

Self-Plagiarism or Fair Use?

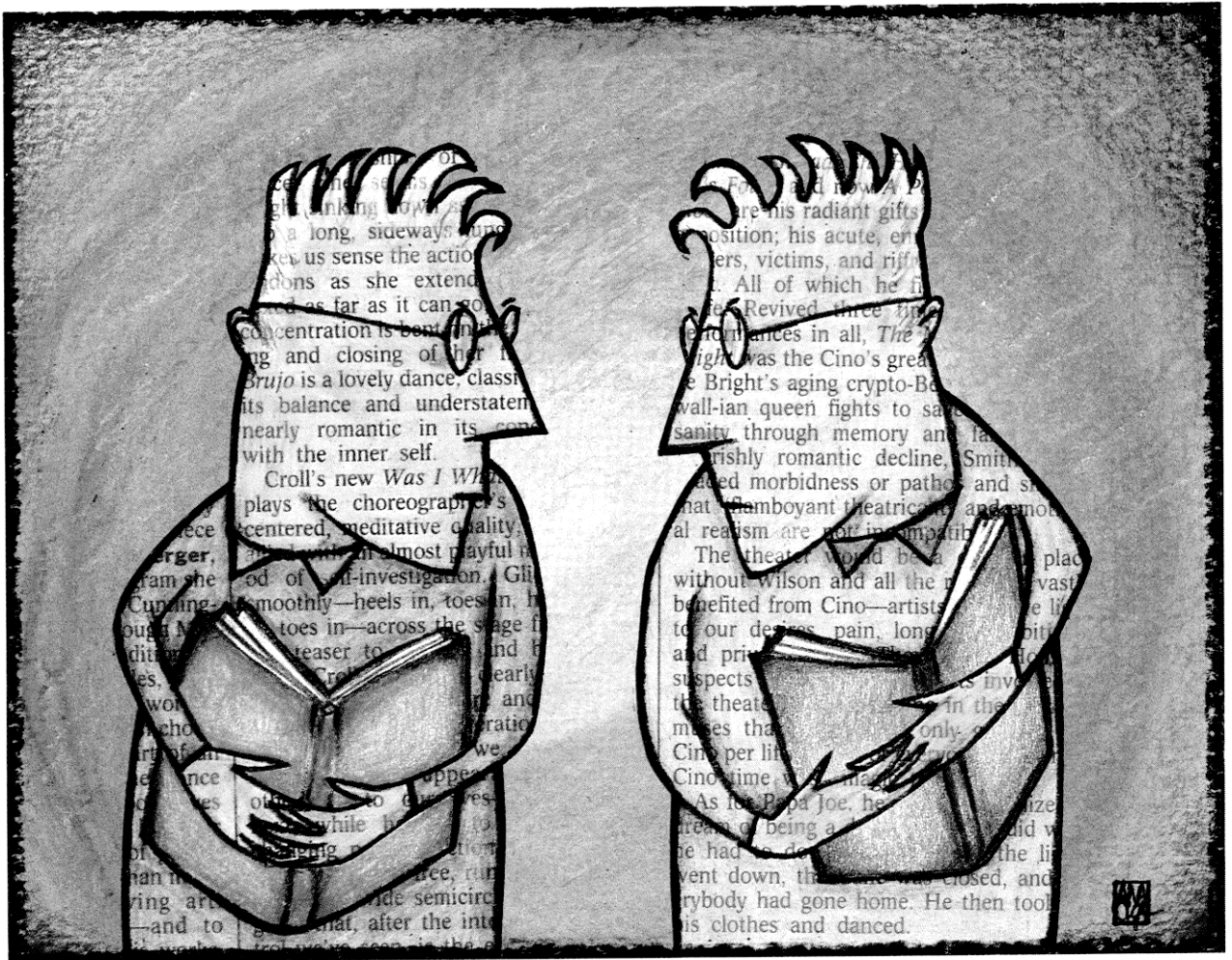


ILLUSTRATION: ANA CICCA

Many *Communications* readers are also authors who, from time to time, must decide whether it is fair to reuse portions of their previous writings in subsequent articles or reports. This can sometimes be a legal question since many publishers, especially in technical fields, require authors to assign copyrights as a condition of publication. Furthermore, in the U.S. and some other countries, employers own copyright interests in articles or reports written by employees within the scope of their employment. In both situations, reusing one's own

