Algorithmic Fair Use

Dan L. Burk†

Legal governance and regulation are becoming increasingly reliant on data collection and algorithmic data processing. In the area of copyright, online protection of digitized works is frequently mediated by algorithmic enforcement systems intended to purge illicit content and limit the liability of YouTube, Facebook, and other content platforms. But unauthorized content is not necessarily illicit content. Many unauthorized digital postings may claim legitimacy under statutory exceptions like the legal balancing standard known as fair use. Such exceptions exist to ameliorate the negative effects of copyright on public discourse, personal enrichment, and artistic creativity. Consequently, it may seem desirable to incorporate fair use metrics into copyright policing algorithms, both to protect against automated over-deterrence and to inform users of their compliance with copyright law. In this Essay, 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. Thus, algorithmic fair use carries with it the very real possibility of habituating new media participants to its own biases and so progressively altering the fair use standard it attempts to embody.

INTRODUCTION

Law, like other human artifacts, is costly to produce, to distribute, and to apply. Like other human artifacts, the marginal cost of law benefits from economies of scale; standardized, one-size-fits-all regulations can be economically produced and promulgated, with perhaps, like a made-to-measure suit, a bit of tailoring at the end of the supply chain by a court or other arbiter. But even moderate judicial tailoring adds enormously to the cost of applied law, and rare instances of bespoke regulation are even more socially costly.\(^1\)

This Symposium examines the proposition that technological advances might dramatically lower the cost of bespoke regulation. The potential for such “personalized law” is dependent on the development of ubiquitous data collection and algorithmic data processing coupled with dramatically lower costs in real-time communication. Applications of these technologies have emerged in numerous areas, including criminal law, immigration, taxation, and contract. In the area of copyright, protection of digitized works is already increasingly mediated by algorithmic enforcement systems that are intended to effectuate the rights of copyright owners while simultaneously limiting the liability of content intermediaries. On YouTube, Google, and many other online platforms, both internet service providers (ISPs) and copyright owners have deployed detection and removal algorithms that are intended to purge illicit content from their sites.

But unauthorized content is not necessarily illicit content. In particular, many unauthorized digital postings may claim legal legitimacy under one or more exceptions to the rights of the copyright holder, most notably under the legal balancing standard known as fair use. Exceptions such as fair use exist to ameliorate the negative effects of exclusive control over expression on public discourse, personal enrichment, and artistic creativity. Consequently, it may seem desirable to incorporate context-specific fair

---


5 See 17 USC § 107.
use metrics into copyright-policing algorithms, both to protect against automated overdeterrence and to inform users of their compliance with copyright law.\(^6\) Fair use was intended to "personalize" copyright to individual contexts; hence the question arises whether old-style statutory personalization can be translated into data-driven, machine-mediated personalization.

In this Essay, I examine the prospects for personalized law, taking the outlook for algorithmic mediation of fair use as a vehicle. A large and growing literature on algorithmic regulation already warns us of the pitfalls inherent in reliance on such technology, including ersatz objectivity, diminished decisional transparency, and design biases.\(^7\) Drawing on this literature, I argue that automated implementation of legal standards is problematic as a practical and technical matter, and these limitations will inevitably serve to shape user expectations regarding the processes they govern. It seems clear that this effect is already occurring in conjunction with automated enforcement of copyright, as the design values embedded in automated systems become embedded in public behavior and consciousness. Thus, algorithmic fair use carries with it the very real possibility of habituating new media participants to its own biases and so progressively altering the fair use standard it attempts to embody. Critical analysis of algorithmic fair use offers a cautionary tale that should give us pause, not only regarding the development of such systems but also regarding the development of algorithmic law generally.

I. COPYRIGHT'S FAIR USE STANDARD

Copyright allows authors to restrict reproduction, performance, and related uses of their original works as a pecuniary incentive.\(^8\) But copyright, like any property right, is never absolute. Jurisdictional copyright systems typically include some number of user privileges or exemptions—circumstances under


which the statute will condone or authorize particular uses of a copyrighted work even if the copyright owner has not done so.\(^9\) These vary between jurisdictions but typically cluster around socially beneficial uses of the work, such as education, news reporting, scholarship, personal enrichment, or public commentary.\(^{10}\) Often known in British Commonwealth countries as “fair dealing” provisions, these exceptions to the authorization of the copyright holder entail a specific laundry list of discrete, statutorily defined circumstances under which a protected work can be used without permission.

In the United States, the Copyright Act\(^{11}\) also includes a number of such discrete statutory carve outs. For example, § 110 of the statute allows otherwise unauthorized performances of certain nondramatic works for classroom instruction, or for religious services, or for the benefit of blind or handicapped persons.\(^{12}\) Section 110 also permits uses that might or might not be judged socially beneficial but that, in any event, were judged by Congress for whatever reason to be statutorily permissible without the authorization of the copyright holder, such as the “performance of a nondramatic musical work by a governmental body or a nonprofit agricultural or horticultural organization, in the course of an annual agricultural or horticultural fair or exhibition conducted by such body or organization.”\(^{13}\)

Additionally, the United States, together with a small handful of other nations, includes in its copyright limitations an indeterminate exception known as “fair use.”\(^{14}\) Codified into the current statute from common law precedent, fair use is not categorically or specifically defined but rather is decided based on adjudicatory assessment of four factors. Roughly speaking, a court determining whether an otherwise infringing use might be fair is to consider how much of the work was taken, what was done with it, what kind of work was subjected to the taking, and

\(^{11}\) Pub L No 94-553, 90 Stat 2541 (1976), codified at 17 USC § 101 et seq.
\(^{12}\) 17 USC § 110(1), (3), (8).
\(^{13}\) 17 USC § 110(6).
what effect the taking likely had on the market for the work.\(^\text{15}\)

Determination as to whether unauthorized use of a copyright work falls under this provision varies from situation to situation depending on the contextual assessment of the four factors.

