Philos. Technol. (2011) 24:463–465
DOI 10.1007/s13347-011-0056-1

COMMENTARY

First, Do No Harm

Vinton G. Cerf

Published online: 5 November 2011

Despite the widely held perspective, which I share, that the Internet and its burgeoning applications have brought enormous benefit in the form of access to information and facilitation of innumerable transactions of all kinds, it is also inescapable that this increasingly pervasive infrastructure can be used in harmful ways. Put another way, this Garden of Eden has its share of snakes. One is rapidly drawn to the writings of Jean-Jacques Rousseau, John Locke, Thomas Hobbes, David Hume, and Hugo Grotius, among others, for insights. Like many other pervasive examples of infrastructure (e.g., roads, waterways, financial transaction systems), there are opportunities for deliberate (and accidental!) harms to befall the users of these systems. Ethics provides us with ways to view these potential hazards and a rationale for their mitigation. The paraphrase of the historic Hippocratic oath is sometimes rendered: "First, do no harm," and this might well be an ethical commitment the users, makers, and operators of the Internet and its applications might undertake.

If we accept this statement as an expression of moral principle, we would have to conclude that use of the Internet to steal, commit fraud, stalk, infect with malware, launch denial of service or other attacks, and so on is a prima facie violation of this principle and should be condemned as immoral. A more nuanced interpretation might extend this notion to include the makers of the software and hardware components of the Internet. Not only would the creation and use of malware be immoral but so would the introduction and use of systems that make no attempt to defend against the various harms undertaken by bad actors.

Humans, by our own admission, are imperfect and thus even a strong commitment to this principle is not a guarantee of protection. It seems to me there are only three ways to deal with potential harms in this environment:

1. Use technology to inhibit harm;
2. Seek to detect and identify harmful actors and take mitigating, including legal action;
3. Use moral suasion when all else fails.

V. G. Cerf (✉)
Google, Inc., 1818 Library Street, Suite 400, Reston, VA 20190, USA
e-mail: vint@google.com

These are not exactly mutually exclusive, as the examples below will illustrate. There are a variety of efforts that might be associated with the first of the categories listed above. Some examples include:

1. Use of non-reusable passwords (e.g., cryptographically generated passwords—sometimes referred to as "two-factor authentication");
2. Introduction of Domain Name System Security (DNSSEC) technology to strengthen the integrity of the mapping of domain names to Internet Protocol addresses;
3. Use of open source software in the expectation that many eyes will help to find and repair vulnerabilities (there are debates about this one).

Examples of the second category may include the use of intrusion detection systems; use of evolving forensic methods to identify the sources of various kinds of attacks, operators of botnets and creators of malware; automatic attempts to detect and flag malware-bearing web sites during the index-creating "crawl" of the World Wide Web and warn users when these sites are encountered. It is important to note that, for legal action to have effect, there must be laws that prohibit bad actions and consequences, if an actor is convicted of violation. The global nature of the Internet also suggests that international reciprocity may be needed (that is, comparable laws and perhaps also extradition agreements) to be effective. Without such reciprocity, bad actors may be able to hide behind a façade of permissiveness beyond the reach of lawful jurisdictions.

The third category, moral suasion, may draw on a number of mechanisms for effectiveness. Among the most forceful might be laws that prohibit unacceptable behavior and come along with penalties; in other words, "You should not do these bad things and, by the way, if we catch you, there will be consequences." There can be economic as well as moral motivators such as reduced insurance premiums for parties practicing good "cyber-hygiene." Parties making commitments to this third category might be rewarded by larger market shares, assuming, of course, that they are trusted and/or can materially substantiate their commitments.

This third category also can be aided by transparency and applies in largest measure to actors providing software, hardware, systems and services in the Internet environment. Clarity of practices can help users to make choices among suppliers of products and services and tends to hold these suppliers accountable for their commitments.

Given the likely impossibility of relying solely on technical means to prevent harm and the uncertainty of commitment of all players to eschew immoral behaviors, we are left with a rather heavy potential reliance on detection and punishment to mitigate harm. Even here, there are ethical tensions. One can imagine legal regimes in which privacy is non-existent. All actions and actors may be visible, and users might be safe; but the loss of all privacy would probably produce a society in which none of us might choose to live. The other extreme, of course, is a society in which privacy is absolute and bad actors impossible to discover leading, again, to a society unacceptable to many. Plainly, an ethical outcome must seek a balance among these tensions to produce a reasonable expectation of safety, privacy, and freedom.

These notions take us to the conclusion that Internet actors (those who make and operate the Internet and its applications) have an ethical responsibility to take steps

to improve the ability of Internet-related technology to protect users from harm, to warn them when they are at risk and to advocate domestic and international regimes to provide recourse when harms peculiar to the Internet environment occur. Perhaps most important, Internet actors must strive to educate users of the Internet, young and old, that bad behaviors in the real world have analogs in cyberspace and moral imperatives dictate that these behaviors should be avoided in the interest of a safer cyber-community.

Although it strikes me as somewhat extreme to argue that access to and use of the Internet should be codified as a "human right"[1] in the sense of the United Nations Universal Declaration of Human Rights,[2] it does seem to me that among the freedoms that are codified, including the right to speak freely, should be the right to expect freedom (or at least protection) from harm in the virtual world of the Internet. The opportunity and challenge that lies ahead is how Internet Actors will work together not only to do no harm, but to increase freedom from harm.

[1] http://www.theatlanticwire.com/technology/2011/06/united-nations-wikileaks-internet-human-rights/38526/
[2] http://www.un.org/rights/HRToday/declar.htm