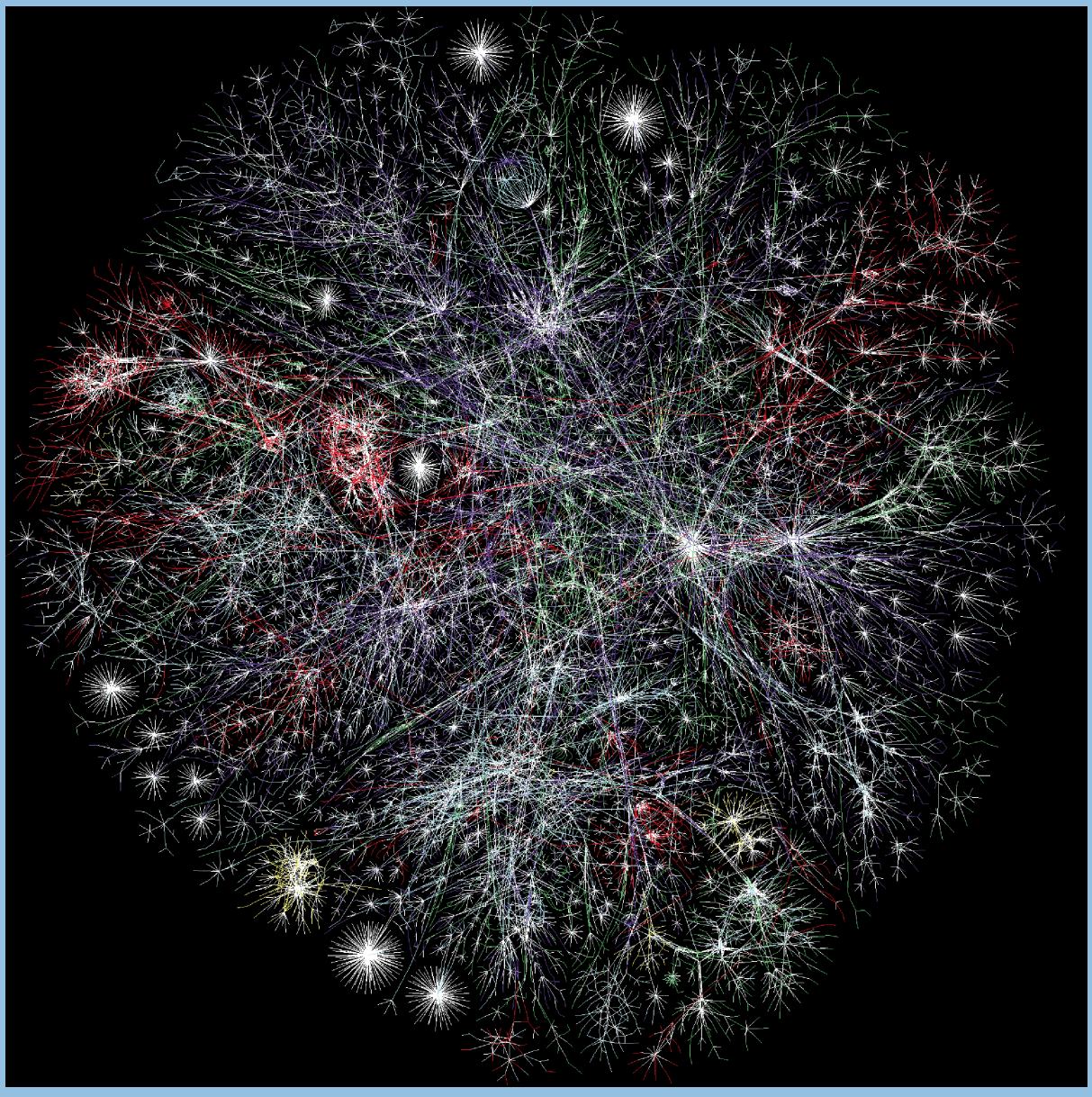
Dædalus

Journal of the American Academy of Arts & Sciences

Fall 2011

Protecting the Internet as a Public Commons

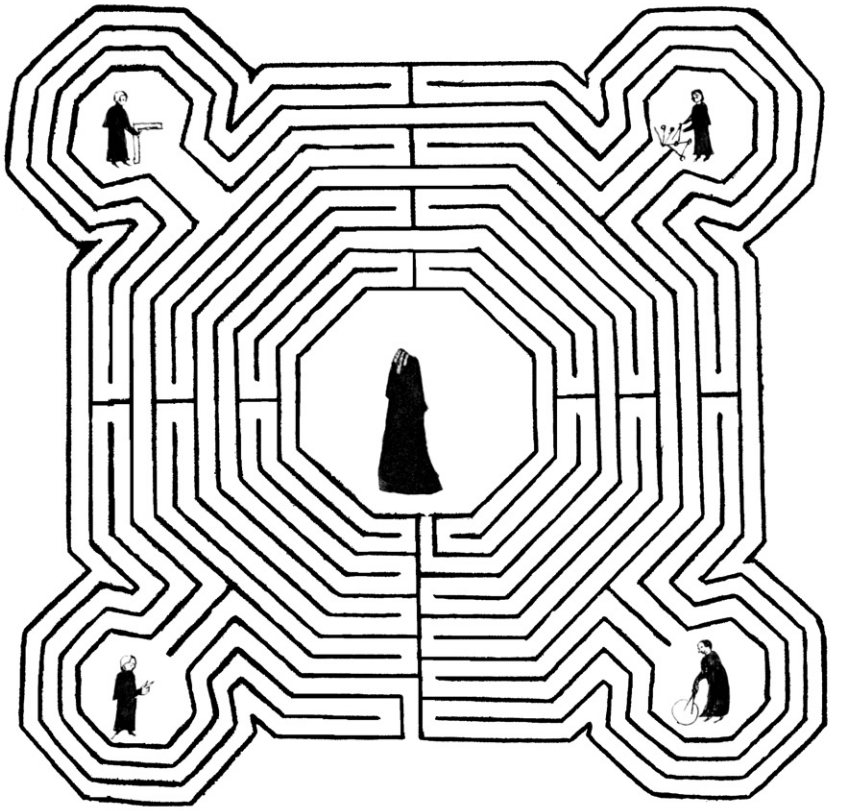


Inside front cover : This representation, from The Opte Project, traces a portion of the routes on the Internet. The Opte Project creates maps to provide an accurate visual model of the connections on the Internet.

Dædalus is designed by Alvin Eisenman.

Dædalus

Journal of the American Academy of Arts & Sciences



The pavement labyrinth once in the nave of Reims Cathedral (1240), in a drawing, with figures of the architects, by Jacques Cellier (c. 1550 – 1620)

Dædalus was founded in 1955 and established as a quarterly in 1958. The journal's namesake was renowned in ancient Greece as an inventor, scientist, and unriddler of riddles. Its emblem, a maze seen from above, symbolizes the aspiration of its founders to "lift each of us above his cell in the labyrinth of learning in order that he may see the entire structure as if from above, where each separate part loses its comfortable separateness."

The American Academy of Arts & Sciences, like its journal, brings together distinguished individuals from every field of human endeavor. It was chartered in 1780 as a forum "to cultivate every art and science which may tend to advance the interest, honour, dignity, and happiness of a free, independent, and virtuous people." Now in its third century, the Academy, with its nearly five thousand elected members, continues to provide intellectual leadership to meet the critical challenges facing our world.

Dædalus Fall 2011
Issued as Volume 140, Number 4

Editorial offices: *Dædalus*, Norton's Woods, 136 Irving Street, Cambridge MA 02138. Phone: 617 491 2600. Fax: 617 576 5088. Email: daedalus@amacad.org.

Library of Congress Catalog No. 12-30299

ISBN 978-0-262-75145-2

Dædalus publishes by invitation only and assumes no responsibility for unsolicited manuscripts. The views expressed are those of the author of each article, and not necessarily of the American Academy of Arts & Sciences.

Dædalus (ISSN 0011-5266; E-ISSN 1548-6192) is published quarterly (winter, spring, summer, fall) by The MIT Press, Cambridge MA 02142, for the American Academy of Arts & Sciences. An electronic full-text version of *Dædalus* is available from The MIT Press. Subscription and address changes should be addressed to MIT Press, Journals Customer Service, 55 Hayward Street, Cambridge MA 02142. Phone: 617 253 2889; U.S./Canada 800 207 8354. Fax: 617 577 1545.

Printed in the United States of America by Cadmus Professional Communications, Science Press Division, 300 West Chestnut Street, Ephrata PA 17522.

Newsstand distribution by Ingram Periodicals Inc., 18 Ingram Blvd., La Vergne TN 37086, and Source Interlink Distribution, 27500 Riverview Center Blvd., Bonita Springs FL 34134.

Postmaster: Send address changes to *Dædalus*, 55 Hayward Street, Cambridge MA 02142. Periodicals postage paid at Boston MA and at additional mailing offices.

Subscription rates: Electronic only for nonmember individuals – $43; institutions – $119. Canadians add 5% GST. Print and electronic for nonmember individuals – $48; institutions – $132. Canadians add 5% GST. Outside the United States and Canada add $23 for postage and handling. Prices subject to change without notice.

Institutional subscriptions are on a volume-year basis. All other subscriptions begin with the next available issue.

Single issues: $13 for individuals; $33 for institutions. Outside the United States and Canada add $6 per issue for postage and handling. Prices subject to change without notice.

Claims for missing issues will be honored free of charge if made within three months of the publication date of the issue. Claims may be submitted to journals-claims@mit.edu. Members of the American Academy please direct all questions and claims to daedalus@amacad.org.

Advertising and mailing-list inquiries may be addressed to Marketing Department, MIT Press Journals, 55 Hayward Street, Cambridge MA 02142. Phone: 617 253 2866. Fax: 617 253 1709. Email: journals-info@mit.edu.

The typeface is Cycles, designed by Sumner Stone at the Stone Type Foundry of Guinda CA. Each size of Cycles has been separately designed in the tradition of metal types.

Introduction

David D. Clark

This issue is concerned with the experience of using the Internet: how its character shapes the user experience and how our collective online participation raises larger societal and political questions. For most of us, the Internet has become indispensable. Whether we are sending email messages or searching the Web, it is a part of our daily lives. It seems to bring powerful benefits, and thus we use it, but it also appears to bring risks, limitations, and frustrations, causing some to react with mixed emotions. People fear loss of privacy and misuse of personal information; they fear the corruption of their computers by malicious software (malware); they fear the possible loss of precious information now stored online; and they resent the complexity of using all this technology. Some people refuse to use computers and the Internet for exactly these reasons, leading us to ask, why is the Internet what it is? What, or who, shapes its character? Are there technical factors that define what can and cannot be done on the Internet? How do the motivations of designers influence the character of the Internet? Are we "locked in" to a constrained set of capabilities, or is the future of the Internet open to many possibilities? Through a variety of essays, this issue explores that set of questions.

DAVID D. CLARK, a Fellow of the American Academy since 2002, is a Senior Research Scientist at the MIT Computer Science and Artificial Intelligence Laboratory. He served as Chief Protocol Architect in the development of the Internet. His current research looks at redefining the architectural underpinnings of the Internet and the relation of technology and architecture to economic, societal, and policy considerations. He is helping the U.S. National Science Foundation organize its Future Internet Design program.

The original design goal of the Internet was modest: to facilitate the remote sharing of expensive computing equipment, at a time when computing was expensive. But even before the Internet became operational in 1983, the notion of its power as a tool for people to interact among themselves had taken hold. The first method to emerge was email, followed by an explosion of options: websites for sharing content, blogs, instant messaging, and "chatting"; shared participation in virtual worlds; social networking sites such as Facebook and Twitter; and so on – a seemingly endless array of tools to interact, collaborate, communicate, and learn.

Questions that center on the user experience are sometimes lost in other debates that arise around the Internet and its future. From a corporate perspective, the Internet is to a large extent driven by commerce, that is, the business of selling. Users are commonly the buyers, and only sometimes (as with auction sites such as eBay or job-search sites) the sellers. Consumers consume: they buy physical objects, which are then delivered to their doors, and they buy virtual products that exist only in digital form, such as music, video content, and movies. Even when users are not actively buying, but are simply "cruising the Web" or using social media, much of what they see is financed by embedded advertising. In this limited view of the Internet, what is needed is a stable and predictable platform that appeals to a set of users affluent enough to have a credit card. But this outlook does not address other aspects of the diverse Internet experience, such as participation in civic discourse or politics, or the simple social process of interacting with friends.

In Washington today, the Internet is increasingly viewed through the lens of security. There is talk about cyber-war, cyber-espionage, and attacks on critical infrastructure. This perspective is not concerned directly with what good might be done online, but with preventing bad outcomes that might cripple the utility of the Internet, for the needs of both the nation and the individual – and of course, for the business of selling.

This issue focuses on the user experience and the Internet as a platform for the wide-ranging endeavors of society because these subjects are sometimes drowned out by the loud voice of selling and the shrill call for security. For many of us, our real hope for the Internet is this broad aspiration, even if it must be financed by commerce. This issue explores the aspects of the Internet that will make it a hospitable platform for socially oriented activities and asks what we can learn from observing how the Internet is used today. A broad view of the Internet takes us beyond the commercial to the space of culture, politics, and – dare we hope – toward a still fragmentary and fragile global civil society.

The positive and negative aspects of email illustrate some of the issues that we must consider in attempting to understand how to make the Internet a hospitable place. Email was the first application that allowed users to interact. In the early days of the Internet, the user community was small and rather homogeneous, and email was an effective mechanism for communication. As the user community expanded, the phenomenon of bulk unsolicited email, otherwise called spam, emerged. The original designers of email were perhaps a bit naive in thinking that all users would be virtuous, polite, and trustworthy, but there were also two conscious design choices that, in retrospect, led to the proliferation of spam. The first was our preference that email be an "open" system whereby any-

one could send to anyone, like the phone system. One could print one's email address on a business card or have it listed in an organization's directory so that others could find and use it. The second was our resistance to requiring verification of one's identity in order to send email messages; we did not like the implications of mandatory identity cards or "Internet driver's licenses." But this openness allowed users to forge their identities. Once spam emerged, we realized that we had no good means to control or discipline spammers because they often operated outside the legal jurisdiction of targeted recipients, and they devised a variety of tricks to avoid detection and deterrence.

A number of important points can be drawn from this example. The first is that *the Internet* and *email* are two different things. One way to explain this difference is by analogy to other systems, perhaps the most accessible being the "information highway" that emerged in the 1990s. The highway analogy is apt in one respect: the Internet itself is a transport infrastructure over which all sorts of applications run, just as a highway is an infrastructure over which all sorts of vehicles run. Thus, it is the Internet that permits email to exist, but it is the particular design of the email application that defines and constrains the user experience. When we talk about *the Internet*, we need to clarify whether we mean only the transport infrastructure or the total experience – infrastructure and applications – that users perceive.

In the same way that email is distinct from the Internet infrastructure on which it sits, the Internet's many different applications (such as email, the Web, games, or Internet telephony) are distinct from each other. Each contains specific design features that create a different context to shape the user experience; different applications can provide very different experiences.

A number of design features influence the collective social experience that the Internet provides. Here, I highlight three that are illustrated in the case of email and that factor into a number of the papers in this volume: *identity*, *trust*, and *controlling bad behavior*. How each of these considerations is approached will to a large extent define the character of the various experiences that we have when we interact with others across the Internet.

In the real world, we have many ways to manage and track identity. We get to know people face to face, and recognition of physical cues is enough to evoke our knowledge of who someone is. In more structured situations, we use identity credentials (such as driver's licenses or passports) or third-party introductions. When we communicate across the Internet, none of the face-to-face cues are available, and we usually lack the more structured credentials. On the Internet, it sometimes seems as if users run around with bags on their heads.

This situation begs the question of why the Internet does not have some sort of built-in identity mechanism that would allow users to be sure of whom they are talking to. The answer is that different applications create different kinds of shared experiences, which have different requirements for identity. If we were to add an identity mechanism to the infrastructure – to the transport layer that defines the Internet – then it would work the same way for all applications, just as the width of a road constrains all vehicles. It would be the same for a consumer completing a banking application and a provider of sensitive information about medical conditions. It would be the same for a dissident in a repressive country, an investigative reporter, or a stockbroker

dealing with a client. That uniformity does not seem to match the needs of Western-style society. In some cases, the communicating parties need strong evidence of identity (for example, when a customer deals with a bank); in others, strong anonymity is crucial (as when we try to protect political speech).

The desire to vary mechanisms by circumstance suggests that tools to manage identity should not be built into the infrastructure layer of the Internet but into the applications themselves. Indeed, some applications contain strong identity tools. For example, banks go to considerable lengths to ensure that they are talking to known clients, and credit card companies act as trusted third parties to identify buyers and sellers in an online purchase. But to impose a uniform strong set of identity mechanisms on the Internet itself would have many negative social consequences.

Consider the above example of email. Given that the prevalence of spam makes email problematic today, should those responsible for email now redesign it so that every sender identifies himself with a valid identity credential issued by a trusted third party? I would argue that this measure would be overkill; it would not match the actual requirements of email as a social interaction tool. When we are introduced to people, we normally do not ask to see their driver's licenses. We use a social process that has been well honed over the ages, called "getting to know them." Over time, we build up a model of who people are and of the extent to which they are trustworthy in the role for which we know them. This process can be used with email as well. Imagine that all email users on the Internet had the option of constructing a credential that certifies who they are. Technically, such a credential is easy to create using encryption, and many people have done so already. From a face-to-face perspective, this approach sounds a little odd: what good does it do for me to have a certificate in which I assert that I am me? In the context of the Internet, it prevents others from impersonating me. If the credential is properly constructed, there is no way for someone else to forge it unless that person breaks into my computer (another issue to be considered). Using such a certificate does not tell you much the first time I send an email message to you. (Of course, if it mattered, you could call me on the phone or otherwise ask me if indeed I am the person who sent the email message.) The second time I send you a message, or the tenth or the hundredth, you can be assured that you are having a conversation with the same person. You are, in the phrasing used above, "getting to know me." This is one possible approach to the problem of identity, and different applications, with different social contexts, will call for different approaches.

Identity, while important, is not an end in itself. Identity is a mechanism that allows us to deal with the other two issues listed above. This brings us to the second issue: *trust*.

As we interact with others in various contexts, we make complex and subtle judgments about trust. We assess whether the parties are trustworthy, whether there are constraints that will limit bad action, and whether we should be confident, cautious, or fearful. We assess strangers on a bus, we decide whether we are at risk of being cheated by a checkout clerk, and we judge whether our friends are more or less trustworthy in different roles. Sometimes we judge wrongly; we may be deliberately misled by a "confidence artist," or we may simply not know someone well enough. But the ability to make and rely on useful assessments

about trust is fundamental to a working society.

On the Internet, the cues we rely on to gauge trust face to face are weak, and our judgments are prone to failure. The interplay between identity and trust is clear. If we cannot know for sure with whom we are talking, if everyone has a bag on his or her head, it makes little sense to assign different levels of trust to different people. A malicious person can (and, in fact, will) pretend to be a good friend. The design of an identity system must take into consideration what sort of identity is needed for appropriate assessment of trust in a particular context. The application designer must know which identity cues would be useful for different applications.

Just as there are many ways to construct and track identity, there are many ways to assess trust. We need the Internet equivalent of being able to tell when we are "going to a bad neighborhood." Will this website infest my machine with malicious software? Will it attempt to steal information about me? Will this merchant defraud me? Many application-specific mechanisms have been put in place to deal with these questions. For example, eBay's reputation system permits buyers and sellers to evaluate each other. Credit card companies not only keep track of buyers and sellers, they cover losses from fraudulent charges. In effect, they act as insurance companies, which relieves the buyers and sellers from having to make as strong a trust assessment as they might have to otherwise. Certainly, it is in the financial interest of credit card companies to provide this service; the sharing of risk allows markets to function, and it is also possible to make a profit through the business of providing insurance.

One of the interesting trends on the Internet is rating sites, where users give ratings of everything from hotels and restaurants to clothing and movies. Current schemes may have flaws, but they signal an important transition from an isolated, individual Internet experience to one embedded in a shared social context. In the real world, most of what we do is rooted in a shared context, but the original image of the "personal computer," for use at home, alone, seemed to decouple our respective experiences. Designers of applications have had to reconstruct that ability to share experiences and generate understanding of the world through interactive processes prevalent offline.

Some "social interaction" schemes can be abused, just as spammers abuse email. Most rating sites do not demand that users give a strong verification of their identity. They may require their users to give some bits of information about themselves, but instead of being identified by name, users choose a "handle" or pseudonym. What is to prevent a user from creating multiple pseudonyms and posting scores of bogus reviews, positive or negative, to change the rating of something?

Perhaps we would be more comfortable with reviews and ratings provided by our friends, people we know and (to an adequate degree) trust. Network-based constructs, such as social networking sites like Facebook, allow us to relate to our friends online. They capture a robust aspect of identity because users (usually) link their online identity to friends that they know in the real world. Given that these sites function on a basis of strong identity, not just pseudonyms, they might serve as the foundation for a rating scheme that allows the user to place a higher degree of trust in the ratings.

Not all users are nice or trustworthy. Internet applications must be constructed to detect and deal with "bad apples."

What are the options? If the law has been broken, perhaps the law enforcement tools that are open to the government can be used. But what if people online are simply rude or disruptive? How can a community protect itself? Spammers disrupt blogs by putting their spam messages into the comment sections of blogs. Disruptive players (called "griefers") interfere with multiplayer games through behavior or tactics that are irritating to other players.

A clear response to such behavior is shunning or expulsion from the community. The question for an application designer is whether the ability to shun or expel a user should be part of the system. Again, the issue of identity is key. If the application requires that users provide a strong indication of identity that is hard to forge or replicate, then a user can be ejected. Games that require users to sign up with a credit card can reject the card, which means that expelled players can return only as many times as they have different cards. A credit card company can refuse to serve a merchant, or refuse to authorize a payment to an overextended purchaser. On social networking sites such as Facebook, if the users have invested a great deal of effort constructing an identity that is linked to the identity of friends, ejection would be a painful punishment. But if a system does not require a user to present a strong form of identity, as many do not, then a user ejected under one pseudonym can obtain another and return.

The construction of online identity is an important aspect of forming a stable community. On the one hand, demands for strong confirmation of "real" identity may chill certain sorts of valid behavior, from political speech to searching for information on sensitive health issues. On the other hand, weak identity may make it hard to detect misbehavior (such as "ballot stuffing" on rating sites) or to eject misbehaving users.

As we begin to explore the experience of using the Internet, we might start by asking: who uses the Internet, and for what purposes? Who does not use the Internet, and why? In his essay, John B. Horrigan draws on survey data he gathered for the Pew Internet & American Life Project, and more recently, for the Federal Communications Commission's National Broadband Plan. Based on recent data, about two-thirds of American homes have broadband access, and people use the Internet for a wide and growing range of purposes, including sending email messages, using the Web, making or researching purchases, gathering news and weather information, watching a video or listening to music and radio, banking, playing games, and connecting with friends using social media tools.

On the other hand, about 22 percent of surveyed homes report that they do not use the Internet at all, citing reasons such as cost, inadequate digital literacy, lack of relevance, or deficient service in their area. The data reveal widespread concern about loss or misuse of personal information; 45 percent of non-users cite fears of bad things that might happen online.

More detailed data from Pew (reported elsewhere in this volume) make clear that the pool of non-users is not homogeneous across society. Non-users tend to be older and of lower socioeconomic status: the poor, the less educated, and the elderly are less likely to partake in the Internet experience. Horrigan observes that as more and more aspects of society move online, the costs of nonparticipation increase, to both non-users and society at large. Nonparticipation online can limit job opportunities – with 80 percent of Fortune 500 companies accepting only

online job applications – as well as access to online government services or health information. Horrigan concludes that society must address barriers to using the Internet, which are not just lack of hardware, but lack of mastery of the increasingly complex skills needed to participate: what he calls *digital literacy*. As the demands for skill level rise, the costs of exclusion may become increasingly significant.

A common fear is the loss of privacy and the misuse of personal information. Today, the rules about privacy are spelled out in often long, confusing "privacy policies" or "consent forms" offered by various providers of network services. In her essay, Helen Nissenbaum rejects this approach. Central to her argument is the observation that cyberspace is not a distinct space with its own distinct norms. Much of what we do on the Web (or on the Net generally) is a reflection of something we do in the real world. Norms from that context, including privacy standards, should be expected to hold in the equivalent online context. But currently there is no recognition of context and implied norms; thus, the privacy consent form must carry the total burden of defining the expectations of the parties who participate. To the extent that a policy tries to capture nuances, it becomes overlong and incomprehensible; to the extent that it aims for brevity and readability, it describes only the general nature of the policy and omits the details that matter in practice.

There are well-understood contexts in which all parties understand the norms that apply. Health care is governed both by laws and by commonly understood norms of behavior. Banking, whether online or offline, is similarly governed by both law and custom. Nissenbaum suggests that many other online behaviors could be understood in terms of prior offline analogues. For example, using a search engine might be analogous to using a library card catalog, which has a strong tradition of freedom from observation and tracking. Even if the online experience is somewhat novel, we can often find real-world analogues.

The crux of Nissenbaum's argument – that the online experience does not take place in a homogeneous and unique context but in a range of contexts that will develop different customary norms and governing laws – can be extended to attributes other than privacy. As I noted above, individual contexts will call for distinct approaches to identity. Coye Cheshire looks at the concept of trust online: how users decide if a service is trustworthy, whether to trust individuals they encounter online, and whether they can rely on the network and the services provided over it. He explores the meanings of *trust* and *trustworthy* in different contexts, observing that in instances where the risk is low, users will be willing to proceed in the face of considerable uncertainty about whether a website, a service, or an individual is trustworthy. A restaurant review may be malicious or hyped, but its accuracy is only minimally consequential for a prospective diner. In cases where the potential risk is high, tools are put in place to minimize uncertainty. Online banking bears a potential high risk, but banks have gone to considerable lengths to remove uncertainty from transactions and give users a high level of confidence that their banking services are trustworthy. Cheshire notes that mechanisms to enforce constraints on behavior (so that users can proceed without developing trust in the other parties) erode trust and the mechanisms by which it arises. Trust can emerge only in a context of ongoing interaction among parties where betrayal is possible. Cheshire's analysis of trust (and the distinction

with *trustworthy*) draws on the one hand from a range of writings on the subject, and on the other hand from experiments involving users in online contexts. He argues that in the future, the Internet will depend on social forms and institutional arrangements as much as technologies and systems. The Internet *is* the real world.

Fear of bad experiences online is an issue for users and a barrier for non-users. Three essays in this volume deal with the problem of system and network security: protecting users, their computers, and the network from attack by malicious parties. The term *security* covers a range of concerns, including attacks by criminals on servers storing sensitive information, attempts to break into and subvert personal computers, and espionage carried out by states and powerful private-sector actors. Again, the essays collected in this volume focus on topics that are relevant to the experience of the individual user; they are less concerned with the potential of cyber-war and more so with the events that give users pause in their daily activities.

Vinton G. Cerf catalogs the many forms that these perils can take, providing insight into the roots of system insecurity as well as institutional approaches to improvement. Hazards include theft of personal information, spam, and the capture of one's computer by a remote operator, who then uses the computer to launch spam attacks on other users, or to flood a target site on the network with traffic to overload and disable it (a so-called denial of service attack). Given that perils can arise from both malicious acts and accidents, Cerf introduces the term *cyber-safety* to widen the scope of our objective beyond the more narrowly defined *cybersecurity*. Using a number of metaphors, including biology (viruses and infection), real-life analogues to understand the online experience (books versus e-books), and comparisons with offline mechanisms of protection (police and fire departments), he sketches the landscape of risk and response.

Deirdre K. Mulligan and Fred B. Schneider propose a new rationale by which society can improve its overall security posture. They first review past approaches to improving online security and examine why these approaches have failed. The attempt to provide security by building entirely invulnerable systems is simply impractical: today's systems are too complex, and the required level of effort would be too costly. Efforts to characterize the security problem as one of risk assessment and management (as we do in the offline world, using tools such as insurance) fail because we lack the tools and methods to quantify online risk. Finally, attempts to improve the landscape of security by using tools of deterrence to discourage misbehavior fail because we lack effective means of attribution and coherent means to pursue attackers across the jurisdictions of different states.

With this analysis as background, Mulligan and Schneider suggest a different way to think about improving security: through a new doctrine they call *public cybersecurity*. Their doctrine views the framework of public health and public health institutions as a model for cybersecurity. Just as good health is a benefit to all of society that must arise from the health of individuals, overall online security will improve by means of the steps individual users take to keep their own user contexts secure. But the benefit to any one user may not seem significant enough to justify his investment of effort and money into improving his own security. Security, like health, is a public issue, not an individual one. Thus, Mulligan and Schneider explore how the analogy of public health can be used to better

understand a large number of online issues, including system development, online surveillance, keeping systems up to date (installing "patches"), and isolation and quarantine of systems. Using the analogy to public health, this new public cybersecurity doctrine envisions a rational balance between the public interest in improved security and the rights of the individual.

L. Jean Camp considers the explicit question of whether and how we can motivate individual users to contribute to improved overall security. She uses two theoretical framings to explore this question. The first is *peer production*, in which users self-organize to create information (or other desired outcomes). She argues that users can be motivated to self-organize in ways that produce better system security, if the security challenge can be framed as a set of discrete tasks for which users can self-select based on skills and proclivity. She offers several examples, involving both technically skilled and ordinary users. The second regards the Internet as a common good, or a virtual *commons*. Using criteria developed by political economist Elinor Ostrom, she explores how the security problem can be framed in a way that allows users to self-regulate the commons. These two theories help model and define the circumstances under which user-centric efforts can be effective.

Several essays explore specific classes of behavior on the Internet. R. Kelly Garrett and Paul Resnick examine the experience of getting news and opinion online, questioning the hypothesis that personalization of news, made possible by the Internet, leads to increased political fragmentation. They reject the necessity of this outcome: personalization can take many forms, they observe, with different implications for social outcomes. Research suggests that people crave opinion reinforcement but do not go out of their way to avoid diverse viewpoints. If news is personalized along ideological lines, mirroring the ideologically segmented world of cable news today, it could indeed lead to increased fragmentation. By contrast, if personalization is used to expose willing readers to a range of viewpoints, selected perhaps for quality and thoughtfulness rather than bias, the result could expose readers to a more balanced selection of material. Research suggests that readers would be open to this sort of personalization. Narrow partisan channels *force* people to choose, but it is not clear that this is what people would prefer if given the choice.

The authors observe that "the technology and how people use it are still malleable; subtle architectural changes could have far-reaching implications for future news consumption patterns. [This] will require effort and creativity. . . . [T]echnologies that produce desirably diverse news streams may not emerge naturally." In understanding the Internet and the experiences that it provides, this observation is critical. As Garrett and Resnick and other essays remind us, the Internet is a built artifact. It is the way it is because people designed and built it to be that way. Thus, the future will be defined by those who choose to step up and design it. Originally, the designers of the Internet and its early applications (such as email) were researchers funded by the federal government. But with the success of the Internet, most of this effort has migrated to the commercial sector. And the motivations of the commercial sector may not perfectly align with one or another vision of preferred social outcomes.

In their essay, Kay Lehman Schlozman, Sidney Verba, and Henry E. Brady explore another specific class of online behavior: participation in the political process. A long-term concern with politi-

cal equality compelled the authors to understand whether the Internet might lower barriers to various sorts of political participation. Using survey data from the Pew Internet & American Life Project, they asked whether the Internet has changed political involvement in fundamental ways.

The survey focused on political participation as a function of socioeconomic status (SES) and age. Observing the striking power and durability of SES-based political inequality, the authors conclude that the Internet is not the "great leveler" that some optimists might hope. Online political behavior shows trends that are similar to traditional offline behavior. Significantly, these trends are not mirrored in other behaviors, such as participation in social networks, where SES is much less a predictor of participation.

As the Pew study confirms, given that those with lower SES are less likely to be on the Internet at all, that group suffers a double barrier: lack of access and the traditional bias against participation. Age is another factor: older populations are less likely to be online; even if older users online seem to be politically active, the overall level of online participation in older populations is low.

Schlozman, Verba, and Brady also look at new forms of political activity that arise online, such as political blogs and social networks. While the survey reveals that younger respondents (under age twenty-five) are heavy users of this technology, the authors do not find strong reasons to conclude that these new technologies may lead to a change in the nature of political participation. But they note that the Internet and its applications are young, and technical design decisions as well as changing user behavior are unpredictable.