Copyright’s multifactor fair use balancing test thus presents a classic example of what has been dubbed a legal standard.\(^\text{16}\) Scholars have long divided legal imperatives into the categories of “rules” and “standards,” the former constituting discrete and defined legal requirements and the latter constituting malleable and fact-dependent directions. These have reciprocal virtues and vices. Rules are simple to understand and enforce but lack nuance and flexibility; standards are flexible and context-sensitive but lack clarity. Institutionally, rules tend to be promulgated ex ante by legislative enactment; standards tend to be determined ex post by courts or other adjudicatory fora. The major institutional costs for rules are typically incurred in development in advance of administration; the major institutional costs for standards are typically incurred during enforcement or administration.\(^\text{17}\)

In an influential discussion of the topic, Professor Carol Rose noted that these are typically not distinct modes of imperative but lie on a continuum, and legal imperatives tend to process between the two.\(^\text{18}\) Because formal rules are too rigid to fairly accommodate unforeseen circumstances, they tend to accumulate exceptions until they begin to resemble standards. At the same time, because standards are expensive to administer, adjudicators begin to develop shortcuts or per se doctrines that are automatically applied when certain recurring circumstances arise, creating de facto rules. Thus, regulation incorporates some combination of ready-to-wear and bespoke regulation, reaping the cost savings from legal economies of scale while attempting to minimize the pinch or the gaps that result from one-size-fits all.

\(^{15}\) 17 USC § 107.


\(^{17}\) See Kaplow, 42 Duke L J at 599–601 (cited in note 16) (discussing how context can change the cost of rule development or standard application).

Fair use and similar standards represent attempts by the institutional legal system to personalize copyright usage by allowing a tribunal to take into account the individualized circumstances of the unauthorized use, after the fact, in rendering a decision on infringement. As with other standards-based legal doctrine, fair use carries with it the disadvantage of ex ante uncertainty; no one can be entirely certain in advance how a court will weigh the four factors, and hence there is always some apprehension that a use may be found infringing rather than fair. Risk averse content users, unable to confidently predict the ultimate decision on their activities, may forgo some socially beneficial uses. But at the same time, this strategy extends copyright exceptions to new or unforeseen scenarios that the legislature would have been unable to anticipate under a discrete “fair dealing” approach.

II. ALGORITHMIC COPYRIGHT

Recent commentary has argued that the doctrinal deployment of rules and standards either has come to an end or will be drastically altered by imminent changes in technological cost structures. This change is expected to be driven by ubiquitous data collection and algorithmic data processing, coupled with dramatically lower costs of communication. The argument postulates a coming world of “microdirectives,” in which automated systems supply citizens with tailored directives, thus capturing both the ex ante advantages of rules and the ex post advantages of standards.

Such speculations likely overstate any foreseeable capability of the relevant technology and certainly understate the role of other social agents in the deployment and implementation of algorithmic systems. Perhaps not surprisingly, this vision of personalized law largely replicates the neoclassical economist’s nirvana of zero transaction costs and perfect information by postulating a world in which data-processing and communication technologies realize the simplifying assumptions of the simplest

---

20 See Casey and Niblett, 92 Ind L J at 1411–12 (cited in note 19).
economic models. As in much of the hypothetical discussion surrounding big data and artificial intelligence, this speculation partakes of the “magic[al]” worldview of trending technology, which promises costless production without the disadvantageous investment of time and resources that technological activity inevitably entails.\footnote{M.C. Elish and danah boyd, Situating Methods in the Magic of Big Data and AI, 85 Commun Monographs 57, 63–64 (2017). See also Malcolm Campbell-Verduyn, Marcel Goguen, and Tony Porter, Big Data and Algorithmic Governance: The Case of Financial Practices, 22 New Polit Econ 219, 220 (2016) (labeling as “techno-utopian” the optimistic view that algorithmic governance will “overcome the imperfections of politics and faulty forms of knowledge”).}

A more grounded framing for algorithmic fair use, then, is to ask whether old-style legal personalization can be translated into data-driven, machine-mediated personalization. Clearly technical practice is already trending in such a direction. Commentators such as Professor Matthew Sag have observed that such algorithmic agents are already commonly deployed to detect and effectively determine cases of digital copyright infringement.\footnote{See Alfred Gell, Technology and Magic, 4 Anthropology Today 6, 9 (1988).} In some cases, such agents are deployed by copyright-intensive industries, such as the recorded music or motion picture industries, to trawl the internet for potentially unauthorized copies of their proprietary works in order to enforce their copyright.\footnote{See Sag, 93 Notre Dame L Rev at 543–44 (cited in note 4).} In other cases these search devices are deployed by intermediaries, such as YouTube, to remove or deter infringing copies so as to avoid contributory liability, meet their obligations as content hosts, and maintain an ostensible public image of vigilance against lawlessness.\footnote{Id at 543–44.}

Sag argues that online algorithmic policing has already changed the nature of copyright enforcement and so effectively changed the nature of copyright infringement.\footnote{Id at 545–46.} Identification and removal of allegedly infringing content is automated, and human oversight or involvement in the removal process is infrequent and perfunctory. Algorithmic removal decisions are seldom challenged due to severe cost asymmetries. Automated identification and removal, whether accurate or mistaken, is relatively cheap, whereas legal and institutional engagement is comparatively expensive. Most removal decisions are effectively final, and

\footnote{Id at 504–05, 543–44.}
all the parties involved—whether users, service providers, or content owners—have altered their legal expectations in light of these realities.28

Copyright enforcement algorithms typically make no provision for user privileges or exceptions, and the removal decision is effectively final before the dispute reaches any forum in which defenses such as fair use might be considered. Thus, far from greater personalization of the copyright notice and takedown procedure, the cost structure of algorithmic content policing has created a largely impersonal process, in which the context-specific factors that should be taken into account in fair use analysis are absent and go unconsidered. The question then arises whether automated copyright policing can and should incorporate determinations of fair use or other statutory exceptions.29

