A.V. ex rel. Vanderhye v. iParadigms, LLC: Electronic Databases and the Compartmentalization of Fair Use

Jonathan Gingerich[*]

DRAFT OF JUNE 16, 2009

I. INTRODUCTION

Scholars almost universally believe that the fair use doctrine is a gigantic mess.[1] Perhaps not a "conceptual disaster area," but close.[2] At least an accident-prone conceptual area. Scholars also like to write about fair use, incessantly. Indeed, they enjoy this so much that district and appellate courts put together can only generate about one fair use opinion for every three law review articles that professors churn out.[3] One move that has been growing in popularity among scholars of late is to attempt to explain this giant mess and impose some order is by arguing that when courts claim to be applying 17 U.S.C. § 107's four-factor test and balancing the factors and subfactors in accordance with the Supreme Court's four post-1976 opinions on fair use, they are actually developing specific categories of uses that count as "fair."[4]

A recent decision by the Fourth Circuit, *A.V. v. iParadigms, LLC*[5] lends support to these scholarly claims about the compartmentalizing of fair use and additionally provides guidance about where the boundaries lie of one of these fair use compartments. The Fourth Circuit's

---

[*] J.D. candidate, Harvard Law School; non-resident research fellow, Jindal Global Law School.

[1] Partially because there is little agreement on why fair use doctrine exists. *See* William Fisher, *Reconstructing the Fair Use Doctrine*, 101 HARV. L. REV. 1659, 1692–95 (1988).

[2] *Cf.* Charles L. Black, Jr., *Foreword: "State Action," Equal Protection, and California's Proposition 14*, 81 HARV. L. REV. 69, 95 (1967).

[3] *See* Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions, 1978–2005*, 156 U. PA. L. REV. 549, 565 n.64 (2008) (finding that from 2000 through 2005, about 3.3 articles show on fair use up in Westlaw's JLR database for every court opinion on fair use in the same time period).

[4] Sort of like what the Europeans do with copyright privileges, only with the common law twist where you figure out what the rule is not by reading a statute but by asking three different juries (or judges) what the rule is, crossing your fingers, and hoping really hard that all three judges say the same thing. Understandably, some of the scholars writing in this vein wish that Congress would do *something* so that we didn't have to keep putting up with this infernal Mansfieldian guessing by the courts. *See, e.g.*, LAWRENCE LESSIG, REMIX 253 (2008) (suggesting that copyright law is "radically out of date" and that Congress should completely overhaul the copyright system).

[5] 562 F.3d 630 (4th Cir. 2009).

move is salutary for the fair use doctrine because it will improve the quantity and quality of expressive activity by making the contours of fair use more certain and allowing for the development of new, socially productive technologies at minimal cost to existing entitlement-holders. However, *iParadigms* also suggests that courts should be cautious in developing fair use categories, because the development of such categories relies to a significant extent on a commitment to particular goals that one believes copyright should serve and may do little to resolve the confusion of the fair use doctrine.

II. FACTUAL BACKGROUND AND PROCEDURAL HISTORY

iParadigms, LLC operates Turnitin Plagiarism Detection Service, a system designed to help college and high school educators detect plagiarized papers. When a school subscribes to Turnitin, it requires its students to submit their papers either directly to Turnitin's website or to course management software integrated with the Turnitin system. Turnitin's software then automatically compares the student submissions with material that is available online, commercial databases of journals, and a database of student papers previously submitted to Turnitin.[6] Schools that subscribe to Turnitin may opt to have their students' papers "archived," which means that they are added to Turnitin's database of student papers. When this happens, Turnitin retains digital copies of the essays to be "archived," but "the employees of iParadigms do not read or review the archived works."[7] To submit a paper to Turnitin, students must agree to an onerous clickwrap agreement that provides that "that the services offered by Turnitin are 'conditioned on [the user's] acceptance without modification of the terms, conditions, and

[6] *Id.* at 634.
[7] *Id.*

notices contained herein,' and that '[i]n no event shall iParadigms . . . be liable for any . . . damages arising out of or in any way connected with the use of this web site.'"[8]

The plaintiffs in the case are students at high schools that subscribed to Turnitin and opted to have its students' papers archived. The schools required students to submit papers through Turnitin, and students who failed to do so would receive no credit at all for an assignment.[9]

Plaintiffs sued, alleging that iParadigms violated their copyrights by archiving their papers in its database without their permission.[10] The district court granted summary judgment to iParadigms, finding that the clickwrap agreement absolved iParadigms of any liability and also finding that the copying of the student papers into the Turnitin database was a fair use under 17 U.S.C. § 107.[11] The plaintiffs appealed to the Fourth Circuit Court of Appeals.

III. THE FOURTH CIRCUIT'S FAIR USE ANALYSIS

The Fourth Circuit conducted a fair use analysis of the Turnitin system and upheld the district court's decision, holding that iParadigms' use of the student writings constituted a fair use. I will discuss the court's analysis of each of the four factors in turn.[12]

A. Purpose and Character of the Use

The court noted that a finding that the use had a commercial purpose is not determinative, and what really matters in the profit/nonprofit analysis is whether a user "stands to profit from

[8] *Id.* at 635 (quoting Joint Appendix at 340).
[9] *Id.*
[10] iParadigms made a counterclaim against one of the plaintiffs for violating the Computer Fraud and Abuse Act. The district court threw out the counterclaim, but the Fourth Circuit reinstated it. *Id.* at 647; *see* Posting of Thomas O'Toole to E-Commerce and Tech Law Blog, http://pblog.bna.com/techlaw/2009/04/fourth-circuits-turnitincom-ruling-brings-more-trouble-for-plaintiffs.html (Apr. 16, 2009) (critiquing the Fourth Circuit's CFAA ruling).
[11] 544 F. Supp. 2d 473 (E.D. Va. 2008).
[12] *See* 17 U.S.C. § 107 (2006).