Lee Sproull looks at a different sort of online activity: *prosocial* behavior; that is, activities intended to help people other than oneself, such as volunteering and supporting charities. (Volunteer activities include service projects, health support groups, peer production of information, and citizen science.) She catalogs various forms of observed online prosocial behavior and provides estimates of its prevalence. She notes that while all major types of online prosocial projects share a small number of attributes that derive from the underlying network technology and communications applications, each context for online behavior is a symbolically differentiated place on the Net, and different people seek out different places. Sproull then discusses the features of the online context that can facilitate or encourage such participation: the modularity and granularity of the task (how the work of one user is scoped, specified, and then aggregated), a site's social structure, and techniques to motivate participants. She also discusses the nuanced role of identity and trust in shaping and motivating the participant experience.

Finally, the essay by Yochai Benkler provides an analysis of two current events that illustrate how various actors, including the government and powerful private-sector players, engage to shape what happens on the Internet. The two cases are WikiLeaks and the ongoing struggle by the holders of copyright to repress the sharing of unauthorized copies of their material. The two cases have much in common: first, they both revolve around attempts to prevent access to material in the context of an open Internet that makes unregulated access the norm; and second, they demonstrate a complex and tangled interplay between the public and private sectors. His concern is that in both these cases, the approaches put in place would allow the blocking of access without the nor-

mal protections of U.S. law. More generally, his essay reminds us that not everyone has the same aspirations for the Internet, and that the future will be shaped by a tussle among those who care enough to advocate for their objectives.

A number of common themes run through the essays in this issue. First, the term *cyberspace* is potentially misleading on two grounds. The term suggests that the Internet is a distinct "space" or "place" to which we go online. Rather, our experience using the Internet is not separate and disconnected from our offline experiences. Much of what we do on the Internet has a close relationship to our offline behaviors. Additionally, the experience of using the Internet is not homogeneous and subject to consistent norms. Various aspects of using the Internet will differ in important ways, including the norms and expectations about behavior; the degree of uncertainty, risk, and significance; and the nature of the resulting interactions.

The Internet, as a low-level platform that supports a range of applications, is not the technology that creates or defines the user experience. The Internet itself, as distinct from the applications that run on it, was originally seen as a neutral platform intended to support as many patterns of interaction as possible. This generality implies that it supports both "good" and "bad" patterns of interaction, and "good" and "bad" experiences. It is the applications that have been built on top of this platform – email, the Web, Facebook, Internet telephony, search engines – that define the user experience. Each application is its own context with its own affordances and constraints.

The Internet is a built artifact. It is designed and engineered. One must not think of the Internet as fixed and exogenous; it can change and evolve, sometimes rapidly. As we consider the limits and benefits of the current Internet, we should not think only like observers or analysts, but also like designers and engineers. By most measures, the Internet is very young, and its designers have much to learn, including how to facilitate predictable, safe, and rich human interaction.

David D. Clark

An important corollary to this last point is that the future character of the Internet will depend in large part on which parties choose to invest in shaping it. To the extent that the Internet is being designed and built by actors with commercial, profit-seeking motives, we may not see the emergence of applications that shape the social experience in ways leading to better civic engagement, prosocial activities, or news sources that offer personalization without polarization. The designers of the early Internet, mostly supported by research grants from the government, may have had different motivations from the designers of today.

Not only is the Internet a built artifact, it is a constraining artifact. In the offline world, people interact in a complex and open-ended environment that offers many different cues and signals for social interaction. In the online world of an Internet application, the context of interaction is bounded and limited by what the application designer has provided. It is a closed system, except to the extent that the application designer has intentionally created tools whereby users can evolve the social context. As a result, we should not view the resulting patterns of human behavior as only socially emergent. Online behavior is shaped by the capabilities and constraints of applications as well as by the socially centered, human factors that influence how people use those applications.

Past experience has taught us that while social interaction mediated by the mechanics of software is a constrained

and limited experience compared to its offline counterpart, the online version has powerful advantages. Thus, we can expect the Internet to be a compelling platform for interaction and engagement. Good design can help mitigate and compensate for limitations. As the essays contained herein remind us, the Internet is not a fixed artifact, but evolving and flexible. There are many possible futures for the Internet, depending on which actors choose to define that future. One of the goals of this issue is to stimulate debate about what that future should be.

What are the Consequences of Being Disconnected in a Broadband-Connected World?

John B. Horrigan

Abstract: The evolution of the Internet in the past decade – from a slow, stationary, and primarily communications-based technology to a mobile, fast technology that supports a range of communication, participatory, and transactional applications – has made access more valuable, and disconnection more consequential, for individuals. This evolution means that stakeholders should embrace a different framing of the digital divide, focusing on the costs of digital exclusion. These costs can fall on an individual, if the Internet is the only way to carry out some tasks, and on society, if expensive and less-efficient legacy systems must be maintained to serve a shrinking minority without access. Whereas the digital divide debate concerns technology scarcity for certain population segments, addressing the costs of digital exclusion is about developing people's capacity to manage today's abundance of digital resources. This essay offers suggestions on a framework to develop tools that will enable individuals to take advantage of the affordances of broadband.

JOHN B. HORRIGAN is Vice President of Policy and Research at TechNet, where he leads research on technology, innovation, and telecommunications policy. Previously, he was part of an FCC team that developed the National Broadband Plan (NBP), designing and conducting the FCC's first national survey on broadband adoption and usage. The survey findings were highlighted in the NBP's first working paper, *Broadband Adoption and Use in America*. He also served as Associate Director of Research for the Pew Internet & American Life Project.

Ten years ago, the debate over the digital divide was mainly a binary one about access: some people had a computer (and therefore dial-up access to the Internet), and some did not. This disparity was widely thought to be unfair because not all members of society could take advantage of the brave new world of the Internet – that is, the easy connectivity to other people and a wealth of information. Those without access were cut off from the benefits of communicating with others (for the most part via email and sometimes on more organized many-to-many groups, such as Listservs).

Stakeholders should embrace a different framing of the digital divide, one that acknowledges the need for a broader range of policy measures to address imbalance in the adoption of broadband Internet. While the digital divide debate should not abandon equity arguments, it should also consider

the costs to individuals *and* to society of having a sizable portion of the population offline. The high and rising cost of digital exclusion makes a lack of Internet access more disadvantageous today than a decade ago. Furthermore, lower levels of engagement with online resources among those *who are* online are more consequential than a few years ago.

Policy measures to address the inequities of the digital divide have rightly focused on providing access to necessary hardware and giving users the computer skills to negotiate the basics of an online session. Some initiatives refurbished old computers and either gave them to people who could not otherwise afford them or sold them at a cut rate. Grant programs – funded by the federal government, states, or localities – were undertaken to support hardware and training for low-income people. Nonprofit organizations also seized the opportunity to bring access to less well-off Americans; some managed to scale their initiatives regionally or nationally.

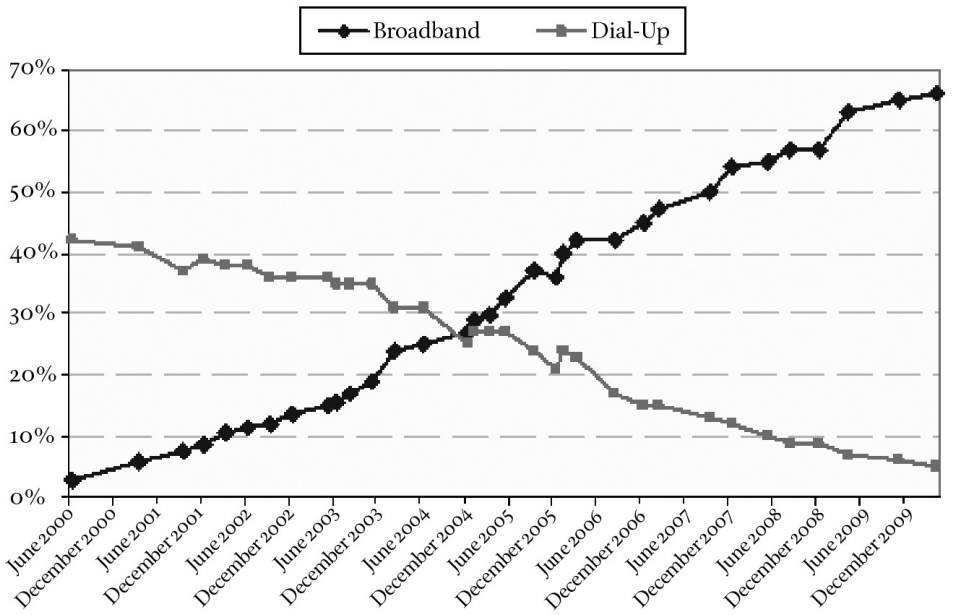
Given the importance of broadband in carrying out everyday tasks, such measures are still crucial, but alone are not sufficient to fully support online adoption and use. In addition, policy-makers and other stakeholders must give people the resources to cultivate a wider set of digital skills, such as media literacy, the ability to access civic information, and proficiency in managing personal privacy. Users must also understand that they have an obligation to educate themselves on the skills needed to participate in an increasingly broadband-mediated world. The Internet's expanding role as a conduit to performing important daily functions has increased the need for tools to help users negotiate an ever more complex online environment.

This essay begins with a data-driven discussion of how online access has changed in the past decade. It reviews trends in Internet and broadband adoption, as well as mobile adoption, while also examining the varying levels of engagement with digital tools across the user population. Second, the essay reviews the literature on technology adoption and user behavior, focusing on the role of skills in user engagement with online resources as well as factors that encourage or inhibit skill development. Finally, after arguing that user support systems are necessary in today's online world, the essay offers suggestions to stakeholders on how to create such systems.

As surveys from the Pew Research Center's Internet & American Life Project have documented, online access and use in the United States have evolved over the past decade. Notwithstanding a recent slowdown in the growth of broadband adoption, far more Americans today connect to the Internet using an "always on" high-speed connection than was the case just a few years ago (see Figure 1).

Data gathered at the end of 2009 and in the first quarter of 2010 show that some two-thirds of Americans have broadband access at home. The Federal Communications Commission (FCC) survey conducted in connection with the National Broadband Plan (NBP) found in November 2009 that 65 percent of Americans had broadband; a Pew survey conducted in April 2010 found that 66 percent of Americans had broadband at home. Data gathered by the National Telecommunications & Information Administration (NTIA) in 2009 and 2010 census surveys show similar trends. In 2009, the NTIA reported that 63.5 percent of households had high-speed Internet access, a figure that had grown to 68 percent by October 2010.[1] More Americans are Internet users than home broadband users, as some Internet users have access at a place other

Figure 1
Trends in Home Broadband and Dial-Up Adoption, June 2000 to December 2009



Source: Based on the author's calculations using data available from the Pew Internet & American Life Project, http://www.pewinternet.org/Data-Tools/Download-Data.aspx.

than home (such as work or a library), and some still use dial-up Internet connections. Overall, according to the NTIA, 72 percent of Americans were Internet users in October 2010.

Many U.S. policy-makers have become particularly focused on how broadband adoption in the United States compares with other countries. According to a June 2010 survey by the Organisation for Economic Co-operation and Development (OECD), the United States places fourteenth in the world for fixed broadband subscriptions per one hundred inhabitants, trailing the Netherlands, Denmark, Switzerland, and Korea (the top four countries), but also France, Germany, the United Kingdom, and Canada. This middling rank, roughly the position the United States has maintained for the past several years, contrasts with the U.S. standing a decade ago. In 2001, the OECD placed the United States at fourth in the world (using the metric of the number of lines per one hundred people).[2]

Some observers have criticized the way the OECD develops its ranking and the wisdom of policy-makers who pay so much attention to this metric.[3] Nonetheless, these ratings help policy-makers frame broadband policy as both urgent and relevant to America's economic competitiveness. Although competitiveness is not the focus of this essay, it plays a significant part in U.S. broadband policy.

The growth in broadband adoption has been accompanied by any number of innovations that have made online life more useful, fun, satisfying, and consequential for users. People sent email messages from a desktop computer a decade

ago, and although many of us still do, many also send text messages from a smartphone, tweet on the go, and routinely share information on Facebook. People purchase items online, rate what they buy, gather news and analysis, seek out health and medical information, educate themselves, and engage in many other activities. Figure 2 presents snapshots of what people do online and how use has changed in the past decade.

A greater share of Americans perform a wider range of online activities than a decade ago. About twice as many people get political news and information online today than in 2000, and more than twice as many adult Americans have bought something online and banked online. Activities that were unheard of in 1999, such as blogging or social networking, are mainstays in the daily routines of large numbers of Americans. People also access the Internet in more varied ways than a decade ago.

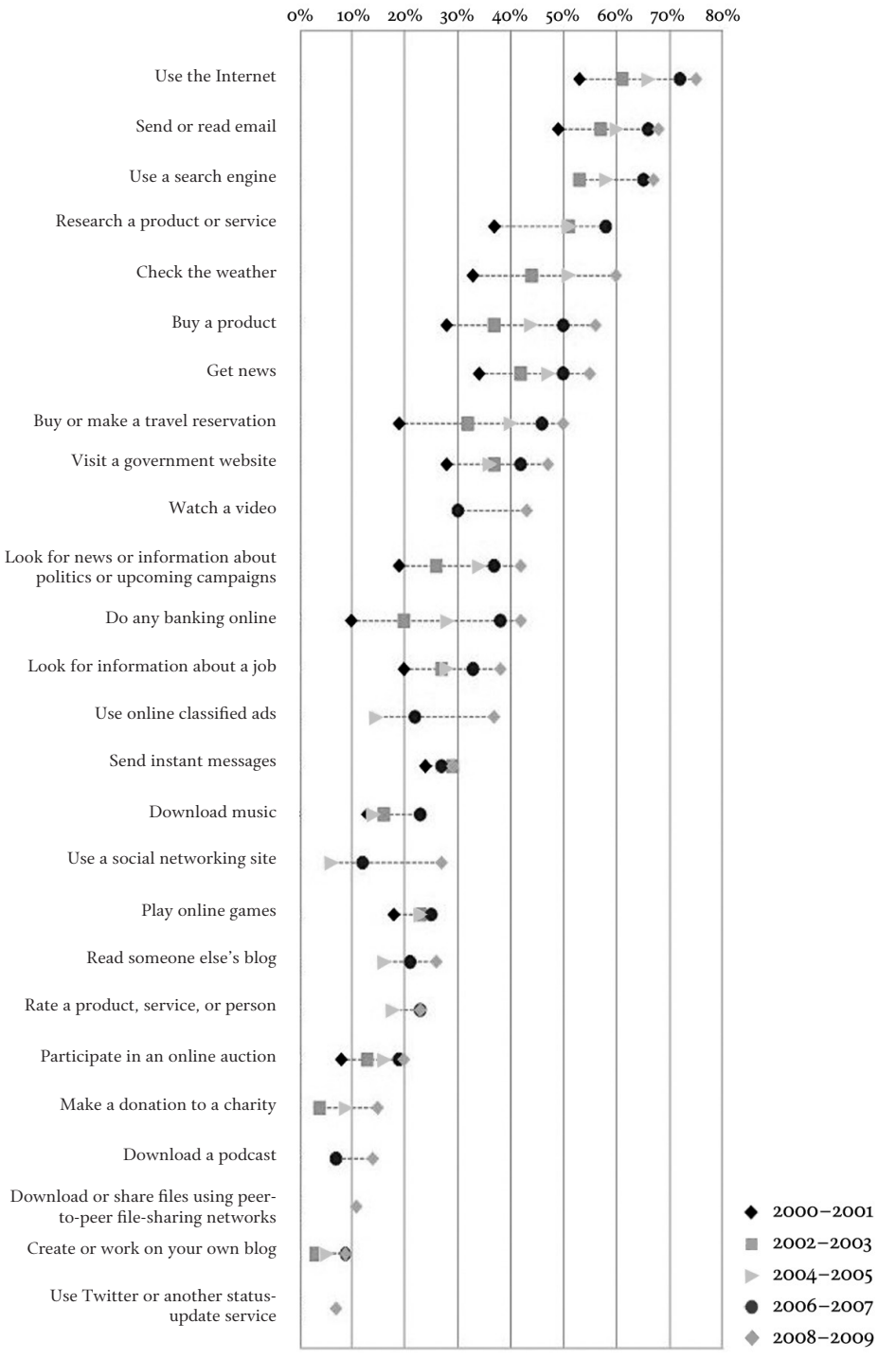
Access is no longer stationary and tethered. With the advent of laptops, netbooks, tablets, and smartphones – in conjunction with the development of wireless networks – access is also portable and ubiquitous. In 2000, slightly more than half (53 percent) of Americans had cell phones, a number that had grown to 86 percent at the beginning of 2010. The capabilities of these devices have increased dramatically: two-thirds of cell phone users send text messages, and 40 percent use their phones to access the Internet.[4] Today, more than half (52 percent) of Americans have laptop computers; many use them to access the Internet through wireless connections.[5] Finally, tablets and e-book readers are beginning to gain a foothold among early adopters. Some 5 percent of adults report having an e-book reader (such as a Kindle or Sony Digital Reader) and 5 percent report having a tablet computer (such as an iPad).[6]

However, individuals still have qualms about online life. A 2008 Pew survey showed that online users had substantial reservations when asked about how companies might use personal information gathered from people using cloud computing applications. The survey found that 68 percent of those who use online applications and storage services would be very concerned about companies using information gathered in the course of such activities to serve targeted ads.[7] The FCC survey conducted for the NBP (mentioned above) found that many Americans – especially those without home broadband access – reported concerns about the security of personal information online. Some 45 percent of all Americans said they strongly agreed with the proposition that it is too easy for their personal information to be stolen online – a figure that stood at 57 percent among those without broadband at home. The survey also found a link between fears about the security of personal information and lower levels of use of online resources.[8]

Notwithstanding user worries about some aspects of online life, the Internet has evolved, in ten years' time, from a slow access technology used by a minority of the population to a technology that, for most of the nation's households, is fast and mobile. It is no longer merely a new way to communicate. A broadband-driven Internet is more consequential than the dial-up Internet – and in some ways is more complex to use – but it also affords collaboration, creativity, and participation in ways unimaginable a generation ago.[9]

These data on access and use help illuminate the cost of digital exclusion, which is reinforced in academic literature. Computer scientist Rahul Tongia and communications scholar Ernest Wilson have ar-

Figure 2
Percentage of American Adults, including Internet Users and Non-Users, who have Engaged in Various Online Activities, 2000 to 2009[10]



gued that the costs of not being online rise faster than the growth of the network. Tongia and Wilson's formulation accomplishes in a model the reframing that I am arguing stakeholders must fold into their narratives. As more people use broadband-connected networks, the cost to those *not* on the network rises at more than exponential rates when even a few people are excluded.[11]

According to the NBP, the cost of digital exclusion is characterized by a decline in offline alternatives as online ones take precedence.[12] Job searching is a good example. When home broadband penetration was at half its current level, people typically had plenty of ways to search for jobs. Online ads were certainly one way, but ads in the print edition of newspapers were a viable alternative; job networking could be conducted online as well as in person; and job searchers could submit applications online or by traditional means. Today, print newspaper ads are less useful sources than in the past, and – importantly – it is often impossible to apply for a job unless one applies online. Some 80 percent of Fortune 500 companies accepted *only* online applications for jobs in 2010.[13] Evidence also suggests that home broadband access can lessen the incidence of "job discouragement" in the labor market, as Internet access may spur some people to stick with a job search rather than exit the labor market.[14]

The basic act of getting the daily news is another example. A decade ago, online news websites were certainly popular, but the daily newspaper was still a robust source of news. More recently, changes in the economics of the news business have resulted in the downsizing of newsrooms. In 2009, the daily newspaper had 27 percent fewer newsroom employees than in the beginning of the decade; in fact, 25 percent of the loss in newsroom employment has occurred since 2006.[15] News is still very much available offline but has, in important ways, migrated to the Internet; traditional media outlets provide additional content online, and the advent of pure play websites (whether independent blogs or news aggregation websites, such as the Huffington Post or the Drudge Report) has supplied new ways to obtain information and opinion. A recent Pew Research Center report found that in a typical day, 36 percent of Americans use both digital and traditional sources of news, and 39 percent use only traditional sources. Although Pew considers online sources as a supplement to Americans' news diets, the easy availability of online news, whether on a computer or a handheld device, has resulted in Americans spending more time getting news in 2010 than at any point in the past decade.[16]

Interacting with government is also different than a decade ago. With more government services provided online, more people are taking advantage of them. It is often cheaper and easier for all parties to carry out government-citizen transactions online rather than via traditional outlets. For means-tested benefits programs (such as food stamps, school lunches, or Temporary Assistance for Needy Families), a family might have to fill out between six and eight applications, often at the cost of hours away from work. Agencies can save on procurement costs using cloud computing, with procurement savings of up to 50 percent.[17] Renewing a license online is easier than going to a government agency. If fewer citizens go in-person to an agency for a particular type of transaction, it may also be cheaper for government – and ultimately taxpayers – to support fewer physical locations for certain types of services.

A final example – online health and medical information – shows the cost of

digital exclusion with respect to participation in important decisions. Offline sources for health care information have not gone away in the past decade, but there has been an explosion in online forums by which patients share with one another information about their conditions. This type of networking, along with other online sources of medical and health information, helps patients become more actively engaged in health care choices. The upshot of this dynamic is that more people look online for health and medical information today than a decade ago. In 2000, 25 percent of Americans had used the Internet for such information, a figure that rose to 61 percent by 2009.[18] A lack of access to information means that people lose out on an opportunity to be more empowered and participatory in health care decisions. There is also evidence that broadband can improve the efficiency with which health care providers deliver care to patients.[19]

In the above examples, the cost of digital exclusion is twofold:

1) For individuals, those without access to broadband either miss out on information (such as job opportunities) or must use more costly means to accomplish certain tasks.

2) For society, institutions have to support the costs of legacy (and often more expensive) means of delivering services in order to meet the needs of a shrinking minority of people who do not have access to broadband, or who have access but lack the skills to use it.

Although there have been attempts to quantify the costs of digital exclusion, those estimates (derived in the United States and the United Kingdom) are best understood as a springboard for further research.[20] Yet digital exclusion and its costs are important conceptual handles for the ongoing debate about equity in the digital age. Mark Cooper, director of research for the Consumer Federation of America, notes that because broadband networks and functionalities continue to advance, and therefore "raise the level of skill necessary to use the network," the disconnected may become "even worse off" than they are today.[21] The link between the evolution of the broadband environment and skills acquisition suggests that stakeholders interested in access gaps must take skills into account when developing programs.

As consequential as having broadband access is (or seems to be) these days, it is important to resist the temptation to think about broadband as something everyone *has* to have. A necessity is not a requirement. Broadband is often – with reason – viewed as a necessity equivalent to electricity; this comparison, in turn, is frequently used to justify interventions to promote broadband. The discourse, however, often conflates access and use. Do we want everyone to have access to the infrastructure that provides service? Although the answer is clearly "yes," we must acknowledge that as a practical matter we mean "nearly everyone," as there are some places where the cost of deploying infrastructure would be exorbitant.

Do we want everyone to use broadband? However well crafted policies to increase access may be, attaining 100 percent home-broadband adoption is not a realistic goal. Even the telephone – widely accepted as indispensable in modern life – does not reach everyone; the FCC reports that in 2010, 96 percent of Americans had telephones, a figure that accounts for a small minority without this apparently indispensable tool.[22] Research that examines this issue shows challenges in having everyone adopt broadband. The FCC's November 2009 survey shows that 10 per-

cent of the adult population is "digitally distant," meaning they have neither the financial means nor skills and attitudes about computers that might draw them to online use. Another 7 percent are "digitally uncomfortable"; despite the fact that many have computers, most people in this group question the relevance of online access.

Although projections of future adoption prospects of non-users are inherently uncertain, the data suggest that some 17 percent of Americans – about half of current non-adopters – face significant hurdles to adoption. Thus, in the medium term, a U.S. adoption rate above 85 percent is unlikely.[23]

Even with the real and potential benefits of broadband use, not everyone will embrace it; among those who do, not everyone will prefer to use broadband in the ways favored by planners and visionaries. It is imperative, therefore, to keep in mind user preferences and abilities in thinking about how to take advantage of broadband's economic and social benefits. Encouraging broadband use should be about "demand pull" (devising ways to lure people to use) as opposed to "technology push" (making broadband use effectively required).

Despite anticipated challenges to attaining ubiquitous deployment and universal adoption of broadband, Congress charged the FCC to develop the NBP and to devise strategies for increasing broadband adoption. Partly driven by this objective, researchers have sought to understand not only who is not online, but why those without high-speed access at home choose not to have service.[24] A plurality of the 35 percent of U.S. adults who do not have broadband at home cites cost as the *main* reason they do not have broadband at home. Close to two-thirds of non-adopters point to factors other than the monthly fee or cost of hardware as the reason they are not home-broadband users.

According to the FCC, the barriers to home-broadband use include:

- *Cost:* 36 percent of non-adopters named cost as the main barrier, with 15 percent citing the monthly price for service and 10 percent citing the cost of a computer.
- *Digital literacy:* 22 percent cited factors that indicate digital literacy problems, such as unfamiliarity with a computer or worries about bad things that could happen to them online.
- *Lack of relevance:* 19 percent said they did not need additional speed (that is, to upgrade from dial-up) or that online content was not relevant to them.
- *Lack of available infrastructure:* 5 percent said they could not get broadband where they live.
- *Other reasons:* 18 percent said they could use the Internet all they wanted at work (3 percent), identified a combination of reasons (4 percent), or cited reasons that did not fall into a discrete category (11 percent).

Though cost looms large as a barrier to adoption, it is not an isolated factor; non-adopters' expressed concern over their level of digital skills clearly is significant. When given the chance to identify more than one reason for not having broadband at home, half of non-adopters pick more than one. Specifically:

- 66 percent cite cost (that is, monthly access fee or cost of computer);
- 52 percent cite relevance (that is, the Internet is a waste of time or they do not think there is relevant content online for them); and
- 47 percent cite digital literacy (that is, they worry about bad things that can

happen online or they are not comfortable with computers).[25] More specifically, 46 percent of non-Internet users cite lack of comfort with computers as a reason why they are not online, and a similar number (45 percent) say they do not use the Internet because they worry about "all the bad things that can happen online."

It is worth noting that the FCC's findings about barriers to broadband differ from those found by the NTIA. When asked to identify the main reason they do not have broadband, respondents in the NTIA survey cited the following reasons:

- 38 percent: don't need it – not interested;
- 26 percent: too expensive;
- 18 percent: no computer or inadequate computer;
- 6 percent: other reasons;
- 4 percent: can use it somewhere else;
- 4 percent : not available in area; and
- 3 percent: lack of confidence or skills.[26]

These discrepancies have to do with how each survey framed its questions. The FCC survey had a different set of categories and undertook a two-step process in which respondents were asked first to list barriers they face and then to identify the main one. The NTIA survey asked a single question. However the question is framed, barriers to broadband adoption clearly have a cost component, but the overlapping concepts of digital skills or literacy and perceived lack of relevance are significant parts of the picture.

Also at issue are barriers that inhibit use of broadband among those who have service. Communications scholar Eszter Hargittai has found variation in levels of online skills that, among young adults, predict the kinds of activities they engage in online. Other research that has developed typologies of Internet users finds, among other things, that a variety of factors – skills, attitudes about information technology, demography – shape how intensively people use information and communications technologies (ICTs).[27] This variation in online usage patterns is to be expected; not everyone needs to be (or wants to be) a fiber-to-the-home/iPad-toting tech gear head. Pew's 2009 report classifying the technology usage patterns in the general population found that 61 percent of adult Americans – many with broadband and cell phones – are not heavy tech users. Some are happy that way, while many have tepid usage patterns due to lack of digital skills.[28] Finally, adapting a method Hargittai developed to measure user skills, the 2009 FCC survey for the NBP found that 29 percent of broadband users had low levels of computer and Internet skills.[29] If users' skill levels or worries about Internet use cause them to turn away from the worthwhile affordances of online life, then developing tools to address these issues may be socially beneficial.

Although cost is most often cited as the main reason for non-adoption, other factors matter: namely, limited digital skills and the perception among non-adopters that access is irrelevant to them. A combination of these reasons often contributes to the decision not to adopt. The following two sections explore the barriers of digital skill and literacy and users' perception that the Internet is not relevant to them. Alleviating cost pressures is obviously important but will not be discussed here; addressing this issue would mean delving into the complex area of reforming the Universal Service Fund and the Lifeline/Link-up programs, which currently provide cost relief for telephone service. The NBP recommends that these programs be updated to reflect the

reality of broadband, and proposals have been advanced on how to proceed.[30]

It is hardly novel to suggest that online engagement is underpinned by the requisite literacy and skills to use ICTs. *Digital skills* refer to the basic competencies of knowing how to operate a computer or initiate an online session. These capabilities do not necessarily require a high level of expertise, but they do demand sufficient knowledge to communicate to others the nature of tech problems when users cannot resolve technical difficulties themselves. *Digital literacy* is a multidimensional concept describing higher-order faculties that, according to media literacy scholar Renee Hobbs, "encompass the full range of cognitive, emotional and social competencies that include the use of texts, tools and technologies; the skills of critical thinking and analysis; the practice of message composition and creativity; the ability to engage in reflection and ethical thinking; as well as active participation through teamwork and collaboration." This level of proficiency extends beyond knowing how to use technology to knowing how to use it to carry out important tasks in everyday life, whether that is reading the news, obtaining information about a medical condition, or making a purchase. Many strategies to promote digital literacy try to make the value proposition to broadband adoption more attractive by opening people up to the "pleasures and power" of using the Internet to become better informed, while recognizing that "people cannot be forced to engage with the public life of the community."[31]

The process of conveying to people the usefulness of broadband typically has a social dimension. Classic studies on technology diffusion explicitly highlight how the social system helps spur technology adoption.[32] People learn about a new product from the people around them; their social networks, in other words, play a key role in helping people discover the utility and usability of an innovation. In an empirical study of computer adoption using 1997 data, economists Austan Goolsbee and Peter Klenow found that the presence of learning externalities influenced the decision to have a computer at home. People were more likely to have a computer if they lived in areas where others had computers and if a large share of family and friends had one. Moreover, email users, as distinct from users of specific types of software, were the most likely to have computers; this fact suggests the manner in which communicative features of the technology motivated adoption.[33] Eszter Hargittai has found that online skills are not evenly distributed among young adults – a cohort often thought to be uniformly tech savvy – and that socioeconomic background is a predictor of digital skill levels for this group.[34] Finally, in a nationally representative sample in the United States, I found that measures of online skills are correlated with the number of online activities that people engage in.[35]

Non-adopters' perceptions about broadband's lack of relevance are also an important factor in their decision not to procure home high-speed service. The notion of broadband's relevance is embedded in a user's context. A non-broadband-using senior citizen may have traditional media habits and not find online news worthwhile. It is possible that being made aware of distinctive characteristics of online news content might make broadband access more appealing. Similarly, a Hispanic individual with limited English-language skills may not have broadband – but may choose to get it if she is introduced to Spanish-language online content that is compelling. Finally, others might

buy broadband if they discovered various economic upsides: namely, the money-saving potential of online comparison shopping or the convenience of carrying out transactions online.

Whereas promoting digital literacy is very much a social process, conveying broadband's relevance to people has a greater emphasis on imparting information. The distinction means that promoting broadband's relevance requires a traditional marketing-campaign approach. Simply learning about the existence of Spanish-language content or the money-saving possibilities of online shopping might be the catalyst that draws people to use. Effective means of conveying the value of broadband to those who question its relevance are likely to emphasize "targeted partnerships that understand both the needs of and the distribution channels relied on" by these non-adopters.[36] Because people's perception about relevance depends on context, these partnerships are likely to be successful if they work from the ground up, that is, if they are local in nature and are embedded in non-adopters' communities.

Strategies for linking the promotion of digital skills and relevance to policy initiatives are not well developed. Programs to bolster broadband adoption – some publicly funded, others spearheaded by nonprofit organizations, and many partnering with the private sector – have existed for more than a decade. Although many have been subject to review, a rigorous assessment of methodologies used would shed light on the difference the programs have made, compared to what might have occurred in their absence. In an exhaustive examination of adoption-promotion efforts, economists Janice Hauge and James Prieger conclude that apparently successful programs take a comprehensive approach to training.[37] Such initiatives provide discounts on access and equipment, but also training on how to use the Internet, with a focus on demonstrating the relevance of specific online activities and guidance on how to carry out certain online tasks.