III. AUTOMATING FAIR USE ANALYSIS

Such questions are implicated in decisions like the opinion of the Ninth Circuit Court of Appeals in *Lenz v Universal Music Corp.*30 Lenz posted to the video platform YouTube a twenty-nine-second clip of her baby dancing to the Prince song “Let’s Go Crazy,” which is heard playing distantly in the background audio of the clip.31 The unauthorized use of the music was detected by the recording label, resulting in a demand under the Digital Millennium Copyright Act32 (DMCA) that it be removed from the platform and leading to a countersuit by Lenz over the propriety of the demand.33 A major legal question in the case was whether Universal had a “good faith belief” that the clip was infringing before demanding

---


30 815 F3d 1145 (9th Cir 2016).

31 Id at 1149.


33 *Lenz*, 815 F3d at 1150.
that it be taken down from the platform. Arguably, the formation of such a belief would require some consideration of fair use or other copyright exceptions because the use could not be infringing if excused by such exceptions.

The court held that consideration of fair use was required before demanding removal of online content. But clearly with the use of automated detection and removal algorithms in mind, the court continued: “We note, without passing judgment, that the implementation of computer algorithms appears to be a valid and good faith middle ground for processing a plethora of content while still meeting the DMCA’s requirements to somehow consider fair use.”

Perhaps not surprisingly, the court later withdrew this particular passage of dicta from the published opinion. The record label’s copyright enforcement search and judgment in *Lenz* was done manually, and it is unclear whether fair use consideration can in fact be automated. In 2001, Professor Julie Cohen and I argued in the context of secured copyrighted content that, because fair use standards could not be programmed into technical protection systems, some type of human oversight or institutional infrastructure would be required to ensure continued access for such uses. Prominent computer scientists similarly expressed their deep skepticism that fair use could be programmed into a technical system. Much of this skepticism was based on the ability or inability to translate inchoate legal imperatives into executable computer code. These challenges would include not only the limitation of human programmers to define the parameters and

---

34 Id at 1151.
35 Id at 1153.
36 *Lenz v Universal Music Corp*, 801 F3d 1126, 1135 (9th Cir 2015). This passage was withdrawn and superseded by *Lenz*, 815 F3d at 1148.
37 *Lenz*, 815 F3d at 1149.
characteristics of legal texts but also the inherent limitations of computer languages, their operating environments, and the capabilities of the hardware available to execute coded instructions.\footnote{See id; Mulligan and Burstein, \textit{Digital Rights Management} at 144 (cited in note 39).}

In particular, the ex ante indeterminacy of a legal standard such as fair use, which in the institutional operation of the law constitutes a benefit, presents a challenge for operational machine coding.\footnote{See Elish and boyd, 85 Commun Monographs at 73 (cited in note 22) ("Because computational systems require precise definitions and mathematically sound logics, sociocultural phenomena that are typically nuanced and fuzzy are rendered in coarse ways when implemented into code.").} Rule-oriented legal imperatives may better lend themselves to automated instructions. It is perhaps not too far-fetched to imagine a programmable exception of the fair dealing laundry list sort—although even for supposedly discrete statutory exceptions, concepts like “educational use” or “news reporting” might be unexpectedly tricky to reduce to computable code. But one can, for example, imagine programming a system to determine, perhaps on the basis of geolocational data and scraped calendaring or advertising data, whether a nondramatic musical work is being performed at an agricultural fair.\footnote{See note 13 and accompanying text.} It is far more difficult to envision how one might program a system to determine whether a given use has a relevant degree of impact on the actual or potential market for the work being used or whether an excerpt from the work is so significant as to constitute the “heart” of an author’s creation.\footnote{See \textit{Harper & Row, Publishers, Inc v Nation Enterprises}, 471 US 539, 564–65 (1985) (holding that an unauthorized publication of a short excerpt constituting the “heart” of a biography weighed against fair use).}

Thus, the prospects for deploying what Fred von Lohmann has called “a judge on a chip” are at best remote.\footnote{See Mulligan and Burstein, \textit{Digital Rights Management} at 139 (cited in note 39), quoting Fred von Lohmann, \textit{Reconciling DRM and Fair Use: Preserving Future Fair Uses?} \textit{*1} (Computers, Freedom, and Privacy Conference, 2002), archived at http://perma.cc/UF8R-VHZP. See also Armstrong, 20 Harv J L & Tech at 108–20 (cited in note 38) (reviewing limitations of a range of embedded fair use architecture possibilities).} Current machine learning techniques attempt to sidestep such difficulties by creating routines that recognize data patterns and by allowing the routine to operate according to the values in the pattern found, rather than attempting to specify values in advance.\footnote{See Elish and boyd, Commun Monographs at 63 (cited in note 22) (describing the current movement toward machine learning techniques); Marion Fourcade and Kieran}
raises the possibility that algorithmic fair use parameters might not have to be explicitly defined and coded. Empirical investigation of the corpus of fair use decisions from federal courts suggests that fair use outcomes are neither random nor unpredictable but may follow particular patterns of judicial decision-making. One can imagine that a neural network or other machine learning system could detect these or other patterns in the data surrounding past cases, matching them to similar patterns in the data surrounding future fair use incidents, situations, and scenarios without formal programming definition of the fair use factors. Such a system might provide the kind of fair use assessments envisioned by the Ninth Circuit, if not prior to the actual use, at least in conjunction with online copyright enforcement decisions.

While such algorithmic decision-making would lower the immediate cost associated with fair use balancing and may be, as the court observes, the only feasible way to deal with petabytes of online content, it cannot be expected to eliminate the costs associated with fair use determinations. As I suggest in this Essay, it would at best reallocate such costs. Law and attendant legal institutions are embedded in a complex web of sociotechnical actors; technological realignment of costs in one section of the network inevitably results in realignment throughout other sections of the network. Costs do not disappear; they are redistributed. As the saying goes, there is no free lunch, and pressing down on the network at one point inevitably causes protrusion at some point.