exploitation of the copyrighted material without paying the customary price."[13] The more transformative a work is, "the less will be the significance other factors, like commercialization."[14] A use is transformative if it uses the quoted material in a different manner or for a different purpose than the original.[15] In arguing that commercial purpose is not determinative of fair use, the Fourth Circuit beat back *Sony's* claim that "every commercial use of a copyrighted material is presumptively an unfair exploitation of the monopoly privilege."[16] *Sony* only requires that a finding of commercial use be weighed against other factors, and it does not establish a *per se* rule that commercial use bars a finding of fair use.[17] Thus, the Fourth Circuit agreed with the district court that the "commercial aspect was not significant in light of the transformative nature of iParadigms' use."[18]

The Fourth Circuit then considered whether iParadigms' use was transformative, and found that it was. First, although Turnitin "stores the work unaltered and in its entirety," the use can be (and is) still transformative.[19] A use does not need to "alter" or "augment" a work in order to be transformative. It can, instead, "be transformative in function or purpose without altering or actually adding to the original work."[20] The court analogized iParadigms' use to Google's use of copyrighted images in *Perfect 10*, which was "'highly transformative' even though the images themselves were not altered, in that the use served a different function than the images served."[21] Similarly, iParadigms' use of plaintiffs' works "had an entirely different function and purpose

---

[13] *iParadigms*, 562 F.3d at 638 (quoting Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 562 (1985)).
[14] *Id.* (quoting Campbell v. Acuff-Rose Music, Inc. 510 U.S. 569, 579 (1994)).
[15] *Id.* (citing Pierre Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1111 (1990)).
[16] *Id.* (quoting Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417, 451 (1984)).
[17] *See id.* at 639 (quoting American Geophysical Union v. Texaco, Inc., 60 F.3d 913, 921 (2d Cir. 1994)).
[18] *Id.* ("[T]he 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." (quoting *Campbell*, 510 U.S. at 578–79)).
[19] *Id.*
[20] *Id.* (citing Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1165 (9th Cir. 2007)).
[21] *Id.*

than the original works . . . ."[22]

Second, the court considered and rejected the plaintiff's argument that iParadigms' use was not transformative because Turnitin failed to effect its purpose. They claimed that it is easy to circumvent the Turnitin system by paraphrasing an original work, and that often the system fails to detect even verbatim copying.[23] Citing *Campbell's* refusal to evaluate the quality of a parody, the court held that "the question of whether a use is transformative does not rise or fall on whether the use perfectly achieves its intended purpose."[24] The plaintiffs did not dispute that the Turnitin system could detect *some degree* of plagiarism, and whether the plagiarism detection system could be better is, the court held, irrelevant to the fair use analysis. Thus, the court concluded that archiving students' papers in the Turnitin database was transformative and found that the first factor favored the defendants.

B. Nature of the Copyrighted Work

The court then proceeded to the second factor of the fair use analysis, and agreed with the district court that this factor favored neither party. The court noted that, "'fair use is more likely to be found in factual works than in fictional works,' whereas 'a use is less likely to be deemed fair when the copyrighted work is a creative product.'"[25] However, "if the disputed use of the copyrighted work 'is not related to its mode of expression but rather to its historical facts,' then the creative nature of the work is mitigated."[26] Although the plaintiffs argued that the second factor favored them because the works they had submitted to Turnitin were highly creative fiction and poetry, the district court found that Turnitin only used the works for comparative

[22] *Id.* The court did not specify in its opinion what, exactly, it found to be the functions of the students' works or of the Turnitin database.
[23] *Id.* at 639–40.
[24] *Id.* at 640 (citing Campbell v. Acuff-Rose Music, 510 U.S. 569, 582 (1994)).
[25] *Id.* (quoting Steward v. Abend, 495 U.S. 207, 237 (1990)).
[26] *Id.* (quoting Bond v. Blum, 317 F.3d 385, 395 (4th Cir. 2003)).

purposes, which did not relate to the creative core of the works and did not diminish the incentive for the students to be creative.[27] If anything, Turnitin actually *enhanced* incentives for creativity by deterring efforts to plagiarize other students.[28]

The plaintiffs argued that because the works were unpublished, the second factor should weigh in their favor. *Harper & Row* held that authors have the "right to control the first public appearance of [their] undisseminated expression."[29] The fact that a manuscript is unpublished, the Fourth Circuit noted, does not by itself prevent a finding of fair use. In *Bond v. Blum* the Fourth Circuit found that the introduction of an unpublished novel into a court proceeding constituted a fair use.[30] "Where . . . the use of the work *is not related to its mode of expression* but rather to its historical facts and there is no evidence that the use of Bond's manuscript in the state legal proceedings would adversely affect the potential market for the manuscript, one cannot say the incentive for creativity has been diminished in any sense."[31] As in *Bond*, Turnitin's service is unrelated to the particular expression in the essays but rather used the papers for analytic purposes and did not adversely affect the market for the manuscripts. Further, the court found, "iParadigms' use of plaintiffs' works did not have the '*intended purpose*' or 'incidental effect' of supplanting plaintiffs' rights of first publication."[32] iParadigms did not publicly disseminate or display the students' writing, and did not transmit the essays to any party