An assessment of existing training and promotion efforts based on rigorous social-scientific study would be useful in designing policy initiatives. Even in its absence, the above discussion highlights the set of digital skills and literacies individuals need to take part in a broadband-connected world:

- *Computer skills*. This category includes the ability to use a computer – whether it is a desktop, laptop, or handheld device – as well as the capacity to upgrade one's skills. The hardware for modern connectivity is continually evolving, and users must be capable of learning and understanding access technologies.
- *Media literacy*. Media literacy refers to the ability to find trustworthy information online about issues that matter to the user personally or to his community. According to the Knight Commission report *Informing Communities*, media literacy enhances the information capacity of individuals, enabling them to sort through the wide range of choices for news and information in today's media environment.[38]
- *Civic engagement*. The ability to find reliable information about elections and government online is arguably an outcome or goal of media literacy, but it is worthy of special attention nonetheless. Data show that a large share of Americans not only use the Internet for news about politics, but also use online tools to carry out transactions with government agencies. Such tools can be particularly beneficial for low-income Americans, who may be more

time pressed in dealing with necessary government transactions than other people.

- *Managing personal information online*. Users need to understand the policies on personal privacy that govern websites and the ever-shifting rules in this realm. Given that concerns about the privacy and security of personal information – for both broadband users and non-users – inhibit the adoption and use of broadband, more information on how websites use the information people share online may help alleviate user concerns.

In each of these examples, what it means to be "media literate" or what it takes to effectively manage one's personal information will evolve over time.

If these skills are necessary for participation in a digital society, how should the gaps in skills provision be addressed? What initiatives should various stakeholders – government, the media, nonprofit organizations, and individuals – undertake? It is impossible to draw bright lines on these questions. Legislation will not outline what it means to be computer literate, but government – as well as other actors, such as libraries – may have a part in supporting training initiatives. Government is also likely to take a lead role in thinking about people's privacy rights in the digital age. Private news media companies may promote online media literacy, although public media and foundations could undoubtedly play a strong part here as well. Civic engagement cuts across all four categories. Given that good citizenship is a worthwhile social norm, the individual has an obligation to engage with his community, whether by voting, volunteering, or giving to charity. The media – public or private – may also provide information that promotes civic engagement, and government can play a role in designing applications that permit users to connect with government effectively. Personal information online is another area where all stakeholders may contribute.

It is worth emphasizing that individuals have an important obligation in each of these areas. The character of the Internet – interactive, collaborative, participatory – gives individuals opportunities for autonomy in the online commons, which in turn suggests a responsibility to cultivate the commons. That means maintaining one's own skills, but also, arguably, supporting efforts to help others develop their skills.

Work remains to be done to turn ideas about digital literacy and promoting relevance into policy and practice. As a starting point, the NBP recommends the development of an online portal that would allow users to look for information on how to upgrade their online skills and even follow lessons there on skills development.[39] Nonprofits already play a role in this space, as entities such as Common Sense Media, One Economy, and Computers for Youth either promote online skills or provide information about the nature of online content. Such initiatives typically have limited geographic scope, meaning that, as valuable as they may be, they do not scale nationally. Furthermore, there is no clearinghouse for best practices for access and training programs – a gap the NBP recommends be filled. With the resources for broadband provided in the 2009 economic stimulus bill, as well as the attention that the NBP has drawn to issues pertaining to broadband adoption and deployment, the time may be ripe for stakeholders in the policy community to consider, in a systematic way, how to meet the training needs of broadband users. That said, coordinating the different stakeholders – companies, new entities

devoted to expanding broadband access in communities, long-standing community-based organizations that have relationships with underserved populations – remains a significant challenge.

A central theme of this essay is how to make a technology that is currently not part of some people's lives both available and attractive to them. But the context for this discussion is digital abundance. When something is abundant – as means of online access and quantity of applications are – those without it become more at risk of falling outside the social, cultural, and economic mainstream. The accelerating pace of innovation in broadband will only exacerbate this risk. The issues discussed in this essay – developing ICT skills and promoting relevance – are not going away. As the "Internet of things" ushers in an era of home-energy management, and with advances in the electronic delivery of health care services, the benefits of connectivity will grow over time. Today's digital abundance will seem like light fare in a few years, and those left behind will forgo a range of benefits.

As the pace of change quickens, promoting access and managing abundance will become different sides of the same coin. To engage with the Internet at even a moderate level, an individual transitioning from *non-user* to *user* may quickly find himself with a host of new hardware with which to cope: a computer, displays in the house for smart meters, and perhaps a mobile app on a smartphone that helps manage a medical condition. Understanding how to use these technologies, why they are relevant, what the tools are to protect personal information, and how to avoid being overwhelmed will challenge newcomers to broadband. Moreover, these questions will confront many who already use Internet devices and services. For these reasons, providing resources for the development of skills to negotiate cyberspace – for new and current users alike – is a challenge for all participants in the broadband ecosystem.

ENDNOTES

1 National Telecommunications & Information Administration, *Digital Nation: Expanding Internet Usage*, NTIA Research Preview (Washington, D.C.: U.S. Department of Commerce, February 2011), http://www.ntia.doc.gov/reports/2011/NTIA_Internet_Use_Report_February_2011.pdf.

2 Daniel K. Correa, "Assessing Broadband in America: OECD and ITIF Broadband Rankings" (Washington, D.C.: The Information Technology & Innovation Foundation, April 2007), http://www.itif.org/files/BroadbandRankings.pdf.

3 For the former point, see Scott Wallsten, *Understanding International Broadband Comparisons* (Washington, D.C.: Technology Policy Institute, May 2008), http://www.techpolicyinstitute.org/files/wallsten_international_broadband_comparisons.pdf. On the latter point, see John B. Horrigan, "Why It Will Be Hard to Close the Broadband Divide," Pew Internet & American Life Project (Washington, D.C.: Pew Research Center, August 2007), http://pewresearch.org/pubs/556/why-it-will-be-hard-to-close-the-broadband-divide.

4 Aaron Smith, "Mobile Access 2010," Pew Internet & American Life Project (Washington, D.C.: Pew Research Center, July 2010), http://www.pewinternet.org/Reports/2010/Mobile-Access-2010.aspx.

5 Aaron Smith, "Americans and Their Gadgets," Pew Internet & American Life Project (Washington, D.C.: Pew Research Center, October 2010), http://www.pewinternet.org/Reports/2010/Gadgets/Overview.aspx.

[6] Kathryn Zickuhr, "Generations and Their Gadgets," Pew Internet & American Life Project (Washington, D.C.: Pew Research Center, February 2011), http://www.pewinternet.org/Reports/2011/Generations-and-gadgets.aspx.

[7] John B. Horrigan, "Use of Cloud Computing Applications and Services," Pew Internet & American Life Project (Washington, D.C.: Pew Research Center, September 2008), http://www.pewinternet.org/Reports/2008/Use-of-Cloud-Computing-Applications-and-Services.aspx.

[8] John B. Horrigan, "Broadband Adoption and Use in America," Omnibus Broadband Initiative (OBI) Working Paper Series No. 1, Federal Communications Commission, February 2010, http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-296442A1.pdf.

[9] See, for example, Henry Jenkins, *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*, MacArthur Foundation White Paper (Chicago: John D. and Catherine T. MacArthur Foundation, 2006).

[10] Pew Internet & American Life Project surveys, 2000–2009. These data represent the average total participation in each activity over each two-year period. Therefore, they may not exactly reflect the combined yearly tracking numbers. In addition, the wording of some survey questions may have changed slightly over time.

[11] Rahul Tongia and Ernest Wilson, "Turning Metcalfe on His Head: The Multiple Costs of Network Exclusion," paper presented at the 35th Telecommunications Policy Research Conference, Vienna, Virginia, September 2007, http://web.si.umich.edu/tprc/papers/2007/772/TPRC-07-Exclusion-Tongia&Wilson.pdf.

[12] See Brian David, "The Cost of Digital Exclusion," The Official Blog of the National Broadband Plan, March 9, 2010, http://blog.broadband.gov/?entryId=236662.

[13] Digital Impact Group and Econsult Corporation, "The Economic Impact of Digital Exclusion," March 2010, http://www.digitalimpactgroup.org/costofexclusion.pdf.

[14] George S. Ford, "Internet Use and Job Search: More Evidence," Phoenix Center Perspective No. 10-02 (Washington, D.C.: Phoenix Center for Advanced Legal & Economic Public Policy Studies, January 2010), http://www.phoenix-center.org/perspectives/Perspective10-01final.pdf.

[15] "State of the News Media 2010," a report of the Pew Project on Excellence in Journalism, http://www.stateofthemedia.org/2010/newspapers-summary-essay.

[16] Pew Research Center for the People & the Press, "Americans Spending More Time Following the News" (Washington, D.C.: Pew Research Center, September 2010), http://pewresearch.org/pubs/1725/where-people-get-news-print-online-readership-cable-news-viewers.

[17] Federal Communications Commission, *Connecting America: The National Broadband Plan* (Washington, D.C.: Federal Communications Commission, 2010), 286, 290, http://www.broadband.gov/download-plan.

[18] Susannah Fox, "The Social Life of Health Information," Pew Internet & American Life Project (Washington, D.C.: Pew Research Center, June 2009), http://pewinternet.org/Reports/2009/8-The-Social-Life-of-Health-Information.aspx.

[19] Federal Communications Commission, *Connecting America*, 202, box 10-3.

[20] The $55 billion cost of digital exclusion in the United States, derived by Econsult and the Digital Impact Group, often focuses on gross costs, not net costs. That is, the estimates concentrate on savings to using digital means for, say, health care information, without fully accounting for costs that may arise.

[21] Mark Cooper, "The Challenge of Digital Exclusion in America: A Review of the Social Science Literature and Its Implications for the National Broadband Plan" (Washington, D.C.: Consumer Federation of America, January 2010), 2, http://fjallfoss.fcc.gov/ecfs/document/view?id=7020384062.

22 Federal Communications Commission, "Telephone Subscribership in the United States: Data Through March 2010" (Washington, D.C.: Federal Communications Commission, August 2010), http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-301241A1.pdf.

23 Horrigan, "Broadband Adoption and Use in America," 32.

24 See ibid. for data reported in the following two paragraphs.

25 Ibid., 31.

26 National Telecommunications & Information Administration, "Exploring the Digital Nation: Home Broadband Internet Adoption in the United States" (Washington, D.C.: NTIA, November 2010), 17.

27 Charlene Li and Josh Bernoff, *Groundswell: Winning in a World Transformed by Social Technologies* (Boston: Harvard Business Press, 2008). See also John B. Horrigan, "The Mobile Difference," Pew Internet & American Life Project (Washington, D.C.: Pew Research Center, March 2009), http://pewinternet.org/Reports/2009/5-The-Mobile-Difference--Typology.aspx.

28 Horrigan, "The Mobile Difference."

29 Eszter Hargittai, "An Update on Survey Measures of Web-Oriented Digital Literacy," *Social Science Computer Review* 27 (1) (February 2009): 130–137. See also Horrigan, "Broadband Adoption and Use in America," 18.

30 Blair Levin, *Universal Broadband: Targeting Investments to Deliver Broadband Services to All Americans* (Washington, D.C.: Aspen Institute, 2010).

31 Renee Hobbs, *Digital and Media Literacy: A Plan of Action* (Washington, D.C.: Aspen Institute, 2010), xi.

32 Everett Rogers, *Diffusion of Innovation*, 4th ed. (New York: The Free Press, 1995).

33 Austan Goolsbee and Peter J. Klenow, "Evidence on Learning and Network Externalities in the Diffusion of Home Computers," *Journal of Law and Economics* 45 (October 2002): 317–343.

34 Eszter Hargittai, "Digital Na(t)ives? Variation in Internet Skills and Uses among Members of the 'Net Generation,'" *Sociological Inquiry* 80 (1) (2010): 92–113.

35 Horrigan, "Broadband Adoption and Use in America."

36 Levin, *Universal Broadband*, 24.

37 Janice A. Hauge and James E. Prieger, "Demand-Side Programs to Stimulate Adoption of Broadband: What Works?" *Review of Network Economics* 9 (3) (2010): article 4.

38 Knight Commission, *Informing Communities: Sustaining Democracy in the Digital Age*, April 2010, http://www.knightcomm.org/read-the-report-and-comment.

39 Federal Communications Commission, *Connecting America*, 177.

A Contextual Approach to Privacy Online

Helen Nissenbaum

Abstract: Recent media revelations have demonstrated the extent of third-party tracking and monitoring online, much of it spurred by data aggregation, profiling, and selective targeting. How to protect privacy online is a frequent question in public discourse and has reignited the interest of government actors. In the United States, notice-and-consent remains the fallback approach in online privacy policies, despite its weaknesses. This essay presents an alternative approach, rooted in the theory of contextual integrity. Proposals to improve and fortify notice-and-consent, such as clearer privacy policies and fairer information practices, will not overcome a fundamental flaw in the model, namely, its assumption that individuals can understand all facts relevant to true choice at the moment of pair-wise contracting between individuals and data gatherers. Instead, we must articulate a backdrop of context-specific substantive norms that constrain what information websites can collect, with whom they can share it, and under what conditions it can be shared. In developing this approach, the paper warns that the current bias in conceiving of the Net as a predominantly commercial enterprise seriously limits the privacy agenda.

HELEN NISSENBAUM is Professor of Media, Culture, and Communication and Senior Fellow in the Information Law Institute at New York University. Her books include *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (2010), *Academy & the Internet* (edited with Monroe E. Price, 2004), and *Computers, Ethics & Social Values* (edited with Deborah G. Johnson, 1995).

The year 2010 was big for online privacy.[1] Reports of privacy gaffes, such as those associated with Google Buzz and Facebook's fickle privacy policies, graced front pages of prominent news media. In its series "On What They Know," *The Wall Street Journal* aimed a spotlight at the rampant tracking of individuals for behavioral advertising and other reasons.[2] The U.S. government, via the Federal Trade Commission (FTC)[3] and the Department of Commerce,[4] released two reports in December 2010 depicting the Net as a place where every step is watched and every click recorded by data-hungry private and governmental entities, and where every response is coveted by attention-seekers and influence-peddlers.[5]

This article explores present-day concerns about online privacy, but in order to understand and explain on-the-ground activities and the anxieties they stir, it identifies the principles, forces, and values behind them. It considers why privacy online has been vexing, even beyond general concerns over privacy; why predominant approaches have persisted de-

spite their limited results; and why they should be challenged. Finally, the essay lays out an alternative approach to addressing the problem of privacy online based on the theory of privacy as contextual integrity. This approach takes into consideration the formative ideals of the Internet as a public good.[6]

Setting aside economic and institutional factors, challenges to privacy associated with the Net are similar to those raised in the past by other information systems and digital media due to their vast capacities for capturing, stockpiling, retrieving, analyzing, distributing, displaying, and disseminating information. In a flourishing online ecology, where individuals, communities, institutions, and corporations generate content, experiences, interactions, and services, the supreme currency is information, including information about people. As adoption of the Internet and Web has surged and as they have become the primary sources of information and media for transaction, interaction, and communication, particularly among well-off people in technologically advanced societies, we have witnessed radical perturbations in flows of personal information. Amid growing curiosity and concern over these flows, policy-makers, public-interest advocates, and the media have responded with exposés and critiques of pervasive surreptitious tracking, manipulative behavioral advertising, and fickle privacy commitments of major corporate actors.

In *Privacy in Context: Technology, Policy, and the Integrity of Social Life*,[7] I give an account of privacy in terms of expected flows of personal information, modeled with the construct of *context-relative informational norms*. The key parameters of informational norms are actors (subject, sender, recipient), attributes (types of information), and transmission principles (constraints under which information flows). Generally, when the flow of information adheres to entrenched norms, all is well; violations of these norms, however, often result in protest and complaint. In a health care context, for example, patients expect their physicians to keep personal medical information confidential, yet they accept that it might be shared with specialists as needed. Patients' expectations would be breached and they would likely be shocked and dismayed if they learned that their physicians had sold the information to a marketing company. In this event, we would say that informational norms for the health care context had been violated.

Information technologies and digital media have long been viewed as threatening to privacy because they have radically disrupted flows of personal information, from the corporate and governmental databases of the 1960s to the surveillance cameras and social networks of the present day. The Net, in particular, has mediated disruptions of an unprecedented scale and variety. Those who imagined online actions to be shrouded in secrecy have been disabused of that notion. As difficult as it has been to circumscribe a right to privacy in general, it is even more complex online because of shifting recipients, types of information, and constraints under which information flows. We have come to understand that even when we interact with known, familiar parties, third parties may be lurking on the sidelines, engaged in business partnerships with our known parties. Information about us that once may have languished in dusty file cabinets is now pinpointed in an instant through search queries by anyone anywhere. In these highly *informatized* (that is, information-rich) environments, new types of information infuse our every action and relationship.

We are puzzled by the new and different types of information generated online, some of it the by-products of our activi-

ties, including cookies, latencies, clicks, IP addresses, reified social graphs, and browsing histories. New and different principles govern the flow of information: information we share as a condition of receiving goods and services is sold to others; friends who would not violate confidences repost our photographs on their home pages; people around the world, with whom we share nonreciprocal relationships, can see our houses and cars; providers from whom we purchase Internet service sell access to our communications streams to advertisers. Default constraints on streams of information from us and about us seem to respond not to social, ethical, and political logic but to the logic of technical possibility: that is, whatever the Net allows. If photographs, likes and dislikes, or listings of friends pass through the servers of a Facebook application, there is no telling whether they will be relinquished; if an imperceptible (to the ordinary user, at least) JavaScript code, or "beacon," is placed by a website one visits and enables the capture of one's browser state, so be it; if Flash cookies can cleverly work around the deletion of HTTP cookies, no harm done.

The dominant approach to addressing these concerns and achieving privacy online is a combination of *transparency and choice*. Often called notice-and-consent, or informed consent, the gist of this approach is to inform website visitors and users of online goods and services of respective information-flow practices and to provide a choice either to engage or disengage. Two substantive considerations explain the appeal of this approach to stakeholders and regulators. One is the popular definition of a right to privacy as a right to control information about oneself. Transparency-and-choice appears to model control because it allows individuals to evaluate options deliberately and then decide freely whether to give or withhold consent. How well it actually models control is not a question I pursue here because whatever the answer, there remains a deeper problem in defining a right to privacy as a right to control information about oneself, as discussed at length in *Privacy in Context*.[8]

A second consideration is the compatibility of notice-and-consent with the paradigm of a competitive free market, which allows sellers and buyers to trade goods at prices the market determines. Ideally, buyers have access to the information necessary to make free and rational purchasing decisions. Because personal information may be conceived as part of the price of online exchange, all is deemed well if buyers are informed of a seller's practices collecting and using personal information and are allowed freely to decide if the price is right. The ideal market assumes free and rational agents who make decisions without interference from third parties, such as government regulators. Doing so not only demonstrates respect for key actors, but also allows the market to function efficiently, producing the greatest overall utility.

However, there is considerable agreement that transparency-and-choice has failed.[9] Privacy advocates, popular media, and individuals have become louder and more insistent in pointing out and protesting rampant practices of surreptitious as well as flagrant data gathering, dissemination, aggregation, analysis, and profiling; even industry incumbents and traditionally pro-business government regulators admit that existing regimes have not done enough to curb undesirable practices, such as the monitoring and tracking associated with behavioral advertising and predatory harvesting of information posted on social networking sites. Why exactly the existing transparency-and-choice, or notice-and-consent,

approach has failed – and what to do about it – remains hotly disputed.

For many critics, whom I call *critical adherents*, the fault lies with the ubiquitous regime of offering privacy to individuals on a "take it or leave it" basis. A range of thoughtful commentaries on the subject, including those in the FTC and Department of Commerce reports mentioned above, have drawn attention to weak instantiations of choice, while others have highlighted problems with notice.[10] Because to choose means to deliberate and decide freely, the near-universal practice of modeling choice as "opt out" can hardly be said to model the ideal consumer making purchasing decisions in the ideal competitive marketplace. A deeper ethical question is whether individuals indeed freely choose to transact – accept an offer, visit a website, make a purchase, participate in a social network – given how these choices are framed as well as what the costs are for choosing not to do so.[11] While it may seem that individuals freely choose to pay the informational price, the price of not engaging socially, commercially, and financially may in fact be exacting enough to call into question how freely these choices are made.

Privacy policies as enactments of notice fare no better. That almost all privacy policies are long, abstruse, and legalistic adds to the unrealistic burden of checking the respective policies of the websites we visit, the services we consider and use, and the content we absorb. Compounding the burden is an entity's right to change its policy at will, giving due notice of such change, ironically, within the policy itself and therefore requiring interested individuals to read it not once but repeatedly. Unsurprisingly, ample evidence reveals that people do not read privacy policies, do not understand them when they do,[12] and realistically could not read them even if they wanted to.[13] This is not merely a matter of weakness of the will.

For critical adherents to transparency-and-choice, these observations point to the need for change, but not revolution. Such critics have suggested correctives including better mechanisms for choice, such as reframing policies in terms of "opt in" rather than "opt out" and locating moments of choice at times when users might be able to pause and think. They also advocate increasing transparency: for example, stipulating shorter policies that are easier to follow, along the lines of nutritional labels. Suggestions also apply to the content of policies. Whereas in the past, online actors were entreated simply to have policies, current correctives would require adherence to fair information principles.[14] The details of these suggestions are beyond the scope of this essay, as are questions about how privacy policies and practices should be monitored and enforced. This is because (as I argue below) the consent model for respecting privacy online is plagued by deeper problems than the practical ones noted so far.

I am not convinced that notice-and-consent, however refined, will result in better privacy online as long as it remains a procedural mechanism divorced from the particularities of relevant online activity. Take the example of online behavioral advertising, which quickly reveals an inherent flaw with the notice-and-consent approach.[15] To begin, consider what might need to be conveyed to users to provide notice of what information is captured, where it is sent, and how it is used. The technical and institutional story is so complicated that probably only a handful of deep experts would be able to piece together a full account; I would hazard that most of the website owners who contract with ad networks providing targeted advertising services are not among such experts. Even if,

for a given moment, a snapshot of the information flows could be grasped, the realm is in constant flux, with new firms entering the picture, new analytics, and new back-end contracts forged: in other words, we are dealing with a recursive capacity that is indefinitely extensible.[16] As a result of this complex and shifting landscape, users have been prone to conflate (in a convenient but misleading way) tracking with targeting. Further, the complexity makes it not only difficult to convey what practices are followed and what constraints respected, but practically impossible.[17]

For critical adherents to notice-and-consent, these types of cases exemplify the need for brief and clear policies that capture the essence of privacy practices in ways ordinary people can grasp. I view this as a futile effort because of what I call the *transparency paradox*. Achieving transparency means conveying information-handling practices in ways that are relevant and meaningful to the choices individuals must make. If notice (in the form of a privacy policy) finely details every flow, condition, qualification, and exception, we know that it is unlikely to be understood, let alone read. But summarizing practices in the style of, say, nutrition labels is no more helpful because it drains away important details, ones that are likely to make a difference: who are the business associates and what information is being shared with them; what are their commitments; what steps are taken to anonymize information; how will that information be processed and used. An abbreviated, plain-language policy would be quick and easy to read, but it is the hidden details that carry the significance.[18] Thus the transparency paradox: transparency of textual meaning and transparency of practice conflict in all but rare instances.[19] We seem unable to achieve one without giving up on the other, yet both are essential for notice-and-consent to work.

Adherents may persist, pointing to other arenas, such as health care and human subject research, in which a similar transparency paradox appears to have been overcome. In health care, informed consent protocols are commonly accepted for conveying risks and benefits to patients undergoing surgery, for example, or to subjects entering experimental treatment programs, even though it is unlikely they fully grasp the details. In my view, these protocols work not because they have found the right formulation of notice and the authentic mechanism for consent but because they exist within a framework of supporting assurances. Most of us are terrible at assessing probabilities and understanding risks of side effects and failed procedures; we are extremely poor at visualizing the internal organs of our bodies. It is not the consent form itself that draws our signature and consigns us to the operating table, but rather our faith in the system.[20] We trust the long years of study and apprenticeship that physicians undergo, the state and board certifications, peer oversight, professional codes, and above all, the system's interest (whatever the source) in our well-being. We believe in the benevolence of institutions of higher learning and, in large part, their mission to promote human welfare. Far from perfect, and subject to high-visibility breaches, the systems that constitute these safety nets have evolved over centuries; they undergird and warrant the consent agreements that patients and subjects confront every day. In the online environment, by contrast, individual consent agreements must carry the entire weight of expectation.

Picking holes in the transparency-and-choice (informed consent) approach, problematic as it is, is not the end point of

my argument. As it is, it may be the best approach for this interim period while the supporting assurances to shore it up are developed. Such assurances are not achieved by fiat, but may require decades for relevant institutional forms and practices to progress from trial and error to a balanced settling point. The theory of contextual integrity offers a shorter and more systematic path to this point by invoking learned wisdom from mature systems of informational norms that have evolved to accommodate diverse legitimate interests as well as general moral and political principles and context-specific purposes and values. The promise of this path is not merely that the equilibriums achieved in familiar contexts may provide analogical guidance for online realms; rather, the path acknowledges how online realms are inextricably linked with existing structures of social life. Online activity is *deeply integrated* into social life in general and is *radically heterogeneous* in ways that reflect the heterogeneity of offline experience.

By now, the story is familiar: about the advent of ARPANET and, out of this, the Internet, email as the unanticipated "killer app," the handoff of management from government to private industry, and emergence of the Web as the dominating platform for most ordinary people's experience of the Net. Along the way, the Internet has progressed from an esoteric utility for sharing computer resources and data sets, intended for use by relatively few specialists, to a ubiquitous, multifunctional medium used by millions worldwide.[21] As it has progressed through these stages, it has been conceptualized through a series of influential ideations:[22] from information superhighway,[23] enabling swift flows of information and commerce;[24] to cyberspace, a new frontier immune from the laws of any land; to Web 2.0, a meeting place overflowing with services and content, much of it generated by users themselves.[25]

Each of these ideations captured salient aspects of the vast socio-technical system that I have been calling "the Net" as it developed through progressive phases, and as it continues to do so today. Indeed, the Net is characterized by enormous malleability, both over time and across applications. Although the brute technical substrate of digital media – architecture, design, protocol, feature sets – may constrain or afford certain activities, it does so no more than, say, gravitational force, which similarly constrains and affords human activity while leaving plenty of room for variation. For example, the Net may have seemed *essentially* ungovernable until China asserted control and territorial borders quietly reemerged. Yet even that maneuver is incomplete, leaving intact exhilarating pockets of autonomy.[26]

A snapshot of today's Net, conceived as an abstraction of technical layers and social (economic and political) systems, operates as infrastructure, bustling spaces, and medium. Whether "online," "in cyberspace," "on the Internet," or "on the Web," individuals engage in banal practices such as banking, booking travel, and shopping, in many instances doing so with the same institutions and companies they could call on the telephone or visit at a physical location. Other activities – viewing movies, listening to recordings, reading literature, talking on an IP phone, seeking information, communicating via email, worshipping, and some forms of shopping – are transformed in their migration to the Net. In many instances, these transformations are not merely experiential but reflect institutional innovations, such as online churches and dating services, virtual universities, websites such as Amazon, Netflix, Mayo Clinic, WebMD, eBay, and E*TRADE, and

programs and services such as e-government, e-zines, e-vites, e-readers, iTunes, and iShares.

Even greater novelty and more fundamental transformations are found in the activities, practices, and institutional and business forms built on top of these offerings, including meta-engines that aggregate, index, organize, and locate sites, services, goods, news, and information; examples include kayak.com, Google search, Google news, and Yelp.com. Web 2.0 has wrought an additional layer of changes, notably in production, creativity, and social life. These changes include interacting via social networks, networking on platforms, and facilitating peer-production and user-generated content by way of innumerable individual and small-group blogs, wikis, and personal websites; repositories of global scale such as Wikipedia, IMDb, Flickr, MMORGs (massively multiplayer online role-playing games), and YouTube; the online patient-support community PatientsLikeMe; as well as mash-ups, folksonomies, crowdsourcing, and reputational systems.

Questions about protecting privacy online, particularly when framed as questions about *online privacy*, suggest that "online" is a distinctive venue, sphere, place, or space defined by the technological infrastructures and protocols of the Net, for which a single set of privacy rules can, or ought to, be crafted. I resist this notion. However exhilarating the vision of cyberspace as a new frontier, experience reveals no insulated domain divorced from "real life" and deserving distinctive regulation. The Net does not constitute (drawing on the terminology of contextual integrity) a discrete context. It is not a single social realm, but the totality of experience conducted via the Net, from specific websites to search engines to platforms and on up into "the cloud," crisscrossing multiple realms. Activities online, mediated by the Net ("on" the Web), are deeply integrated into social life: they may be continuous with brick-and-mortar correlates or, at the very least, have the power to affect communications, transactions, interactions, and activities in those realms (and vice versa). Not only is life online integrated into social life, and hence not productively conceived as a discrete context, it is *radically heterogeneous*, comprising multiple social contexts, not just one, and certainly is not just a commercial context where protecting privacy amounts to protecting *consumer* privacy and commercial information.[27] To be sure, the contours of technology (architecture, protocol, design, and so on) shape what you can do, say, see, and hear online, but while alterations, or disruptions due to particular characteristics of the Net, impose puzzles and pose challenges for social contexts, they do not warrant sui generis, uniform, cross-cutting rules determined by the medium. Instead, the contexts in which activities are grounded shape expectations that, when unmet, cause anxiety, fright, and resistance.[28]

Answering questions about privacy online, like those about privacy in general, requires us to prescribe suitable, or appropriate, constraints on the flow of personal information. The challenge of privacy online is not that the venue is distinct and different, or that privacy requirements are distinct and different, but that mediation by the Net leads to disruptions in the capture, analysis, and dissemination of information as we act, interact, and transact online. The decision heuristic derived from the theory of contextual integrity suggests that we locate contexts, explicate entrenched informational norms, identify disruptive flows, and evaluate these flows against norms based on general ethical and political principles as well as context-specific purposes and values.

To be sure, locating *contexts online* and explicating the *presiding norms* is not always straightforward (in the same way that it is not when dealing with unmediated social spaces). Some of the more familiar cases, however, may provide insight into the task. Whether you transact with your bank online, on the phone, or person-to-person in a branch office, it is not unreasonable to expect that rules governing information will not vary according to medium. In the United States, banks and other financial institutions are governed by privacy rules formulated by the FTC, which was given this authority by the Gramm-Leach-Bliley Act.[29] Auxiliary information (for example IP address or clickstream), the artifacts of online transaction, should not simply be deemed "up for grabs" just because that information was not explicitly considered in rules formulated before online banking became common. Instead, it should be held to the same standards that guided financial privacy in the first place.

Similarly, while expectations of visitors to Bloomingdales.com, NYTimes.com, and MOMA.org may be affected by corresponding, preexisting brands, they are also shaped by the respective social contexts that these entities inhabit, including the types of experiences and offerings they promise. Accordingly, Amazon.com, which came on the scene as an online bookstore with no brick-and-mortar precursor, is nevertheless recognizable, akin to, say, the Moravian Book Shop in Bethlehem, Pennsylvania, which was founded in 1745 and is believed to be the oldest continually operating bookstore in the United States.[30] As Amazon.com expanded into other arenas, selling and renting DVDs, for example, one would assume personal information flows generated in these transactions to be regulated by constraints expressed in the Video Privacy Protection Act of 1988[31] in the same way that West Coast Video must adhere to the Act. Whether laws applicable to brick-and-mortar video rental stores actually apply to online video rental providers such as iTunes and Amazon seems uncertain; still, the requirements of contextual integrity, which anchors privacy rules in social contexts and social roles, would imply that they should.

These examples merely scratch the surface of the Net's remarkable heterogeneity. Online offerings range from specialized information providers and distributors, such as MayoClinic.com and WebMD; federal, state, and local government portals, providing services and information directly to citizens; and structured repositories of user-generated content, such as Wikipedia, YouTube, Flickr, Craigslist, and social networks, including Facebook. Religious denominations around the globe have online presences, ranging from The Holy See, claiming to be the "official" Vatican website,[32] to online churches,[33] offering in-home, Web-based religious engagement that replaces or supplements regular church attendance. This list does not capture the fluidity and modularity of existing offerings, which include combinations and permutations (mash-ups) constrained only by human creativity and the technological limits of the moment. Many popular websites, for example, combine modules of enterprise-generated content with user-generated feedback, or storefronts with varieties of social networks, political content with open blogs, and more.

To the extent that the Net is deeply embedded in social life, context-specific informational norms may be extended to corresponding online activities. Thus, privacy rules governing financial institutions, for example, would extend to E*TRADE even though it operates primarily via an online portal. Online offerings and experiences may defy existing norms, how-

ever, as they incorporate some of the novel forms mentioned above. In such circumstances, the theory of contextual integrity directs us beyond existing norms to underlying standards, derived from general moral and political considerations as well as the ends, purposes, and values of respective contexts.[34] Novel activities and practices, which implicate different types of information, expanded groups of recipients, and altered constraints on flow are evaluated against these standards.