Healy, Seeing Like a Market, 15 Socio-Econ Rev 9, 24 (2017) (observing that artificial intelligence research abandoned the idea of machines that can think in favor of machines that can learn).


48 See Elkin-Koren, 64 UCLA L Rev at 1096–97 (cited in note 6) (speculating about this type of fair use algorithm). Ironically, because machine learning techniques inevitably involve the digital reproduction of training content, algorithmic fair use may be dependent on the fair use doctrine in order to acquire and process the materials needed to learn fair use. See Benjamin L.W. Sobel, Artificial Intelligence’s Fair Use Crisis, 41 Colum J L & Arts 45, 80–81 (2017) (discussing the dependence of artificial intelligence learning on fair use); Amanda Levendowski, How Copyright Law Can Fix Artificial Intelligence’s Implicit Bias Problem, 93 Wash L Rev 579, 619–29 (2018) (same). See also James Grimmelmann, Copyright for Literate Robots, 101 Iowa L Rev 657, 665–67 (2016) (arguing that large-scale robotic scanning or copying is permissible as a fair use).


other point. Rather than imagining that costs vanish, it is imperative to estimate where, and to whom, and in what form they will occur.

IV. THE SOCIAL COST OF ALGORITHMS

There already exists a fairly large body of literature attempting to determine, predict, and assess the social impact and cost of algorithmic governance. Professor Tarleton Gillespie, in an influential article, encapsulates and categorizes several social effects that are already apparent in algorithmic deployment.\(^{51}\) Such effects entail hidden or unexpected costs of algorithmic deployment:

- **Patterns of Inclusion:** “Big data” does not simply mean a lot of data; data must be collected, structured, and groomed for processing. The explicit or implicit biases of these procedures, including the choice of what data are included or excluded before algorithmic processing, are determinative of algorithmic output.\(^ {52}\)

- **Cycles of Anticipation:** Data processing routines are structured with particular audiences and purposes in mind; they are tailored and retailed according to predicted uses. Such predictive designs determine who is likely to find the output useful, and the characteristics of the user pool recursively shape future updates to the algorithm.\(^ {53}\)

- **Evaluation of Relevance:** Presentation of algorithmic output necessarily entails assignment of relevance; assigned relevance is meaningful only when adopted by users. Data processing routines thus effectively enact policy choices through their determination of what is relevant or irrelevant.\(^ {54}\)

- **Illusion of Objectivity:** Design and execution of algorithmic processes are typically hidden from their audience. Machine-generated outputs thus appear to materialize without human bias, often creating the unwarranted perception of impartiality and objectivity. This perception further obscures the origins and the biases of the algorithm, lending it unwarranted authority.\(^ {55}\)

---

52 See id at 169–72.
53 See id at 172–75.
54 See id at 175–79.
• *Patterns of Entanglement*: Audiences will inevitably alter their behavior under the influence of the algorithms they depend on, and these behavioral changes then impact the data and data relationships that form the inputs to the same algorithms, a mirrored parallel to the cycles of anticipation in design.\textsuperscript{56}

• *Production of Calculated Publics*: Presentation of algorithmic results to an entangled public reshapes the public’s sense of self, propriety, and purpose. But the audience expectation embedded in algorithmic systems may be taken up by other institutions—by courts, schools, businesses, legislatures—reinforcing both the social position of the algorithm and its assumptions about its audience.\textsuperscript{57}

In this Essay, I am primarily concerned with the latter two issues, although these concatenated effects are deeply intertwined and the preceding four topics are undoubtedly also matters of serious concern. There is no question that the open and hidden biases introduced in the construction of algorithmic systems are bound to have an effect on their social relevance.\textsuperscript{58} As Professor N. Katherine Hayles points out, the products of data processing have no inherent meaning; they require some explanatory narrative that lends them significance.\textsuperscript{59} Thus, as Professor Malte Ziewitz explains, algorithms are developed in terms of the problem they are expected to address, and so their design is inevitably framed in terms of a particular narrative about the algorithm’s purpose.\textsuperscript{60} Such “ontological gerrymandering” manipulates the boundary between the problematic and unproblematic by presenting selected assumptions about the problem as given or by cloaking their presence altogether.\textsuperscript{61}

It is therefore somewhat alarming to read legal commentators who confidently assert that “human decision makers are

\textsuperscript{56} See id at 183–88.

\textsuperscript{57} See id at 188–91.


\textsuperscript{59} N. Katherine Hayles, *How We Think: Digital Media and Contemporary Technogenesis* 176 (Chicago 2012).

\textsuperscript{60} Malte Ziewitz, *How to Think about an Algorithm: Notes from a Not Quite Random Walk* *10 (draft discussion paper, Sept 29, 2011), archived at http://perma.cc/S3CZ-866V. See also Hayles, *How We Think* at 176 (cited in note 60) (observing that database outputs acquire meaning only with narrative explanation).

flawed and biased. The biases and inconsistencies found in individual judgments can largely be washed away using advanced data analytics.” On the contrary, the observation that “[r]aw data is . . . an oxymoron,” famously coined by Professor Geoffrey Bowker in his influential work on scientific classification, has become something of a catchphrase among critical analysts of big data and its attendant algorithmic processes. The data are always cooked, before algorithmic processing and certainly during algorithmic processing, as indeed they must be in order to be useful in any way. The question is never whether the data are biased but rather how, by whom, and for what purposes.

