The Automation of Originality

Neville Holmes, *University of Tasmania*

When originality is automated, what becomes of personality?

Originality is the spice of personality. Personal originality shows itself in two ways: in the expression of ideas and in the ideas expressed; in brief, in behavior and in thought.

Curiously, this concept is reflected in the distinction between the two main aspects of what is nowadays called intellectual property. Copyright relates to the expression of ideas, patents to useful industrial ideas however expressed or implemented (*Computer*, May 2004, pp. 92, 90-91). It is also reflected in the distinction between *data* and *information* embodied in the international standard vocabulary that the computing profession is supposed to respect but doesn't (*Computer*, May 2001, pp. 96, 94-95).

COPYRIGHT

Copyright inheres to the creator of a literary, dramatic, musical, or artistic work. In Australia at least, a computer program has been held to be a literary work. Adaptations of such work might or might not infringe copyright, and might or might not attract copyright. It depends on the degree of originality the creator brings to the work.

In copyright, originality is a tricky idea. On the one hand, two people taking a photograph of the same scene each have copyright to their own photograph even if taken at roughly the same time. On the other hand, copyright gives the right to a translation or other adaptation of the copyrighted work. A translator requires licensing by the original copyright holder before the translation is published, at least while that original copyright persists (librarycopyright.net/wordpress/punbb/viewtopic.php?id=1466).

There are all kinds of outs, though, many springing from the basic distinction between the expression of an idea or fact and the idea or fact itself. The compilation of an examination paper has been held to be an original work, provided it is not wholly copied from another work.

The question of originality becomes moot when digital technology is used to create works for publication. Is the user of a computer to carry out a translation sufficiently the originator to hold copyright to the translation? If so, then someone independently carrying out the same translation in the same way would independently hold copyright.

Does someone who uses a computer to merge and massage musical scores to produce a work quite distinct from its sources have independent copyright in that work? If an academic uses a search engine to pull together a set of examination papers from which a program selects a set of questions to construct another and distinct examination paper, does that academic hold copyright to the paper? Or deserve it?

Such questions seem complex to an observer, but they are not merely theoretical. Questions about copyright have recently arisen over the use by Professor Philip Parker of his computer and the Web to put together more than 100,000 works on Amazon—my quick look there gave 108,190. The issue of copyright was raised, not altogether convincingly, of puzzle books and thesauri for Australian languages (theage.com.au/articles/2008/12/18/1229189813966.html).

The use of digital technology to manipulate and generate expressive works of various kinds is only in its infancy. People will need even less personal originality to create distinct literary, musical, and other such works with their computers, and this will change the whole idea of copyright. Incidentally, digital technology will also have a great effect on any legal proceedings relating to copyright (economist.com/business/displaystory.cfm?story_id=12010377).

PLAGIARISM

The use of digital technology has implications for originality at levels lower than copyright. For example,

Published by the IEEE Computer Society

I find myself using the phrase *digital technology* quite often, partly in revolt against the ugly initialism IT. If I want to abbreviate it I could shorten it to *digitology*, except that Google tells me that this is in use already, and if the use is commercial then the work might be registered as a trademark, which just might bring trouble.

The alternative is to adapt *dactylogy* from the term *dactylology*, which means alphabetic finger signing. It seems, courtesy Google again, that the shorter version has recently and mistakenly come into use with the same meaning, but not commercially, making it safe to use here.

the publication of *bigini*, cheap Italian translations of Latin school texts, in wartime Italy. "To outwit the teacher it was necessary to copy selectively, even incorporating some deliberate mistakes in one's work."

The Web with its various search engines is a gigantic bigino, the search engines providing translation. Ten years or so ago when I used to mark student essays on technological issues I would often notice abrupt changes in style, especially in the submissions of students who didn't speak English natively. A search engine would quickly find the source of the plagiarism.

contrast, research publications put together arguments that cite or openly quote related work, often densely, to explain the prior art that the research of the author(s) extends. Peer review in one form or another is supposed to ensure the extension's originality.

Here the originality lies in the ideas or facts themselves, rather than in their expression, and they are the focus of the peer review. With the rapid increase in publication volume, and the extension in fields of research, the effectiveness of peer review has become increasingly questioned. Of course, e-mailing and videoconferencing could help with reviewing, but not in a major way.

> **With the rapid increase in publication volume, and the extension in fields of research, the effectiveness of peer review has become increasingly questioned.**

Search engines are useful for tracking down words and phrases, but if they are used for finding longer texts, the use of those texts might well be a form of theft called *plagiarism*. This term has a long history, as described by Thomas Mallon in his interesting and thoroughly documented book *Stolen Words* (Ticknor & Fields, New York, 1989), though he goes overboard about political plagiarism, particularly in the case of then-Senator Joseph Biden of Delaware.