Applying this reasoning to online filing of income tax returns is fairly straightforward. In the United States, rigorous confidentiality requirements governing individual tax records, impervious even to certain types of law enforcement needs, have developed over the past 150 years.[35] Although present-day code, formalized in the 1970s, may have little to say about e-filing specifically, we would not expect auxiliary information generated through online interactions to be "up for grabs," freely available to all comers. Even in the absence of explicit rules, guidance can be sought from the values and purposes that have yielded existing confidentiality rules for information in traditional paper-based tax records. In the Disclosure and Privacy Law Reference Guide, the IRS asserts that "there must be public confidence with respect to the confidentiality of personal and financial information given to us for tax administration purposes. . . . The confidential nature of these records requires that each request for information be evaluated in light of a considerable body of law and regulation which either authorize or prohibit disclosure."[36] This connection was acknowledged as far back as 1925, when Secretary of the Treasury Andrew Mellon remarked, "While the government does not know every source of income of a taxpayer and must rely upon the good faith of those reporting income, still in the great majority of cases this reliance is entirely justifiable, principally because the taxpayer knows that in making a truthful disclosure of the sources of his income, information stops with the government. It is like confiding in one's lawyer."[37] A presumption of strict confidentiality is derived from values and purposes – public compliance, trust, confidence in government – that prohibit all sharing except as allowed, on a case-by-case basis, by explicit law and regulation.

The more challenging cases confronting us include forms of content, service, and interaction that are specific to the Net or that do not have obvious counterparts elsewhere. Search engines such as Google and Bing, essential for navigating the Web, may constitute an important class of cases. And while sites such as Mayo Clinic and WebMD might seem similar enough to familiar reference resources, health information sites are prodigious,[38] offering everything from personalized and interactive services that allow users to sign in and pose questions about their particular problems, to personal health record repositories that provide space to store health records (for example, Microsoft Health Vault), and to social networking sites devoted to communities of fellow sufferers (for example, PatientsLikeMe).

Without denying that the Net has yielded much that is novel and strange, including new types of information and new institutional forms, online activities themselves are strangely familiar: connecting with a friend, collaborating on a political mission, applying for a job, seeking religious or spiritual sustenance, pursuing educational opportunity, catching up on local and world news, or choosing a book to read, music to enjoy, or movies to watch. Although searching on Google is different from looking up material in a library catalog, in part because the contents of the

Web are quite different from the contents of a library, there is similarity in these two activities: both may include the pursuit of research, knowledge, and intellectual enrichment. In all such activities, liberal democratic societies allow great freedom, unconstrained by the watchful gaze or approbation of authorities, just as they allow citizens to seek political or religious information or affiliation. Just as with the offline environment, we would expect the same standards to prevail online, dictating that online footprints should not be recorded and registered in order to minimize risk of interference, by either human or machine.

The interest in privacy online that the FTC and Commerce Department have recently shown is a positive development, particularly because it acknowledges a growing concern over privacy and amplifies public discussion of the wildly unrestrained collection of personal information by nongovernmental actors. Their interest has been limited, though, by a focus on protecting privacy online as, predominantly, a matter of protecting *consumers* online and protecting *commercial* information: that is, protecting personal information in commercial online transactions.[39] Neither agency has explicitly acknowledged the vast landscape of activity lying outside the commercial domain. As long as public discourse about privacy online takes the marketplace and commerce as proxies for the whole, conceptions of privacy will be inadequate. We need to take full account of the radical heterogeneity of online activity and practice.

One might argue that the Net *is* almost completely commercial, pointing to the prevalence of private payment as the means supporting online activity. Aside from government presences, the Net is almost wholly owned by private, for-profit entities, from the underlying physical infrastructure to network service providers, providers of utilities and applications, and retailers of goods, content, and services. Furthermore, commercial advertising managed through the complex ecosystem of ad networks, ad exchanges, and analytics and marketing companies has emerged as a dominant business model for supporting online content and services. This model prevails in a variety of online locations, from large corporate websites, to personal blogs such as Noob Cook, a site with seventeen trackers, or Dictionary.com, with nine trackers from advertising companies, such as Doubleclick, Media Math, Microsoft Atlas, and others.[40] Wikipedia remains a rather remarkable standout, supported by the nonprofit Wikimedia Foundation and sporting no trackers.[41]

By this logic, the Commerce Department and the FTC would be precisely the governing bodies to have oversight of online activity, and norms of the competitive, free marketplace would make the most sense for regulating it. Yet private payment, whether through direct charges for goods, services, access, or participation, or through income from advertising, does not on its own signal complete surrender to marketplace norms. According to political philosopher Elizabeth Anderson, many functions in society straddle boundaries between the commercial and noncommercial. How they are supported is not decisive but rather how they measure up to standards of quality or excellence. Private payment as a form of support does not require total concession to marketplace norms; instead, we expect functions such as education, health care, religion, telecommunication, and transportation, whether privately paid for or not, to meet independent ideals. As Anderson warns, "When professionals sell their services,

they enter into market relations that impose norms on their activities which potentially conflict with the norms of excellence internal to their professional roles."[42] But we expect more from professionals – from doctors, lawyers, athletes, artists, church ministers, and teachers – than the pursuit of profit. People pay for medical care at private practices and hospitals, for instance, and for education at a variety of institutions. In these and other cases, in which complete surrender to marketplace norms would result in corrupt and impoverished practice, Anderson advocates a proper balance of market norms with internal standards of excellence.

This point might seem obvious, but certain brands of free-market capitalism make it easy to confuse the quest for profit with the pursuit of internal standards of excellence.[43] When Sergey Brin and Larry Page first launched the Google search engine, they regarded commercial influences as contrary to a search engine's core mission as a performance-driven tool serving individuals' interests in locating information on the Web. Eschewing advertising, they wrote in the appendix of a 2007 paper, "The goals of the advertising business model do not always correspond to providing quality search to users.... We believe the issue of advertising causes enough mixed incentives that it is crucial to have a competitive search engine that is transparent and in the academic realm."[44] In other, less visible cases, similar concerns may be raised: for example, Amazon's purchase of IMDb, a website of information about movies, developed and initially maintained by volunteers. Even in the case of the familiar lending library, originally conceived in the United States by Benjamin Franklin as publicly funded, many functions have been taken over by private, for-profit companies online.[45]

The point is not to see Brin and Page or other developers as "sellouts." Confounding sources of support with guiding norms obscures our recognition of the internal standards of excellence that we can hold search companies to even as they seek commercial support, independent of their performance in the marketplace.[46] The same argument holds for content vendors and information services providing, in the private sector, many services also provided by public libraries. I am not suggesting that there is consensus, or that questions about internal norms of excellence are easily settled; endless struggles over what constitutes a good newspaper, school, or health care system attest to this. But they also reveal a strong belief that beyond profit, such standards are at play and are socially important.

Recent attention given to the challenge of protecting privacy online is a positive development. Although success is hamstrung by the foot-dragging of those whose power and profit are served by unrestricted flows of personal information, it is also limited by underdeveloped conceptions of privacy and the role it plays in sustaining the Net as a public good, capable of serving diverse interests. Early portrayals of cyberspace as a new frontier, different, distinct, and out of the reach of traditional law, have for the most part been abandoned, yet no other single vision has captured public imagination in quite the same way. This is unsurprising given the Net's massive growth as a complex infrastructure, content delivery system, and media space. The lack of a reigning public vision has meant that controversial moral and political matters are settled more often by technical affordances than by clearly articulated public moral and political principles. For privacy this has been devastating, as the Net, constructed

through an amalgamation of the sciences and technologies of information, computation, and networking, affords radical disruptions of information flows. With economic and social incentives stacked against constraints on flow, burdening individuals with the full weight of protecting their privacy online through notice-and-choice is unlikely to yield success.

My preferred alternative builds from the vision of life online as heterogeneous and thickly integrated with social life. Despite distinctive qualities of online movement, transactions, activities, relations, and relationships, when abstracted from particulars these retain fidelity with the fundamental organizing principles of human practice and social life. Drawing on work from social theory and philosophy, the framework of contextual integrity conceives of these spheres as partially constituted by norms of behavior, among them norms governing the flow (sharing, distributing) of personal information. We should not expect social norms, including informational norms, simply to melt away with the change of medium to digital electronic any more than from sound waves to light particles. Although the medium may affect what actions and practices are possible and likely, sensible policy-making focuses on the actions and practices themselves, with an eye to their function within social spheres and their standing in relation to entrenched social norms.

Two broad recommendations follow from the argument thus far. First, in our online activity we should look for the contours of familiar social activities and structures. For much that we do online – banking, shopping, communicating, and enjoying culture and entertainment – this is not a difficult task. Where correspondences are less obvious, such as consulting a search engine to locate material online, we should consider close analogues based not so much on similarity of action but on similarity of function or purpose. Consulting a search engine, in this regard, is akin to conducting research, seeking information and association, searching a library catalog, and pursuing intellectual enlightenment. Time spent on social networks, such as Facebook, is an amalgam of engagement with personal, social, intimate and home life, political association, and professional or work life. As we locate correspondences, we bring into view the relevant governing norms. If I am right about how search engines are used and for what purposes, then the governing norms would be strict confidentiality with regard to Web search histories and perhaps, as practiced by many public libraries, the prompt expunction of such records to minimize risks of leakage or mandated handovers as well as the temptation of future sharing for financial gain. At present, Google, unlike other search providers, has expressed a commitment to maintaining a barrier between identifiable search records and other records it accumulates with user profiles. Although this decision adheres to the spirit of the conception of Web search I have urged, questions remain about the efficacy of their approaches to de-identifying search logs and the fact that the commitment can be revoked at any moment, as was Google's commitment to forgo behavioral advertising.[47]

This view of online privacy also implies that contexts, not political economy, should determine constraints on the flow of information. Companies merge and acquire other companies for many different reasons: for example, to strengthen and expand their range of holdings, to gain market dominance in a particular area, or to establish control over vertical chains of necessary resources. Among the valu-

able assets that motivate acquisitions are databases of personal information, as demonstrated (presumably) by Google's acquisition of Doubleclick and Choicepoint's systematic acquisitions of smaller, special-purpose data holders.[48] But databases of personal information shared in one context, under constraints of the relevant informational norms, should not be treated as just another asset, akin to buildings, furniture, and supplies. The privacy policies of large diverse companies, such as Walt Disney, General Electric, Google, Citigroup, Viacom, and Microsoft, however, reveal porous boundaries among subsidiaries, with little acknowledgment of a need to account for patterns of information flow within a single company. Online conglomerates are no different as they strive to achieve vertical integration by controlling the raw materials of their industry, namely, information. Against these trends, we must establish respect for the boundaries of context and associated informational norms.

There is little doubt that when communicating with the public, corporations understand the importance of acknowledging the integrity of contexts. Even though to corporate investors a company might boast diverse informational assets, to the public it generally identifies units that are socially meaningful. It may be that in these acts of self-presentation, companies acknowledge the contextual heterogeneity of their offerings and therefore open the door to corresponding context-specific norms. By calling its online offering a university, a shoe store, a church, a medical center, a friendship network, or a bank, a company gives users a way to understand the services or activities that take place there, and it invites evaluation against respective norms, whether these are embodied in law or simply arise from reasonable expectations.[49] Staying true to these self-portrayals requires companies to commit to partitioning information holdings along contextual contours rather than along lines of corporate ownership.

There is no denying the transformative effects of digital technologies, including the rich and teeming online activity they have spawned. Recommending that we locate familiar social contexts online and, where it makes sense, connect activities and offerings with them is not to dispute this. Instead, the aim is to reveal relevant standards of excellence. As social contexts, activities, roles, and rules migrate online, respective context-specific values, ends, and purposes serve as standards against which information-sharing practices can be evaluated as legitimate or problematic. It is important to keep in mind that privacy norms do not merely protect individuals; they play a crucial role in sustaining social institutions.[50] Accordingly, restraints on search engines or social networks are as much about sustaining important social values of creativity, intellectual growth, and lively social and political engagement as about protecting individuals against harm. Benjamin Franklin knew as much when he insisted on privacy protection for the U.S. mail, not only to protect individuals but also to promote a meaningful social role for the service. We should expect no less for email and IP telephony.

My second recommendation applies to online cases without straightforward social precedents. As discussed earlier, social forms online sometimes enable configurations of actors, information, activities, and experiences that are unfamiliar, at least prima facie. In these cases, I suggest starting with ends, purposes, and values and working from there back to norms. A politician's website that allows citizens to "talk back" comprises an unusual platform for which no preexisting rules apply to, say, digital footprints left

by visitors to the site. Here, the right approach is not an opportunistic information grab. Although this may serve immediate needs of an imminent political campaign, it does not serve the purposes of encouraging frank political discussion, which is understood to flourish in environments of great freedom.[51] If people expect to be monitored, if they anticipate that their recorded views will be shared with particular third parties for money or favors, they are likely to be more watchful, circumspect, or uncooperative. The issue, however, is not how particular practices affect individuals, but the implications for particular purposes and values. Circumspection and cooperativeness are productive of certain ends but not others. Working backward from these values, we develop rules for situations in which there appear not to be any obvious candidates.

Growing momentum to confront the problem of privacy online is a welcome development. It would be a mistake, however, to seek remedies that make privacy online something distinct. Protecting privacy is a matter of assuring appropriate flows of personal information, whether online or offline, and disruptions in information flow, enabled by information technologies and digital media, can be equally disturbing, whether online or off. Because much of what happens online is thickly integrated with social life writ large (and vice versa), solving the privacy problem online requires a fully integrated approach. I have articulated one step toward this goal, resisting the suggestion that, with regard to privacy, the Net is virgin territory where it falls to the parties to construct terms of engagement for each transaction. Given how deeply rooted are our expectations of right and wrong concerning the sharing of information about ourselves and others, it is no wonder that over time intricate systems of norms have developed to govern all domains of social life.

To adapt these systems to social relations and contexts that have expanded into digital media spaces, we must make explicit much that has operated implicitly, and in the process reject entrenched norms that no longer promote the achievement of moral and political values as well as context-specific ends. To leave the protection of privacy online to negotiations of notice-and-consent is not only unfair, it is to pass up a critical public policy opportunity that will have ramifications for the shape and future of the Net. If pursued conscientiously, the process of articulating context-based rules and expectations and embedding some of them in law and other specialized codes will yield the safety nets that buttress consent in fields such as health care and research. With these precautions in place, plenty of room would still remain to express personal preferences and to maintain a robust role for informed consent.

ENDNOTES

[1] This essay has benefited from opportunities to present at the Center for Law, Technology, and Society, University of Ottawa; the Center for the Study of Law and Society, University of California, Berkeley; and the Center for Internet and Society, Stanford University, where questions and comments led to significant improvements and refinements of the argument. I am grateful for valuable feedback from David Clark and NYU's Privacy Research Group, expert guidance from Cathy Dwyer and Foster Provost, and sterling research assistance from Jacob Gaboury and Marianna Tishchenko. This work was supported by grants AFSOR: ONR BAA 07-03 (MURI) and NSF CT-M: Privacy, Compliance & Information Risk, CNS-0831124.

[2] Jennifer Valentino-Devries, "What They Know About You," *The Wall Street Journal*, July 31, 2010, http://online.wsj.com/article/SB10001424052748703999304575399041849931612.html.

[3] U.S. Federal Trade Commission, Preliminary FTC Staff Report, "Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers," December 2010, http://www.ftc.gov/os/2010/12/101201privacyreport.pdf.

[4] U.S. Department of Commerce, "Commercial Data Privacy and Innovation in the Internet Economy: A Dynamic Policy Framework," December 2010, http://www.ntia.doc.gov/reports/2010/IPTF_Privacy_GreenPaper_12162010.pdf.

[5] Compare these depictions to earlier accounts of the Net as a new frontier of freedom and autonomy: for example, David R. Johnson and David G. Post, "Law and Borders – The Rise of Law in Cyberspace," *Stanford Law Review* 48 (1996): 1367; John Perry Barlow, "Electronic Frontier: Coming into the Country," *Communications of the ACM* 34 (3) (March 1991).

[6] In this essay, I draw most of my examples from the World Wide Web because almost all the controversial privacy concerns that have captured public attention have stemmed from Web-based activity and because the online experiences of ordinary people occur mostly on the Web. I will use the term *Net* when observations made about the Web seem pertinent to other Internet applications and services.

[7] Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford, Calif.: Stanford University Press, 2010).

[8] Ibid., chap. 5.

[9] Federal Trade Commission, "Protecting Consumer Privacy in an Era of Rapid Change" and Department of Commerce, "Commercial Data Privacy and Innovation in the Internet Economy."

[10] Fred Cate, "The Failure of Fair Information Practice Principles," in *Consumer Protection in the Age of the "Information Economy,"* ed. Jane K. Winn (London: Ashgate Publishing, 2006).

[11] Ian Kerr, "The Legal Relationship Between Online Service Providers and Users," *Canadian Business Law Journal* 35 (2001): 1 – 40.

[12] Joseph Turow, Lauren Feldman, and Kimberly Meltzer, *Open to Exploitation: American Shoppers Online and Offline* (Philadelphia: Annenberg Public Policy Center, University of Pennsylvania, June 1, 2005), http://www.annenbergpublicpolicycenter.org/Downloads/Information_And_Society/Turow_APPC_Report_WEB_FINAL.pdf.

[13] Lorrie Faith Cranor and Joel Reidenberg, "Can User Agents Accurately Represent Privacy Notices?" The 30th Research Conference on Communication, Information, and Internet Policy (TPRC2002), Alexandria, Virginia, September 28 – 30, 2002.

[14] U.S. Department of Health, Education, and Welfare, *Records, Computers, and the Rights of Citizens*, Report of the Secretary's Advisory Committee on Automated Personal Data Systems, July 1973, http://aspe.hhs.gov/datacncl/1973privacy/tocprefacemembers.htm.

[15] Solon Barocas and Helen Nissenbaum, "On Notice: The Trouble with Notice and Consent," *Proceedings of the Engaging Data Forum: The First International Forum on the Application and Management of Personal Electronic Information*, Cambridge, Massachusetts, October 12 –

13, 2009; Vincent Toubiana, Arvind Narayanan, Dan Boneh, Helen Nissenbaum, and Solon Barocas, "Adnostic: Privacy-Preserving Targeted Advertising," *Proceedings of the Network and Distributed System Symposium*, San Diego, California, February 28 – March 3, 2010.

16 Counting ad servers alone, a list current as of April 2011 shows 2,766 unique entries; see http://pgl.yoyo.org/adservers/formats.php (accessed April 13, 2011).

17 Barocas and Nissenbaum, "On Notice," and Toubiana, Narayanan, Boneh, Nissenbaum, and Barocas, "Adnostic."

18 Vincent Toubiana and Helen Nissenbaum, "An Analysis of Google Log Retention Policies," *The Journal of Privacy and Confidentiality* (forthcoming).

19 For example, personal information is shared with no one and destroyed after each session.

20 Deborah Franklin, "Uninformed Consent," *Scientific American*, March 2011, 24 – 25.

21 See Katie Hafner and Matthew Lyon, *Where Wizards Stay Up Late: The Origins of the Internet* (New York: Simon and Schuster, 1999); Tim Berners-Lee, *Weaving the Web: The Past, Present and Future of the World Wide Web by Its Inventor* (London: Texere Publishing, 2000); and Janet Abbate, *Inventing the Internet* (Cambridge, Mass.: MIT Press, 2000).

22 On this transformation of the Internet through the "tussle" of interested parties, see David Clark, John Wroclawski, Karen Sollins, and Robert Braden, "Tussle in Cyberspace: Defining Tomorrow's Internet," *Proceedings of the ACM SigComm 2002 Conference*, Pittsburgh, Pennsylvania, August 19 – 23, 2002, published in *Computer Communications Review* 32 (4) (October 2002).

23 Al Gore, "Infrastructure for the Global Village: Computers, Networks and Public Policy," special issue, "Communications, Computers, and Networks," *Scientific American*, September 1991, 150 – 153.

24 John Perry Barlow, "The Economy of Ideas," *Wired*, March 1994, 84.

25 Clay Shirky, *Here Comes Everybody: The Power of Organizing Without Organizations* (New York: Penguin, 2009).

26 Samantha Shapiro, "Revolution, Facebook-Style," *The New York Times*, January 22, 2009, http://www.nytimes.com/2009/01/25/magazine/25bloggers-t.html.

27 Federal Trade Commission, "Protecting Consumer Privacy in an Era of Rapid Change," and Department of Commerce, "Commercial Data Privacy and Innovation in the Internet Economy."

28 Compare this notion to Mark Zuckerberg's claim that norms change due to the contours of Facebook's privacy policies; see Bianca Bosker, "Facebook's Zuckerberg Says Privacy No Longer a 'Social Norm,'" The Huffington Post, January 11, 2010, http://www.huffingtonpost.com/2010/01/11/facebooks-zuckerberg-the_n_417969.html.

29 U.S. Federal Trade Commission, Gramm-Leach-Bliley Act 15 U.S.C., Subchapter I, sec. 6801 – 6809, November 12, 1999, http://www.ftc.gov/privacy/glbact/glbsub1.htm; Adam Barth, Anupam Datta, John Mitchell, and Helen Nissenbaum, "Privacy and Contextual Integrity: Framework and Applications," *Proceedings of the IEEE Symposium on Security and Privacy*, Berkeley, California, May 21 – 24, 2006.

30 The Moravian Book Shop now has its own online portal, http://www.moravianbookshop.com/ (accessed April 13, 2011).

31 The Video Privacy Protection Act 18 U.S.C. sec. 2710, 1988, http://www.law.cornell.edu/uscode/html/uscode18/usc_sec_18_00002710----000-.html.

32 See http://www.vatican.va/phome_en.htm (accessed April 13, 2011).

33 See http://www.online-churches.net (accessed April 13, 2011).

34 Nissenbaum, *Privacy in Context*, esp. chap. 8.

[35] David Kocieniewski, "IRS Sits on Data Pointing to Missing Children," *The New York Times*, November 12, 2010, http://www.nytimes.com/2010/11/13/business/13missing.html.

[36] Internal Revenue Service, "Disclosure and Privacy Law Reference Guide," Publication 4639 (10-2007) Catalog Number 50891P, 1–7.

[37] Hearings on Revenue Revision 1925 Before the House Ways and Means Committee, 69th Cong., 1st sess. 8–9 (1925).

[38] A Google search on "HIV status," performed January 11, 2011, yielded more than 7.5 million results.

[39] Federal Trade Commission, "Protecting Consumer Privacy in an Era of Rapid Change," and Department of Commerce, "Commercial Data Privacy and Innovation in the Internet Economy."

[40] Noob Cook is presented as being written by "a girl who likes to cook"; see http://www.noobcook.com/about/ (accessed February 25, 2011).

[41] Establishing the number of trackers on a website is highly inexact. Various utilities offer this service, for example, Ghostery; numbers vary depending on the approaches they adopt. Not all approaches recognize all types of trackers. Further, these numbers also vary from time to time because websites may frequently revise their underlying policies and business arrangements. (I am indebted to Ashkan Soltani for clarifying this point.)

[42] Elizabeth Anderson, *Value in Ethics and Economics* (Cambridge, Mass.: Harvard University Press, 1995), 147.

[43] Milton Friedman, "Can Business Afford to be Ethical?: Profits before Ethics," in *Ethics for Modern Life*, ed. Raziel Abelson and Marie-Louise Friquegnon, 4th ed. (New York: St. Martin's Press, 1991), 313–318.

[44] Sergey Brin and Lawrence Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine," *WWW7/Computer Networks* 30 (1–7) (1998): 107–117; quotation taken from Web version, http://infolab.stanford.edu/~backrub/google.html (accessed February 26, 2011). See also Alex Diaz, "Through the Google Goggles: Sociopolitical Bias in Search Engine Design," in *Web Searching: Interdisciplinary Perspectives*, ed. Amanda Spink and Michael Zimmer (Dordrecht, The Netherlands: Springer, 2008).

[45] Robert Ellis Smith, "Ben Franklin's Web Site: Privacy and Curiosity from Plymouth Rock to the Internet," *Privacy Journal* (2000): 34, 51.

[46] Lucas Introna and Helen Nissenbaum, "Shaping the Web: Why the Politics of Search Engines Matters," *The Information Society* 16 (3) (2000): 1–17; Frank Pasquale and Oren Bracha, "Federal Search Commission? Access, Fairness, and Accountability in the Law of Search," *Cornell Law Review* 93 (2008): 1149; Toubiana, Narayanan, Boneh, Nissenbaum, and Barocas, "Adnostic."

[47] Toubiana and Nissenbaum, "An Analysis of Google Log Retention Policies."

[48] See "Choicepoint," EPIC-Electronic Privacy Information Center, http://epic.org/privacy/choicepoint/ (accessed April 13, 2011).

[49] As a practical matter, standards for communicating contexts will be needed through interface design. See, for example, the work of Ryan Calo and Lorrie Cranor.

[50] Nissenbaum, *Privacy in Context*, esp. chap. 8–9.

[51] Danielle Citron, "Fulfilling Government 2.0's Promise with Robust Privacy Protections," *George Washington Law Review* 78 (4) (June 2010): 822–845.

Online Trust, Trustworthiness, or Assurance?

Coye Cheshire

Abstract: Every day, individuals around the world retrieve, share, and exchange information on the Internet. We interact online to share personal information, find answers to questions, make financial transactions, play social games, and maintain professional and personal relationships. Sometimes our online interactions take place between two or more humans. In other cases, we rely on computers to manage information on our behalf. In each scenario, risk and uncertainty are essential for determining possible actions and outcomes. This essay highlights common deficiencies in our understanding of key concepts such as trust, trustworthiness, cooperation, and assurance in online environments. Empirical evidence from experimental work in computer-mediated environments underscores the promises and perils of overreliance on security and assurance structures as replacements for interpersonal trust. These conceptual distinctions are critical because the future shape of the Internet will depend on whether we build assurance structures to limit and control ambiguity or allow trust to emerge in the presence of risk and uncertainty.

COYE CHESHIRE is an Associate Professor in the School of Information at the University of California, Berkeley. He conducts experiments and field research on social exchange, trust, collective behavior and interpersonal relationships in computer-mediated environments. His publications include *eTrust: Forming Relationships in the Online World* (edited with Karen S. Cook, Chris Snijders, and Vincent Buskens, 2009) and "Trust and Transitions in Modes of Social Exchange" (with Alexandra Gerbasi and Karen S. Cook), *Social Psychology Quarterly* (2010).

The Internet is an extraordinary tool for human interaction, connecting millions of people with various information technologies and systems across a global communication network.[1] Despite the advantages of the Internet for communication breadth and efficiency compared to offline interactions, there are legitimate concerns about what it means to trust others in a shared Internet commons. For example, the social cues we rely on to detect risk and uncertainty in the physical world are often unreliable when we do not know who is behind the digital curtain of anonymity. We interact online to share personal information, find answers to questions, make financial transactions, play social games, and maintain professional and personal relationships. Sometimes our online interactions take place between two or more humans using different tools and services, including social networking websites, chat clients, messaging sys-

tems, blogs, and forums. In other cases, individuals are not directly involved in information transfers; we allow computer systems to act as agents, managing and exchanging our information according to specific rules and heuristics.

In each type of interaction mentioned above, risks and uncertainties complicate our decisions. Should we post pictures from a recent party to a social networking site? Can we believe the highly negative evaluation of a local restaurant on a review site? Are our medical search histories safe with our favored search engine companies? In some cases, the stakes for imparting our trust may be relatively low, such as when we evaluate the validity of an email chain letter or choose a restaurant based on a few anonymous reviews. However, when real financial and material assets are involved and there is ambiguity about a given outcome, trust is imperative. Valuable assets constitute risk (or what is at stake) in an interaction; the numerous sources of ambiguity about an outcome create uncertainty.

The combination of risk and uncertainty is critical for understanding trust in any situation, whether online or offline. Some situations might carry high risks, such as financial assets or personal reputation, but little or no uncertainty about the outcome. In other circumstances, nothing of significant value is at stake, but the outcome is indeterminate. I agree with those who argue that trust is relevant only when risk and uncertainty exist together.[2] In this view, trust is not simply the *absence* of risk and uncertainty. More accurately, trust is a complex human response to situations that are rife with risk and uncertainty.

Given the many risks and sources of uncertainty that exist in online interactions, it is unsurprising that trust remains a principal topic in social and technical studies of the Internet. In addition to the interpersonal communications that occur between humans on the Internet, the Internet infrastructure itself (computers, protocols, and connections) depends on reliable and secure relationships between systems. The Internet is a complex arrangement of layered trust relationships, including small networks among hardware systems, protocols that often presume reliable relationships between these networks, human administrators who configure and maintain the disparate networks, and, finally, the end users who operate personal computers and devices to interact with other humans and systems.

This essay examines key concepts such as *trust*, *trustworthiness*, *cooperation*, and *assurance* in online environments. The distinction between interpersonal trust (human to human) versus system trust (human to system) is particularly important as we begin to think about the future of trust on the Internet. To many social scientists, it makes little or no sense to treat systems as autonomous entities in the same way we regard humans. Likewise, some computer scientists who worry about the security and reliability of systems may see little in common between the uptime or security of an Internet server and the ambiguities of friendships and online social relationships. However, there is much to gain from considering different meanings of trust in the context of specific risks and various sources of uncertainty. The interests of those who endeavor to build a more reliable and secure Internet may not be so different from those who use the Internet for countless social purposes. In each case, the concept of trust can help define and describe the complexity of anticipating imminent outcomes and behaviors in the presence of uncertainty.

Some scholars argue that the term *trust* is accurate only when applied to interper-

sonal, human relationships. According to this view, interpersonal trust develops over time in a direct relationship when one party believes the other party has incentive to act in her interest or take her interests to heart.[3] Eventually, an enduring relationship can develop, consisting of either indirect or direct interactions. The continually fulfilled expectation of trustworthy behavior from another person over time defines a trust relationship.

Intuitively, we separate the deeper sense of trust that we evoke when we speak of family, friends, and other individuals with whom we share close ties from the term's more mundane usages (as in, for example, the statement "I trust the mail carrier to pick up on time"). The same issue extends to online environments. On popular online social networking services, such as Facebook, we might not trust every friend in the same way or in the same circumstances. One way to disentangle the complexity of interpersonal trust is to consider carefully what constitutes a trusting relationship in a specific context. Trust is a useful and meaningful term when we consider the possibility that not all friendships and acquaintances are, in fact, trust relations.

Political scientist Russell Hardin describes significant trust relationships as those in which each party encapsulates the other's interests.[4] For example, consider two workers who collaborate on a complex project. If the individuals trust one another, then they each believe the other is trustworthy enough to perform a certain type of task in a competent way. One person may entrust her work for modification or enhancement by the other and continue to do so in a reciprocal fashion over time. In Hardin's encapsulated-interest framework, the project is important and valuable to both individuals, and each understands the value to the other. This example also emphasizes the fact that trust is largely dependent on the specific situation of social interaction. We might trust various people in different contexts, but trust very few (if any) people in every context.

On the Internet, it may be difficult or impossible to develop encapsulated interest between individuals with ephemeral communications and very little at stake (for example, public chat rooms or threaded conversations on a range of topics). However, in popular forms of Internet communication such as online dating, trust is intrinsic to every aspect of the user experience. Beyond its popularity as a tool for finding romantic relationships, online dating provides a unique window into the complexities of building online interpersonal trust over time through different forms of communication. Online dating is largely about learning to use the affordances of online communication channels with low personal risk, with the purpose of finding individuals who are, among other things, sufficiently trustworthy to meet in person.[5]

Online dating interactions typically follow trajectories that begin with highly uncertain, but low-risk, online interactions with others (such as online messaging and inquiries about possible interest). Interactions are initially uncertain because individuals know only the version of each other presented in their dating profiles. Individuals have a high degree of control over how much they want to risk revealing about themselves in different channels of communication (including instant messages, email messages, online chats, telephone conversations, and face-to-face interactions). If the process of learning about a potential date is successful, individuals may continue their communication offline, where physical and personal risks are arguably more numerous.

Online Trust, Trustworthiness, or Assurance?

Although *trust* and *trustworthiness* appear interchangeably in common vernacular, they are distinct concepts. *Trustworthiness* is a characteristic or property of an individual; *trust* is an attitude or belief we have about those who are trustworthy.[6] Unlike interpersonal trust, an assessment of trustworthiness does not require prior firsthand communication or experience.