[27] *Id.* at 641.
[28] *Id.* at 640 (quoting A.V. v. iParadigms, LLC, 544 F. Supp. 2d 473, 483 (E.D. Va. 2008)). This is a somewhat odd statement for the court to make, as courts have generally accepted that proof that a use of a copyrighted work benefits the rights holder is no defense to infringement. *See* Iowa State University Research Foundation, Inc. v. American Broadcasting Companies, Inc., 621 F.2d 57, 62 (2d Cir. 1980).
[29] *iParadigms*, 562 F.3d at 640 (quoting Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 555 (1985)).
[30] 317 F.3d 385 (4th Cir. 2003).
[31] *iParadigms*, 562 F.3d at 641 (quoting *Bond*, 317 F.3d at 395–96).
[32] *Id.*

other than the teacher to whom the works were submitted. Thus, there was no basis for finding that "iParadigms' use of the plaintiffs' papers undermined the first right of publication."[33]

C. Amount and Substantiality of the Portion Used

The court then analyzed the third factor, finding that this factor also favored neither party. Generally, the greater the amount of the copyrighted material copied, the less the likelihood that its use is a fair use. However, courts must also consider the "'quality and importance' of the copyrighted material used"[34] Thus, the extent of permissible copying depends on the purpose and character of the use being made of the copyrighted material. This factor overlaps with the first factor.[35] Because iParadigms' did not attempt "to use the expressive content in the plaintiffs" works but used them for a completely unrelated purpose, the fact that Turnitin copied the entirety of archived student essays did not tilt this factor to the plaintiffs.[36]

D. Effect on the Potential Market

Finally, the court considered the fourth factor. The court began its analysis by noting that this factor is "the single most important element of fair use"[37] because "a use that has no demonstrable effect upon the potential market for, or the value of, the copyrighted work need not be prohibited in order to protect the author's incentive to create."[38] This factor, therefore, should ensure that fair use is structured to promote welfare. Copyright is unnecessary except to the extent that it causes authors to create works that they would not otherwise create.[39]

The Fourth Circuit adopted its test from the Second Circuit, focusing "not upon 'whether the secondary use suppressed or even destroys the market for the original work, but [upon]

---

[33] *Id.*
[34] *Id.* at 642 (quoting Campbell v. Acuff-Rose Music, 510 U.S. 569, 587 (1994)).
[35] *Id.*
[36] *Id.* (quoting *Bond*, 317 F.3d at 396).
[37] *Id.* (quoting Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 566 (1985)).
[38] *Id.* at 642–43 (quoting Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417, 450 (1984)).
[39] *Id.* at 642 (citing Pierre Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1124 (1990)).

whether the secondary use *usurps the market of the original work*."[40] So fair use can still be found if the use causes consumers not to buy the original work (e.g., biting reviews of a book) as long as it doesn't serve as a market substitute for the work. To some extent this overlaps with the question of whether the use of the work is transformative.[41] The court cited several parody cases to argue that where a fair use lessens the public estimation of the original work, fair use may still be found.[42] And "regardless of whether the defendant used the original work to critique or parody it, the transformative nature of the use is relevant to the market effect factor."[43]

Turnitin did not serve as a market substitute or even harm the market value of the plaintiffs' works, because all of the plaintiffs stated in deposition testimony that they had no plans to sell their original works for submission by other students.[44] The plaintiffs argued that iParadigms interfered with potential markets for their works, because iParadigms' archiving of their essays would impair their sale to other high school students seeking to purchase papers. The court acknowledged that there is such a market and that Turnitin's archiving would make it harder to sell the papers on that market.[45] However, the court found that there would be no market substitution effect — Turnitin does not supplant the plaintiff's works in the paper mill market but instead reduces demand by making it easier for teachers to detect plagiarism.[46]

---

[40] *Id.* at 643 (quoting NXIVM Corp. v. The Ross Institute, 364 F.3d 471, 482 (2d Cir. 2004)).
[41] *Id.* The court does not, however, seem to think that the fourth factor overlaps with the second factor. The court suggests that Turnitin simply "lower[s] the public's estimation of the original" but does not provide a market substitute. *Id.* at 644. The value of the papers on the paper mill market, however, seems to be closely tied to the papers being unpublished — an unpublished paper seems to be worth much more than a published one. This strongly suggests that iParadigms' use *did* undermine the plaintiffs' rights to first publication, but this does not enter the court's analysis of the market impact factor.
[42] *Id.* at 643.
[43] *Id.* (citing Suntrust Bank v. Houghton Mifflin Co., 268 F.3d 1257, 1274 n.28 (11th Cir. 2001)).
[44] *Id.* at 644.
[45] *Id.*
[46] *Id.* The court does not state exactly how much protection of future markets the plaintiffs should get. A possible future market would be the sale of licenses for the inclusion of papers in an anti-plagiarism database, which the Fourth Circuit did not discuss. *See* Samuel J. Horovitz, Note, *Two Wrongs Don't Negate a Copyright: Don't Make Students Turnitin if You Won't Give it Back*, 60 FLA. L. REV. 229, 253 (2008) ("[T]he lack of licensing of papers for

The plaintiffs also argued that Turnitin would hurt their ability to submit their own work as college essays or to literary journals that might use Turnitin to verify originality.[47] However, the Fourth Circuit concluded that anyone familiar with Turnitin would understand that matches from such submissions resulted from the authors' earlier submissions rather than plagiarism.[48]

Thus, the first factor strongly favored iParadigms (and helped it on the other factors), the second factor was neutral, the third factor was neutral, the fourth factor tilted to iParadigms, and the fourth circuit found that iParadigms' use was a fair use.[49]

IV. RECONSTRUCTING THE COURT'S COMPARTMENTALIZATION OF FAIR USE

A. Disaggregating Fair Use

Two recent articles by prominent copyright scholars suggest that we should try to disaggregate fair use doctrine.[50] Courts, they say, see fair use cases as falling into certain recurring categories, and courts tend to adopt particular rules for particular categories. These categories, rather than the four § 107 factors, largely shape how courts rule in fair use cases.[51] Pamela Samuelson and Paul Goldstein both tend to think that this compartmentalization is a good thing.[52]

---

use in plagiarism detection may be more a product of Turnitin's longstanding conduct than of the infeasibility of such a market.").