prose may raise copyright issues. Other times, the question is an ethical one. Even if an author has retained all copyrights, there are times when it may be unethical to publish multiple versions of the same article as though each was a distinct contribution to the literature in the field when this isn't really so.

This column will explore the legal and ethical landscape of self-plagia-

rism and fair use. Self-plagiarism is sometimes both unlawful and unethical. Other times it is unethical but not unlawful. There are also times when reuse of one's own material is fair, both as a matter of law and as a matter of ethics.

Although it is possible to infringe a copyright in a work you have created, the standards of copyright infringement when you reuse your own prose will, I believe, be somewhat different (and more generous) than if someone else reused the same quantum of your prose in their works. Courts



Pamela Samuelson

have generally been quite sympathetic to authors whose subsequent works employ components of their earlier works except when authors try to usurp markets for their work that have previously been assigned to someone else.

Similarly, as a matter of ethics, a writer will less often be considered a plagiarist for reusing portions of his or her own previous work than if he or she had reused the same quantum of someone else's prose without attribution. Interestingly, when one addresses self-plagiarism as a matter of ethics, it has a different and somewhat more complex flavor than if one addresses it as a matter of law.

As in so much of human affairs, there is no firm rule that applies to all situations. Rather, people must exercise good judgment when reusing their prose. Those who reuse too much are far more likely to suffer opprobrium from violating community norms than to be slapped with a lawsuit for copyright infringement. We should all be moved to be fair about reusing portions of our previous works out of concern for being good members of our fields rather than because we do not want to have to deal with lawyers and the hassles of litigation.

My Own Dilemma

Being a writer for both legal and technical audiences, I grapple with reuse of prose questions quite often. The content of articles for each audience often overlaps substantially. With a legal audience, I can take for granted certain legal concepts (such as summary judgment) that I need to explain to technical audiences. On the other hand, lawyers don't usually understand what algorithms are (as is clear from the Patent & Trademark Office's continued attempt to distinguish between mathematical and nonmathematical algorithms), whereas I can take for granted that *Communications* readers will know exactly what this term means.

I also feel obliged to write more formally to legal audiences than to technical ones. So if I write an article on the same topic to each audience, I will need to rewrite a lot of the prose to conform to the style that each of my audiences expects. For

legal audiences, I generally have to have lots of footnotes (sometimes running into the hundreds) to document or expand upon virtually every assertion the legal article makes. Because of this, when I write articles on the same topic, one for a legal audience and one for a technical audience, each will have a different tone, a different level of detail, and sometimes a different theme, and yet there are often paragraphs or sequences of paragraphs that can be bodily lifted from one article to the other. And, in truth, I lift them.

Oddly enough, the legality of lifting my own prose will depend, in part, on which article is written first. When I write an article for a law review, I almost always retain the copyright and simply license the review to publish the article, usually with a pledge not to republish it without a formal acknowledgment of the review's first publication of the article. (Law reviews are published by law schools as nonprofit centers in order to bring prestige to the school and train very bright students in writing and precision in citation. So they don't care very much about copyright assignments.) If I rewrite a legal article for a technical audience and reuse some parts of the first article, I obviously don't have to worry about getting permission.

If instead I first write for a technical journal, I don't own the copyright for the article because I have to assign it in order to get the article published. (Many technical journals are published by profit-making institutions. Because they undertake the expense of publication and distribution, they want to control the market for the works they publish, and so require copyright assignments. Even nonprofit publishers of technical articles, such as ACM, will tend to require assignments of copyrights as a matter of course.) This means that reuse of my prose from a technical article into a legal article at least requires that I face the permissions question.