A common way to infer the trustworthiness of another in online environments is through explicit or implicit third-party reputation information. If we receive online advice from someone we do not already know, even basic information about others' experiences with the individual can help us make an informed decision about the credibility of his or her claims. Explicit reputation information includes ratings, reviews, and other assessments of his or her qualities based on prior personal experience. Alternatively, implicit reputation information includes signals, cues, and traces from an individual's prior actions that might correlate with future behavior (for example, his or her frequency of activity in an online forum or accuracy of grammar and spelling). Online reputation information is not a perfect solution to the problem of trust because it is vulnerable to exploitation, deception, and misinterpretation. Despite these challenges, stable reputation systems can ameliorate concerns about risk and uncertainty, leading to online cooperation and higher assessments of trustworthiness.

Many different issues can affect judgments of trustworthiness, including the nature of the situation and perceptions about another person's intentions and motivations. When we assess an individual's trustworthiness, we essentially decide whether to take a risk with that person. Typically inconsequential features and characteristics become essential for inferring trustworthiness, even if our assumptions about the link between certain characteristics and behavior are imperfect. This observation extends to the online world of social interaction, where email addresses, domain names, and other features imply trustworthiness in the absence of other available attributes.[7]

Given that relational trust requires ongoing, experiential information about another person, nonrepeated interactions between individuals with no prior communication are not based on trust: they are acts of risk-taking. In interpersonal online interactions, an act of risk-taking does not guarantee reciprocity or the development of trust. In fact, experimental research in computer-mediated environments demonstrates that risk-taking among anonymous actors often goes unreciprocated.[8] Numerous online interactions are fleeting, one-time communications. Examples include question-and-answer forums, threaded comments on websites, blog commentaries, and public online chat rooms. In these situations, assessments of trustworthiness can be essential, even if relational trust is not possible. For this reason, to call for more trust in anonymous online interactions is misleading. Relational trust is not even possible or desired in these situations, but the ability to infer *trustworthiness* is essential.

Signaling one's intentions through an initial act of risk-taking is critical in online interactions when incomplete information (anonymity or pseudonymity) makes it difficult to assess the trustworthiness of others. Risk-taking can act as a signal when individuals intentionally give up something of value without any explicit form of assurance. For example, taking a disproportionate amount of risk in an online transaction by sending money before an item has shipped sends a clear signal that the buyer believes the seller is trustworthy. Willful acts of on-

line risk-taking can also send a more general signal about an individual. As media and social communications scholar Judith Donath argues, "[R]isk taking is another behavior that may seem irrational, but when viewed as a signal, can be seen as a way of claiming a high level of fitness. . . . [P]osting revealing or culpable material online arguably has become another forum for signaling imperviousness to danger and repercussions."[9] Those who show that they are willing to take a chance online open themselves to risk; however, the long-term payoff is a kind of online social intelligence that rewards risk-taking with new opportunities.[10]

If individuals cooperate with one another over time, is it accurate to say that the individuals are engaged in a trusting relationship? There are roughly two schools of thought on the link between trust and *cooperation*, but both perspectives agree that the two concepts are separate. Scholars such as political scientist Robert Putnam argue that trust is required to produce cooperation, which in turn helps create productive societies.[11] A competing viewpoint maintains that trust exists at the interpersonal level to produce "social order and to lower the costs of monitoring and sanctioning that might be required if individuals were not trustworthy."[12] In this latter view, cooperation can (and often does) exist independently of trust.

Conceptualizing the difference between trust and cooperation requires an understanding of motivations and the context in which individuals interact. For example, elected officials on different sides of the political spectrum may cooperate on legislation to further their own (perhaps competing) interests. Similarly, individuals who agree to exchange valued goods and services through popular online classified advertisement websites (such as Craigslist) may cooperate by successfully completing an agreement or exchange. In both examples, cooperation can exist without requiring interpersonal trust. However, "trust is most likely to emerge in contexts in which the parties find themselves in ongoing relationships,"[13] argue sociologist Karen Cook and her colleagues. Trust may not be required in cooperative interactions but becomes increasingly important as relationships persist.

Experimental research clearly demonstrates that acts of trust are distinct from acts of cooperation. In one laboratory study of social exchange in a computer-mediated system, my colleagues and I measured entrusting behavior as the number of valued items a participant entrusted to another.[14] Individuals could keep the valued items or return them to the entruster (who would then receive double the entrusted value in return, paid by the experimenter). Cooperation, therefore, was the decision to return rather than keep the entrusted items. The results of the study show that individuals entrust very little to their partners when they exchange with a new, random partner on every interaction opportunity. However, when individuals interact with the same cooperative partner over time, they tend to increase slowly the amount entrusted. Thus, cooperation over time leads to larger, higher-risk entrustments. The occurrence of ongoing risk-taking in the presence of uncertainty makes trust both possible and relevant. It is the *process* of risk-taking with the same individual over time that separates one-time assessments of trustworthiness from interpersonal trust.

When organizational or institutional mechanisms exist to protect individuals from harm (such as betrayal), individuals tend to rely primarily on the third-party protections *instead* of building mutual

trust. This inclination leads to a trust paradox because the *assurance* structures designed to make interpersonal trust possible in uncertain environments undermine the need for trust in the first place. Individuals have a strong incentive to act cooperatively when robust monitoring and assurance structures are present, but cooperation in these cases has more to do with sanctions and other negative outcomes than interpersonal trust. In essence, strong forms of online security and assurance can supplant, rather than enhance, trust.[15] A different way of viewing the trade-off between trust and assurance requires returning to the preconditions of risk and uncertainty that underlie interpersonal trust relations. Effective assurance systems minimize or eliminate one or more forms of uncertainty, such as the ambiguity of another's intention to follow through with a prior agreement. The source of uncertainty then shifts to the assurance system, thereby making trustworthiness and reliability of the institution or organization the salient relationship.

Strong assurance systems are widely portrayed as a solution to the problem of trust in Internet environments. For example, third-party verification systems for buyers and sellers who use Internet auctions help reduce malfeasance and fraud.[16] In addition, a leading suggestion for addressing the reliability of routing information on the Internet involves the concept of a *trust anchor*, or an authoritative body that provides assurances about data authenticity between Internet networks.[17] However, my research with Ashwin Mathew, Ph.D. candidate at the University of California, Berkeley, demonstrates that the complexities of Internet routing are currently resolved through interpersonal trust relationships among network administrators.[18] Ongoing trust relationships enable the network administration community to act collectively and dynamically to keep information flowing properly through the disparate networks that form the larger Internet. Replacing human interpersonal relationships with authoritative trust anchors would provide a centralized point of decision-making, but also of potential failure.

The choice between assurance structures or emergent trust relationships without structural assurances on the Internet is largely about short-term versus long-term goals. Assurances lessen uncertainties, while trust thrives where uncertainties are abundant. Although individuals cooperate as if they trust one another when third-party assurance structures guarantee interactions, this behavior is ultimately unstable. When and if institutional assurance structures fall short, individuals realize that the original source of security is gone, and interpersonal relationships must be forged once again amid uncertainty.

Experimental research on computer-mediated trust relationships supports the claim that assurances may solve short-term problems of cooperation at the expense of building long-term trust. In laboratory research on trust and forms of social exchange, my colleagues and I find that individuals build high levels of trust when partners cooperate in highly uncertain environments, compared to those who cooperate in low-uncertainty environments (namely, when interactions are ensured by a third party). However, when individuals who have built trust with their partners without assurances shift into a new interaction environment with assurances, trust significantly decreases between the pair even when cooperation remains high.[19] As individuals learn to rely on assurance structures, they do so at the expense of interpersonal trust. The key implication is that if assurances fail, cooperation and trust will fail as well.

My discussion of trust and related concepts has thus far been limited to interpersonal, human-to-human interactions. However, many researchers and practitioners primarily focus on the relationship between humans and the technologies, systems, and interfaces that we use on the Internet. Online trust has been the focus of a wide variety of research in Internet studies, computer-mediated communication, human-computer interaction, computer-supported cooperative work, and related fields. Among the more technical areas in Internet research, the term *trust* routinely refers to several related but distinct concepts, including *credibility*, *security*, *surety*, and *reliability*. Unfortunately, these numerous meanings create an overabundance of complex models that are largely incompatible because of the vast conceptual differences among key terms.

Trust in an information system primarily involves individuals' expectations about whether the system will operate in a predictable manner and provide reliable outputs. While many may feel that a hard-line distinction between human-to-human and human-to-system trust is unnecessary, the potential difference in meaning for terms such as *intention*, *agency*, *cooperation*, and *choice* is difficult to ignore when considering programmed systems versus human actors. For example, philosopher Annette Baier argues that relational, interpersonal trust depends on the possibility of betrayal by another person.[20] Information systems and computer programs appear to lack the agency and consciousness to choose freely to betray the trust that users place in them.

In many situations, there may be little perceived difference in the evaluation of risk and uncertainty when interacting with a system or with another person. From the perspective of humans who believe that they trust systems, evaluating the trustworthiness of a system is very much like assessing the trustworthiness of a person because an individual will use his or her prior experiences with the system (and other similar systems) to create an estimate of risk and uncertainty. However, the interaction lacks the *relational* dynamic that is essential to interpersonal trust. Even if an individual develops very complex heuristics for "trusting" systems, it is often a unidirectional evaluation.

Unlike complex social relationships in which humans must continually assess one another's motivations, intentions, and behaviors before and after an interaction, it is difficult (and currently implausible) for an online system to be endowed with the sentience and free will to shift between different motivations and intentions unless these are programmed, defined, or changed by other humans. As computer scientist John Seely Brown and historian and social theorist Paul Duguid demonstrate through numerous examples, people tend to treat bots and information agents as if they are human, even when we clearly know they are not.[21] In addition, the experimental work of communications scholar Cliff Nass and his colleagues provides further empirical evidence that people anthropomorphize computers and programmed systems that respond like humans.[22]

It is also possible that "trust-like" behavior in computers and systems is basic categorization and simplification of the sort that linguist George Lakoff describes as fundamental to human understanding.[23] According to Lakoff, individuals lump similar concepts together based on personal experiences, a process that is often good enough for routine sense making. We might feel betrayed by a computer even if we fully comprehend that the computer did not really *choose* to

act a particular way. Thus, equating trust in a computer system with trust in another person can seem completely reasonable from an individual's perspective.

Distinguishing between interpersonal trust and human-system "trust" may seem somewhat pedantic, especially outside of scholarly debate. However, when we delve deeper into the realm of building enduring trust relationships and constructing trustworthy networks and systems, the distinctions between concepts like *interpersonal trust* and *system trust* become essential for understanding when and why trust fails. When a human betrays a friend's trust, the friend knows who is culpable, and the consequences are often clear for both parties. When a system "betrays" a human's trust, assigning blame can have enormous repercussions: should we blame the system itself, those who programmed the system, the organization that hosts the service, a quality control person, the director of the organization, or the organization in general? Can the system *learn* or do anything differently after failed trust? Put simply, we need to be able to attribute an outcome to someone or something if we hope to understand and respond to failures of trust in systems.

In late January 2011, Egyptian authorities successfully shut down the international Internet access points that allow traffic to pass to and from systems in Egypt. Shutdowns of this type occur from time to time on a smaller scale for technical and political reasons, but this event was largely unprecedented in terms of range and scale. The explanation for the Internet connectivity blackout in Egypt emerged from a straightforward political decision: the autocratic Egyptian government hoped to quell information sharing and coordination among its citizens amid the growing talks of protests against the current regime.[24]

The Internet blackout in Egypt was not really about system trust, nor was it about direct, interpersonal trust between individual citizens and government officials. It ultimately concerned the Egyptian government's lack of trustworthiness in the eyes of the Egyptian people as a result of various actions over time. The government's decision to interrupt Internet access furthered the image of an untrustworthy regime, leading to contempt that spurred calls for revolution.[25]

The event in Egypt is an example of what sociologist Niklas Luhmann views as a type of system trust in governments and other organizational machinations. In Luhmann's macro-level view of system trust, individuals trust the *structures* of human interaction and organization.[26] However, I believe that the Internet shutdown in Egypt is a model example of why trust placed in information systems often has less to do with the actual systems and more to do with the humans, organizations, and governments that maintain or control them. In contrast to Luhmann's position, Russell Hardin argues that trust in macro social structures (governments and other human systems) is largely implausible. Hardin believes that the complexity and scope of any large government makes true interpersonal trust impossible; indeed, each citizen cannot realistically build firsthand experience with every politician or official. Thus, the relational view of trust maintains that it is more important "that government be trustworthy than that it be trusted."[27] Hardin's position is consistent with Luhmann's opinion if we view each as a statement about *trustworthiness* rather than *trust*.

Creating and maintaining social order and trust between individuals are fundamental problems for human interaction, whether online or offline. Many existing and emerging online systems enable

social interaction, collective action, and interpersonal communication on a scale that was unthinkable before Internet use became commonplace. The Internet *is* the real world, but we must remain cognizant of the complexities of trust and social interaction in computer-mediated environments. Our ability to understand and articulate the differences between trust, trustworthiness, cooperation, assurance, and related concepts is not a purely bookish problem: the possibilities and limitations of trust and social interaction on the Internet will depend entirely on how we design online communication technologies in the context of the surrounding global political and institutional environment. Will we continue to facilitate interpersonal interaction and embrace the complexity of emergent social uses of online information technologies? Alternatively, will we build structures to control, limit, and secure online social interactions – and accept the potential trade-offs for trust in exchange for assurance?

Clarifying the differences between forms of online trust is crucial because current and future policy-makers justify strategies and decisions with outcomes and recommendations from trust and security research. In spite of attempts to create trust by eliminating doubt and minimizing peril, there is no quick way to build meaningful trust without overcoming real risk in the presence of uncertainty. Assurance structures are undoubtedly an important part of complex systems; sufficient evidence shows that they can encourage cooperative behavior – especially when all other options fail. However, if we attempt to stamp out all online uncertainty and risk through security measures and centralized assurance structures, we may inadvertently create a *less* trusting Internet environment in the long term.

ENDNOTES

[1] I am deeply grateful to Ashwin Mathew for his insightful comments and suggestions on an earlier draft of this essay.

[2] Russell Hardin, *Trust and Trustworthiness* (New York: Russell Sage, 2002).

[3] Karen S. Cook, Russell Hardin, and Margaret Levi, *Cooperation without Trust?* (New York: Russell Sage, 2005).

[4] Russell Hardin, "Conceptions and Explanations of Trust," in *Trust in Society*, ed. Karen S. Cook (New York: Russell Sage, 2001), 3 – 39.

[5] Andrew Fiore and Coye Cheshire, "The Role of Trust in Online Relationship Formation," in *Trust and Technology in a Ubiquitous Modern Environment: Theoretical and Methodological Perspectives*, ed. Dominika Latusek and Alexandra Gerbasi (Hershey, Pa.: IGI Global, 2010), 55 – 70.

[6] Carolyn McLeod, "Trust," in *The Stanford Encyclopedia of Philosophy*, ed. Edward Zalta (Stanford, Calif.: Stanford University Press, 2008).

[7] Judith Donath, "Identity and Deception in the Virtual Community," in *Communities in Cyberspace*, ed. Mark A. Smith and Peter Kollock (London: Routledge, 1998).

[8] Toshio Yamagishi, Masafumi Matsuda, Noriaki Yoshikai, Hiroyuki Takahashi, and Yukihiro Usui, "Solving the Lemons Problem with Reputation: An Experimental Study of Online Trading," in *eTrust: Forming Relationships in the Online World*, ed. Karen S. Cook, Chris Snijders, Vincent Buskens, and Coye Cheshire (New York: Russell Sage, 2009), 73 – 108.

[9] Judith Donath, "Signals in Social Supernets," *Journal of Computer Mediated Communication* 13 (1) (2007): article 12.

[10] Coye Cheshire and Judd Antin, "None of us is as lazy as all of us: Social Intelligence and Loafing in Information Pools," *Information, Communication & Society* 13 (4) (2010): 537–555.

[11] Robert Putnam, "Bowling Alone: America's Declining Social Capital," *Journal of Democracy* 6 (1995): 65–78.

[12] Cook, Hardin, and Levi, *Cooperation without Trust?*, 1.

[13] Ibid., 4.

[14] Karen S. Cook, Toshio Yamagishi, Coye Cheshire, Robin Cooper, Masafumi Matsuda, and Rie Mashima, "Trust Building via Risk Taking: A Cross-Societal Experiment," *Social Psychology Quarterly* 68 (2) (2005): 121–142.

[15] Helen Nissenbaum, "Will Security Enhance Trust Online, or Supplant It?" in *Trust and Distrust within Organizations: Emerging Perspectives, Enduring Questions*, ed. Roderick M. Kramer and Karen S. Cook (New York: Russell Sage, 2004), 155–188.

[16] Brad Stone, "EBay Says Fraud Crackdown Has Worked," *The New York Times*, June 14, 2007, http://www.nytimes.com/2007/06/14/technology/14ebay.html.

[17] Matt Lepinski and Stephen Kent, working paper, "An Infrastructure to Support Secure Internet Routing," 2010, http://tools.ietf.org/html/draft-ietf-sidr-arch-11.

[18] Ashwin J. Mathew and Coye Cheshire, "The New Cartographers: Trust and Social Order within the Internet Infrastructure," *Proceedings of the 38th Research Conference on Communication, Information and Internet Policy (Telecommunications Policy Research Conference)*, George Mason University School of Law, Arlington, Virginia, 2010.

[19] Coye Cheshire, Alexandra Gerbasi, and Karen S. Cook, "Trust and Transitions in Modes of Social Exchange," *Social Psychology Quarterly* 73 (2) (2010): 176–195.

[20] Annette Baier, "Trust and Antitrust," *Ethics* 96 (2) (1986): 231–260.

[21] John Seely Brown and Paul Duguid, *The Social Life of Information* (Cambridge, Mass.: Harvard Business School Press, 2000).

[22] Cliff I. Nass, Jonathan Steuer, and Ellen R. Tauber, "Computers are Social Actors," *Proceedings of the Special Interest Group on Computer Human Interaction Conference on Human Factors in Computing Systems: Celebrating Interdependence*, Boston, Massachusetts, 1994.

[23] George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: University of Chicago Press, 1987).

[24] Matt Richtel, "Egypt Cuts Off Most Internet and Cell Service," *The New York Times*, January 28, 2011, http://www.nytimes.com/2011/01/29/technology/internet/29cutoff.html.

[25] Mansoura Ez-Eldin, "Date With a Revolution," *The New York Times*, January 30, 2011, http://www.nytimes.com/2011/01/31/opinion/31eldin.html.

[26] Niklas Luhmann, *Trust; and Power: Two Works*, trans. Howard Davis, John Raffan, and Kathryn Rooney; ed. Tom Burns and Gianfranco Poggi (New York: Wiley, 1979).

[27] Hardin, *Trust and Trustworthiness*, 152.

Safety in Cyberspace

Vinton G. Cerf

Abstract: Safety in cyberspace continues to be an elusive objective. This essay explores various metaphors to aid thinking about the means by which safety might be increased. Notions such as cyber-fire-departments or cyber-police-departments as well as models drawn from public health scenarios are considered. The legal frameworks in which safety can be improved and international agreements adopted toward this end are briefly discussed. Users can also contribute to their own safety by adopting various practices that reduce vulnerability to cyber-infection and compromise.

VINTON G. CERF, a Fellow of the American Academy since 1995, is Vice President and Chief Internet Evangelist for Google Inc. Previously, he served at MCI, the Corporation for National Research Initiatives, the Defense Advanced Research Projects Agency, and as a member of the Stanford University faculty. He co-invented the architecture and basic protocols of the Internet and is the co-recipient of the U.S. National Medal of Technology and the Japan Prize.

The term *cyberspace*, first coined by novelist William Gibson[1] in 1984, has been absorbed into daily language and extended through the prefixing of *cyber-* to an extraordinary number of words. The effect is to imbue ordinary terms with the mystery and otherworldliness of the Internet, World Wide Web, and other artifacts formed through the use of computers and their interconnection. As the Internet and its World Wide Web application have expanded in geographic scope and use, especially by the general public, we have witnessed the emergence of a wide range of practices in the online environment that mirror their counterparts in the physical world. Of course, the Internet itself is realized in a physical way, but it also forms the basis for a virtual environment fashioned of bits rather than atoms. It is this artificial environment that Gibson labeled "cyberspace."

To distinguish various activities that entail the use of computers from those that involve the physical world, it has become common to label the computer-oriented activities as "cyber-activities" or, sometimes, "e-activities" or "e-objects." We speak of "email" or "e-books" and "e-commerce" as if they occupy a place in a virtual universe. We use our commonplace, real-world activities as models for their cyberspace counterparts. In many respects,

these models help us appreciate the significance and utility of the networked, online equivalents. Thus, "digital signature" is used to refer to a way of cryptographically "signing" a digital document, giving the sense that the digital signature is the equivalent of one rendered with pen and ink or other marking instrument.

While this terminology is convenient, it often has limited applicability and may even mislead our intuitions about the cyber-equivalents of real-world objects, practices, actions, and events. For example, the "owner" of a printed book is free to sell, loan, hypothecate, destroy, fold, spindle, or mutilate the work without intervention by the copyright holder. The owner is not generally free to reproduce and distribute copies unless such rights have been granted by the copyright holder. Now consider an electronic (digital) book that might be downloaded into an electronic book reader, laptop, "netpad," or other display device. The copy of the book held in the display device may not be transferable to another party, in part for technical reasons (the formats of various book reading devices may be incompatible, for example, or there may not be a mechanism for transfer between even compatible devices) or because such transfers are not authorized by the copyright holder.

While some electronic book distribution services permit users to hold multiple copies of a purchased book if the copies are placed solely on devices owned by the user, transfers among reading devices owned by distinct parties are atypical. This example alerts us to the potential hazards of uncritically adopting models to guide our thinking about policies applied to activities undertaken in cyberspace. Not every familiar action in the real world has an accurate analogue in cyberspace.

Although there is no single, agreed upon definition of *cybersecurity*, it is important to recognize that the term covers more than notions of security specific to the Internet. We have many devices and appliances that rely on computers and communication systems that are not necessarily part of the Internet. That these systems are vulnerable in varying degree seems unquestionable, even when they are not part of or even episodically connected to the Internet. Various kinds of digital media (DVDs, thumb drives, memory sticks) can be the bearers of bad software (malware) or potential infection resulting from a visit to a malware-bearing website.

Broadly interpreted, cybersecurity refers to safety from deliberate or accidental harm in connection with the use of computer-based systems, whether networked or not. How many of us have experienced a hard-disk failure and suffered consequential damage, despite the likelihood that the failure was random, neither induced by any malicious attacker nor triggered by a network-based attack? Then there are the seemingly ever-present "bugs" in software that cause data loss or other detrimental failures. For the purpose of exploring some of the potential risks to consumer use of computing systems and responses to them, the term *cyber-safety* seems somewhat more apt. This term may help induce a broader mindset for thinking about the conditions needed to protect the users of software systems from malfunction and harm, even when the harm is not the consequence of intended malice.

In the twenty-five years since the Internet was operationally launched (January 1, 1983) and Gibson's book appeared (ironically, in 1984), a number of metaphors have been applied to frame thinking

about software and the challenges it poses. Not long after personal computing became popular, hackers found ways to make computers ingest software that users did not intend to have loaded into their machines. The general term for this unintended ingestion, *infection*, is drawn from the obvious biological metaphor. The bad software was often called *malware*, shorthand for "malevolent software." In the period before the Internet became widely available, a common means of propagating malware was to place it on physical media such as floppy disks, CD-ROMS, and, later, memory sticks and thumb drives. When targets "read" these media, the malware would enter their computers, where it would carry out its function regardless of user awareness or intent.

The computing community adopted several terms for such malware, including *virus*, *worm*, and *Trojan horse*. Viruses are fragments of software that propagate through some vector (physical media, attachments to electronic mail, or downloads from websites). Worms are programs that propagate themselves, often through the Internet or by installing themselves on physical media. Trojan horse software may be drawn into a target computer by direct action of a hacker or through a virus or worm. The Trojan horse software typically stays resident and hidden until conditions cause it to activate (for example, by receipt of a message from some Internet source, or by some local condition in the computer in which it is lodged).

In recent years, among the most common ways for personal computers to become infected is through the use of browsers that help users "surf" the World Wide Web. In its earliest incarnation, the World Wide Web served up Web pages that were mostly text and imagery, but the last decade has seen the development of high-level programming languages such as Java[2] or JavaScript[3] that can be interpreted ("executed") by or through browsers to increase the level of interactivity and functionality of a Web page. When running on certain browsers and operating systems, these high-level language tools can be used to introduce viruses, worms, and Trojan horses into a target computer. Just as we are vulnerable to infection in the biological real world, our computers are vulnerable to becoming infected with malware in the cyber-world.

The objectives of infection may vary from entertainment without harm to considerable damage and data loss and everything in between. In some cases, the purpose is to obtain information possessed by the user of the computer that has become infected: passwords, usernames, financial account information, and corporate information are examples of the targets for this kind of attack. In other cases, the purpose may be to take control of the computer and use it to mount denial-of-service attacks against other computers that are reachable on the same network (not necessarily the Internet, per se) or to generate huge quantities of spam (unwelcomed commercial email). Sometimes, the target is extortion ("Pay me or I will cause all your information to be wiped out") or some other coercive purpose ("I have found incriminating email messages or images and will publish them unless you take the following actions"). Human ingenuity and malfeasance are alive and well in the cyber-world.

Many of the assets used by Internet attackers are harvested from the world's nearly two billion Internet users: namely, their laptops and desktops. In addition, Web servers have been compromised. The mechanisms of compromise vary, and this essay is not intended to catalog even a fraction of them. In principle, users of the

World Wide Web are frequently subject to compromise when they visit infected server sites that inject malware into their computers simply as a consequence of visiting a particular website.

The compromised machines, especially laptops and desktops, become "zombies" controllable by the attacker. Collections of zombie machines, called "botnets" (for "robot networks"), could include hundreds of thousands to millions of computers. The "botnet generals" control these assets remotely, using them to launch denial-of-service attacks, generate spam email, or carry out other nefarious actions. Ironically, many of these actors have no desire to disable the Internet. Rather, they are parasites seeking monetary gain from their control of these resources. Botnets are often rented out for use by third parties. More ironic, botnet generals jealously guard their ill-gotten resources. To defend against a takeover by a competing botnet, their resident malware sometimes tries to detect attempts by other infecting sources to capture the machine. One wonders whether it is possible to invent a "good botnet" we could all join that would protect us from the "bad guys" – rather like a vaccination! Employing such an idea would be more proactive than using current virus detection software, which generally tries to recognize malware before it is activated (executed) and usually relies on static signatures of previously recognized viruses, worms, and Trojan horses.

It is important to note that the creators and propagators of malware may wish to keep their targets unaware of infection – remaining invisible and unidentifiable as the source of control – in order to use the infected computers to generate income. Their purpose may not necessarily be to destroy or disrupt but to abuse access to computer resources they lack the legal or moral rights to use.

In keeping with the metaphor of infection, we might consider what steps consumers can reasonably take to minimize the risk that their computer-based devices become infected with malware. That is, how can they practice good *cyber-hygiene*? Commercial software systems called "anti-virus packages" examine incoming files and messages arriving via email, on physical media, or through browsers to try to detect and filter out known or suspected malware. While not 100 percent capable of detecting every infection, such systems provide a partial defense against the ingestion of malware.

Anti-viral defenses cannot be static because the makers of malware evolve their software in response to these defenses. Users of anti-viral software must install regular updates to stay current with known viruses, worms, and Trojan horses. More recently, these commercial offerings have begun trying to detect malware never seen before (so-called zero day attacks) by detecting various forms of anomalous behavior.

Search companies such as Google also attempt to identify websites housing malware that might be employed to infect a user's computer. As the search engine forms the index of the World Wide Web by "crawling" through all the Web pages on the Internet, the indexing software also looks for possible malware on each site and page visited. If malware is suspected, the crawling software can make note of it. For example, if a user of the Google search service clicks on a hyperlink leading to a site that Google suspects is infected, a bright-red warning page shows up on the screen, notifying the user of the potential hazard of visiting that website.

If a website operator believes that the malware report is incorrect, he or she is invited to visit the StopBadWare site[4] for consultation and assistance to verify the

presence and assist in the removal of any malware.

Users can also contribute to increased cyber-safety by observing good safety and security practices. Passwords should be long enough and sufficiently complex to eliminate guessing as a means of "cracking." Incredibly, some users choose the word *password* to "protect" their access to online services. Users should avoid the use of birth dates, addresses, common words, or even common word combinations as passwords. For example, random, pronounceable passwords that alternate consonants and vowels, mixed together with digits and special symbols (such as "horif@mi837" or "5atogesi#37"), can form useful bases for stronger protection.

In the past decade or so, new mechanisms for strong authentication have emerged. Devices with small liquid crystal displays are used to generate cryptographically random six- to ten-digit numbers serving as temporary passwords that expire within tens of seconds. Even if these passwords are detected by malware or seen by other users, they are not reuseable. The use of such techniques is sometimes referred to as "two-factor authentication" because users must offer not only a username and conventional password before access to a service or system can be authenticated, but also the cryptographically generated value from a physical device they have in their possession.

People do forget their passwords. One of the values of the two-factor scheme is that the random password generator does not have to remember anything; it continually generates new passwords (in synchrony with the cooperating server). Some online services prompt users to create "secret" questions to be answered with "secret" answers. Examples include "mother's maiden name" or "name of your pet." One of the problems with such methods is that the answers may be easily discovered through a World Wide Web search. Users need to be very creative about their choices of "secret" questions and answers. It might even be sensible to use incorrect information (for example, don't use your mother's real maiden name). Of course, that tactic requires users to remember the false information, which may be harder than remembering the password they forgot!

Some online services will send users a new password (or their old one) as an email message to an email address that users have configured into their account information for that service. That system is flawed by the fact that if the email account has been compromised, it can be used as the avenue through which many other accounts can be penetrated. An attacker can visit the target site and assert that "I have forgotten my password," prompting the service site to send password information by email, which can then be intercepted by the attacker who has, by some means, obtained the user's email-access password(s).

A smarter practice may be to establish backup email accounts used rarely and perhaps predominantly for password recovery; but generally, allowing passwords to be sent by email can compromise security. Some systems rely on mobile phones, sending a text message with a new password. In other cases, though they are more expensive to implement, true out-of-band methods (postal mail or telephone calls) may be more effective; however, their disadvantages include delay for access or the problem of determining whether the caller in need of a new password is, in fact, whom he or she claims to be!

A commonly reported problem in cyberspace is identity theft. The thief manages to discover sufficient information through Web-surfing, possibly adding information gathered by other means (from *Who's Who* publications or alumni magazines, for

example) to make a creditable (no pun intended) application for loans, credit cards, bank checking accounts, and other financial instruments. This problem suggests that consumers should be careful in choosing the information that they share, for instance, in social networking applications, blogs, personal Web pages, email messages, and other online means of communication. Financial institutions are hard-pressed to balance consumer safety with convenience and utility. Some brokerage houses, for example, will not accept email orders (out of concern for timeliness and reliability), and others will insist on voice or fax confirmations of orders.

Consumers must be thoughtful in disclosing personal information on the Internet. The organizations they belong to must be equally careful in deciding what to provide online (such as whether board or staff biographies should include email addresses, telephone numbers, or information about family members).

In their essay for this issue, Deirdre K. Mulligan and Fred B. Schneider explore in greater detail the public health analogy for cyber-safety. The notions of worms, viruses, infections, vaccinations, immune systems, and the like all derive from a biological metaphor for the relationship between software and the engines that interpret it. The idea of public health as a model for defense against computer malware has much to offer, at least as an organizing principle for considering responses to threats to cyber-safety and security.

The public health metaphor also suggests the many distinct but interacting "cyber-life-forms" in the software ecology. Interactions among these cyber-life-forms may occur through networks, including the Internet, or by physical transfers of data using memory sticks, thumb drives, CDs, DVDs, and other devices. Humans inject data into these systems through keyboards, cameras, tracking pads, microphones, and an increasing array of sensor systems. The idea that interacting software systems may produce something as complex as a biological ecology should give us pause as we think about protecting our society from deliberate or accidental malfunctions of the complex software systems that we depend on so heavily.

Protecting public health may involve quarantining, vaccination, and other preventative mechanisms; exhortations to manage diet and exercise regularly; or treatment of chronic illnesses through repetitive consumption of medication. When applied to software systems, however, many questions arise regarding proper analogies in the cyber-environment. What does it mean to quarantine a computer or computer-based system? Who can decide to quarantine and how is it enforced? What software vaccinations (that is, anti-virus software) are necessary? Which are effective? How are these validated? Is there an equivalent to the Food and Drug Administration and, if not, should there be? These and many other questions arise when bio-notions are applied to the cyberspace ecology. Establishing institutional practices and guidelines for safety in cyberspace will likely require thoughtful legislation and regulatory response if the metaphor is to prove practical and useful.

