learning has succeeded, such as computer vision, viewing recommendations, and even in comparison to machine translation, where the operative unit was the sentence, not a concept that could be distributed across a document. But techniques presently in use do find patterns in data to build more abstract features, and then use the same process to build more abstract features. It may be that such automated processes can provide the conceptual blocks necessary.

---

281 *Bookstore Battles Internet Amazon Over Use of its Name: Amazon Bookstore Inc. v. Amazon.com Inc.*, 2 ANDREWS COMPUTER & ONLINE INDUS. LITIG. REP. 5 (1999) ("Amazon Bookstore, a Minneapolis bookstore that caters to women, seeks both injunctive relief and damages for what it says is Amazon.com's 'unauthorized and improper use' of the trademark 'Amazon.'").

282 Amy Goetzman, *The Stuff of Herstory: Original Amazon Bookstore to Close*, MINNPOST (June 5, 2008), https://www.minnpost.com/arts-culture/2008/06/stuff-herstory-original-amazon-bookstore-close/ ("The real Amazon received a small cash settlement from the online store, and used the money to keep the business going a little longer."); Keith Bradsher *Apple Settles an iPad Dispute in China*, N.Y. TIMES (July 2, 2012), https://www.nytimes.com/2012/07/02/business/global/apple-settles-an-ipad-trademark-dispute-in-china.html ("A Chinese provincial court said on Monday that Apple had settled a lawsuit there by agreeing to pay $60 million for the legal rights to use the iPad trademark in China, according to Xie Xianghui, a lawyer for the Chinese company involved."); Timothy O. Stevenson, *The Importance of Trademark Clearance Searches in Brand Development*, MARTINDALE (Feb. 27, 2015), https://www.martindale.com/intellectual-property-law/article_Smart-Biggar-Fetherstonhaugh_2193454.htm ("Only then did the agency discover that Motel 6 in fact owns a registered trademark in Canada for 'We'll leave the light on for you,' and consequently the slogan in the Yukon tourism commercials was quickly changed to 'Come to my Yukon - We'll light the way.'").

In addition, tools drawn from knowledge engineering (the branch of artificial intelligence that of late has been eclipsed by machine learning) may extract concepts from such data as judicial opinions. Such tools would include new methods of knowledge representation and automated tagging.

If the data questions are overcome, machine learning provides intriguing possibilities, but also faces challenges from the nature of fair use law. Artificial neural networks have shown formidable performance in classification. Classifying fair use examples raises a number of questions. Fair use law is often considered contradictory, vague, and unpredictable. In computer science terminology, the data is "noisy." That inconsistency could flummox artificial neural networks, or the networks could disclose consistencies that have eluded commentators. Other algorithms such as nearest neighbor and support vectors could likewise test the consistency of legal reasoning by analogy. Decision trees may be simpler than other approaches in some respects but could work on smaller data sets (addressing one of the data issues above) and provide something that machine learning problematically lacks: transparency. Decision trees disclose their decision-making process, where neural networks are opaque black boxes. Finally, unsupervised machine learning could be used to explore fair use case law for patterns, whether they be consistent structures in its jurisprudence, or biases that have played an undisclosed role.[283] Any possible patterns found should be treated as possibilities, pending testing by other means.

It may well be that early generation fair use algorithms are like spam filters or machine translation. Such software does not read and interpret the meaning of the text at issue, rather it relies on the power of networks to find statistical linkages that permit high-probability assessment of results. Whether later generations of software will be able to conduct fair use assessments in the same manner that lawyers do depends on the improvement of machine learning in identifying abstractions and representing knowledge.

283 Potential bias is a pervasive issue in the applications of machine learning and data. *See, e.g.*, Charles A. Sullivan, *Employing AI*, 63 VILL. L. REV. 395, 397 (2018) ("Alternatively, the data may be accurate as far as it goes but problematic insofar as it incorporates prior discrimination, as when past performance evaluations are tainted by conscious or unconscious bias."); Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633 (2017). On the other hand, algorithms may avoid human biases. *See* Cass R. Sunstein, *Algorithms, Correcting Biases*, 86 SOC. RES. 499 (2019).