[47] *Id.* at 644–45.

[48] *Id.* at 645.

[49] *Id.* The court declined to address the question of whether the clickwrap agreement was enforceable, in light of its decision on fair use. *Id.* at 645 n.8.

[50] Paul Goldstein, *Fair Use in Context*, 31 COLUM. J.L. & ARTS 433 (2008) (arguing that successfully developing a general theory of fair use is impossible and that we should instead understand fair use analysis more contextually); Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537 (2009) (attempting to "unbundl[e] fair uses" and developing a taxonomy of the different categories that courts rely on in fair use analyses). *But see* Fred Von Lohmann, *Fair Use as Innovation Policy*, 23 BERKELEY TECH. L.J. 829, 857 (2008) (arguing that Congress has been more able to "act incrementally on an industry-by-industry and technology-by-technology basis" than have the courts).

[51] Goldstein, *supra* note 50, at 438 ("First, fair use is a quintessentially pragmatic doctrine, one that proceeds from fact to fact, case to case, with expedience, not theory, as its guiding influence. . . . Second, although expedience rules, fair use cases tend to present themselves in recurring categories. . . . Third, each of these recurring categories possesses its own equities and efficiencies that courts attend to, more or less, in resolving particular cases.").

[52] *See, e.g., id.* at 443; Samuelson, *supra* note 50, at 2619. A number of other prominent copyright scholars have argued that further disaggregation of fair use would be desirable as a normative matter. *See, e.g.,* WILLIAM FISHER,

Goldstein's article is, in large part, a polemic against general theories of fair use.[53] Arguing that the four factors of § 107 "are an abstracted, antiseptic version of the real world" that do not "constitute a theory of fair use any more than tent poles constitute a tent,"[54] he argues that it is more realistic to think that judges who decide fair use cases implicitly (or even explicitly) categorize different uses depending on the context in which they occur and treat uses that fall into contextual fair use categories as privileged.[55] Certain categories of cases, including parody cases and new technology cases have equities, efficiencies, and politics of their own and these contextual factors largely determine how courts will rule, rather than the text of § 107. Thus, in *Sony*, a new technology case, "only the fourth factor mattered, and then only because the Court finessed the factor's circularity with a rule based on burdens of proof,"[56] while in *Campbell*, a parody case, "the four factors matter[ed] . . . only to the extent that they respond[ed] to the unique considerations" of parody cases.[57] Goldstein claims that judges' tendency to rely on categories of cases to decide whether particular uses of copyrighted works are fair uses explains empirical findings that lower courts "often ignore Supreme Court decisions and leading circuit authority."[58] In determining fair use, district courts rely on common law reasoning, analogizing and fitting cases that they decide into existing categories of fair (or unfair) uses or creating new categories, as the balance of equities, efficiencies, and political economy may require.

PROMISES TO KEEP, 199–258 (2004) (arguing that copyright for digital media should be replaced with a privilege to copy such media and a government run reward system to stimulate creation of music and film.

[53] Articles propounding such general theories include, says Goldstein, William Fisher, *supra* note 1; Wendy Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600 (1982); and Pierre Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105 (1990).

[54] Goldstein, *supra* note 50, at 437.

[55] *Id*. at 438–43.

[56] *Id.* at 439.

[57] *Id.* at 440.

[58] *Id.* at 441 (discussing Beebe, *supra* note 3).

Samuelson pursues a similar line, but develops an ambitious taxonomy of fair use categories. Samuelson argues that fair use is "both more coherent and more predictable than many commentators have perceived."[59] Samuelson argues that fair use cases crop up in "policy clusters" and that by analyzing prior decisions about cases in a particular policy cluster we can predict with considerable accuracy whether another use that falls into the same cluster is likely to be fair.[60] Somewhat less aggressively than Goldstein, Samuelson argues that policy clustering is not a substitute for analysis of the § 107 factors but an additional element of fair use jurisprudence that complements the statutory factors and also informs how they are applied to particular contexts.[61] Divvying up fair use cases into "policy-relevant" clusters will provide a better description of what district courts actually do when they decide fair use cases,[62] and will, normatively, lead courts to better results when deciding fair use cases.[63] Samuelson then develops a taxonomy of the policy clusters into which fair use cases fall, classifying them in the overarching headings of "Free Speech and Expression Fair Uses," "Authorship Promoting Fair Uses," "Uses that Promote Learning," "Foreseeable Uses of Copyrighted Works Beyond the Six Statutorily Favored Purposes,"[64] and "Unforeseen Uses."[65]

---

[59] Samuelson, *supra* note 50, at 2541.
[60] *Id.* at 2542.
[61] *Compare id.* ("Policy-relevant clustering is not a substitute for appropriate consideration of the statutory fair use factors, but it provides another dimension to fair use analysis that complements the four-factor analysis and sharpens awareness about how the statutory factors, sometimes supplemented by other factors, should be analyzed in particular contexts.") *with* Goldstein, *supra* note 50, at 442–43 (suggesting that Courts will decide like cases, *i.e.*, uses that fall in the same category, alike, with the possible exceptions of special treatment for uses that fall into more than one category and copyrighted works that originate in Berne Convention or WTO countries).
[62] Samuelson, *supra* note 50, at 2542–43 (arguing that this unbundling helps to explain the Beebe's empirical findings, *see supra* note 3, that lower courts do not consistently follow Supreme Court or circuit court authority when applying the § 107 factors).
[63] *Id.* at 2543 ("[U]nbundling will provide courts with a more useful and nuanced tool kit for dealing with the plethora of plausible fair uses than can be achieved merely by focusing on the four factors set forth in the statute.").
[64] The six uses explicitly favored in § 107 are criticism, comment, news reporting, teaching, scholarship, and research. *See* 17 U.S.C. § 107 (2006).
[65] Samuelson, *supra* note 50, at 2538; *see id.* at 2544–46.