Two other metaphors have emerged in the past two decades: *cyber-police-* and *cyber-fire-departments*, which both fall under the rubric of *cyber-first-responders*.

Many kinds of attacks in cyberspace have analogies in the real world: stalking, fraud, denial-of-service, identity-theft, theft of goods or services, possession or sale of stolen goods, counterfeiting, drug

trafficking, and so on. In addition, many conventional crimes use the Internet as an aid to perpetration. However, it is not always apparent that a cyber-attack is, in fact, deliberate or a violation of law. Moreover, laws are not uniform, varying among countries, within countries, and among intra-national jurisdictions. Indeed, in the case of the Internet or cyberspace more generally, jurisdiction may be a slippery notion, owing to the highly geographically distributed nature of the parties involved.

If the metaphor of a cyber-police-department is to be of value in organizing a means for protecting citizens from assault in cyberspace, many questions will have to be answered. What actions are considered violations of law? In which jurisdictions? Are there applicable extradition treaties? Are there agreements for international or interjurisdictional cooperation among law enforcement agencies? Applying existing law to cyber-crimes across many jurisdictions would require a great deal of work. It is also important to recognize that not all incidents that appear to be attacks are the result of deliberate intent. Cyber-law-enforcement will need tools and nuanced facilities to distinguish crimes from accidents or harmful but unintended mistakes.

Now imagine that your home or office building is ablaze. Your first instinct is to call not the police department but the fire department. The objective is to put out the blaze as quickly as possible. The private sector's potential inadequacy to respond to a serious fire is captured by the image of a homeowner standing before his burning residence holding a garden hose. To fight the fire, he needs someone with a big hose, a pump, and a lot of water. A cyber-fire-department might be called to help respond to a cyber-attack that the victim is not able to cope with using locally available tools.

This metaphor has some useful features. The fire department's primary job is to put out the fire. After the fact, it may try to determine the cause of the blaze. If it appears to be deliberate, that is, arson, law enforcement agencies may then be called on for further action. Moreover, the fire department often will offer to inspect buildings for fire code violations or hazards and may also participate in the development of fire codes to help protect the community from poorly designed buildings.

This metaphor, too, raises many questions. Who is permitted to call the cyber-fire-department to fight a cyber-attack? Is it purely voluntary? Can a company call the cyber-fire-department as an anti-competitive tool to disrupt a competitor's business? What is the cyber-fire-department allowed to do to systems that are under attack? The conventional fire department is allowed to break doors, windows, roofs, and walls in its effort to quell a blaze and/or rescue endangered parties. What about the cyber-fire-department? What authorities would need to be granted, and by whom, to allow for effective operation? These and many other questions remain unanswered and are made more complex by the possibility that a cyber-fire is burning across international or other jurisdictional boundaries.

In the commercial sector, perhaps the closest relative to cyber-emergency-response is the customer service department of consumer equipment outlets. The Apple Store comes to mind as an example. At Google, there are Tech Stops populated by engineering experts who analyze malfunctions, debug configurations, and help consumers return to productive use of their computer-based equipment.

Many of the metaphors for dealing with cyber-emergencies draw on the idea

of first responders. It may be that the notion of first response is applicable to various situations that will later evolve into much more complex, potentially internationally coordinated responses. For the most part, very few institutions have been crafted on the basis of the models suggested above; thus, their utility as guides for increasing safety in cyberspace remains to be explored.

Responses to risks in cyberspace fall into three broad categories:

1) technical responses to prevent harm and preserve safety;

2) post-hoc detection and punishment regimes; and

3) moral suasion.

We know that none of these approaches, nor even any combination, can absolutely guarantee the preservation of safety and protection from harm. All are capable of mitigation in some degree. For this reason, most responses to potential hazards, threats, and vulnerabilities are treated as risk-management problems.

It is not within the scope of this essay to identify every technical effort bent on preventing harm in cyberspace, but a few illustrative examples may be helpful. In the realm of standards creation, the Internet Engineering Task Force (IETF)[5] and the World Wide Web Consortium (W3C)[6] have begun to introduce mechanisms for improving the security of the Internet's infrastructure. For the most part, these mechanisms operate without the need for user intervention. For example, many Web-based services operate by encrypting the data flowing between the user's browser and the serving website. These protected "tunnels" are set up automatically, generally without user action. Employees using organizational resources remotely (whether from home or from Starbucks) may be required to use what are called Virtual Private Networks, created by end-to-end encryption of the data exchanged. The W3C has developed encrypted versions of the HyperText Transport Protocol (HTTP and HTTPS) that use underlying Internet security standards to provide confidentiality.

Recently introduced mechanisms for securing the Domain Name System – called Domain Name System Security Extensions (DNSSEC)[7] – have been implemented by the Internet Corporation for Assigned Names and Numbers (ICANN) and other agencies that manage Internet domain names. (An example of a domain name is www.icann.org.) Again, operating invisibly to users, the system allows the user's laptop or desktop software to verify that it has the correct Internet address for the destination computer that the user is seeking to reach. These methods mitigate sophisticated attacks that alter the integrity of the directory of locations found in cyberspace.

There are ongoing efforts to create additional standards and practices to protect the routing system in the Internet from malfunction or deliberate misrouting. As in most of the previously mentioned tactics, users are largely oblivious to and do not need to take action to benefit from these methods.

In the Internet environment, one of the earliest institutional responses to the problem of cyberspace hazards was the formation of the Computer Emergency Response Team (CERT) at the Software Engineering Institute of Carnegie Mellon University.[8] The Information Processing Techniques Office of the U.S. Defense Advanced Research Projects Agency funded the creation of CERT in response to the so-called Morris worm.[9] CERT initially focused its efforts on identifying vulnerabilities in the UNIX operating system and

its many derivatives. It has since broadened its mandate to include facilitating the formation and coordination of national, local, or private Computer Security Incident Response Teams (CSIRTs), including the formation of the US-CERT.[10]

A Forum of Incident Response and Security Teams (FIRST)[11] emerged from this early initiative, linking many of the incident-response team systems together in an information sharing, and sometimes incident-response coordinating, community.

National law enforcement agencies around the world have expanded their scope of operation and attention to include cyber-crime and forensic assessment of security incidents that may have inimical origin. The American FBI has set up an extensive cyber-crime response system,[12] as have many other federal government agencies. In addition to the intelligence agencies, the U.S. military has formed a new Cyber Command (USCYBERCOM)[13] as part of the U.S. Strategic Command. Its mission statement reads:

> USCYBERCOM plans, coordinates, integrates, synchronizes, and conducts activities to: direct the operations and defense of specified Department of Defense information networks and; prepare to, and when directed, conduct full-spectrum military cyberspace operations in order to enable actions in all domains, ensure US/Allied freedom of action in cyberspace and deny the same to our adversaries.

In carrying out this mission, the Cyber Command coordinates its efforts across all military departments and other security agencies, including the U.S. Department of Homeland Security, among many others.

A few private firms have attempted to offer *cyber-insurance*: that is, policies that pay off when some terrible cyber-event occurs. Such offerings could potentially play a role in managing risk more effectively. In a mature industry, the insurance company would have a wealth of cases of incursions and other bad events and have both the means and the incentive to identify weaknesses in existing practices. Just as a company might offer fire insurance only if a home security system or smoke alarms are installed, a company might offer cyber-insurance only if a firewall and an intrusion detection system are in operation.

In general, insurance companies would find it financially attractive to discover and encourage various methods of risk management for computer installations and insist that appropriate measures be used either as a requirement for receiving insurance at all, or as a requirement for receiving lower rates.

Note that insurance companies are well placed to determine whether proper procedures have been followed after an event has occurred; thus, costly monitoring of compliance might not be necessary on an ongoing basis. Audit trails of good practice would likely be a natural consequence of thoughtful implementation of good security in business settings.

In this model, insurance companies take on a role as knowledge repositories for risk management techniques, along with an incentive system for ensuring accuracy of their knowledge: they have to pay damages only when they are wrong. Historically, insurance companies operated in this manner in the development of building codes; there is no reason why they could not play a similar role in the future.

At this point, the constraining factor seems to be expertise. Insurance companies, and the actuaries who work there, know very little about the relevant economic risks. There have been noteworthy attempts by economists and computer engineers to share expertise about appro-

priate forms of risk management, but this field is in its infancy.

Even a cursory search of the World Wide Web turns up many examples of national and international efforts to define cybercrime, create response regimes, and develop tools to defend against attacks.

It is worth considering that in other domains of discourse, efforts have been made to minimize or even inhibit the "militarization" of commonly shared resources. For example, Article IV of the 1967 Outer Space Treaty[14] specifically rules out the placement of nuclear or other weapons of mass destruction in space. In his lengthy treatment of this subject,[15] Detlev Wolter of the UN Institute for Disarmament Research sets the stage for further elaboration of international protections against the harmful use of outer space. In a similar vein, the international UN Convention of the Law of the Sea[16] facilitates common agreements on how the world's oceans, and events on and under them, are to be treated by signatories to the convention.

One can readily imagine the potential for a similar convention regarding uses and practices in cyberspace. One might find it necessary to focus at first on the common, publicly accessible Internet because the term *cyberspace* covers much more virtual ground than the Internet alone. Reaching common agreements about unacceptable behaviors or practices on the Internet could form a basis for reciprocal cooperation in the enforcement of national laws or the protection of the world's networked citizens from abuse and harm. This is not to overlook the many concerns about definitions, permitted actions, coordinating mechanisms, and the like that such an endeavor would create.

The ability to detect abuse and make use of forensic tools to identify perpetrators is the second of the three legs of cyber-safety named above. International cooperation seems appropriate and even necessary if we are to achieve any measure of security and safety in our use of computer-based systems.

Apart from defense against abusive practices, multilateral treaties can create a basis for improved electronic commerce and an increasing sense of safety, or at least protection, in the online environment. We might try to agree on the means by which digital signatures can be treated as the legal equivalent of ink signatures on a paper contract. The technology of and the rules by which cyber-certificates are issued to validate identity in cyberspace could enhance consumer, business, and government confidence in cyberspace. Cooperation among law enforcement agencies, financial institutions, and other commerce-enabling entities has the potential to significantly improve cyber-safety and the growth of cyber-transactions.

Finally, one can imagine that an additional element of such multilateral treaties might include commitments for educating the general public, the private sector, and governments at all levels about best practices and behaviors promoting safety.

There is much to be gained through voluntary practices that improve safety in cyberspace, including the use of strong authentication mechanisms, anti-virus practices, good cyber-hygiene, and international cooperation on improving safety and security in cyberspace. Absent from this essay is a discussion of the general problem of software "bugs" and the ways in which they can create unsafe conditions and even have fatal consequences. Much more research is warranted to make software more reliable.

In addition to seeking formal mechanisms and agreements that promote the

general welfare of citizens dependent on the Internet and computer systems, informal cooperation mechanisms among service and software providers also can provide powerful means of response. It is in the realm of law enforcement and diplomacy, however, where formality can enable the protection of cyber-safety. All these tools will be needed if we are to realize the benefits and minimize the hazards of the increasingly complex, powerful, but potentially brittle virtual worlds we have created in cyberspace.[17]

ENDNOTES

1 William Gibson, *Neuromancer* (New York: Ace Books, 1984).

2 See http://www.java.com/en.

3 See http://www.java.com/en/download/faq/java_javascript.xml.

4 See http://www.stopbadware.org.

5 See http://www.ietf.org.

6 See http://www.w3.org.

7 See http://en.wikipedia.org/wiki/Domain_Name_System_Security_Extensions.

8 See http://www.cert.org/meet_cert.

9 See http://www.symantec.com/connect/articles/brief-history-worm.

10 See http://www.us-cert.gov/index.html.

11 See http://www.first.org.

12 See http://www.fbi.gov/about-us/investigate/cyber/cyber.

13 United States Strategic Command, U.S. Cyber Command fact sheet, October 2010, http://www.stratcom.mil/factsheets/Cyber_Command.

14 See http://www.state.gov/www/global/arms/treaties/space1.html.

15 Detlev Wolter, Common Security in Outer Space and International Law (Geneva, Switzerland: United Nations Institute for Disarmament Research, 2006), http://www.unidir.org/pdf/ouvrages/pdf-1-92-9045-177-7-en.pdf.

16 See http://www.un.org/depts/los/index.htm.

17 David Clark, Hal Varian, and Harry Wingo were most helpful in reviewing and recommending changes and additions to earlier drafts of this essay.

Doctrine for Cybersecurity

Deirdre K. Mulligan & Fred B. Schneider

Abstract: A succession of doctrines for enhancing cybersecurity has been advocated in the past, including prevention, risk management, and deterrence through accountability. None has proved effective. Proposals that are now being made view cybersecurity as a public good and adopt mechanisms inspired by those used for public health. This essay discusses the failings of previous doctrines and surveys the landscape of cybersecurity through the lens that a new doctrine, public cybersecurity, provides.

DEIRDRE K. MULLIGAN is an Assistant Professor in the School of Information at the University of California, Berkeley, where she is also a Faculty Director of the Berkeley Center for Law and Technology.

FRED B. SCHNEIDER is the Samuel B. Eckert Professor of Computer Science at Cornell University.

*(*See endnotes for complete contributor biographies.)*

Governments, businesses, and individuals are growing increasingly worried about the security of networked computing systems. This concern is justified. Press reports of successful attacks grow ever more frequent: cross-site scripting used to pilfer consumers' passwords, large-scale breaches of corporate customers' personal information, distributed denial-of-service attacks on websites, cyber espionage aimed at classified documents, and attacks on civil critical infrastructures.

Consequently, computer scientists and their funders are investing heavily in technological means for improving cybersecurity. But technological solutions are useless if they are not deployed or if operating practices allow attackers to circumvent them. Policy must create incentives for system developers, operators, and users to act in ways that enhance rather than weaken system security. Moreover, neither technologists nor policy-makers have the luxury of starting with a clean slate. All must labor in the shadows of legacy networks and end systems that are not secure (nor easily made so) and in the context of extant policy that reflects societal values from a time when dependence on networked information systems was minimal.

Enhanced levels of cybersecurity can create tensions over cost, function, convenience, and societal values such as openness, privacy, freedom of expres-

sion, and innovation. Absent a widely accepted doctrine, evaluation of proposals for improvement is difficult, and debate about their adoption can be neither compelling nor conclusive. The utility of a doctrine is thus determined by the extent to which it offers a framework for resolving these tensions while not imposing, ignoring, or ruling out possible technical or policy solutions.

We thus conclude that a prerequisite for achieving enhanced cybersecurity is articulating a *cybersecurity doctrine*, which specifies *goals* and *means*.

- *Goals* define what system properties will be preserved as well as what policies will be enforced, for whom, at what costs (monetary expenses as well as costs to convenience and compromised societal values), and against what kinds of threats. Goals might be absolute, or they might specify a range of permissible trade-offs. In allowing trade-offs, we acknowledge the political nature of cybersecurity and the need for conversations among those affected when goals are set.
- *Means* might involve technological, educational, and/or regulatory *measures*. We should expect means to include policy that creates incentives – which might range from market-based to coercive – that foster adoption and/or deployment of the measures proposed.

Through incentives provided as part of its means, a cybersecurity doctrine can address barriers to market production of cybersecurity that reflect a lack of will rather than a lack of ability, as others have aptly noted.[1] Incentives can also prompt continued improvement to address the constantly emerging landscape of threats and the new needs that arise as a growing range of applications is being migrated to networked information systems.

Our doctrine of *public cybersecurity*, the subject of this essay, is rooted in the thesis that cybersecurity is a public good. The doctrine focuses on the collective interest rather than on any single individual's or entity's computer, network, or assets. It can be compared with public health, another public good, and we make this comparison below.

We begin by analyzing the limitations of various cybersecurity doctrines that have been proposed. Next, we discuss in detail our new doctrine of public cybersecurity. Then we describe how to support public cybersecurity, starting with approaches to building systems that have fewer vulnerabilities. Subsequent sections explore approaches for managing insecurity: diversity, surveillance, installation of patches, isolation, and the role of intermediaries. Finally, we put this work into a larger perspective and offer some conclusions.

The advent of time-sharing in the 1960s meant that computations on behalf of multiple users were interleaved on a single computer. Each user's computation and data thus had to be protected from misbehavior by programs run by other users. Confronted by a problem born of technology, engineers of early time-sharing systems sought solutions in technology. Therefore, early cybersecurity doctrine focused on developing new technology. Societal values could be and were ignored because the doctrine respected the shared values of the small population that used these early computing systems.

Technological solutions for creating the needed isolation were beyond early capabilities, especially when users could be motivated, capable adversaries bent on disrupting another user's computation or stealing information. Improved technology, however, is not the only way to solve

problems that technology has created, and subsequent cybersecurity doctrines focused on policy to leverage those technological solutions that were at hand. But these doctrines, too, were unsuccessful. Even if they had succeeded, they ultimately would have been inadequate because the problem was changing.

Computer systems were becoming pervasive, which had two broad consequences. First, the information technology sector became a significant economic force, raising concerns about the freedom to innovate and success in the marketplace. Second, computer systems increasingly touched the lives of ordinary people. Citizens' records were stored electronically, and information technology allowed workers to be more efficient. Eventually, computer networks and the Web changed how people shopped, communicated, socialized, and engaged in politics. As a result, privacy and other societal values have become crucial considerations in developing cybersecurity doctrine.

Because their effectiveness has been limited, it is instructive to review the three doctrines – prevention, risk management, and deterrence through accountability – that have dominated cybersecurity thinking for the past fifty years. In particular, analyzing the measures each doctrine proposes offers insights into properties that affect whether a cybersecurity doctrine fails or succeeds.

Doctrine of Prevention. The goal of this doctrine is to render systems completely free of vulnerabilities. Absent vulnerabilities, attacks are not possible, so the system is secure. Such *absolute cybersecurity*, though a worthwhile undertaking, is unlikely ever to be achieved.

To secure systems that incorporate humans as users and operators, we would need some way to prevent social engineering attacks and intentional insider malfeasance. Here, prevention requires overcoming the frailty of humans, which is likely to involve more than technology.

If we ignore human involvement in a system, then the problem is different but no less challenging. Software systems today are too large and complicated to be verified using formal logics. Researchers, assisted by computers, have been able to devise formal proofs for small systems[2] (those with fewer than ten thousand lines of code), and software producers regularly employ automated checking for analyzing specifications as well as for relatively simple properties of large bodies of code. For the reason that smaller code bases are more amenable to formal verification, techniques to reduce the size of the code bases for certain key systems are also being explored.[3] But revolutionary advances are needed before formal verification could be used to validate the entire code base that runs a desktop system or server. Thus, this approach to prevention is not a practical solution for the near term.

System testing is the clear alternative to ensure that a system has no vulnerabilities. Tests, however, can reveal only the presence of vulnerabilities – not their absence. Demonstrating the absence of vulnerabilities requires exhaustive testing; the amount of work involved is prohibitive even for small components, much less for large systems.

Formal proofs and testing are performed relative to some expectations about what the system must do and the environments in which it will operate. In other words, the doctrine of prevention establishes the absence of vulnerabilities only for settings where certain assumptions hold. Unfortunately, reasonable assumptions about the environment today might subsequently be invalidated. Attacks evolve in sophistication in response to better defenses. Threats emerge to exploit new opportunities for disruption

that are created when cyberspace provides access to new forms of value. Therefore, a system that is deemed to be secure might not remain so for long.

In light of this dynamic, expectations about the environment must be periodically revisited and, if necessary, revised. Thus, the doctrine of prevention involves a recurring expense. That expense is inconsistent with the business model employed by many of today's software providers, which favors reuse and extension of existing hardware and software in order to lower the cost of producing new systems.

The adoption of mandatory standards can be viewed as a way to support the doctrine of prevention because implementing standards increases the chances that what is built and/or deployed will have fewer vulnerabilities. Some standards concern functions an artifact must or must not support; some govern its internal structure; others prescribe the process by which the artifact is constructed or maintained; and still others stipulate qualifications the personnel who are involved in creating the artifact must have. Examples include the Department of Defense Trusted Computer System Evaluation Criteria, or TCSEC[4] (also known as the Orange Book); its successor, the Common Criteria for Information Technology Security Evaluation[5]; security provisions in information privacy laws[6]; the Federal Information Security Management Act[7]; and the Voluntary Voting System Guidelines.[8] Current market activity suggests that such mandates show value in some areas. However, a correlation between the absence of vulnerabilities and compliance with standards has not yet been documented. The stated goal for the doctrine of prevention is unlikely to be achieved through these measures.

Doctrine of Risk Management. Absolute cybersecurity is cost prohibitive, but fortunately, it is unnecessary for most systems. The doctrine of risk management stipulates a more modest goal: that investments in security reduce expected losses from attacks. To adopt this doctrine is to admit that all vulnerabilities are not equal, that one should focus only on vulnerabilities whose exploitation (i) is sufficiently likely to occur based on perceived threats and (ii) could enable expensive (by some cost measure) system compromises. In contrast to the doctrine of prevention, the objective of defending a smaller body of code against a more restricted set of threats is likely within our capabilities. Moreover, maintaining that steady state would require fewer assumptions about the environment to be revisited periodically.

In theory, the doctrine of risk management seems sensible. But a lack of information about vulnerabilities, incidents, and attendant losses makes actual risk calculations difficult. Companies and individuals do not know how to value (i) confidentiality of information, (ii) integrity of information, or (iii) the pain of dealing with recovery from an attack's effects (bad credit ratings, for example). These costs appear to vary tremendously. Also, people have difficulty reasoning about extremely low-probability events. Finally, when costs are borne by third parties, investment incentives materialize only well after a breach occurs, and causation is difficult to discern, much less prove.

Accurate information about threats and attacks may not be publicly available because those with that knowledge fear tarnishing their reputations or compromising their intelligence methods and/or sources. Even were that information accessible, deployment of replacement systems and upgrades would alter the functions that systems perform and the set of relevant vulnerabilities, which, in turn,

could lead to new attacks. These differences mean that the past is not a good predictor of the future. As a consequence, actuarial models cannot be constructed, and insurance to transfer risk is impossible to price in a way that ensures profits to the policy underwriter.[9]

Were there a method to analyze a system mechanically and obtain a quantity that indicates just how secure that system is, we could form a basis for assessing what is gained from specific investments in cybersecurity. At present, such cybersecurity metrics do not exist. Neither can investments be justified by quantities derived entirely from empirical observations. The absence of detected system compromises could indicate that investments in defenses worked, that attacks have not been attempted, or that the compromise escaped notice (as in theft of confidential information, for example). Whether prior security investments were well targeted is impossible to know; such ambiguity leaves security professionals to justify investments based solely on non-events.

Risk management approaches are further confounded by externalities that arise from the emergent nature of cybersecurity in networks. Individuals and entities employing these tactics can neither fully reap the benefit of their security investment nor entirely control their vulnerability through investments.[10] For example, a single compromised system anywhere in a network can serve as a launch point for attacks on other systems connected to that network. Thus, local investment in defenses not only provides local benefits but also benefits others; likewise, underinvestment in defenses elsewhere in the network could facilitate local harms. Absent coordination, the sole logical strategy would be to invest in insurance (if it were available); only with insurance can an entity reap the entire benefit of its investment.[11] However, this strategy does nothing to improve security, and as noted above, viable long-term business models for insurance do not exist today.

The outlook for risk management is not entirely bleak. In the policy arena, state security breach notification laws[12] are a form of risk management intervention. Significant costs are incurred to notify individuals and to manage the adverse publicity surrounding reportable breaches. These potential costs act as a proxy for the costs of security failures to customers, forcing companies to internalize previously externalized security failures. The price tag on breaches also means that these laws have created a set of data to use in calculating risk and return on investments. Nonetheless, current laws focus on only a narrow set of breaches and, as a result, might artificially skew investments.

Doctrine of Deterrence through Accountability. This doctrine treats attacks as crimes; it focuses on infrastructure to perform forensics, identify perpetrators, and prosecute them. In theory, attacks are deterred by increasing the chances that perpetrators will be found and convicted.[13] Implementations of this doctrine require strong authentication technologies and surveillance of network activity. Robust forms of user identity would allow us to overcome the loose binding that exists today between individuals and machines.[14]

Absent an effective means for retribution, this doctrine has no teeth, and fails as a result. Moreover, punishment of perpetrators of cyber-attacks is not always feasible in today's global environment. Attribution of actions by machines to individuals is complicated, agreement about illegality is illusive, and cross-border enforcement requires more cooperation than is likely to emerge between nations. Recent attacks against U.S. and other systems suggest that we cannot ignore

non–nation-state actors that engage in terrorism and large-scale financial crimes. The very features that make the Internet a profitable environment for criminals – worldwide reach, connectedness, neutral treatment of packets, and weak binding of machines to individuals – make it difficult for law enforcement to identify and catch perpetrators. Other features of the international landscape complicate efforts to bring them to justice.[15]

Conceptual obstacles also limit the effectiveness of the doctrine. First, the doctrine is punitive. Like most criminal law, it is aimed primarily at using punishment to produce both general and specific deterrence. This approach does little to keep networks up and running when they are under siege, nor does it prompt proactive security investments. Second, the doctrine could require individuals to sacrifice privacy and, in the extreme case, abandon the possibility of anonymity and the protections for freedom of speech and association that it affords.

Nevertheless, many attacks are indeed carried out by criminals plying their trade. We are likely to benefit if criminal activity in cyberspace faces the risk of retribution that we employ to deter crime in the physical world. Thus, the doctrine of deterrence through accountability has value. But in cyberspace, unlike in the physical world, terrorists or state actors are difficult to distinguish from common criminals.[16] Deterrence through accountability is not necessarily effective against these transnational threats; other doctrine is also required.

Cybersecurity is non-rivalrous and non-excludable; by definition, therefore, it is a *public good*. It is non-rivalrous because one user's capacity to benefit from the security of a networked system does not diminish the ability of any other user to enjoy the same advantage. It is non-excludable because users of a secure system cannot easily be excluded from the benefits security brings. Measures intended to foster the production of public goods thus constitute a sensible starting point in our search for doctrines that promote cybersecurity.

Economists define a *common good* as one that is rivalrous and non-excludable. The sea, outer space, and air are examples. Insofar as common goods are inherently different from public goods, doctrines for common goods are likely unsuitable for enhancing cybersecurity. Indeed, this irrelevance is apparent in laws for protecting common goods, which typically aim to ensure rights of equal use, and in the mechanisms these laws introduce, which are intended to manage the depletion and inequitable consumption by first-comers or more sophisticated users. The production of cybersecurity bears little relation to these issues.

Public health – the prevention of disease and promotion of good health in populations writ large – is a public good. It is non-rivalrous because having a healthy population implies a lower prevalence of disease, which in turn decreases the chances any member will fall ill. It is non-excludable because no one can limit an individual's ability to profit from the health benefits that living among a healthy population brings.

The essential characteristics of public health law are a focus on the health of the population as a whole and the singular role of governments in that enterprise.[17] To discharge these responsibilities, the law authorizes various agencies to engage in a broad set of activities, including:

- Public education about the causes and effects of disease, as well as methods of prevention. Education empowers individuals to act in ways that optimize their own health, which in turn furthers public health.

- Creation and use of methods for the prevention and treatment of specific diseases. Methods could involve (i) providing subsidies to procure care needed by those who could not otherwise afford it[18] or (ii) imposing specific health standards as eligibility requirements for receiving various societal benefits (for example, public education, which we discuss below).
- Identification and management of disease and infected individuals though surveillance, information gathering, and analysis. Methods include mandatory reporting requirements for certain diseases and conditions, mandatory testing or screening for others, symptom surveillance that seeks to identify obscure public health threats in masses of routine records, and mandatory treatment.

The interests of individuals and the public often align; public health law speaks to the points of conflict. It offers frameworks to mediate tensions between the rights of individuals as sovereigns over their physical bodies and the obligation of the state to protect the population as a whole.

For example, public health mandates that children be vaccinated because, in a generally healthy population, such vaccinations cannot be justified based on the benefit to the individual. In fact, the optimal choice for any given child might be to avoid vaccination and thus avoid the risk of side effects. Mandatory vaccination creates *herd immunity*, which benefits the collective by reducing the total number of hosts available to carry a disease, thereby decreasing the risk to individuals who have not been vaccinated. However, if too many individuals act in self-interest and eschew vaccinations, then the herd immunity that gives some protection to the unvaccinated may disappear. This is a "tragedy of the commons" whereby individuals acting rationally leave everyone worse off.

Every state in the United States conditions a child's attendance at school on satisfying some specified regimen of vaccinations. In addition, vaccine manufacturers are indemnified from liability for side effects users might experience. Specifically, the Vaccine Injury Compensation Program (VICP) provides certain payouts from public coffers to children injured as a result of a vaccine. VICP also offers a compensation mechanism outside seeking damages from the vaccine manufacturers, thereby establishing an environment conducive to both the production and willing use of vaccines. While the program cannot fully compensate for negative health consequences from vaccinations, it is an important component of the overall public health strategy.

Public health is a logical outgrowth of disease detection and prevention mechanisms, which transformed societal perception of health from a primarily private concern to a concern of the collective. Ultimately, this development led to the perception of public health as a public good that the government should enable.[19] Health now becomes intertwined with societal values. For example, we impinge on societal values when we introduce mandatory reporting and surveillance systems that (i) alert individuals at specific risk so they can be tested and treated and (ii) allow isolation, quarantine, and even mandatory treatment to be imposed. At the same time, public health interventions aim to minimize their intrusiveness because of the chilling effect that may have on access to health care; for example, we see anonymous HIV testing and needle exchange.

This public health framework (of laws, agencies, and measures) applies equally well to weaponized pathogens. This is not

to suggest that motive is irrelevant in considering public health strategies. For example, weaponized pathogens may change more quickly than those that evolve in nature, and certainly, the transmission vectors may differ when pathogens are used as weapons. But the basic tools of public health – public education (to minimize exposure and facilitate early detection), investments to create means for prevention and treatment (antidotes and vaccines), and surveillance and analysis (facilitating isolation and quarantine as defenses) – still apply.

Both public health and cybersecurity aim to achieve a positive state (health or security) in a loosely affiliated but highly interdependent network. The former is a network comprised primarily of people existing in an environment over which they have some limited control; the latter is a network of people, software, and hardware (for communications, storage, and processing). Given that the positive state is ultimately unachievable, both struggle with how to manage in its absence as well as with how to work toward attaining it. Success ultimately depends not only on technical progress but on reaching a political agreement about (i) the relative value of a public good in comparison to other societal values and (ii) the institutions granted authority to resolve conflicts (and the methods they use).

We define a *doctrine of public cybersecurity* to be any cybersecurity doctrine whose goals are (i) to produce cybersecurity and (ii) to manage insecurity[20] that remains, where political agreement balances individual rights and public welfare. There is no single doctrine of public cybersecurity, for the reason that there are different meanings attached to "cybersecurity" and "insecurity."[21] Also, different choices of measures and incentives result in different doctrines of public cybersecurity. Notice, though, that none of the doctrines discussed above has all the elements we require for a doctrine of public cybersecurity.[22]

The analogy to public health inspires cybersecurity measures such as prevention, containment, mitigation, and recovery – that is, strategies that direct resources toward production and preservation of cybersecurity. But modern public health doctrine does not compensate victims of disease; thus, a parallel doctrine of public cybersecurity would not focus on restitution. Indeed, restitution is economically efficient only when attacks are infrequent, and that assumption cannot realistically be made today.

Furthermore, modern public health does not punish victims of disease, but there is some nuance. Using quarantine to limit the spread of disease benefits the collective by depriving an individual of certain freedoms. Such a response could be considered a "harsh consequence," which is one definition of "punishment." By analogy, a doctrine of public cybersecurity could dictate responses that deprive individuals of actions, but only if those responses benefit the collective. Punishments solely for retribution could not be part of a public cybersecurity doctrine (because retribution does not benefit public welfare); however, nothing precludes implementing a doctrine of public cybersecurity alongside a cybersecurity doctrine that incorporates retribution. Finally, the parallel with public health also suggests that prevention be preferred to recovery.

With regard to incentives, ensuring that actors contribute to public cybersecurity requires interventions to overcome positive and negative externalities that lead rational individuals to underinvest, as occurs in public health. When incentives are insufficient to motivate private provisioning, the public interest re-

quires making value-ridden choices to interfere with the rights and interests of individuals and organizations. Those choices are embodied in goals that reflect political agreement about how to define the good in question; the socially desirable level that should be maintained, given competing priorities and values; and provisions for determining when the individual's desires yield to the collective's need. For example, an agreement might stipulate that state coercion is permitted only when certain incursions into the rights and interests of individuals are tightly circumscribed.