On at least one occasion, when I intended to convert a technical article into a law review article with relatively few changes—adding footnotes and one substantive section the article hadn't had before—I asked the technical journal for permission to repub-

lish the article. It was granted on condition that the technical journal's first publication of the article be acknowledged. If my plan instead is to reuse a relatively small number of paragraphs—particularly as to descriptive material—in a second article that consists of extensive reworking or substantially different text, I don't think twice about the borrowing. If I've reused more than a few paragraphs, I usually drop a footnote to the article where the prose first appeared. This way, the first publisher gets credit it deserves and I am giving the readers a pointer to another source they might want to investigate if they are interested in the issue.

I will come back to ethical considerations after reviewing how copyright law has dealt with self-plagiarism claims.

Copyright and Self-Plagiarism

On the specific issue about how much reuse of one's own work is legally acceptable, it is more difficult to get guidance from the copyright case law than one might expect because authors have so rarely been sued for infringing copyrights they have assigned to others. This, in itself, is noteworthy given that the phenomenon of author reuse of text is far from uncommon.

Gross vs. Seligman (decided in 1914) seems to be the only case in U.S. copyright history in which the owner of a copyright won an infringement lawsuit against a self-plagiarist. The defendant in that case had posed a nude model in a particular manner and had produced a photograph of her which he entitled "Grace of Youth." Thereafter, he sold the photograph and his copyright interests in it to the plaintiffs. Two years later he posed the same model in the same way and produced a nearly identical photograph. Apart from changes in the model's body because she was now two years older, the new photograph mainly differed from its predecessor in that it featured a smile on her lips and a cherry between her teeth. The new work was entitled "Cherry Ripe." After the defendant tried to commercially distribute the second photograph, the purchasers of the first copyright sued for infringement. The court observed that

The prudent thing to do when considering reusing code that belongs to someone else, even if you wrote it yourself, is not to reuse any of it.

the photographer was certainly free to make further photographs of the same young woman, but having been paid once for his copyright interest in the "Grace of Youth" photograph, it was unfair for the defendant to try to reap commercial rewards for this creation a second time by marketing a photograph so similar to the first that consumers would perceive it to be a copy and a commercial substitute for the first.

A very recent alleged self-plagiarism dispute was the unsuccessful lawsuit brought by Fantasy Inc., a music publisher, against the musician John Fogerty. Fogerty came to fame as a songwriter for the 1960's rock group Creedance Clearwater Revival. His songs for this group included "Run Through the Jungle," rights in which were sold to Fantasy. Fourteen years later, Fogerty wrote another song entitled "The Old Man Down The Road," and Warner Brothers recorded his performance of it. Fantasy sued both Fogerty and Warner for copyright infringement alleging that the second song infringed its rights in the first song because of similarities in the melody. The jury ruled in Fogerty's favor.

In a few other unsuccessful self-plagiarism cases, decided by judges instead of juries, one can detect in the judicial opinions written about them some sympathy with the plight of an author who returns to a familiar theme. In the most famous self-plagiarism case, Dashiell Hammett was sued for copyright infringement because, down on his luck and past the glory days of his career, he had reused Sam Spade and other characters from "The Maltese Falcon" in scripts for a series of radio programs. More than 20 years earlier, Hammett had sold the rights to make motion pictures and other productions based upon "The Maltese Falcon" to Warner Brothers for \$8,500. Warner complained that Hammett's scripts interfered with rights conveyed

under this contract. The court resolved the dispute in Hammett's favor by construing the Warner Brothers contract narrowly. The judges took into account that authors of detective fiction often reuse characters in subsequent works. Unless there was an unequivocal transfer of all rights in a particular character, any doubt should be resolved in Hammett's favor; \$8,500 was not enough, the judges thought, to indicate Hammett's intent to give up all rights in the characters he'd created in the book.