Of particular relevance to *A.V. v. iParadigms, LLC* is one of Samuelson's subcategories of unforeseen uses, "Access to Information-Promoting Fair Uses."[66] Unforeseen uses include uses that Congress did not anticipate when it passed the Copyright Act in 1976, and this category is informed by the policy goals of "promoting competition and innovation in complementary technology industries, furthering privacy and autonomy of users of copyrighted works, and fostering enhanced public access to information."[67] The access to information-promoting fair uses subcategory includes "Internet search engine copying for the purpose of indexing or otherwise making information about protected works more publicly accessible."[68] Cases falling into this category include: *Kelly v. Arriba Soft Corp.*,[69] where the Ninth Circuit held that it was fair use for Arriba Soft to copy images from open pages on the Internet, create thumbnails of those images, and serve the thumbnails to users of its search engine in response to user queries with links to the original images;[70] *Perfect 10, Inc. v. Amazon.com, Inc.*,[71] where the Ninth Circuit held that Google's copying of thumbnails was fair use even when a licensing market existed for thumbnail versions of Perfect 10's images and when Google was making revenue by selling ads next to thumbnails that it served in response to users' searches;[72] and *Field v. Google, Inc.*,[73] where the court held that it was fair use for Google to copy texts freely available on websites, store the copies in its servers, and serve snippets of the websites' contents along with hyperlinks to the websites in response to search queries.[74]

---

[66] *Id.* at 2610.
[67] *Id.* at 2602.
[68] *Id.* at 2610.
[69] 336 F.3d 811 (9th Cir. 2003).
[70] Samuelson, *supra* note 50, at 2611.
[71] 508 F.3d 1146 (9th Cir. 2007).
[72] Samuelson, *supra* note 50, at 2612–13.
[73] 412 F. Supp. 2d 1106 (D. Nev. 2006).
[74] Samuelson, *supra* note 50, at 2613–14.

The relevant factors for courts to consider in these fair use cases, Samuelson says, include (1) whether the use in question actually makes publicly available copyrighted works more accessible, (2) whether the information-access tool makes searches better, (3) whether it is necessary or reasonable for the information-access provider to make copies in order to facilitate access, (4) whether transaction costs for licensing the copyrighted works are too great for a market to be formed, and (5) whether the information-access tool supplants or supercedes the market for the copyrighted work.[75] Furthermore, findings that the information-access tool enhances the market for the plaintiff's work, that the defendant made the information-access tool available in good faith, that the defendant has superior knowledge about market opportunities for the information-access tool, and that the plaintiff made its work openly available on the internet all count in favor of fair use.[76] On the other hand, the commerciality of the defendant's information-access tool (creators of information-access tools need to recoup expenses), the creativity of the plaintiff's work (it is socially valuable to improve access to the plaintiff's work whether it is creative or not), and the plaintiff's willingness to license a particular use (the possibility of enhancing the market for the work by improving access is a more important consideration) should be given little weight in the fair use analysis.[77] A finding that a defendant repeatedly makes copies of entire works should have little weight in fair use analysis, as this copying is often necessary to develop new information-access tools.[78] The information-access-promoting category informs how courts should analyze fair use analyses and apply the statutory

[75] *Id.* at 2614.
[76] *Id*. at 2614–15.
[77] *Id.*
[78] *Id.* at 2615. Samuelson further argues that this sort of iterative copying should be recognized not as a transformative use but as an orthogonal use. *Id.*

factors, but membership in this category is not itself conclusive of whether a potentially infringing use is fair.[79]

B. Disaggregation and iParadigms

iParadigms provides further support for the claims of the unbundlers, because the Fourth Circuit's decision is easier to explain with reference to contextual factors and outcomes in similar cases than with reference to § 107. It also suggests some minor modifications to Samuelson's category of information-access-promoting uses, expanding the category to include uses of all copyrighted works, not just those freely available on the Internet.

The Fourth Circuit's *iParadigms* decision can best be understood as privileging a use that fits into a category of uses that courts have favored in the past, rather than as a straightforward application of the statutory factors and the Supreme Court's glosses on them. While the Fourth Circuit goes through the four factors and analyzes them, it would have been difficult to predict the outcome of the case by looking at these factors alone. For instance, while the court's discounting of the commerciality of iParadigms' use may have predictably followed from the Supreme Court's decision in *Campbell*, it is less than obvious from Fourth Circuit and Supreme Court precedent that iParadigms' use would count as "transformative." For instance, Judge Leval's article on fair use[80] suggests that transformative uses include "criticizing the quoted work, exposing the character of original author, proving a fact, summarizing an idea argued in the original in order to defend or rebut it . . . parody, symbolism, aesthetic declarations, and innumerable other uses."[81] While Turnitin could fall into the "innumerable other uses" category,

[79] *Id.* (noting that peer-to-peer file sharing is information-access-promoting but is not a fair use because of the substantial risk that supplant demand for purchased copies of copyrighted works).
[80] Which the Fourth Circuit quoted when declaring that a transformative use "is one that 'employ[s] the quoted matter in a different manner or from a different purpose from the original,' thus transforming it." A.V. v. iParadigms, LLC, 562 F.3d 630, 638 (quoting Leval, *supra* note 53, at 1111).
[81] Leval, *supra* note 53, at 1111.

it does not seem to fit into any of the enumerated uses, and it seems to have little in common with them, as it does not seem to transform student papers in order to provide "new information, new aesthetics, new insights and understandings."[82] The database is not used to conduct research on student writing but only to check to see if high school and college students are cheating. It may, in some ancillary way, promote originality by deterring plagiarism, but this is a rather indirect effect. iParadigms' copying itself seems to do little to promote authorship.