The significance of most plagiarism is that the offenders offer someone else's text as their own. For short texts—a phrase, a sentence, or even a short paragraph—informal use is much like using a common saying. The opening sentence of this essay came from my memory as something I've once read, but since I can't find it on the Web, how can I acknowledge it? And why should I?

The problem is much more serious when students submit for credit work that isn't their own. This problem is an old one. In his autobiography *Giovinezza* (Ginninderra Press, Canberra, 2007, p. 97), John Maneschi mentions

By now I would think that originality could be simulated by software smoothing out the writing style of an essay, and even shuffling the text around. But dactylogy has other resources for the plagiarist. The Web and its e-mail now support an industry that sells individually written student essays. How long will it take for software to become available that automatically writes essays of a specified length on specific topics incorporating material culled from the Web, using techniques extending those of Professor Parker?

Certainly teachers can, and are, also using software to detect plagiarism, but it's a race of the few against the many. Some months ago, I read an old novel, *College Days at Oxford* by Rev. H.C. Adams (Griffith, Farran, & Co., London, pre-1895), which led me to think that the answer to student plagiarism, at least at university level, might well be a return to face-to-face oral examination.

RESEARCH

Student plagiarism is about faking originality in submitted work. In

Also, the dramatic expansion of research publication on the Web has brought many problems as well as opportunities (guardian.co.uk/education/2008/oct/28/research-internet). So many papers are published nowadays that few of them are read in the ordinary sense. Instead, search engines are used to find relevant prior art. Findings are then dependent on the ranking vicissitudes of those search engines.

For researchers, the academic worth of an article (and hence of its authors) is judged by how often it is cited. Authors therefore write their articles so that search engines will find them more readily, and publishers even produce advice on how to achieve this. Publication is focusing more on advancing the researcher than demonstrating originality of the art.

A side effect of online publishing is that, by and large, citations there seem to be increasingly more recent and numerous, though from fewer publications. The implication is that research is becoming narrower and hence less original (http://sciencemag.org/cgi/content/abstract/321/5887/395).

PERSONALITY

Beyond copyright and research there are many uses of technology that bear on originality and per-

sonality. Dactylogy can be used to manipulate data about people in many ways. Now that photography is primarily digital, pictures of people can be changed easily. Therefore the value of photography as evidence becomes highly questionable because its origin can be so easily concealed.

If digital pictures can be transformed and manufactured so easily, documents kept digitally can be even more easily forged. A recent BBC news item describes a website where people can buy apparently original official personal sick notes and have them within 48 hours (news.bbc.co.uk/2/hi/uk_news/7583830.stm). The site also offers a variety of other personal certificates.

When it comes to ideas or facts rather than works expressing ideas or facts, dactylogy is still effective. For example, the sequencing of genomes is wholly dependent on computation. In such cases, the question then arises of how much human originality is involved. If there is none, as surely can be argued, then there is no ground for a patent to be granted for the original idea as there is no human inventor, and patents can only be granted to humans.

In any case, under the British tradition of patent law at least, scientific discoveries are not patentable, only products and processes. The whole idea of patenting genome sequences is therefore dubious.

Nevertheless, software is increasingly being used to help scientists, engineers, and other professionals develop products and processes. Indeed such use is already widespread.

The patentability of a product or process depends on many factors. Primarily, the product or process must be novel and not obvious to anyone skilled in the prior art. These factors are aspects of originality. The originality must lie in the personality of the inventor as there can be no originality in the operation of the software used by the inventor unless she wrote it herself.

There is another aspect of dactylogy that greatly affects personality, and that is its use in education, both in the home and at school. Personality is personal and social.

There is a tendency to occupy children, at school and at home, with the use of digital machinery of one kind or another. Often this machinery imposes a uniformity of thought and behavior that is the very opposite of originality (see, for example, tinyurl.com/b62wkd). Both parents and teachers must give the highest priority to preventing this uniformity.

Even for adults, the Web is seen as an unprecedented source of knowledge. The problem here is that, if people are to use dactylogy to amplify their originality, surely a wonderful use, then they will need some originality of their own to start with. But when people, especially children in school, have little basic knowledge to start with because of their reliance on the Web, then there can be little originality to amplify. That means banal personality.

As more and more originality comes from dactylogy, what will be left for people to do? A related question that gets far more publicity concerns what is called *The Singularity*, the point at which machine intelligence exceeds human intelligence. When will this happen? (tinyurl.com/by8ry4)

To me, this is like asking, "When will the strength of I-beams exceed the strength of Olympic weightlifters?" Machine intelligence and human intelligence are not comparable (*Computer*, Nov. 2003, pp. 120, 118-119). Machine intelligence is a computational phenomenon, human intelligence a social one. Computing professionals do great social harm when they portray their machinery as anything but tools to amplify human accomplishment and originality. **C**

Neville Holmes is an honorary research associate at the University of Tasmania's School of Computing and Information Systems. Contact him at neville.holmes@utas.edu.au.