When the four factors of the fair use standard are concerned, many of the points at which such design choices must be made quickly become obvious. Determining the impact of the unauthorized use of a work on the actual or potential market for the underlying work requires a model of the market and decisions about the data that properly populate that model. The amount of the work used can be mapped to the percentage of lines or words or pixels or bits taken for a given use, but some weight or significance must be accorded to that number, whether defined by explicit programming values or by algorithmically learned data patterns. The type of work used and the use to which the protected taking is put require some categorization of works and uses. These and a multitude of other design choices made in advance would determine the allowance or disallowance of uses for protected content; algorithms do not make judgments; they are rather the products of human judgment.

The need for human interpretation stems from the disjunction between data representation and reality: the correlations found by data mining algorithms have meaning within the formal

62 Casey and Niblett, 66 U Toronto L J at 437 (cited in note 19). While this rather astonishing assertion lacks any citation to supporting authority, one possible source could be Professor Viktor Mayer-Schönberger and Kenneth Cukier, who have similarly claimed that the completeness of big data sets obviates the messiness of the data. Viktor Mayer-Schönberger and Kenneth Cukier, Big Data: A Revolution That Will Transform How We Live, Work, and Think 33–35 (Houghton Mifflin 2013). This claim has been strongly criticized as inaccurate by a number of subsequent commentators. See, for example, Carl Lagoze, Big Data, Data Integrity, and the Fracturing of the Control Zone, 1 Big Data & Society *5 (July–Dec 2014); S. Leonelli, What Difference Does Quantity Make? On the Epistemology of Big Data in Biology, 1 Big Data & Society *6–8 (Apr–June 2014).

63 Geoffrey C. Bowker, Memory Practices in the Sciences 184 (MIT 2005) (“Raw data is both an oxymoron and a bad idea; to the contrary, data should be cooked with care.”).

properties of the data set but have unknown significance outside
the data set.\textsuperscript{65} Thus, for example, facial recognition algorithms,
employed to further security, law enforcement, immigration
screening, and other purposes, have been much discussed in rela-
tion to algorithmic governance.\textsuperscript{66} But as computer scientist Bill
Smart reminds us, such systems are not in fact “face detectors,”
they are actually “set-of-pixel-values-that-often-correlate-well-
with-the-presence-of-faces-in-the-training-data-that-you-collected-
detector[s].”\textsuperscript{67} Similarly, fair use algorithms would be more accu-
rately understood as something like “patterns-of-numerical-values-
that-often-correlate-well-with-similar-patterns-of-numerical-values-
related-to-judicial-findings-of-fair-use-in-the-training-data-that-
you-collected-detectors.” Patterns detected by a machine evaluat-
ing fair use–related data should not be confused with a legal in-
stitutional determination of fair use.

Thus, again, algorithms do not make judgments; they are the
products \textit{and} the tools of human judgments. Human narrative
may be baked into the system ex ante or it may be assigned to the
output ex post, but at some point someone must put policy and
ideology to work to declare the numbers relevant.\textsuperscript{68} Data analysis
may indicate that certain data occurrences coincide, but the ex-
planation as to why this occurs is a human narrative or categori-
ization, not a technical determination.\textsuperscript{69} Thus, when data mining
(in one famous example) shows a strong correlation between
movements in the S&P 500 Stock Index and the production of but-
ter in Bangladesh, a human decisionmaker is required to desig-
nate the trend as spurious rather than meaningful.\textsuperscript{70}

Equally problematic is the realization that fair use is not a
static concept. Even if the engineering vision of fair use, whether

\begin{itemize}
\item \textsuperscript{65} See Dourish, 3 Big Data \& Society at \textsuperscript{*7} (cited in note 2); Elish and boyd, Commun
Monographs at 70 (cited in note 22) (developing machine-readable code is “not about a
search for meaning, but about the construction and depiction of statistical models”).
\item \textsuperscript{66} See generally Clare Garvie, Alvaro Bedoya, and Jonathan Frankle, \textit{The Perpetual
Lineup: Unregulated Police Face Recognition in America} (Georgetown Law Center on Privacy
\& Technology, Oct 18, 2016), archived at http://perma.cc/L3CC-BNS6 (discussing use of
facial recognition algorithms).
\item \textsuperscript{67} See Elish and boyd, Commun Monographs at 69–70 (cited in note 22) (quoting Smart).
\item \textsuperscript{68} See boyd and Crawford, 15 Info, Commun \& Society at 667–68 (cited in note 7).
\item \textsuperscript{69} See also John Symons and Ramón Alvarado, \textit{Can We Trust Big Data? Applying Philosophy
of Science to Software}, 3 Big Data \& Society at \textsuperscript{*4}–6 (July–Dec 2016) (discussing the epis-
temic difficulties in error correction for big data systems).
\item \textsuperscript{70} See Dourish, 3 Big Data \& Society at \textsuperscript{*7}–8 (cited in note 2).
\end{itemize}
it is in defined programming values or as machine learned patterns, is somehow entirely faithful to the relevant legal doctrines, we are left with the question as to exactly which version of fair use is being instantiated in the machine. The common law evolves, whether from purely judicial reasoning or from judicial riffing off of legislative enactments. Fair use today does not look entirely like it has in the past, either as it did as a common law doctrine prior to its 1976 statutory codification or as it did when codified by Congress. Neither has the codified version remained static, as the Supreme Court has added a variety of judicial glosses, most notably the concept of transformativity. No doubt the official, judicially articulated understanding of the doctrine’s character can be expected to continue to change in the face of developing technical and social circumstances.

Thus, one concern that could stem from the dynamic legal nature of fair use is whether automated instantiation of fair use freezes the standard as of the time it was encoded, so that the law and the algorithm diverge. The algorithm could of course be updated to learn or incorporate shifts in the legal standard. But far from preventing divergence, updating almost assures it. Whether to incorporate new data or to accommodate new equipment, digital processes require continual updating that create unexpected dynamism as the deployed algorithm evolves. Maintenance and upgrades to the system inevitably deviate from the expectations of the original design. This is a source of inconsistency, as are the multiple serendipitous interactions of the particular algorithm with other hardware, software, and devices with their ongoing updates.