The uncertainty of the court's analysis of the first factor spills over to its analysis of the third and fourth factors as well. In analyzing the third factor, the court determined that although iParadigms made complete copies of the plaintiffs' works, the factor was neutral because iParadigms' use was limited in purpose and scope to creating digital records for automated comparisons, which was already found to be transformative.[83] Likewise, the court's finding that the fourth factor tilted toward the defendant rested on its earlier finding that iParadigms' use was transformative. Regardless of whether the defendant's use was parodic or critical, the Fourth Circuit said, the use's transformative nature was relevant to a determination of market effect.[84] Because iParadigms' use was transformative, it did not create a market substitute for plaintiffs' works,[85] and copyright does not protect against harm to a market caused by a transformative use that does not provide a market substitute for the original work.[86] If the court had determined that iParadigms' use was not transformative because it added nothing meaningful to the original work, the court could easily have found that it supplanted the market for plaintiffs' works, significantly reducing the value of their works on the paper mill market. The court acknowledged that such a market existed and there seems to be no reason that iParadigms could

[82] *Id.*
[83] *iParadigms*, 562 F.3d at 642.
[84] *Id.* at 643.
[85] *Id.* at 644.
[86] *Id.*

not have populated its Turnitin database by purchasing essays on the paper mill market, so the court's finding that the Turnitin database was a transformative use seems to have been a necessary, if not sufficient, condition for the court's decision that iParadigms did not supplant the market for the plaintiffs' works.[87]

For similar reasons, the court determined that the unpublished and creative status of the plaintiffs' works did not cause the second factor to favor the plaintiffs. iParadigms' use did not facilitate the public appearance of plaintiffs' works in readable form, so it did not supplant the authors' rights to control the first public appearance of expression and was unrelated to the creative elements of the plaintiffs' work, so it did not impinge on the "core of creative expression."[88] If the court had determined the defendant's use was not transformative, it would have been very difficult to hold that iParadigms' wholesale copying did not supplant the authors' rights to control the first public appearance of their works. (Although the public could not read the essays stored in the Turnitin database, queries could be run against it and the title of the work could be displayed if text contained in the essay matched text in another submission to Turnitin). Without a finding of transformativity it would be similarly difficult to find that iParadigms' copying did not impinge on the creative core of copyright. (Although iParadigms' use of the materials was not itself "creative," the creativity of the students' work was of central importance to its inclusion in the database: iParadigms felt the need to include the defendants' works precisely because the expression contained in them differed from the expression contained in other sources, including commercial electronic databases and the Internet.)

[87] *See id.* ("Undoubtedly, there is a market for students who wish to purchase such works and submit them as their own for academic credit.") While the court cites *ibuytermpapers.com* as an example of web sites that purchase papers from their original authors, there are many sites, such as *duenow.com* that sell term papers
[88] *Id.* at 641.

I do not argue that the Fourth Circuit's application of the statutory test is implausible but that the Fourth Circuit's analysis was far from a foregone conclusion and the statutory factors and Supreme Court precedent alone provide little insight into the appropriate outcome in *iParadigms*. The Court's decision can better be explained by seeing iParadigms' use as falling into Samuelson's category of information-access-promoting fair uses. Indeed, a critical move behind the court's decision that iParadigms' use was transformative was analogizing this case to *Perfect 10, Inc. v. Amazon.com, Inc.* The court relied on a citation to *Perfect 10* to support its claim that a use could be transformative "without altering or actually adding to the original work."[89] Like Google in *Perfect 10*, iParadigms copied the entirety of defendants' works and let users run queries against a database assembled from a large number of such copies. Applying Samuelson's test for information-access-promoting uses, iParadigms' use certainly makes copyrighted works more accessible (it would be difficult or impossible to find student papers without a database like Turnitin),[90] it makes searches better, it is reasonable for the information-access provider to make copies to facilitate access because the moderately comprehensive character of the database is necessary to facilitate plagiarism detection, and transaction costs would likely be substantial enough that iParadigms could not operate its Turnitin database if it had to pay to license all of the papers that it included.[91] The fifth factor in Samuelson's test could tilt to the plaintiffs (although it does not necessarily do so), because Turnitin might

[89] *iParadigms*, 562 F.3d at 639 (citing *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007)). Samuelson suggests on the basis of the district court's opinion in *iParadigms* that it falls into her category of "use in litigation and for other government purposes" (within her broader category of "foreseeable uses of copyrighted works beyond the six statutorily favored purposes") because inputting student papers into the database was "investigatory in nature." Samuelson, *supra* note 50, at 2593 n.388. While making a copy of a work in order to determine if it matches works already in the database might fit this description, copying works to add them to the database seems to fit poorly in this category because this copying is not necessary to investigate the copied work; indeed, iParadigms offers institutions the opportunity to subscribe to Turnitin *without* the "archive" option. *See iParadigms*, 562 F.3d at 634. The iterative nature of iParadigms' wholesale copying along with the fact that it is a private actor rather than a government actor make the information-access-promoting category seem a much more plausible categorization than the "litigation and other government purposes" categorization.