Public health solutions do not always translate into sensible support for public cybersecurity, but the former often inspires strategies for the latter. The examples we explore below illustrate how doctrines of public cybersecurity can be useful for evaluating current cybersecurity proposals. Our choice of examples should not, however, be seen as an endorsement for any particular proposed set of interventions.

Underutilized approaches (formal methods, testing, and improved software engineering processes and standards, for example), developed in part to serve the doctrine of prevention, are effective in producing cybersecurity (by reducing the number of vulnerabilities present in a system), even if they cannot produce absolute cybersecurity. Thus, existing methods could serve as a means for a doctrine of public cybersecurity, just as disease prevention through vaccination and the monitoring of our food and water supplies fosters public health. The question is: What incentive structures would ensure that these methods are used?

Education could play a key role in defect reduction. Knowledgeable developers are less likely to build systems that have vulnerabilities. They are also better able, and thus more likely, to embrace leading-edge preventions and mitigations. There is, however, no agreement about what should be taught. Reaching such a consensus would require a dialogue among universities and practitioners.

Were there an agreed-upon body of knowledge for cybersecurity practitioners, mandatory certification could ensure that practitioners master that material as a condition for practice. But the details of how certification is handled can be subtle. Possession of a certificate does not by itself compel the use of best practices, and it is easy to imagine certified system-builders who cut corners by choice (out of laziness, for example) or by mandate (because management is trying to reduce costs). Moreover, unless the certification process imposes a continuing education requirement to ensure that certificate holders stay current with new developments, it might impede rather than promote the spread of innovation. Even when continuing education is mandated, old habits die hard; for example, physicians who have been shown new methods that are empirically demonstrated to be superior nevertheless tend to stick with familiar practices.[23]

Utilizing techniques to reduce defects during system development and employing better-educated practitioners will mean that systems become more expensive to produce. Today's software-procurement market does not provide developers with compelling incentives to incur those additional expenses. Moreover, purchasers are unable to predict the costs of a system's vulnerability to attack and, without ways to measure a system's security, cannot rationalize paying higher prices. The doctrine of risk management failed for the same underlying reasons.

Law could force system producers and/or purchasers to make the necessary investments. Software distributors cur-

rently disclaim liability beyond the purchase price for damages caused by their products. This practice probably reduces the time and energy that developers devote to eliminating defects, as evidenced by the number of buffer overruns and other exploitable coding errors still being discovered and exploited by attackers. Existing law could, for example, be revised to disallow limits on damages flowing from attacks taking advantage of poor coding practices that lead to buffer overflows and other easily exploited vulnerabilities. Limits on liability could depend on the use of formal methods, type-safe languages, or specific forms of testing (such as fuzz testing[24]). Creation of a class of certified security professionals could also provide the basis for a professional duty-of-care supporting liability for shoddy security.

Furthermore, law could require that software developers adhere to security standards. Alternatively, safe harbor provisions could be created to protect software developers against future findings of liability for those systems built according to specified standards. In fact, the law arguably already mandates that companies follow certain standards regarding personally identifiable information. Through a series of settlement agreements, the Federal Trade Commission established a de facto standard that requires a company collecting and handling the personal information of consumers (i) to establish reasonable security processes and (ii) to mitigate system vulnerabilities that are known in the marketplace and for which mitigations exist. A first step in determining whether law should more broadly mandate the adoption of security standards might be research that identifies connections between security development processes and positive security outcomes.[25]

Monocultures in nature risk extinction from pathogens and are less able to adapt to changing conditions. *Diversity* – of the individuals within each species and by virtue of many species coexisting within an ecosystem – creates a resilient ecosystem. By extension, public health benefits from individuals in a population having different inherent resistance to pathogens and, by virtue of different exposures[26] to diseases, having different immunities.

Although nature abhors monocultures, cyberspace seems to favor them. A collection of identical computing platforms is easier, and hence cheaper, to manage because it demands that users master only one interface and managers make only one set of configuration decisions. In addition, user-training costs are reduced when job transfers do not have the overhead of learning another operating system and suite of applications; in a monoculture, investments in educating system users or managers can be amortized over a larger user base. Finally, a monoculture facilitates networking: interoperability of a few different kinds of systems is far easier to orchestrate than integrating a diverse collection, standards notwithstanding. Mindful of these advantages, both the public and private sectors tend to adopt procurement policies that foster creating computer monocultures.[27]

Methods exist, however, for artificially and automatically creating diversity in software systems without sacrificing the advantages a monoculture provides. These methods involve tools that randomly transform code and/or stored information while preserving its semantics. Once such *artificial diversity* is introduced, internal details of an individual system are no longer predictable. Thus, an attack that depends on knowledge of internal details is more likely, after a small number of instructions, to cause a system crash than to give an attacker con-

trol of that system. In many settings, a system crash is preferable to attacker control. Moreover, a platform that crashes in response to an attack cannot then help spread that attack to other platforms. By (implicitly) signaling to system operators that something is wrong, a crash also creates an opportunity for initiating other means to block an attack's spread.

Like the diversity found in nature, artificial diversity is inherently a probabilistic defense. An attack against any one component might not be derailed by the specific random transformations that were made to that component. Also, by converting some attacks into crashes, artificial diversity can adversely affect a system's availability.

Despite these limitations, artificial diversity facilitates public cybersecurity by providing a means to cope with residual vulnerabilities, thereby supplying a way to manage insecurity. Today, artificial diversity is used often in operating systems but less so in applications[28] (even though, increasingly, it is applications that attackers target). However, the various legal approaches (discussed above) for incentivizing defect reduction during development are equally well suited for incentivizing system producers to support artificial diversity. There is no shortage of incentives at hand for encouraging broader adoption of the measure.

Public health relies extensively on *surveillance*. Data collected through a variety of means enable disease containment and mitigation through:

- dissemination of information that facilitates individual actions;
- isolation and quarantine, which limit the interaction of affected individuals with the rest of the population to avoid exposure to infection; and
- mandatory treatment to reduce danger to the public.[29]

Data collection for public health occurs at many levels. At the lowest level is the inclination of individuals to assess their own well-being. Education equips individuals with a basic level of knowledge about health indicators – normal body temperature, pulse, blood pressure, and respiratory rate – as well as with simple precautions to limit infection and the spread of disease (frequent hand washing, for example). Primary care providers collect other data in conjunction with annual check-ups and, when symptoms require further analysis, at hospitals and other more advanced diagnostic facilities. Each successively higher level is concerned with the overall health of a larger population and thus provides a natural venue for constructing and analyzing larger data aggregations.

By minimizing disclosure of information about an individual's health, public health law strives to reduce one potential deterrent to seeking health care: that is, an individual's fear of being shunned because of a publicized health condition. In general, identifying information should flow away from primary health care providers only in instances where aggregation and/or analysis is necessary to identify significant trends. Even in this case, efforts are undertaken to protect individuals' privacy.

In contrast to public health, cybersecurity is not supported today by extensive coordinated surveillance, yet it would be feasible and advantageous to do so. Low-level indicators about the basic "health" of a computer can be made available by running built-in checking software (such as virus scanners and intrusion detection systems). Each of the Internet Service Providers (ISPs) that constitutes the network has an infrastructure that facilitates

monitoring of events internal to its network as well as interactions with other networks.

Surveillance of network traffic (including volume, distribution over time, and destinations) could be a powerful potential source of information about certain attacks and vulnerabilities. Denial-of-service attacks, for example, have a clear manifestation and a natural mitigation based on traffic filtering by ISPs. However, the source(s) of such attack packets, the target(s), and the intermediaries are likely to span multiple ISPs, which would have to share data and coordinate for mitigation. Unfortunately, data sharing among ISPs today is inhibited by competition and, in some cases, varied interpretations of privacy law.[30] ISPs thus do not always have the situational awareness that would enable them to suppress packets delivering attacks. Widespread sharing of information, however, can introduce a risk by increasing chances that attackers learn about vulnerabilities for specific sites.

Just as there are privacy issues with collecting data about an individual's health, network traffic surveillance raises privacy concerns. The extent to which collecting packets actually impinges on privacy depends on what information is recorded, how long it is stored, how it is used, and who can access the information. For example, real-time responses to protect networks can be accomplished by authenticating machines, a far less politically fraught solution than proposals for "Internet drivers' licenses" and other tight bindings between machines and individuals.[31]

ISP cooperation and information sharing is less likely to raise privacy concerns than the collection of information by centralized government organizations. Yet given that defense of its citizens is a clear responsibility of government, seeing the packets themselves can be invaluable to a government seeking situational awareness about threats in cyberspace. Unfortunately, packet inspection is also easily abused if a government intends to spy on citizens; critics cite this fear (among others) when discussing the Einstein[32] systems recently deployed by the U.S. government for monitoring connections to the Internet at federal civilian agencies. As with public health, political agreement must weigh the expected benefits of surveillance (backed by sound research and field experience) against the risks it poses to other values.

An understanding of the kinds of vulnerabilities found in systems is a form of situational awareness of potentially great value to system builders. In the absence of mandatory reporting requirements for cybersecurity incidents, diverse public and private reporting mechanisms have evolved. The U.S. Department of Homeland Security's Computer Emergency Readiness Team (US-CERT) and the National Institute of Standards and Technology (NIST) Computer Security Division maintain databases of common vulnerabilities. Many organizations contribute, but these are not the only such databases and none provides more than a partial view. Some vulnerabilities never reach the public databases. For example, a private-sector community of "security researchers" report their findings on system vulnerabilities to middlemen, who offer the information for sale to companies that build and sell anti-malware or intrusion prevention/detection products.[33] Yet the ad hoc Conficker Working Group[34] is an example of a rather successful coordinated private-sector activity involving information sharing about risks.

A *patch* is an update that can be applied to an installed system in order to eliminate one or more previously identified vulnerabilities. Exploitation of an unpatched

vulnerability on a computer could target that machine, and an individual's assets contained therein, and therefore be fully internalized as a result. Alternatively, the exploitation could target the machines and systems of others, producing a negative externality.[35] The uncertainty about consequences means that self-interest is not a strong incentive for machine owners to apply patches.

Various policy interventions could raise patch rates. Choosing among them requires additional information about why people and businesses delay or outright fail to apply patches.

- Research might conclude that low patch rates in the consumer market are caused by an underappreciation of the risks. Public education to inform individuals that applying patches improves cybersecurity might dramatically increase patching.
- We might find that individuals lack awareness of vulnerabilities present on their machines. Here, built-in software to check whether all current patches have been applied might suffice for triggering consumers to be more attentive to downloading patches for their machines.
- Feedback about what others do could create new behavioral norms that might lead to better patching practices. Researchers in other areas have found that showing individuals how their behavior compares to others' taps into competitive and/or social consciousness. Simply stating that a significant percentage of others have patched their machines, and are thus doing their part for cybersecurity, might push laggards into applying patches.
- Research might find that individuals or enterprises hesitate to install patches for fear of destabilizing their other software. Greater transparency about the specific configurations and applications software vendors have tested might help individuals overcome their reluctance. The fears of enterprises that depend on homegrown software might be somewhat assuaged by providing test suites to patch developers.[36] As a final safety net, all software could be required to contain mechanisms whereby a patch that has been applied can easily be removed and the system and data restored to the pre-patch state.
- If the impediment to installing security patches is time or expertise, then vendors could mitigate the problem by configuring defaults that automatically download and apply security patches.
- Another reason consumers forgo installing patches could be that they are charged for Internet access in proportion to the amount of bandwidth they use and will incur lower costs by not downloading patches. Here, one solution is to subsidize the bandwidth required for such downloads; another is to introduce tariffs that distinguish between different kinds of traffic.

Those who run pirated software might hesitate to install patches for fear that the installation process would disable the illegal software or detect and report it. Regulations could address this obstacle by prohibiting security-patch installation from implementing functionality in support of license enforcement or any other form of intellectual property protection. Thus, even pirated software could be patched, so that herd immunity can be achieved.

Incentives to apply patches could also have a useful indirect effect. If patch installations are frequent and disruptive, then consumers have reason to prefer products with fewer security vulnerabilities. Consumer demand would then pres-

sure software producers to build and deploy more secure products.

Mandates to apply patches raise concerns about subsidizing the installation of patches[37] and compensating injured parties when patches cause harm. Losses from applying mandated patches, particularly where unacknowledged and uncompensated, will breed suspicion and resistance to patching efforts. Thus, it seems advisable to consider backstop measures, analogous to what VICP provides to incentivize the use and production of vaccines and the process used by the Food and Drug Administration to ensure vaccine efficacy.

Geological features, such as mountains and oceans, have proved valuable in protecting individuals and populations. When natural boundaries are absent, we build our own: fences surround buildings and nations, often with guards to control who is allowed to transit the border. Such boundaries protect activities on one side from activities occurring on the other. A boundary may limit travel in one direction or in both directions; it may be entirely impervious or may selectively limit who or what may pass. Neither is a panacea: an impervious boundary could bar the good with the bad; a selective boundary must employ some kind of *filter*, and that filter might block what should not be blocked or let pass what should.

Firewalls, so-called network guards, intrusion detection/prevention systems, and "air gaps" are examples of mechanisms that implement boundaries in networked systems.[38] Data collected through surveillance can serve as the basis for *signatures*, which are then used to define filters, effectively creating dynamic boundaries. Surveillance thus can lead to automatically imposed quarantines. Given that attacks in networks propagate rapidly, automatic response is especially attractive.

Ideally, we would deploy selective boundaries that block only attacks. In practice, though, filters will be far from perfect.

- Filters that inspect packet payloads (known as *deep-packet inspection*) in addition to checking packet headers are ineffective when packet payloads are encrypted or otherwise obfuscated. Encryption is not used extensively in networks today, but that could easily change. Attackers often use encryption to evade detection. Moreover, what is being spread are often malware variants, where each variant is obfuscated by the application of a different random set of semantics-preserving transformations. It is difficult and often impossible to construct a signature that matches all variants by generalizing from a few.
- A filter might be designed either to (i) block packets and protocols corresponding to known attacks or (ii) pass packets from protocols or conveying content that is known to be normal. Filters that implement (i) are fooled by new attacks (in addition to suffering limitations described above) and those that implement (ii) could block previously unseen protocols and kinds of packets, thereby stifling innovation.
- Whether a packet is part of an attack could depend only on sender intent. Consider a large number of request packets being sent to a Web server. Are many people trying to access the same particularly topical content, or is a denial-of-service attack in progress? Sender intent is the sole differentiator.

There is also a human element to consider. Boundaries and filters must be installed, configured, and managed by human operators, and people make mistakes. Moreover, when such a mistake allows

unimpeded flow, then the error might be difficult to detect until it is too late.

Network providers are understandably reluctant to publicize details of defenses, because revealing that information could help attackers. Yet we see example defenses in today's commercial networks, which create and reference "black lists" of sites whose communications will be ignored and "white lists" of sites that are known to be trustworthy. Some ISPs create a competitive advantage by offering their customers a service whereby suspicious inbound-traffic spikes directed at the customer's site will automatically prompt upstream filtering to block those suspicious packets. As a result, denial-of-service attacks in such networks are more difficult to undertake. Other ISPs monitor each endpoint, disconnecting a given endpoint if outgoing traffic suggests that the endpoint is compromised.[39]

A boundary may be deployed around a system (be it a single computer or a network) that must be protected from attacks, or around a system that is likely to harbor attackers. Different incentives are effective in each case. One natural scenario for direct government investment exists when security boundaries and national ones overlap. Systems in various countries are subject to different laws, typically reflecting a range of societal values. A government might therefore justify installing a boundary whenever systems subject to its laws are connected to systems located in a jurisdiction that allows system behavior the first considers an attack.

Boundaries are more likely to be accepted and work effectively when initiated by the collective rather than by individuals. First, an individual is unlikely to have the necessary authority to mandate changes to defenses on all the remote systems that could be involved in creating a quarantine. Second, the possibility of free-loading limits the incentives for owners or operators of networks or individual systems to make the investments to support enforced isolation. Finally, an agent of the collective, equipped with a broader view of system vulnerabilities, would define better signatures for filters.

An example of such boundaries is found in recent proposals for deterrence through accountability. Some have suggested that the Internet be partitioned into national or multinational enclaves. Those enclaves that serve the population whose network security is of concern (i) run protocols that enable packet-sender tracing and (ii) do not carry traffic from enclaves where packet-sender tracing is not supported or cannot be trusted. The ability to trace attack packets back to an individual machine enables support for accountability in those enclaves that serve the populations the boundary is intended to protect.

Boundaries with sufficiently powerful filters have the potential to intrude on societal values. One concern arises when the defining filters not only block packets that contain attacks but can be configured to block other kinds of packets. Such a filter could be used to prevent data from leaving an enclave, which makes it well suited for protecting confidential information against theft. But content filters also permit government censorship, as illustrated by the firewalls China has installed to protect that nation's computing systems from receiving information in violation of local laws regarding allowed speech. Deployed in the reverse direction, a content filter could block someone from sharing information with others, thereby stifling debate.

So there are trade-offs, with social values and potential benefits for the collective requiring constraints on activities by individuals and businesses. Moreover, no criteria for deciding where a system should be segregated will be infallible.

The result is a complex risk-management decision procedure that society must prescribe, with imperfect information and unknowable consequences.

The public health system leverages health professionals and other institutions to influence individuals' behavior. For example, health professionals educate individuals about the benefits of vaccinations, schools demand conformance with vaccination schedules, and airports screen passengers for symptoms during some infectious disease outbreaks. Intermediaries clearly play an important role in public health strategies.

Intermediaries also have an important part in fostering cybersecurity. For example, many network operators, such as employers and universities, require that all machines on their networks run virus detectors or malware detectors (with up-to-date signature files). These intermediaries could require that all machines are up-to-date on security patches. Similarly, some ISPs have chosen to notify subscribers when a computer appears to be infected.[40] At least one ISP restricts Web surfing until the infected machine is cleaned up, while another ISP reportedly quarantines any compromised machine until it is clean.[41]

ISPs are well positioned to facilitate patching and, by monitoring traffic, to enforce isolation of machines harboring certain malware. Yet they currently have little incentive to engage in such practices because they would then incur the bulk of the security costs, but any costs from infected machines would be more widely dispersed. Moreover, an ISP that disables or limits a machine's access to the Internet will likely bear the burden of assisting that customer as she attempts the necessary repairs. Analysis[42] suggests that the cost incurred by an ISP in fielding a customer's tech-support call approaches the ISP's annual revenue from that customer. Making this sort of monitoring and clean-up a mandatory obligation for ISPs would not only force action but would also prevent consumers from contracting with ISPs that enforce weaker security requirements.

More daunting are the potential costs an ISP might incur from making an incorrect decision to disconnect a customer.[43] To limit spam email, for example, an ISP might block all bulk sending of email. But missives sent by a political organization might then be blocked, resulting in unwanted attention from advocacy groups and the press.[44] While the law is evolving to provide ISPs that take steps to protect security with immunity from suits brought by providers of malware, those users who experience losses after installing required patches or system upgrades, or who suffer because of isolation, might also file legal complaints. In sum, the costs of ISP intervention present a formidable barrier to such action; nonetheless, the law could remove these disincentives.

In a recent proposal,[45] legal scholars Doug Lichtman and Eric Posner argue that expanding ISPs' liability "for violations of cyber-security" would improve cybersecurity because (i) individual attackers are often either beyond the reach of the law or are judgment-proof and (ii) ISPs "can detect, deter, or otherwise influence the bad acts in question." Similar to the way cardholder purchases are monitored by credit card companies, ISPs could detect cybersecurity violations by building profiles of their users and looking for traffic anomalies. But as Lichtman and Posner openly admit, anomaly detection with usable levels of fidelity has eluded cybersecurity researchers for decades. Thus, implementing the proposal is not feasible at present.

Still, a policy holding ISPs liable for the damage caused by infected machines running on their networks might encourage

more diligence in monitoring and fixing their subscribers' machines. The details are subtle and depend on the standard for liability – whether strict, knowledge-based, or otherwise defined. To complicate matters, the policy could have an undesirable outcome: the ISP could undertake less monitoring as a way to avoid its duty to intervene.

Alternatively, governments could provide indirect or direct subsidies to foster cybersecurity-preserving activities by ISPs. For example, creating a centralized service for hosting patches or subsidizing bandwidth to all endpoints could ensure that cost or delay to download a patch would not become an impediment to installing that patch.

Given the decentralized and private provisioning of network resources in the United States and many other countries, understanding the role of intermediaries in driving cybersecurity is essential. As in other areas, such as copyright, the challenge is to establish policies that incentivize desirable behavior while minimizing impact on other values.

Computer scientists have discussed a biological basis for cybersecurity for at least two decades. The thrust of that research is to understand whether computer networks can benefit from implementing defenses similar to those that protect living things. Developers have explored intrusion detection systems that mimic pathogen detection in the human immune system[46] and software defenses based on artificial diversity.[47] A recent Department of Homeland Security white paper[48] describes how a human immune system's response mechanisms might serve as the blueprint for software that defends individual computers and networks against cyber-attacks. Much research remains to be done, however, before those ideas are reduced to running code.

In contrast to the biological metaphor, which focuses on technical measures for blocking cyber-attacks, the analogy between public health and cybersecurity is primarily concerned with new policy and new institutions. Proposals for a Cyber-CDC, for example, have attracted considerable interest.[49] Inspired by the existing Centers for Disease Control and Prevention, the Cyber-CDC is envisioned as a government institution that organizes public- and private-sector strategies to enhance cybersecurity. It would also undertake data collection about threats and attacks, analyze and disseminate that information (perhaps in partnership with the private sector), serve as a repository for technical remedies, and educate the public about best practices, defenses, and remedies.

An IBM white paper[50] broadens the analogy. Borrowing not only from public health but also from public safety, the paper recommends establishing a Cyber Federal Emergency Management Agency and devising a Cyber National Response Framework. Independently, Microsoft's Corporate Vice President for Trustworthy Computing, Scott Charney, has advocated measuring "device health," with device "health certificates" serving as a basis for authorizing device access to network resources.[51] Rather than focusing on institutions, cybersecurity expert Jeffrey Hunker looks to public health as a model for behavioral norms.[52] Individuals would be expected to satisfy certain norms, and government institutions would focus on supporting those norms.

None of the aforementioned work includes a compelling argument for why the analogy to public health is a suitable starting point for a cybersecurity doctrine. Public health informs people's behaviors (seemingly an obvious route to enhanced cybersecurity), but so does religion (which nobody is advocating as a

cybersecurity solution). In formulating our doctrine of public cybersecurity, we use economic theory to justify the shared status of public health and cybersecurity as public goods because economics explains the externalities and incentives that arise in cybersecurity. Viewing cybersecurity as a public good is not new,[53] but we do appear to be the first to employ insights from economic theory to justify the public health model for cybersecurity.

Our public cybersecurity doctrine goes beyond prior work that explores cybersecurity counterparts for institutions and policies that have served public health well. Public cybersecurity is obtained by identifying cybersecurity counterparts to the *goals* of public health – not the *institutions* of public health. First, public health law provides a powerful framework for balancing collective versus individual interests. Second, just as managing disease is an important goal of public health, managing insecurity is an important goal of public cybersecurity. The siren call for the production of "secure" systems and networks must be – and, with public cybersecurity, is – augmented with a mandate to manage the inevitable insecurity that comes from the constant vulnerabilities and adversaries that networked systems face.

The goals of public cybersecurity focus on the collective. Individual high-consequence systems, such as those that control critical infrastructures, are not singled out. Why not focus on the seemingly smaller problem of making only the high-consequence systems secure? For the same reason that the public health system does not focus on keeping only "important" people healthy, isolation is not a realistic proposition for cybersecurity. Public health teaches that it is easier to keep specific individuals healthy when everyone is healthy. The same is true with cybersecurity. If we foster the production of cybersecurity generally, building our networks' capacity to manage insecurity, we will be better able to ensure that our high-consequence systems are secure.

Cybersecurity, like security in so many other contexts, involves trade-offs with other values.[54] Conflicts will arise between public cybersecurity and the interests of specific individuals, entities, and society at large. A cybersecurity doctrine is obliged to provide principles and processes to negotiate and resolve these conflicts. Public health already offers such guidelines to benefit the public good.

First, the state intervenes most drastically when an individual's health decision might directly impact the health of others. The state is generally unable to coerce an individual's decision when the health of only the individual is implicated. Substitute "health" for "security," and we have sensible guidelines for public cybersecurity.

Second, public health guidance, applied to managing the externalities associated with public cybersecurity, suggests the following:

- The state's obligations and abilities to shape and override private choices should turn on the extent to which they have a direct impact on the security of the broader public rather than the security of an individual or entity.
- To facilitate better decisions by individuals, the state should provide information or gentle interventions that influence the perception of risk but that do not supplant the decision-making.
- Where security choices of the individual will impact the security of others, the state should use a wider array of tools to alter behavior.
- Even where state action is permissible, impact on other societal values must be considered in choosing among solutions.

Doctrine for Cybersecurity

- Whenever possible, the state should opt for minimal interventions implemented in a decentralized manner, so as to limit the negative impact they may have on willingness to participate.

Inadequate cybersecurity is the obstacle to success in the information age. Though the problem resides in technologies, the solution involves policies. It requires intervention in the private choices of individuals, hard trade-offs, and political agreements that could span nations. We believe that a doctrine of public cybersecurity can be the basis for those policies. Our doctrine of public cybersecurity establishes a framework for state incentives and coercion that we believe is rational, defensible, and legitimate. It directs the focus of cybersecurity away from the individual and toward the collective. It advocates building systems with fewer vulnerabilities while acknowledging that systems cannot be rid of all vulnerabilities and must therefore be resilient in the face of attacks. If adopted, public cybersecurity will reorient public policy and discourse toward the proper goals of encouraging collective action to produce the public good of cybersecurity and managing the insecurity that remains.

ENDNOTES

* Contributor Biographies: DEIRDRE K. MULLIGAN is an Assistant Professor in the School of Information at the University of California, Berkeley, where she is also a Faculty Director of the Berkeley Center for Law and Technology. She is the Policy Lead for the National Science Foundation's TRUST Science and Technology Center, Chair of the Board of the Center for Democracy and Technology, and Cochair of Microsoft's Trustworthy Computing Academic Advisory Board. Her recent publications include "Privacy on the Books and on the Ground" (with Kenneth A. Bamberger), *Stanford Law Review* (2011); and "Catalyzing Privacy: New Governance, Information Practices, and the Business Organization" (with Kenneth A. Bamberger), *Law & Policy* (2011).

FRED B. SCHNEIDER is the Samuel B. Eckert Professor of Computer Science at Cornell University. He also serves as the Chief Scientist for the National Science Foundation's TRUST Science and Technology Center and is Cochair of Microsoft's Trustworthy Computing Academic Advisory Board. He is a Fellow of the Association for Computing Machinery, the American Association for the Advancement of Science, and IEEE, and is a member of the U.S. National Academy of Engineering and its Norwegian counterpart (NTV). He was awarded a D.Sc. (*honoris causa*) by the University of Newcastle-upon-Tyne in 2003.

Acknowledgments: We benefited from comments on early drafts of this paper and discussions with Marjory Blumenthal, Aaron Burstein, Scott Charney, John Chuang, David Clark, Craig Fields, R. Kelly Garrett, Jens Grossklags, Joshua Gruenspecht, Joseph Lorenzo Hall, Carl Landwehr, Susan Landau, Steve Lipner, Greg Morrisett, Helen Nissenbaum, Shari Pfleeger, Audrey Plonk, Ashkan Soltani, participants at the 2009 Workshop on the Economics of Securing the Information Infrastructure, and attendees at several planning meetings for this issue of *Dædalus*.

This essay is supported in part by the Air Force Office of Scientific Research (AFOSR) grant F9550-06-0019; National Science Foundation grants 0430161, 0964409, CNS-0524745 (ACCURATE), and CCF-0424422 (TRUST); Office of Naval Research grants N00014-01-1-0968 and N00014-09-1-0652; and a grant from Microsoft. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of these organizations or the U.S. government.

[1] Ross Anderson and Tyler Moore, "The Economics of Information Security," *Science* 314 (5799) (October 27, 2006): 610 – 613.

[2] Gerwin Klien et al., "seL4: Formal Verification of an OS Kernel," *Proceedings of the ACM SIGOPS 22nd Symposium on Operating Systems Principles* (SOSP '09), Big Sky, Montana, October 11 – 14, 2009 (New York: Association for Computing Machinery, 2009), 207 – 220.

[3] For example, "prerendering" can reduce the code required to generate the user interface in a voting system, thereby simplifying the vote-entry software and making it more amenable to verification; see Ka-Ping Yee, "Building Reliable Voting Machine Software," Ph.D. dissertation, Department of Computer Science, University of California, Berkeley, Fall 2007.

[4] Department of Defense Computer Security Center, *Department of Defense Trusted Computer System Evaluation Criteria*, CSC-STD-001-83 (Fort George G. Meade, Md.: Department of Defense, August 1983).

[5] *Common Criteria for Information Technology Security Evaluation*, International Organization for Standardization (ISO) Standard 15408, August 1999, http://www.niap-ccevs.org/cc-scheme.

[6] Health Insurance Portability and Accountability Act of 1996 (HIPAA), Public Law No. 104-191, 110 Stat. 1936 (1996) (regulating the use and disclosure of "Protected Health Information"); Gramm-Leach-Bliley Act (GLBA), Title V, Public Law No. 106-102, 113 Stat. 1338 (1999) (codified at 15 U.S.C. sec. 6801 – 6827 [2006]), 15 U.S.C. sec. 6801, sec. 6805 (empowering various agencies to promulgate data-security regulations for financial institutions); Children's Online Privacy Protection Act, 15 U.S.C. sec. 6501 et seq. (prohibiting the collection of personally identifiable information from young children without their parents' consent).

[7] Federal Information Security Management Act of 2002 (FISMA), 44 U.S.C. sec. 3541 et seq.

[8] *Draft Voluntary Voting System Guidelines: Version 1.1* (Washington, D.C.: The U.S. Election Assistance Commission, May 27, 2009), http://www.eac.gov/testing_and_certification/voluntary_voting_system_guidelines.aspx.

[9] Rainer Böhme and Galina Schwartz, "Modeling Cyber-Insurance: Towards a Unifying Framework," working paper presented at the Workshop on Economics of Information Security, Harvard University, Cambridge, Massachusetts, June 2010, http://weis2010.econinfosec.org/papers/session5/weis2010_boehme.pdf.

[10] Anderson and Moore, "The Economics of Information Security."

[11] Jens Grossklags, Nicolas Christin, and John Chuang, "A Game-Theoretic Analysis of Information Security Games," *Proceedings of the 17th International World Wide Web Conference* (WWW2008), Beijing, China, April 21 – 25, 2008.

[12] Security breach notification statutes that require companies to notify individuals when certain personal data have been accessed or disclosed without authorization are in place in forty-six states, the District of Columbia, Puerto Rico, and the Virgin Islands. The first was California's Notice of Security Breach Law, California Civil Code, sec. 1798.29 (2002).

[13] Butler W. Lampson, "Computer Security in the Real World," *Computer* 37 (6) (June 2004): 37 – 46.

[14] South Korea currently requires Internet users to attach their real names and resident identification numbers when they post messages on the Internet. Websites that allow posting must collect and confirm names and resident IDs with a government server; see Se Jung Park, Yon Soo Lim, Steven Sams, Sang Me Nam, and Han Woo Park, "Networked Politics on Cyworld: The Text and Sentiment of Korean Political Profiles," *Social Science Computer Review* (September 21, 2010).

[15] Jason Franklin, Vern Paxson, Adrian Perrig, and Stefan Savage, "An Inquiry into the Nature and Causes of the Wealth of Internet Miscreants," *Proceedings of the 14th ACM Conference on Computer and Communications Security* (CCS '07), Alexandria, Virginia, October 29 – November 2, 2007 (New York: Association for Computing Machinery, 2007), 375 – 388.

[16] William J. Lynn III, "Defending a New Domain: The Pentagon's Cyberstrategy," *Foreign Affairs*, September/October 2010.

[17] The mission of public health was defined by an influential Institute of Medicine committee as "fulfilling society's interest in assuring conditions in which people can be healthy"; Institute of Medicine, *The Future of Public Health* (Washington, D.C.: National Academies Press, 1988), 7. A rich description of the legal framework is set out in Lawrence O. Gostin, *Public Health Law: Power, Duty, Restraint* (Berkeley: University of California Press, 2001).

[18] For example, after several outbreaks of measles among primarily unvaccinated children, a federal law was passed to provide free vaccines to certain groups of children and funds to states for supporting efforts to enhance vaccination levels.

[19] Institute of Medicine, *The Future of Public Health*, 3.