To be sure, one might argue that judicial determinations of fair use factors requires the same set of judgments, and no matter what a judge may articulate in her written opinion, the actual process of judicial reasoning is never fully transparent. But as Professor Lawrence Lessig long ago pointed out, when technical design is effectively legal regulation, the major difference between legal code and computer code may be that the latter type of regulation devolves policy choices from the hands of publicly accountable officials to those of largely invisible and unaccountable

---

72 See Dourish, 3 Big Data & Society at *8–9 (cited in note 2) (discussing the co-evolution of algorithms and data sets as implemented).
73 Id at *9.
74 See Casey and Niblett, 66 U Toronto L J at 437 (cited in note 19).
software engineers. Or as Professor McKenzie Wark suggested, technology is merely politics by different means; when speaking of the technical or of the political, one is speaking of the same systems viewed through different lenses. Political, ideological, and even unconscious biases are well understood to permeate traditional legal codes developed in the legislative arena, but they are equally present in deployment of computer code developed in the technical arena.

V. INSTITUTIONAL INFRASTRUCTURE REDUX

To guard against intentional or unintentional algorithmic error, the natural suggestion is to require some degree of human oversight. And thus, the suggestion of automated fair use assessment, notwithstanding any changes in machine learning technology, circles back to the finding by Professor Cohen and myself nearly twenty years ago that automated fair use systems require human institutional oversight. But once again we are confronted with the observation of Professor Paul Dourish and others that algorithms may be best regarded as “convening” objects that interact with a complex ecology of hardware, software, social

---


76 McKenzie Wark, #Celerity: A Critique of the Manifesto for an Accelerated Politics ¶ 3.7 (Speculative Heresy, May 14, 2013), archived at http://perma.cc/2YKW-X5NA. See also Langdon Winner, The Whale and the Reactor: A Search for Limits in an Age of High Technology 29 (Chicago 1986) (“[T]echnological innovations are similar to legislative acts or political foundings that establish a framework for public order.”).

77 See Elkin-Koren, 64 UCLA L Rev at 1098 (cited in note 6) (suggesting such mixed oversight in the context of automated removal decisions).

78 See Burk and Cohen, 15 Harv J L & Tech at 59 (cited in note 38). See also Erickson and Mulligan, 92 Proceedings of the IEEE at 993–94 (cited in note 40) (discussing the prospects for integrating algorithmic and human oversight of fair use). Professor Yeung argues that such institutional oversight is a generalized requirement for regulation by design. See Yeung, Toward an Understanding of Regulation by Design at 93–94 (cited in note 28).

institutions, and human actors. The social impact of algorithmic fair use depends on the assemblage of actions and actors that are tied together in such an infrastructure. We should therefore consider closely how any putative fair use detector came into existence and consider what entities have the motivation and the resources to construct such a system.

Designing, maintaining, repairing, gathering, curating, and updating a database and its attendant algorithm are not costless activities. They are, to the contrary, likely to be expensive as standalone activities, or might constitute marginal costs related to the investment in a larger undertaking. Copyright algorithms are currently deployed, as the Ninth Circuit underscores in its *Lenz* dictum, in order to manage the overwhelming job of policing digital content. In the vision of algorithmic fair use casually articulated by the Ninth Circuit, the fair use algorithm might constitute part of a good faith effort by content owners to evaluate likely infringement. Alternatively, one could imagine service providers, such as Google or Facebook, creating and deploying algorithmic fair use as part of their effort to comply with their responsibilities under the Copyright Act and to justify their decisions to remove or allow content on their platforms. Far less likely is any scenario in which the users or consumers of copyrighted content deploy a fair use algorithm, or even in which fair users would have any hand in designing or crafting the systems that assess the applicability of the exemption to their activities.

Given the inordinate cost associated with reviewing online content for infringement, what type of human oversight might we expect from the likely originators of automated fair use assessment? Would human oversight guard against Type I or Type II error? False algorithmic fair use positives are the likely concern of content holders, whereas false algorithmic fair use negatives are most detrimental to the public good. Screening for both types of error would effectively mean human review of every algorithmic

---

81 See Elish and boyd, Commun Monographs at 69 (cited in note 22).
82 *Lenz*, 801 F3d at 1135.
83 See notes 30–31 and accompanying text.
84 See J. Neyman and E.S. Pearson, *The Testing of Statistical Hypotheses in Relation to Probabilities a Priori*, 29 Mathematical Proceedings of the Cambridge Phil Society 492, 497–98 (1933) (labeling false positives and false negatives rejecting null hypotheses as Type I or Type II errors, respectively).
decision, negating any cost advantage from the algorithmic review, and so is utterly implausible. The statute considered in *Lenz* clearly contemplates human decision-making in the formation of a “good faith belief,” ⁸⁵ and judicial enforcement of that expectation might require human review before a takedown demand; but this does not address algorithmic policing of uploads, downloads, or access. ⁸⁶ Certainly, content owners are no more likely to engage expert human oversight of fair use analysis than they do now for automated decision-making regarding content blocking or removal.

In other contexts, the idea of third party audits for algorithmic decisions has been advanced. ⁸⁷ But ex ante assessment of intended or unintended biases will prove difficult or impossible. Transparency of algorithmic systems is obscured in at least three different interlocking dimensions. ⁸⁸ First is the explicit or intentional obscurity stemming from trade secrecy and protection of confidential business information—to the extent that algorithms are commissioned or developed by commercial entities, they may attempt to shield proprietary aspects of the technology from misappropriation or competitive copying. ⁸⁹ A second barrier to transparency stems from the esoteric nature of the technology, requiring technical expertise to understand its workings. Even if the code is openly available, judges and lay consumers are unlikely to understand how the algorithm operates, so any assessment of the technology’s operation or suitability is at best reliant on expert interpretation and translation of the algorithm’s features into understandable lay terms.