[90] *See infra* note 93 and accompanying text (discussing the "publicly available" element of Samuelson's first factor).

[91] *Cf.* Samuelson, *supra* note 49, at 2614.

supersede the paper mill market for the defendants' papers. But four of the five primary factors that Samuelson enumerates appear to favor iParadigms. If creating a new database through iterative wholesale copying is a transformative use, as it was in *Perfect 10*, *Kelly*, and *Field*, the court's finding of a fair use in *iParadigms* appears much more likely.

iParadigms advises a modification to Samuelson's information-access-promoting uses, suggesting that the information-access promoting category extends beyond uses that catalog materials that are publicly available on the Internet, and extends beyond uses that clearly enhance the market for copyrighted works.[92] While *iParadigms* certainly is similar to *Perfect 10* in many ways, there are several important differences between the two cases. Unlike the rights holders in *Perfect 10*, the students don't stand to gain anything from the inclusion of their works in the database while rights holders with images linked to on Google might experience increased site traffic, and the authors in *iParadigms* might not want their writings to be publicly available in any way, while in *Perfect 10*, images were already published before their inclusion in Google.[93] Finding iParadigms' use to be fair even though it enabled iParadigms to extract profits from students' works for which they were not compensated, reduced the value of students' works on the paper mill market, and allowed the creation of a database composed of works not otherwise publicly available makes sense *if* the goal is to encourage the creation of a new information-access tool that we otherwise would not have. Without a finding of fair use, it would be difficult to create a useful database of student papers to detect plagiarism, because if such use were not privileged students who wished to sell their papers on the paper mill market,

[92] As Goldstein argues, the universe of categories that guides judicial results in fair use cases is not static. Goldstein, *supra* note 50, at 440.

[93] Another contrast is that unlike Google, Turnitin is not only propriety, but available only to subscribers. *iParadigms*, by omitting any discussion of this, suggests that limited accessibility of a propriety information-access tool will not preclude a finding of fair use.

facilitating plagiarism, would be unlikely to license their papers to the Turnitin database.[94] Limiting the scope of the information-access-promoting category to works already publicly available on the Internet would do nothing to stimulate productive thought expressed to the public, as the works submitted by students to Turnitin will almost certainly be created even if the authors are unable to monetize them (because they get academic credit for them) and, except for the papers that go on the paper mill market, are unlikely to be available for the public to make productive uses of them.[95] *iParadigms* thus significantly expands the range of copying that can be privileged in Samuelson's information-access-promoting category of uses.[96]

V. THE FOURTH CIRCUIT'S REASONING ADVANCES AND IMPROVES FAIR USE DOCTRINE

As Samuelson argues, the unbundling of fair use evidenced by *iParadigms* is largely salutary for the copyright system: it provides somewhat greater certainty in determining whether particular uses are privileged under § 107 than does reference to the statutory factors alone and it may help to ensure that courts promote the "Progress of Science and useful Arts"[97] by helping to ensure that copying necessary for the creation of innovative information access and processing tools, like the Turnitin database or the Google Book Search,[98] will be privileged.

---

[94] If iParadigms had to match the prices on the paper mill market to procure papers for its database, the expense would likely be too great for the business to break even. A May 25, 2009 search on *AcaDemon.com* for papers on "copyright" turned up a six-page paper selling for over sixty dollars, and which was not licensed for commercial use, suggesting that the costs of licensing a paper for inclusion in a proprietary database would be even higher.

[95] *See* A.V. v. iParadigms, LLC, 562 F.3d 630, 636 n.5.

[96] This modification to Samuelson's category of information-access-promoting uses suggests that if the Google Book Search lawsuit proceeded to litigation rather than settling, as presently seems likely, the scanning of copyrighted books and display of snippets surrounding search terms from these books in response to user queries would be a fair use. While the examples on which Samuelson relies in constructing her category concern exclusively materials that are freely available on the internet, *iParadigms* suggest that even copying of materials that are not otherwise public, much less openly accessible to anyone with a web browser can be fair if it is an orthogonal use facilitating the creation of a new type of database. *See* Edward Wyatt, *Writers Sue Google, Accusing it of Copyright Violation*, N.Y. TIMES, Sept. 21, 2005, at C3; *see also* James Grimmelmann, The Google Book Search Settlement: Ends, Means, and the Future of Books, American Constitution Society for Law and Policy Issue Brief, April 2009, at 15, *available at* http://www.acslaw.org/files/Grimmelmann%20Issue%20Brief.pdf (describing the proposed settlement and suggesting that it is still possible that the court hearing the lawsuit will reject it).

[97] U.S. CONST. art. I, § 8, cl. 8.

[98] *See supra* note 96.

However, it also suggests several ways in which we should be cautious about the disaggregation of fair use. First, it suggests that the gains in certainty that accompany "unbundling" may not be as great as they first appear. Users that wish to make particularly novel uses of copies are unlikely to fit precisely into any of the categories that Samuelson has laid out; in fact, it is precisely because they are innovative that their uses are not quite like other uses on which courts have ruled. Users like iParadigms might have somewhat greater assurance that their copying will be privileged when they see that their use is close to an existing category of fair use, but a good deal of uncertainty remains. [99]