A more mundane case in which a literate judge attempted to elevate to a higher plane was *Schiller & Schmidt vs. Nordisco Corp.* S&S sued Nordisco for copyright infringement because a former S&S employee had set up Nordisco as a competing office supply business and had laid out portions of Nordisco's catalog in much the same manner as he had done when he'd been in charge of the S&S catalog. The judge perceived the similarities as unsurprising: "We might...suppose the evidence of copying of the layouts compelling were it not that the creator of the copyrighted work and the infringer were the same person. Although one can and often does copy one's own work (and so may be an infringer if one doesn't own the copyright in it), one is also more likely to duplicate one's own work without copying that another person would be likely to do so. If Cezanne painted two pictures of Mont St. Victoire, we should expect them to look more alike than if Matisse had painted the second, even if Cezanne painted the second painting from life rather than from the first painting." Along similar lines, someone faced with the task of laying out an arrangement of items he had previously worked with will naturally be drawn to the same arrangement a second time without necessarily copying from the first. The same person doing both tasks will generally not perceive the same

range of alternative possibilities that a host of other people might.

The closest thing to a self-plagiarism dispute in the software copyright case law is the *Brown Bag Software vs. Symantec* case. Curiously, the defendant was not the alleged self-plagiarist, but rather the company that had bought rights in a program from the same person who had previously written the program the plaintiff owned. John Friend wrote an outlining program called "PC-Outline," rights in which were sold to Brown Bag. He subsequently wrote another outlining program called "Grandview," rights in which were sold to Symantec. Brown Bag sued Symantec for copyright infringement because of the similarities in the user interfaces and features of the two programs. The court decided that the similarities were largely attributable to the fact that the two programs performed the same functions and ruled that no infringement had occurred. The self-plagiarism issue didn't get discussed in the opinion.

This, believe it or not, pretty much exhausts what U.S. copyright law has to say about self-plagiarism. Of course, more self-plagiarism lawsuits may have been brought and settled. Some publishers who could win against a self-plagiarist may decide not to because there's not enough money at stake to justify litigation.

A Different Standard for Author Reuse?

Some copyright lawyers might assert that even though there aren't many self-plagiarism cases, there are plenty of cases in which plagiarism has been found to be an infringement. A self-plagiarist, they might assert, will stand in no better position than any other plagiarist when facing copyright infringement claims. I would disagree, at least in contexts other than employer/employee copyright disputes.

If a stranger reuses two pages of

someone's prose in an article and claims to be its author, copyright infringement would almost certainly be found. If, however, the author reused the same two pages in a subsequent article because, for instance, it concisely summarized some baseline work which readers needed to understand in order to comprehend the breakthrough being reported in the second article, it is almost inconceivable that copyright infringement would be found, even if the author had assigned his or her copyright in the article to the first journal. Copyright infringement would be even less likely if the person did not receive any compensation for publication of the article and if the work was in a field in which it was virtually impossible to get any article published without assigning copyrights to the journal as a condition of publication.

Although it seems not to have been raised in any of the self-plagiarism cases, copyright law's fair use defense would likely provide a shield against many potential publisher claims of copyright infringement against authors who reused portions of their previous works. Under U.S. law, four factors are generally considered in making determinations as to whether a reuse of copyrighted material is fair or infringing. Those factors are: the purpose and character of the defendant's use, the nature of the copyrighted work, the amount taken, and the potential for harm to markets for the copyrighted work arising from the defendant's actions. The more noncommercial and research-oriented the purpose of the reuse, the more factual (as compared with fanciful or artistic) the nature of the copyrighted work, the smaller the quantum reused in relation to the work as a whole, and the lower the potential for meaningful economic harm to the owner of the copyright arising from the defendant's activities, the more likely author reuse of his or her own material would be considered fair use.