Third, and closely related to the point regarding updates: the complexity of the algorithm in operation creates opacity. Even if the system is entirely open to inspection by experts, the experts are unlikely to understand how it operates. ⁹⁰ Because machine

---

⁸⁵ *Lenz*, 801 F3d at 1135.
⁸⁶ See generally Bridy, *Copyright’s Digital Deputies* (cited in note 4). As Professor Annemarie Bridy has documented in some detail, the statutory notice and takedown procedure considered in *Lenz* has morphed into a network of voluntary filtering, blocking, and removal practices that are cheaper and more convenient for the businesses involved. Id at 195–98.
⁸⁷ See generally, for example, Pasquale, *The Black Box Society* (cited in note 3) (arguing for audits of search engine and financial accounting algorithms).
⁸⁹ This is a central and signature concern of Professor Frank Pasquale’s analysis. See Pasquale, *The Black Box Society* at 4 (cited in note 3).
⁹⁰ Id at 6–7.
learning codes for routines and leaves the routine to develop values, it is often impossible to predict or even to know what the machine has learned.\(^91\) Additionally, the algorithm itself is embedded in a larger technical structure, including other software and hardware components that will affect its operation, often in unexpected or inscrutable ways.\(^92\) Thus, when Google’s image recognition algorithm infamously labeled pictures of African American people as “gorillas,” it might have been due to some unconscious racial bias in the training data, or it might have been due to some kind of bias in the system design, or it might have been an unfortunate but inadvertent occurrence caused by a random technical glitch somewhere in the system.\(^93\) But the most significant lesson from the debacle, whatever the origin of the offensive output, may be that the only solution was to block the system from labeling any image as a “gorilla” because Google’s technical staff simply had no ability to locate, isolate, or remedy the source of the problem.\(^94\)

VI. SELF-FULFILLING ALGORITHMS

Because the technical limitations I briefly sketch above are opaque, there is a strong tendency for them to remain invisible and unconsidered, and the lack of apparent limits lends to the machine a magical aura of automated objectivity. And yet the limitations will be there, preventing automation of what we now call fair use. While we should be deeply concerned with these inevitable biases that attend algorithmic design and implementation as well as with the distracting myth of algorithmic objectivity, my primary concern here is with their combination to produce recursive biases that change public practice and so change social meaning. As I note above, this type of effect is already seen in the algorithmic copyright policing of online content, in which the algorithmic removal action has become a de facto finding of infringement, the public has begun to internalize such outcomes, and formal copyright law may be incorporating those expectations into its weft.\(^95\) As one video creator has described the development of an online guide to moviemaking:

\(^91\) See Burrell, 3 Big Data & Society at *11 n 17 (cited in note 89).
\(^92\) See id at *5.
\(^93\) See id at *7.
\(^94\) See Marion Fourcade and Kieran Healy, Categories All the Way Down, 42 Hist Soc Rsrch 286, 293–94 (2017).
\(^95\) See Sag, 93 Notre Dame L Rev at 503 (cited in note 4).
You could make a video that meets the criteria for fair use, but YouTube could still take it down because of their internal system (Copyright ID) which analyzes and detects copyrighted material.

So I learned to edit my way around that system.

Nearly every stylistic decision you see about the channel—the length of the clips, the number of examples, which studios’ films we chose, the way narration and clip audio weave together, the reordering and flipping of shots, the remixing of 5.1 audio, the rhythm and pacing of the overall video—all of that was reverse-engineered from YouTube’s Copyright ID.

I spent about a week doing brute force trial-and-error. I would privately upload several different essay clips, then see which got flagged and which didn’t. This gave me a rough idea what the system could detect, and I edited the videos to avoid those potholes.  

Whatever form algorithmic fair use might take would likely become a similar social, legal, and creative default.

The effective loss of any user exceptions in current online copyright enforcement might be seen to favor incorporation of some approximation of fair use into policing algorithms, however far it may depart from the actual legal grant to users, on the theory that occasional and biased user access is better than none. I have become increasingly chary of such interventions, however well-intentioned. We have some historic experience with the effect of fair use approximations in the context of old-fashioned, nonautomated legal formulas. In the context of the 1976 revision of the US Copyright Act, educators, publishers, and other stakeholders met to discuss the application of fair use standards to the

---

96 Tony Zhou, *Postmortem: Every Frame a Painting* (Medium, Dec 2, 2017), archived at http://perma.cc/U5WU-M6ZZ. Hat tip to Professor James Grimmelmann for pointing out this example.

photocopying of classroom materials. After considerable discussion, the groups reported to the relevant congressional committee that they had agreed on certain guidelines for photocopying.

The guidelines were not enacted into law, nor endorsed or approved by Congress, although they were discussed in committee reports. Rather, the guidelines were effectively an agreed-upon metric, conformity with which would be considered by the group to be “fair” and so excused from infringement liability. For example, the guidelines specified that material taken without authorization must be brief and offered definitions of permissible “brevity” for various types of works literary, such as:

- A complete poem if less than 250 words and if printed on not more than two pages or from a longer poem, an excerpt of not more than 250 words.
- Either a complete article, story or essay of less than 2,500 words, or an excerpt from any prose work of not more than 1,000 words or 10 percent of the work, whichever is less, but in any event a minimum of 500 words.
- Each of the numerical limits stated above may be expanded to permit the completion of an unfinished line of a poem or of an unfinished prose paragraph.

Note that, under the statutory test, these metrics might or might not be deemed “fair.” Depending on the circumstances, 250 words from a poem might fall within the statutory determination of “fair,” or 250 words might be too much. Certainly, many uses not recognized within such guidelines would be fair under the statute. The guidelines comprised a set of simple, discrete, quantitative (and, not coincidentally, eminently programmable) substitutions of private rules for the statutory standard, offering certainty rather than flexibility.

