DOI:10.1145/3726533

Steven M. Bellovin

Inside Risks
Computer Science and the Law

Making a case for stronger influence and overlap of technology and law.

There were three U.S. technical/legal developments occurring in approximately 1993 that had a profound effect on the technology industry and on many technologists. More such developments are occurring frequently.

The three developments were technically unrelated. One was a bill before the U.S. Congress for a standardized wiretap interface in phone switches, a concept that spread around the world under the generic name "lawful intercept." The second was an update to the copyright statute to adapt to the digital age. While there were some useful changes—caching proxies and Internet service providers (ISPs) transmitting copyrighted material were no longer to be held liable for making illegal copies of protected content—it provided an easy way for careless or unscrupulous actors, including bots, to request takedown of perfectly legal material. The third was the infamous Clipper chip, an encryption device that provided a backdoor for the U.S.—and only the U.S.—government.

All three of these developments could be and were debated on purely legal or policy grounds. But there were also technical issues. Thus, one could argue on legal grounds that the Clipper chip granted the U.S. government unprecedented powers, powers arguably in violation of the Fourth Amendment to the U.S. constitution. That, of course, is a U.S. issue—but technologists, including me, pointed out the technical risks of deploying a complex cryptographic protocol, anywhere in the world (and many other countries have since expressed similar desires). Sure enough, Matt Blaze showed how to abuse the Clipper chip to let it do backdoor-free encryption, and at least two other mechanisms for adding backdoors to encryption protocols were shown to have flaws that allowed malefactors to read data that others had encrypted.

What are the implications for copyright law if a system has to be trained on more or less everything available on the Internet?

These posed a problem: Debating some issues intelligently required not just a knowledge of law or of technology, but of both. That is, some problems cannot be discussed purely on technical grounds or purely on legal grounds; the crux of the matter lies in the intersection.

Consider, for example, the difference between content and metadata in a communication. Metadata alone is extremely powerful; indeed, Michael Hayden, former director of both the CIA and the NSA, once said, "We kill people based on metadata." The combination of content and metadata is of course even more powerful. How-

ever, under U.S. law (and the legal reasoning is complex and controversial), the content of a phone call is much more strongly protected than the metadata: who called whom, when, and for how long they spoke. But how does this doctrine apply to the Internet, a network that provides far more powerful abilities to the endpoints in a conversation? (Metadata analysis is not an Internet-specific phenomenon. The militaries of the world have likely been using it for more than a century.) You cannot begin to answer that question without knowing not just how the Internet actually works, but also the legal reasoning behind the difference. It took more than 100 pages for some colleagues and I, three computer scientists and a former federal prosecutor, to show how the line between content and metadata can be drawn in some cases (and that the Department of Justice's manuals and some federal judges got the line wrong), but that in other cases, there is no possible line[1]

Newer technologies pose the same sorts of risks. Consider today's hottest technology, generative AI. What are the implications for copyright law if a system has to be trained on more or less everything available on the Internet? Does the Berne Convention cover it? Who is liable if an erroneous answer (sometimes incorrectly called a "hallucination") libels someone? I have repeatedly asked one such system for my biography. It has consistently gotten my major, alma mater, year of graduation, published books, and so forth, wrong, even though those answers are readily available on my own website. Others have had similar experiences. This was only laughable, not defamatory—but what if it were libelous?

Many issues involve international law or conflicting laws between different jurisdictions.

Many issues involve international law or conflicting laws between different jurisdictions. How does one balance freedom of speech, a core U.S. value, with the very understandable desire to ban pro-Nazi speech in much of Europe? Who should reconcile the different legal standards? What should ISPs do? What is the effect on everyone else, if, say, a search engine in a NATO country decides it has to suppress information on Tank Man at Tiananmen Square, but perhaps only in certain countries? How do those policies in turn interact with virtual private networks, content distribution networks, and more? How do those location-sensitive answers interact with privacy legislation?

There are many more very difficult questions at the border of law and technology. Is Internet voting a good idea? There are obvious technical risks, but there is the social good of increasing turnout and the technical challenge of preserving ballot secrecy. There is also the technical ability for people to

CREDIT TK

verify their vote was counted, but only if sophisticated cryptographic methods are used to cast votes. Computer crime? What is unauthorized access? Does iterating through sequence numbers in a URL violate the law? How unpredictable must customer IDs be? Is cryptanalyzing a bad pseudorandom number generator illegal? Should it be, if that ability is used to gain access to customer profiles and thus to violate their privacy? How bad must that generator be, if cracking it is to be legal?

The risks can be civil. I know of an incident where a corporate takeover was stymied because the lawyers involved did not understand IP address-based geolocation. That is, their lack of technical knowledge caused them trouble, and they were not even aware of what they did not know.

All of these questions pose considerable risks to society. Lawyers alone cannot answer them; for the most part, they do not know the technology. But technologists alone cannot answer them, because for the most part they do not know the law. Besides, the Internet and other forms of technology are international; what is legal in one place may not be in another—and how to find the boundary in a service offering is itself a difficult question. We need people who understand all of this. More significantly, we need people who can keep current in both fields, because both change.

There is progress. There are annual legal workshops on privacy and cybersecurity; they welcome technical papers, too. *Communications* publishes the Legally Speaking column by Pamela Samuelson and the Law and Technology column by James Grimmelmann on important legal developments; in addition, ACM conducts a regular conference (Symposium on Computer Science and the Law). But the challenge remains: educating people who understand not just technology, but also law, policy, ethics, and all in an international context. (For part of that, a broad, liberal education is necessary.) Combined majors will help, but for many students it is difficult to fit in enough courses in both fields. That said, without such people, we are all at the mercy of systems mandated by well-meaning legislators who do not understand the technical risks of their proposals. C

We need people who understand all of this. More significantly, we need people who can keep current in both fields, because both change.

Reference

1. Bellovin, S.M., Blaze, M., Landau, S., and Pell, S. It's too complicated: How the Internet upends Katz, Smith, and electronic surveillance law. *Harvard J. of Law and Technology 30*, 1 (Autumn 2016); https://bit.ly/4iBIf6p

Steven M. Bellovin (smb@cs.columbia.edu) is a professor of computer science and affiliate law faculty at Columbia University, New York, NY, USA.