One reuse scenario that may be fairly typical for readers of *Communications* is borrowing some paragraphs or a page or two from a previous article in a subsequent research paper where the author was not compensated for publication of either article.

In making informal inquiries about self-plagiarism among friends and colleagues, I have heard reports that some people use a "30% rule" (i.e., a rule of thumb that if one reuses no more than 30% of one's prose in another article, that's OK). This strikes me as a grey zone, and I would certainly not recommend any greater reuse than this, and very likely would recommend less than that, unless one has sought permission for the reuse.

It is surely true that a 30% rule would be inappropriate as applied to a book an author agreed to publish with a particular firm after having assigned copyright in a first book to another publisher. The market for the first book would likely be impaired by the author's reuse of 30% of the text of the first book in the second. In such a case, a reusing author would likely be liable for copyright infringement unless the first publisher had acceded to the second publication.

Even more certain would be the likelihood of infringement if a programmer reused, at a second firm, 30% of the code written while employed at a first firm, particularly if the two firms were direct competitors in the marketplace. The prudent thing to do when considering reusing code that belongs to someone else, even if you wrote it yourself, is not to reuse any of it. The same advice (no reuse) would apply in many other instances in which an employer owned the copyright in works created by its employees.

Law, Ethics, and Plagiarism

Upon reflection, one can detect that copyright law has a somewhat different view of the sin of plagiarism (and consequently of self-plagiarism) than does ethical thought. Indeed, copyright discourse rarely uses the term "plagiarism." A far more common pejorative in copyright cases is the term "pirate," which obviously has quite a different connotation than "plagiarist."

The sin in plagiarism to which copyright law responds is chiefly an economic one: the infringer has deprived (or threatened to deprive) the owner of the copyright of some of the financial benefits that owning the copyright is supposed to bring. As a

consequence, the copyright owner can seek as damages whatever profits were lost as a result of the infringement. The copyright owner can also get a court award to receive from the infringer any of the latter's profits that are attributable to the infringement. If there is a likelihood of recurrence of the infringement, the copyright owner can generally also get a court order to stop any further infringing activity. And the costs of the litigation, including attorney fees, can generally be recouped from an infringer if one wins a lawsuit.

The sin in plagiarism to which ethics responds is, by contrast, mainly a species of misrepresentation. The wrong in plagiarism lies in misrepresenting that a text originated from the person claiming to be its author when that person knows very well that it was derived from another source, knows also that the reader is unlikely to know this, and hopes to benefit from the reader's ignorance. The plagiarist tries to take undeserved credit for an accomplishment that belongs to another person. And, if there is a second sin in plagiarism from an ethical standpoint, it is the sin of laziness. It isn't all that difficult to say something in one's own words. If one hasn't done the work necessary to receive the praise he or she seeks, one shouldn't get it. The reason that academic and research-oriented communities react so negatively to plagiarism is that if this practice is not sharply sanctioned, plagiarists will undermine the community's confidence in the integrity of the edifice of knowledge that lies at the heart of academic and research enterprises.

It is worth noting that the extent of plagiarism that can cost someone a loss of tenure or a precipitous drop in one's standing in a research field because of its violation of community norms will often present too little potential for economic harm to be regarded as a copyright infringement. Or it may arise from copying facts or ideas from someone else's work which copyright law would not punish, for copyright protects only the "expression" in a work, not its facts or ideas. This too illustrates how different are the perceived sins of plagiarism from the standpoint of law and ethics.

Law, Ethics, and Self-Plagiarism

Legal and ethical attitudes toward self-plagiarism are likewise somewhat distinct. In regulating self-plagiarism, the law will chiefly be concerned with whether the economic interests of the copyright owner are being unduly interfered with by a second work that incorporates some of the author's previous work. If so, infringement will be found; if not, infringement is unlikely.