Despite the fact that they were not legally required, and copyright users were likely entitled to more than the guidelines offered,

---


100 Agreement on Guidelines for Classroom Copying in Not-for-Profit Educational Institutions with Respect to Books and Periodicals, HR Rep No 94-1476, 98th Cong, 2d Sess 68 (1976).

101 Id at 68–69.

102 Id at 69.
the guidelines quickly found considerable purchase with copyright users, who were often advised by their employers or by professional societies to remain within the guidelines in order to avoid the uncertainties of the actual statute’s multifactor calculus. This is perhaps not surprising, as the discrete metrics of the guidelines were easier to communicate and to understand than the inchoate factors of the actual legal test. Somewhat more surprisingly, the guidelines began to show up in infringement litigation, were cited by copyright owners as marking the boundaries of fair use, and were adopted by some courts as indicative of fair use. Fair use analysis is costly in terms of judicial resources, and the guidelines offered a ready-made rule for some judges to use.

In short, implementation of algorithmic fair use will inevitably, and probably detrimentally, change the nature of fair use. Much as in the historical case of the fair use guidelines, we should expect that the deployment of any algorithm purporting to assess fair use would engage strong incentives toward the adoption of a quick and easy substitute for a complicated legal test. Adoption might be explicit, as in the case of the fair use guidelines, or tacit, as courts and the public internalize the activity of the algorithm. Indeed, our experience to date with algorithmic systems suggests that the incentives toward de facto definition of fair use as equivalent to its automated doppelganger would be much stronger. In practice then, whatever choices or biases, inclusions or exclusions, expectations or oversights were engineered into the algorithm would become a self-fulfilling prophecy as to the nature of fair use.

Thus, the problem is not so much the concern advanced by Professor Roger Brownsword that regulation such as fair use by design forecloses a population’s moral and normative choices, as the concern is that the moral and normative choices embraced by the population are informed, manufactured, and ultimately

---

103 See Crews, 62 Ohio St L at 662–63 (cited in note 99) (summarizing judicial uses of the guidelines).


distorted by the architecture of regulation. Indeed, Professor Mireille Hildebrandt has suggested that algorithmic technologies cannot support, and may be iminimal to, the public values that are fundamental to democratic and civil society.\footnote{Mireille Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* 184–85 (Edward Elgar 2015).} Machine learning may seem a novelty, and technology may have changed, but basic human nature and institutional practice have not. Careful consideration of these and related effects are necessarily part of any realistic assessment of algorithmic fair use—or indeed of any movement toward automated governance.

## CONCLUSION

The conclusion compelled by our current understanding of algorithmic governance is stark but real. I have outlined two possible roads ahead, and on neither of them does it appear that viable fair use survives intact in the algorithmic twenty-first century.\footnote{But see Elkin-Koren, 64 UCLA L Rev at 1100 (cited in note 6) (advocating the development of fair use algorithms for the twenty-first century).} Failure to incorporate fair use into copyright enforcement algorithms likely means the de facto loss of the fair use exception, making it available only as a rarified defense to the few litigants who can afford to persevere until favorable judicial review. However, the alternative of attempting to incorporate fair use into enforcement algorithms threatens to degrade the exception into an unrecognizable form. Worse yet, social internalization of a bowdlerized version of fair use deployed in algorithmic format is likely to become the new legal and social norm. We can of course try to shoehorn some type of infringement forgiveness into enforcement algorithms, and we might even label such mechanized user latitude “fair use,” but it will not resemble fair use or serve the goals of fair use, in any sense that we now know them.

Essentially, because fair use cannot be automated, algorithmic fair use simply cannot be fair use at all. There is perhaps some cold comfort, when drawing upon the very deep literature examining algorithmic governance, to realize that this situation is not unique either to fair use in particular or to copyright in general. The reality of what Professor Jack Balkin has called the “Algorithmic Society”\footnote{Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 UC Davis L Rev 1149, 1151 (2018). See also Ian Bogost, *The Cathedral of Computation* (The Atlantic, Jan 15, 2015), archived at} is that these processes are operating
across vast swaths of legal governance, from privacy to consumer welfare to freedom of speech.\textsuperscript{109} In the little corner of the world concerned with copyright, it has been clear for some time that copyright in the information age is probably not fulfilling its mandate of encouraging authors while promoting human flourishing;\textsuperscript{110} the dysfunction inherent in algorithmic copyright is only the latest sign of a system in dissolution.

It should, therefore, come as no surprise that the fair use component of copyright is no more amenable to automation than is the overall copyright system itself. Moreover, I have focused here only on the difficulties of algorithmic fair use, but similar difficulties would attend automation of the idea/expression distinction,\textsuperscript{111} exhaustion,\textsuperscript{112} functionality,\textsuperscript{113} and other doctrines that likely do far more than fair use to control the shape and scope of copyright. Rather than attempting to salvage the accustomed analog copyright balance by substituting some changeling form of fair use for the familiar doctrine, the reality of algorithmic governance may instead mean radically rethinking the goals of copyright as a whole.


\textsuperscript{110} The literature grappling with this problem is now immense, but a sampling of major works would include Hector Postigo, \textit{The Digital Rights Movement: The Role of Technology in Subverting Digital Copyright} (MIT 2012); John Tehranian, \textit{Infringement Nation: Copyright 2.0 and You} (Oxford 2011); Lawrence Lessig, \textit{Remix: Making Art and Commerce Thrive in the Hybrid Economy} (Penguin 2008); Tarleton Gillespie, \textit{Wired Shut: Copyright and the Shape of Digital Culture} (MIT 2007); Jessica Litman, \textit{Digital Copyright} (Prometheus 2001).

\textsuperscript{111} See Pamela Samuelson, \textit{Reconceptualizing Copyright’s Merger Doctrine}, 63 J Copyright Society 417, 419 (2016).