Furthermore, relying on categorical analysis to inform determinations may not provide a way to overcome the problems that arise from the incommensurability in policy aims that undergird copyright law and the fair use doctrine. As William Fisher has argued, there are a wide array of policy goals that copyright law ought arguably to promote, including maximizing social welfare, assuring creators of fair returns for their labors, protecting the personhood interests of creators and users of expressive material, and promoting the realization of a just and attractive culture.[100] Of these four perspectives, promoting social welfare and promoting a just and attractive culture seem like they could support (though would not necessarily require) a finding of fair use in *iParadigms*. There is an argument that treating iParadigms' copying as a fair use helps generate a culture that does not tolerate plagiarism.[101] Similarly, there is an argument that privileging iParadigms use allows for the development of a socially beneficial

[99] As I argued *supra* in Part IV.B, fitting iParadigms' copying into Samuelson's information-access-promoting category requires making at least one significant modification to the category.
[100] Fisher, *supra* note 1, at 1686–91; *see also* William Fisher, *Theories of Intellectual Property*, *in* NEW ESSAYS IN THE LEGAL AND POLITICAL THEORY OF PROPERTY 168, 169–173 (Stephen Munzer ed., 2001).
[101] On the other hand, it is not obvious that a feature of a just and attractive culture is intolerance for plagiarism, and it is possible that the use of Turnitin might "unfairly punish subconscious appropriation and . . . chill expression or intimidate students." Posting of Ashley Gorski to Harvard Law School Intellectual Property: Advanced (Spring 2009) Discussion Board (Apr. 19, 2009, 17:34:42 EST) (on file with the author).

information-access tool that would not be developed if iParadigms were forced to negotiate licenses for works that it wished to include in the database.[102]

However, it is difficult to envision plausible arguments in support of fairness or personhood interests that would support a finding of fair use. Fairness theory, often based on Lockean accounts of property rights, often has difficulty providing determinative answers about the extent of the property rights that authors should acquire to their works,[103] but it tends to be at its strongest ebb in situations where a copier is able to reap profits from the wholesale copying of an author's work while the author gets nothing in return.[104] Uses such as iParadigms' use, which relies on iterative copying of the entirety of works of authors who receive no compensation but allows iParadigms to reap significant rewards from the higher fees that it can charge for use of its plagiarism detection service than it could without a database of student papers,[105] seem to fare poorly under almost any fairness-based account of copyright. A pushback is that there is no real loss for the students and it does not really hurt them in any way for their works to be included in the database while iParadigms did the work to assemble software for the database and develop software to use to analyze the database. However, this argument seems to be a stretch, since iParadigms can charge more for its plagiarism detection service than it could without its database. iParadigms did not do the work to populate the database — the students did that.

---

[102] However, it is possible that privileging iParadigms' use may reduce the incentives for authorship. The court believes that grades will provide sufficient incentives for students to write creative essays, poems, and short stories even in the absence of financial rewards from the paper mill market, but this might not be the case. What if a paper is written for a completion grade only? In that case, without grade incentives, might there be students who would work harder and write a much better paper if they knew they could sell it and make a profit off of it? Would the case come out differently if college students had brought suit? What about college seniors with some high degree of expertise and some reasonably good chance of publishing term papers in scholarly periodicals? These are plausible scenarios where the possibility of financial remuneration would enhance incentives for creative authorial production.

[103] *See* Fisher, *supra* note 1, at 1719–39.

[104] This leaves aside the issue of the clickwrap contract to which the students agree, but the Fourth Circuit did not reach the issue of whether this contract was valid, nor did it consider this contract in conducting its fair use analysis. *See* A.V. v. iParadigms, LLC, 562 F.3d 630, 645 n.8.

[105] iParadigms advertises its massive database of student papers as an important selling point of its plagiarism detection services. *See* Horovitz, *supra* note 46, at 247 n.112 (arguing that "Turnitin's chief selling point, more than its plagiarism detection software" is arguably its database of student papers).

Similarly, personhood based accounts of intellectual property seem to weigh against a finding of fair use. The plaintiffs wrote fiction and poetry to which they might have strong personal attachments and they are now denied the ability to fully control the dissemination of these writings. Their very personal reflections were shoved into a database without their permission. There are very few possible personhood interests that could favor Turnitin, a commercial enterprise run by a for-profit corporation. One possibility is that Turnitin helps ensure that other students don't violate the original authors' personhood interests, but the fact that Turnitin does not notify a student-author when another paper has been submitted that appears to have been copied from the first author's paper undermines this argument,[106] as does the inability of student authors to opt out of the Turnitin database even if they strongly disagree with the aims of database.[107] The Fourth Circuit does not discuss personhood interests in *iParadigms*, and it is hard to imagine how this decision could further such interests.

This incongruity between the outcome in *iParadigms* and goals that a sizable number of commentators believe that copyright law ought to promote suggests that compartmentalizing fair use analysis into smaller categories cannot fully avoid one of the most significant problems that "general" theories of copyright face.[108] Creating categories of uses that ought to be analyzed in the same fashion may require choosing some policy goals that we believe copyright ought to promote rather than others in certain circumstances. This is not to suggest that the outcome in *iParadigms* was necessarily wrong, or to deny that compartmentalizing decisions about fair use may reduce the difficulty of making hard choices about which policy goals to pursue. It is a

[106] *See id.* at 259.
[107] In this sense, privileging iParadigms' copying seems dangerously close to dragooning students into speaking in a particular way and appears to undermine the goal of promoting free expression, an important objective behind fair use for Samuelson. *See* Samuelson, *supra* note 50, at 2546–47.
[108] *See* Fisher, *supra* note 1, at 1691–92; *see also* Goldstein, *supra* note 50, at 434 (suggesting that it is impossible to identify a single overarching policy goal that all fair use decisions do, or even ought to, serve).

reminder, however, that the sources of chaos surrounding the fair use doctrine are deep seated and cannot be surmounted simply by adopting a more contextualized understanding of fair use.