From an ethical standpoint, the misrepresentation of self-plagiarism may be more subtle than that of plagiarism, but it is still present. The self-plagiarist does not, of course, misrepresent the identity of the author of the work, but implies that the work the reader currently sees is new and original and not copied from previous work. Just as the plagiarist breaks community norms by claiming someone else's work as his or her own, the self-plagiarist may violate community norms by claiming a work as a new contribution to the field when it isn't.

One reason that many professional journals have peer review procedures for judging whether to accept articles for publication is to guard against multiple publications of the same material by the same author. Journals rely on those familiar with the literature in the field to say whether there is enough new, original and important material to justify publication of the submitted piece. This may work less well than the journals would hope [1]. Journals generally seek to cultivate a reputation for bringing important new contributions to the field. If a journal acquires a reputation as a recycling bin, subscriptions are likely to fall off, and innovative authors will not want to publish there. For some of the same reasons, many journals require authors to attest that the material submitted is original material as a condition of publication.

The National Science Foundation takes self-plagiarism so seriously that it has designated not only plagiarism but also self-plagiarism—that is, multiple publication of the same scientific work in more than one journal—as a serious deviation from accepted practices and as actionable misconduct [2]. A self-plagiarist can also make

him- or herself vulnerable to scathing attacks in later commentaries if others in the field take the time and trouble to research previous writings and point out earlier uncited publications that the author had hoped would go unnoticed [3].


It is, in truth, a far more complex matter to determine what is self-plagiarism versus what is fair use than to determine what is plagiarism and when is it fair use (almost never).

A number of factors can be cited to explain, and perhaps excuse, reuse of portions of one's previous work that would be implausible in a plagiarism context. For example, one might assert that the previous work needs to be restated in order to lay the groundwork for the new contribution in the second work; or that portions of the previous work must be repeated in order to deal with new evidence or arguments; or that the audience for each work is so different that publishing the same work in different places was necessary to allow the message to get out; or that it is an accepted practice in a field to do particular kinds of republications (i.e., turning a conference paper into a journal article or a book chapter); or that an author will sometimes say things in much the same way without realizing it because that is how the author thinks about the issue; and/or that the author thinks he or she said it so well the first time that it makes no sense to say it differently a second time. (Even so, it may be good professional conduct to cite the earlier work when reusing portions of it. Publishers are less likely to be upset by a little reuse if they can get some credit by the citation.)

When I reuse portions of my prose from a technical to a legal article or vice versa, I tend to rely on the different audience rationale. Techies don't read law review articles, and lawyers don't read technical journals. So if I reuse a little prose from one place to another, it is both a matter of economy and of furthering my goal of serving as a bridge between the technical and legal communities on issues of law that affect computing professionals. On the other hand, as intelligent agent technology improves and comes to be used by lawyers as well as by computing professionals, this justi-

fication may no longer be as persuasive because each community may have access to the other's literature in ways that, for the most part, they do not today.

Conclusion

As bad as plagiarism and self-plagiarism can be, we should not forget that reuse of creative material is something humans have been doing for thousands of years and is not always a bad thing. Our culture enjoys the familiarity that reuse can sometimes bring [4]. A little self-plagiarism never hurt anyone. As George Bernard Shaw once said, "I often quote myself. It adds spice to the conversation." 

Acknowledgments

My thanks to Jonathan Grudin for the suggestion to write this column and for articulating some of the issues with which it might grapple and to Ann Okerson for bringing some materials on this subject to my attention.

References

1. Peters, D.P. and Ceci, S.J. Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral Brain Sci.* 5, 187 (Jun. 1982).
2. Anderson, R. M. The federal government's role in regulating misconduct in scientific and technological research. *J. Law Tech.* 3, 121 (1988).
3. Brogan, M. L. Communicating ideas: The politics of scholarly publishing. *College & Research Libraries*, 5, 453 (Sept. 1992).
4. Editorial. Summer recycling. *Christian Science Monitor*. July 29, 1993.

Pamela Samuelson is a professor of law at the University of Pittsburgh School of Law.