[20] Systems that employ technical means to enable continued operation in the face of attacks are sometimes called *intrusion tolerant*. A sampling of specific techniques for achieving intrusion tolerance is discussed in Jaynarayan Lala, ed., *Foundations of Intrusion Tolerant Systems* (Los Alamitos, Calif.: IEEE Computer Society, 2003). In this essay, the term *managing insecurity* is intended to denote something broader, admitting nontechnical means as well as intrusion tolerance techniques.

[21] Helen Nissenbaum, "Where Computer Security Meets National Security," *Ethics and Information Technology* 7 (2) (June 2005): 61–73.

[22] The doctrine of prevention is not concerned with managing insecurity; the doctrine of risk management and doctrine of deterrence through accountability are not concerned with producing cybersecurity. None concern trade-offs of individual rights for public welfare.

[23] Deborah G. Mayo and Rachelle D. Hollander, *Acceptable Evidence: Science and Values in Risk Management* (New York: Oxford University Press, 1994).

[24] In fuzz testing, a system is exposed to random inputs of unexpected kinds. This form of testing reveals inadequacies in the input validation routines of a system. Several classes of attacks are blocked by implementing stringent input validation.

[25] For a case study comparison of four vulnerability reduction techniques, see Koen Buyens, Bart De Win, and Wouter Joosen, "Empirical and Statistical Analysis of Risk Analysis-Driven Techniques for Threat Management," *Proceedings of the Second International Conference on Availability, Reliability and Security* (ARES '07), Vienna, Austria, April 10–13, 2007 (Washington, D.C.: IEEE Computer Society, 2007), 1034–1041. For a theoretical comparison of two high-profile development processes (Microsoft SDL and the Open Web Application Security Project's Comprehensive, Lightweight Application Security Process [CLASP]), see Johan Grégoire, Koen Buyens, Bart De Win, Riccardo Scandariato, and Wouter Joosen, "On the Secure Software Engineering Process: CLASP and SDL Compared," *Proceedings of the Third International Workshop on Software Engineering for Secure Systems* (SESS '07), Minneapolis, Minnesota, May 2007 (Washington, D.C.: IEEE Computer Society, 2007). An argument in favor of evidence-based practices appears in Daniel Jackson, Martyn Thomas, and Lynette Millett, eds., *Software for Dependable Systems: Sufficient Evidence?* (Washington, D.C.: The National Academies Press, 2007).

[26] Vaccination works by causing exposure to a relatively benign form of the disease against which protection is being sought.

[27] A notable example is "Implementation of Commonly Accepted Security Configurations for Windows Operating Systems," policy memorandum M-07-11 (Washington, D.C.: U.S. Office of Management and Budget, March 22, 2007), http://www.cio.gov/documents/FDCC_memo.pdf. The memorandum lists the few versions of Windows that certain civilian federal agencies are permitted to use.

[28] David Ladd et al., *The SDL Progress Report* (Microsoft Corporation, 2011), 23–24, http://www.microsoft.com/downloads/en/details.aspx?FamilyID=918179A7-61C9-487A-A2E2-8DA73FB9EADE&displaylang=en.

29 The general rule allows individuals to refuse treatments. Some states mandate treatments for communicable diseases, such as tuberculosis, that pose a danger to the public.

30 Michel J.G. van Eeten and Johannes M. Bauer, "Economics of Malware: Security Decisions, Incentives and Externalities," OECD STI Working Paper 2008/1 (Organisation for Economic Co-operation and Development Directorate for Science, Technology and Industry, May 2008), http://www.oecd.org/dataoecd/53/17/40722462.pdf.

31 See David D. Clark and Susan Landau, "Untangling Attribution," *Harvard National Security Journal* 2 (2) (2011), http://harvardnsj.com/2011/03/untangling-attribution-2.

32 "Einstein Intrusion Detection System: Questions That Should Be Addressed" (Washington, D.C.: Center For Democracy and Technology, July 28, 2009), http://cdt.org/security/20090728_einstein_rpt.pdf.

33 The risk that such vulnerabilities might be sold or disclosed to irresponsible or hostile parties is cause for concern, but this market structure flourishes in the absence of alternatives.

34 See http://www.confickerworkinggroup.org.

35 For example, unpatched machines can be co-opted by attackers into a botnet. Such collections of remotely controlled machines are used today for a variety of illicit activities, including generation of spam email and distributed denial-of-service attacks.

36 However, some enterprises regard their software and data as proprietary. They might not be comfortable providing test suites to patch developers.

37 Richard Clayton, "Might Governments Clean Up Malware?" *Proceedings of the Ninth Annual Workshop on Economics and Information Security* (WEIS10), Cambridge, Massachusetts, June 7–8, 2010.

38 The term *air gap* originally referred to isolation caused when no wires are connected to a given component. With the advent of wireless networking, the term's meaning has broadened to denote isolation caused when the Laws of Physics ensure no signal can reach the component.

39 *Malicious Software (Malware): A Security Threat to the Internet Economy*, Ministerial Background Report DSTI/ICCP/REG(2007)5/FINAL, Organisation for Economic Co-operation and Development, June 2008, http://www.oecd.org/dataoecd/53/34/40724457.pdf.

40 In October 2009, Comcast announced a trial of an automated service that "warn[s] broadband customers of possible virus infections." The Comcast Constant Guard issues a pop-up notice directing customers to resources to rid their machine of infection; see Elinor Mills, "Comcast Pop-ups Alert Customers to PC Infections," CNET News, October 8, 2009. A similar service, the Qwest Customer Internet Protection Program, displayed a Web page warning to customers with options for removing the detected infection for free; customers were obliged to do so before they were allowed to return to surfing the Web. Similarly, an older, now discontinued SBC service quarantined computers until they were cleaned.

41 Ibid.

42 Clayton, "Might Governments Clean Up Malware?" 5 n.2.

43 For a discussion of some of the issues raised by an ISP's decision to cut off broadband access due to infection, see George Ou, "Comcast Heading the Right Direction on Cybersecurity," Digital Society, October 9, 2009, http://www.digitalsociety.org/2009/10/comcast-heading-th-right-direction-on-cybersecurity.

44 See http://www.eff.org/wp/noncommercial-email-lists-collateral-damage-fight-against-spam.

45 See generally, Doug Lichtman and Eric Posner, "Holding Internet Service Providers Accountable," *Supreme Court Economic Review* 14 (2006): 221, 233–234.

[46] Stephanie Forrest, Alan S. Perelson, Lawrence Allen, and Rajesh Cherukuri, "Self-Nonself Discrimination in a Computer," *Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy*, Oakland, California, May 16 – 18, 1994.

[47] Stephanie Forrest, Anil Somayaji, and David H. Ackley, "Building Diverse Computer Systems," *Proceedings of the Sixth Workshop on Hot Topics in Operating Systems*, Cape Cod, Massachusetts, 1997.

[48] *Enabling Distributed Security in Cyberspace: Building a Healthy and Resilient Cyber Ecosystem with Automated Collective Action* (Washington, D.C.: Department of Homeland Security, March 23, 2011), http://www.dhs.gov/xlibrary/assets/nppd-cyber-ecosystem-white-paper-03-23-2011.pdf.

[49] *National Cyber Leap Year Summit 2009 Co-chairs Report*, September 16, 2009, http://www.nitrd.gov/NCLYSummit.aspx.

[50] Daniel B. Prieto and Steven Bucci, "Meeting the Cybersecurity Challenge: Empowering Stakeholders and Ensuring Coordination," IBM U.S. Federal White Paper (Somers, N.Y.: IBM Corporation, February 2010).

[51] Scott Charney, "Collective Defense: Applying Public Health Models to the Internet," white paper (Redmond, Wash.: Microsoft Corporation, 2010), http://www.microsoft.com/security/internethealth.

[52] Jeffrey Hunker, "U.S. International Policy for Cybersecurity: Five Issues That Won't Go Away," *Journal of National Security Law & Policy* 4 (2010): 197 – 216.

[53] Benjamin Powell, "Is Cybersecurity a Public Good? Evidence from the Financial Services Industry," Working Paper #57 (Washington, D.C.: The Independent Institute, March 14, 2005); Bruce H. Kobayashi, "An Economic Analysis of the Private and Social Costs of the Provision of Cybersecurity and Other Public Security Goals," Working Paper #26 (Arlington, Va.: George Mason University School of Law, 2005), http://law.bepress.com/gmulwps/gmule/art26; Brent Rowe and Michael Gallagher, *Private Sector Cyber Security Investment Strategies: An Empirical Analysis* (Research Triangle Park, N.C.: RTI International, March 2006).

[54] Nissenbaum, "Where Computer Security Meets National Security."

Reconceptualizing the Role of Security User

L. Jean Camp

Abstract: The Internet is not the only critical infrastructure that relies on the participation of unorganized and technically inexpert end users. Transportation, health, waste management, and disaster preparedness are other areas where cooperation between unorganized citizens who lack experience with the domain has increased resiliency, reduced social costs, and helped meet shared goals. Theories of community-based production and management of the commons explain this type of cooperation, both offline and online. This essay examines these two complementary approaches to organizing the cybercitizen for cybersecurity. Cybersecurity discourse has reasonably focused on centralized parties and network operators. From domain name registrars to network service providers, solutions are sought through incentives, regulation, and even law enforcement. However great the ability of these centralized entities to implement change, the end user plays a crucial role. The Internet must remain open to enable innovation and diffusion of innovation; thus, the end user will continue to be important. What is the role of the citizen in cybersecurity? What socio-technical characteristics might enable a system that encourages and empowers users to create a secure infrastructure?

L. JEAN CAMP is a Professor in the School of Informatics, Adjunct Professor of Telecommunications, and Adjunct Professor of Computer Science at Indiana University. Her publications include *Trust and Risk in Internet Commerce* (2000), *Economics of Information Security* (edited with Stephen Lewis, 2004), and *Economics of Identity Theft: Avoidance, Causes, and Possible Cures* (2007).

How can cyberspace be augmented or organized so that security is more widely produced at home by citizens who lack technical expertise? Answering this question is critical to governance of the Internet. When one person or machine is not secure, any or all of the people connected to the Internet potentially pay a cost. The average user receives spam because other average users allow their machines to host spammers. Securing cyberspace is an inherently cooperative venture.

A growing body of work illustrates how classes of goods are constructed by a collective (*community-based production*) and how shared resources are managed (*managing the commons*). When viewed from the first of these two perspectives, security is a good that can be cooperatively produced. When viewed from the second, computer security appears to be a common good that can be consumed

and preserved, but not produced, cooperatively. The Internet as a commons can be compromised if too many people accept a high level of insecurity. In both cases, requirements for the nature of the good, whether public or private, must be defined. Cybersecurity is a good with significant private incentives; in the same way that no one seeks to become ill, no one wishes to be the victim of identity theft. Cybersecurity may also have a tipping point, after which a herd effect motivates action or adherence. On the network as in the realm of public health, herd immunity is needed to prevent epidemics.

In this essay, I describe resource management as community-based production and as the management of the commons. I suggest that the underlying requirements for each of these approaches may already exist or could be created.

Community-based Production. Community-based production refers to the self-organization of community members to create a good or service.[1] This model of social production differs from "crowdsourcing," in which a single entity, such as a firm, handles the organization and management of a collective effort.[2] I begin by considering the case of community-based production of security information by self-correcting experts, then argue for computer security as a good that could be enhanced, though not fully implemented, by community-based production.

In his 1937 article "The Nature of the Firm," economist Ronald H. Coase explained why individuals self-organize into firms with high degrees of specialization.[3] Adam Smith's *The Wealth of Nations* illustrated the value of specialization with the famed example of pin production.[4] Arguments for producing a particular good or service through this approach span centuries and cannot be adequately cataloged in this essay. In contrast, intellectual theories and reproducible analyses that explain and describe community-based production as sustainable (and preferable in some cases) have emerged more recently.

Community-based production can drive the creation of any good that possesses the following characteristics: modularity, low capital requirements for entry into production, low marginal cost of production, and well-defined interfaces or interactions resulting in low cost of integration. The concept of community-based production was first defined in relation to the economics of software that runs the Internet. Called open-source software, it must be shared, readable by all Internet users, and cost-free in order to prevent barriers to connection.[5] Finance scholar Josh Lerner and economist Jean Tirole have described how individuals self-identify by contributing to tasks for which they are uniquely suited.[6] They argue that increased income is not the sole incentive for the production of open-source software; rather, social intangibles such as respect, reputation, and the knowledge that one is the first to solve a particular problem are significant motivators. Contributions must be visible to the community for these incentives to function. In the realm of cybersecurity, neither the benefits of securing a machine nor the costs of failing to do so are visible, even to a machine's owner, thus reducing the incentive to take action. Nonetheless, there are multiple examples of loosely coordinated online communities that produce security information, including *vulnerability sharing* and *probe networks*.

A vulnerability is an error in coding or installation that enables unauthorized access to electronic resources. More specifically, it is a technical flaw allowing unauthorized access or use, where the

relationship between the flaw and access allowed is clear and has been documented to have been used to subvert a machine. Information about vulnerabilities can be held secret, shared openly, or sold to a company.

Probe networks are sets of networked computers that have no services and thus offer no legitimate reason to connect to them. The connections they receive are from massively parallel attacks, in which attackers are trying every feasible IP address. A social network of system operators run the computers to share statistics on these broadcast attacks, identify patterns as well as anomalies, and provide real-time distributed-network monitoring.

Vulnerability sharing has long been community-produced. The email list Bugtraq provides a mechanism to report vulnerabilities and rate the resulting report. Users of the list contribute knowledge, discuss vulnerabilities, and devise patches. A large amount of computer security information is generated and validated by these volunteers with no obvious incentives; they do not have property rights to the information they provide or to the products that their critiques help improve. Like any informal system of social reputation, earning credibility is difficult, time-consuming, and uncertain. In contrast, when information production is managed by a firm, there is no basis on which to argue that such self-organization would occur. The HP service TippingPoint is a firm created to purchase vulnerability information, thereby constraining the distribution of this information. Thus, the application of property rights to vulnerability information decreases social welfare.[7] Yet Bugtraq's open system continues (with eight postings on April 26, 2011, for example).

The DShield project, one example of a probe network,[8] is a community of system operators who run network monitors or "probes" to detect intrusions and attacks. Information gathered by network probes is shared among and collectively analyzed by members. The DShield project enhances the reputation of contributors by identifying them as such; increases the perceived value of its parent, Euclidean Consulting; and enables the identification of Internet Service Providers (ISPs) that are responsible for the worst attacks. DShield also provides a free self-help service that allows individuals to verify whether their IP addresses are associated with attacks. In this model, expert members of the community provide and generate security information at no charge and without claiming property rights.

Contributors to network probe systems are valuable to the entire Internet community and provide aggregate data that no individual could produce. However, there is no monetary incentive to participate. An individual who includes his or her own data in the aggregate may potentially benefit from a marginal increase in the ability to detect an early worm attack. Nonetheless, for the purpose of determining the pervasiveness of attacks or having an early warning system in place, the self-optimizing choice is to ensure that one's network neighbors participate without participating oneself.

While DShield provides a public list of malicious sites, other organizations have a more traditional, centralized production method. For example, WebSense offers security-based blocking of sites that it locates using its own honeypots. (A honeypot is a computer on the network designed to be attractive to attackers by a combination of weak security and tempting file names. For example, the machine might be running old versions of software with known vulnerabil-

ities and have fake files with names such as "creditcardnumbers.xls" or "Accounts AndAddresses.db.") Individuals who do not subscribe to WebSense do not obtain the lists of compromised sites. Thus, WebSense offers security production in the more conventional model: that is, via a firm.

The above examples of community-based production by highly educated network administrators, engineers, and researchers describe collective activities of security experts and network engineers committed to a more resilient Internet, not the creation of security information by non-experts. Is the latter practical? Which types of information goods can a community of non-experts feasibly create, under what conditions, and by whom? To what extent can community-based production be generalized?

Legal scholar Yochai Benkler's research on the production of trusted information describes the goods and services that can be effectively produced through community-based means.[9] In certain contexts, this model has many advantages over firm-based production. By altering the modularity, granularity, and cost of integration of the good that is produced, the cooperative model can shift the distribution of production costs to those most able and willing to bear them.

Challenges that can be effectively addressed by community-based production are modular rather than continuous; that is, they are characterized by clearly delineated decision points and explicit requirements for participants. For example, while prosecuting organized online crime is a broad objective best addressed by a highly organized governing body or firm, the task of identifying malicious websites is amenable to community-based solutions because each website is distinct and the task is well-defined. The task is to determine not if the website is providing good and correct information but if it is a masquerade site (also called a phishing site) or if it is distributing malicious code, thereby infecting visitors' computers. Masquerading sites are correctly recognized by experts more often than by nontechnical users. However, if users are given simple information, such as their own browsing history, and are told that this is a first visit to a site, they may detect that the site claiming to be Bank of America is a phishing site, based on their knowledge of having previously visited Bank of America's actual website. That is, users know that a site they have never visited before is not one where they should enter their banking passwords. The computer recognizes that the site is not part of the user's history; the user does not. Alerting the user to a first-time visit before he transmits his password would be a straightforward change in technical communication. Similarly, malware sites have very short life spans, and computers might be programmed to indicate to the user if a site is a day old, if it is certified by a rarely used certification site in a remote locale, and whether the user has ever visited similarly risky sites. Such alerts could enable the user to make an informed decision.

The cost of integrating individual distributed efforts must also be limited for this model to work. Benkler mentions software as an example in which integration can occur with a well-documented application program interface; his counterexample is aircraft construction, which requires an exacting physical integration of many components. What elements of cybersecurity, like aircraft, require regulation and coordination? Which can be enabled from the bottom up? Design parameters change when systems are implemented to enable peer, as opposed to firm or governmental, production. In

terms of participant requirements, considerations include lifestyle (for example, whether the user must be connected to the Internet for 85 percent of the day) and the type of response required (for example, whether a system simply reacts to alerts or requires constant monitoring and an endless attention span).

The identification of malicious websites illustrates the potential uses and limits of community-based production. Whereas legitimate banks can be appropriately identified in a centralized fashion, by the Federal Depository Insurance Corporation and the National Credit Union Administration, the rate by which unknown websites are increasing prevents any centralized entity from identifying all of them. Yet online behavior in the user community suggests a potential solution. The vast majority of sites that individuals visit in a browsing session are the ones they visited in previous weeks. One study of browsing history over a four-week period found that within a social network of only ten people, more than 99 percent of all participants' clicks led to previously visited sites.[10] Only one in a hundred clicks brought individuals to a site identified as unknown. (For the average individual in the study, 95 percent of clicks led to familiar places.) By reconceptualizing the global issue as a community problem, the study uncovered new and potentially untrustworthy sites without compromising user privacy.

Attackers have long used social networks to enhance attacks. The "I love you" virus was the first malicious use of address books; today, attackers harvest Facebook and the comment sections on blogs. As a result, centralized solutions must address the vast heterogeneity of the Web. Given the sheer scope of the challenge, the identification of individual websites as new, and thus suspect, is best done by community members. Yet the long-term efficacy of the community model depends on the capacity of centralized institutions to coordinate the identification and takedown of malicious sites in a timely manner – that is, before the sites build reputations. Both community-based and centralized production are necessary; neither is adequate on its own.

In assessing whether websites or individual users are trustworthy, community-based production can incorporate implicit, behavior-driven ratings or explicit, personal recommendations or selections from trusted parties. The amount of time an individual spends using, and therefore contributing to, various resources is another implicit measure of trustworthiness. Social trust reduces technical complexity[11] and can alter the nature of cumulative risks taken, in terms of system failure, privacy, and even threats inherent in system operation. Community production achieves a governance system that either could not be accomplished by a centralized agency or could not be accomplished without very large-scale investment of capital.

Community production recognizes the incentive individuals have to maintain their own information. Individuals' ability to protect themselves against the risks they take – that is, the capacity to shift costs to those with the greatest incentive to bear them – applies to malicious sites and many types of machine subversion. However, one difficulty of cybersecurity is that while the individual bears the cost of some machine subversions, in some cases the costs are borne by others. Again, the creative use of social networks and incentives can be applied here to help develop a more robust and resilient infrastructure.

The Commons without Tragedy (or Government). A modest estimate suggests that 5 percent of machines connect-

ed to the Internet are under the control of malicious parties.[12] Commonly, large numbers of machines, called "zombies," are brought under the control of a single centralized entity, forming a "botnet." Zombies do not necessarily steal the information of affected machine owners. Rather, access to the machine provides a launchpad for attacks, a storage resource, and a safe space away from virtual home. The fact that the malicious controllers of these machines are centralized has triggered centralized solutions, such as coordinated law enforcement. However, the subversion of these machines works from an economic perspective because of the large supply, high connectivity, and very low marginal cost of hitting tens of millions of machines to find hundreds of thousands that can be subverted. One possible solution for the cyber-commons is a ground-up approach that would induce secure behavior in communities and subnetworks.

Any ground-up approach will depend on preexisting social trust and risk communication to create subnetworks that seek to change individual and group behaviors. A range of powerful authentication technologies has yet to be applied to the challenges of securing devices (including proximity authentication, for example, which works only for devices that are physically collocated). Virtual neighborhoods created from secure group formation and physical neighborhoods authenticated by proximity are examples of possible subnetworks where effective interaction design combined with social transparency can enable neighborhood self-defense.

Political economist Elinor Ostrom has illustrated that effective governance of shared resources can emerge under certain conditions. In a 2003 summary in *Science*, Ostrom and colleagues list five conditions: 1) the monitoring of resources and their use, 2) moderate rates of change in social and economic conditions as well as user populations, 3) the existence of social capital in the community, 4) the ability to exclude outsiders from the community, and 5) user-supported enforcement of norms.[13] In theory as well as in practice, creating these conditions requires the development of secure systems that are designed as social networks.

It is also important to consider the relationship between social networks and herd immunity. If policy cannot change the behavior of all users, what category or number of people must be encouraged to change? Can visibility of low-security choices be leveraged to create the transparency necessary for self-governance of an Internet-security commons?

Monitoring Resources. Monitoring resources in a shared domain is one of the simplest but most underused governance mechanisms. Information monitoring has a trivial effect on the information infrastructure, and simply providing information can be a potent agent of change. Individuals are rarely capable of monitoring their own network experience, and yet there are few available interfaces for monitoring network resources. Even organized efforts have difficulty measuring resources available to individuals, with some comparisons of national broadband networks measuring nothing more than the parameters set by the machines for sharing bandwidth.[14] Individuals often do not know their own resources and, arguably, never have. The 2003 spread of the Slammer worm is an example of the challenges home users face in monitoring their own resources. As Slammer attacked SQL servers, most people were unaware that they might be vulnerable even if they knew of the worm's existence. Announcements specified that the worm attacked Microsoft

SQL Server 2000, but how many users knew that their PCs, in fact, ran an SQL server? Any technically useful report could have been construed by the average user as acknowledgment that the worm did not apply to him or her. Today, few individuals who have broadband are aware that they have a Web server in their homes. Yet every wireless hub has a simple Web server that enables the owner to initialize and configure the device. Any individual who sets up a wireless router using a browser may be unaware that there must be some server code running on the device. Transparency may have improved in terms of an individual's knowledge of his or her own resources, but there is no evidence that awareness has improved.

Currently, some ISPs may notify individuals when the ISP detects or is notified that a machine on the ISP's network is subverted. Though an estimated 5 percent of all machines on the Internet have been subverted,[15] even the most aggressive ISP responses have offered recovery services to just 1 percent of subscribers.[16] For this reason, both end users and network service providers have limited awareness of the existence of botnets. One problem for a network service provider is the heterogeneity of the network population.

Concerns about individual privacy and close monitoring of network behaviors also limit network monitoring, as organizations that have been found to practice unwarranted monitoring have faced, at the least, a media backlash.[17] For this reason, close monitoring by an agency may be impractical, but the homogeneity and consistency of individual behavior is an asset to the extent that home users can observe the actions their machines take on their behalf. Detecting a change in website visitation behavior across the entire network is very difficult. Detecting a change by a single computer, which has very few (human and therefore nonrandom) users, is a problem that is easier to solve. While the individual machine is fairly consistent, the network is not. When a machine suddenly becomes inconsistent, a home user will have better information on why that may or may not be suspicious. For example, a network service provider may observe changes in traffic behavior without knowing if it is because there is a family reunion taking place (and thus a dozen teenagers are on the wireless) or a machine has been subverted; the individual user, on the other hand, can easily distinguish between the two scenarios. Thus, while recovery services may need to be centralized, network monitoring and communication with individual users function best when decentralized. The global network is dynamic: it changes rapidly, constantly reconfigures routes, and is profoundly heterogeneous. The individual, by contrast, is a notoriously poor source of entropy. Monitoring resources at the end point means leveraging the innate homogeneous humanity of the single user, as opposed to simply bemoaning the fact that humans produce weak, nonrandom passwords.

Home users face a plethora of add-ins, add-ons, and an ever-expanding lexicon of attacks and defense. A more productive approach would be to provide individuals with a single narrative and a clearly marked path to risk mitigation and recovery. Users could be informed of radical changes on one machine in the house either by other machines in the house or by the machines that participate in the social networks I describe below. Current technology is not designed to communicate, in an effective, carefully timed, and educational manner, the particular risks to which a user might be exposed; nor does it automatically change settings to respond to personal contexts

(work, play, banking) or technical ones (public wide area network, protected workplace, patched or unpatched).

Market forces, property rights, and even assigned identifiers can solve some of the incentivization problems related to computer security. However, an effort to control and enforce behaviors on the population in a "war on computer insecurity" risks being both ineffective and expensive. In contrast, making risks and decisions visible to individuals, thus enabling them to monitor their own machines, is a technological challenge that can be met without violent metaphors or intrusive monitoring.

Moderate Rates of Change. The notion that the Internet is open to all is a canard; exclusion is now and always has been practiced online. The earliest form of exclusion was email lists. The existence of some of these lists was secret (particularly from professors in some schools). Some lists added members by invitation only; others allowed open subscription as well as banning; and still others embraced a simple no-holds-barred approach. For example, a group of mothers whose children share a birth date have followed each other and stories of their respective offspring for twenty years; their email list is highly exclusive. The same models apply today in the blogworld. Blogs can be completely open, allowing unmoderated (and immoderate) anonymous comments. More often, they are slightly restrictive, allowing open readership with member comments or moderated anonymous comments. Many reserve access for members only, especially when members have strong shared experiences (such as surviving abuse) or have tired of defending the existence of the group itself and simply wish to discuss the topic at hand (for example, feminist blogs that exclude men's rights activists). Such closed blogs can be read only if an application for inclusion is accepted.

The second-oldest form of the closed online community is arguably the chat room, a service built on the idea of the single-identity provider. The chat room functioned primarily because of a large, AOL-installed user base and AOL's centralized governance ability. Chat rooms are also called "pull technology," meaning the individual must actively log in to participate. Before the chat room, there were closed mailing lists.

The Internet has long been applauded for its openness. Yet the network enables the creation of spaces that can be closed or even invisible to others. Social networking has enabled exclusion since the first days of email, when reply-all became reply-to-sender for the occasional snark. The implementation of stable Internet communities is widely managed by those communities across an array of platforms. Even with difficult interactions and exploitive privacy requirements, the story of the Internet is one of community formation. The ability to form communities, whether bound by physical location, shared interest, or sheer random selection of the moment, illustrates that the second condition for management of shared resources can be met. An unknowable number of groups – from high school classmates to their mothers, who have been chatting online since the first positive pregnancy result – meets the requirement of "moderate rates of change in social and economic conditions as well as user populations."

Social Capital. Discussions of security in economic terms – that is, as financial capital – have been active for the past ten years.[18] Security as social capital, however, is rarely considered. Most related technologies define security as an individual ef-

fort and presume that information is an individually owned resource.

A 2000 paper introduced the idea of computer security as a good with externalities.[19] Since then, models of various components of security-related externalities have been widely explored. In economic terms, the current crisis in computer security is a market failure. There is some agreement that components of cybersecurity are a public good.[20] Private security decisions have a public externality, as the cost of an insecure system is accrued by other systems that are subsequently infected as malicious computer code, such as viruses and worms, spreads. Various solutions to this problem have been proposed, including liability,[21] insurance markets for business risks,[22] and enforcement mechanisms for ISPs.[23]

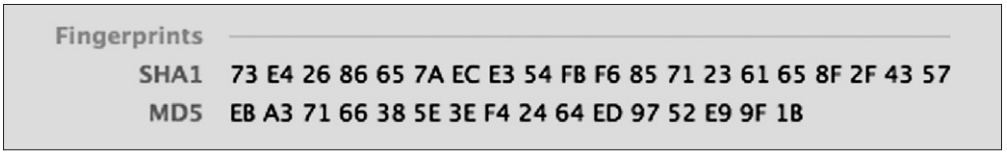
Security defined as a good has both public and private elements, but security proposals have tended to focus on the private aspect. In terms of the public good, solutions have emphasized monetary liability or insurance. Although proposals for liability could function for larger actors in the security market, where decisions (at least in theory) are driven by cost-benefit analysis, this approach would likely backfire in the realm of vulnerable home users. In fact, some proposals could increase the potential risks for individuals without providing any mechanisms for enabling them to avoid these risks. Increased liability, as a means to encourage individuals to cease insecure behaviors, is unlikely to be highly effective if individuals are unaware of being engaged in such behaviors. Again, consider that 5 percent of home machines may be infected. A regime that criminalizes users for having an insecure home machine would immediately transform some 5 percent of the online population from law-abiding citizens to enemies of the state. It is difficult to imagine an external attack that would similarly hinder or harm so many Americans.

A commonly proposed solution to an externality is the creation of a property interest that would enclose the information security space. A functional market would require adequate information, the ability to process this information, and a sufficient attention span. Currently, security technologies provide none of these requirements. For example, when a public key certification is identified as invalid, the user receives two incomprehensible hash values for calculation and comparison but no information about the source of the warning or the certificate (see Figure 1). Only the rare mathematical genius could compare these two values in any useful way. Yet this is the only information presented to the end user to help him or her determine if the certificate should be trusted.

Is the certificate invalid because it is signed by a university rather than a more widely recognized corporate provider of certificates? Is it invalid because it expired yesterday? Or is it signed by a previously unknown, and therefore likely to be malicious, party? Is it signed by a leading certificate provider for large corporate entities, or a rarely used provider favored by marginally legal organizations? Certificate providers have social capital and reputations that are well known by those who read security literature and manage networks. Experts in the field know that some certificate providers are more trustworthy than others. Good reputation is a form of social capital, but unless this is visible to typical users, it cannot be an effective part of collective decision-making.

Similarly, ISPs have widely varying records for the protection of individuals. Some ISPs simply let user computers drop onto blacklists, never contacting the owner, while others notify and assist cus-

Figure 1
An Example of Hash Values Provided as a Warning of a Potentially Compromised or False Public Key Certificate



tomers whose computers are apparently infected. There is no place to locate this information. Individuals as well as organizations have histories and social capital on the network. This information exists and should be made much more widely available. To paraphrase Attorney Samuel Warren and Associate Supreme Court Justice Louis Brandeis,[24] that which is whispered in the halls of North American Network Operators Group (NANOG)[25] should be shouted from the rooftops.

Better information monitoring would allow individuals to know whether their machines are responsible for spam. Were this information more readily visible, neighbors and friends would know, too. Computer insecurity, were the costs to others apparent, could become as socially unacceptable as littering: that is, it would exist, but to a much less egregious extent. Given that computer crime is driven by profit more than pride, making insecurity anomalous rather than ubiquitous may be adequate to stem the tide.

Exclusion. For policy-makers' purposes, exclusion from a community is permanent. Permanent exclusion requires permanent, single identities. Yet exclusion has long been possible without the costs and stochastic risks inherent in single identifiers. (This consideration complements the above discussion of stable communities.) A policy based on single, true names, in theory implemented by government, is in no way the best approach. Instead, stable pseudonyms in specific communities are more than adequate. Social capital requires a social context, and the nation is too large a context to be workable for any but the most famous personalities or dangerous criminals.

The use of social networks and the explosion of social communication illustrate the capacity of those on the network to implement change through reputation mechanisms of all levels of openness. While Facebook consistently alters users' control of their own profiles, there has consistently been a wide range of Facebook mechanisms that provide user control. Anyone who has ever "unfriended" another person has experienced the power of exclusion on the Internet. Similarly, the new social network system from Google, G+, enables more detailed control, with groups of people in categories such as "friend," "acquaintance," or "family."

The major functions of security must include exclusion and control. As noted above, the most straightforward solution involves a centralized entity that has the authority to issue – and therefore revoke – accounts, enabling access control. Proposals for federated or trusted identities all follow the same logic: a single identity will enable access control and ensure responsible behavior. A similar logic applied to the recent cancelation of G+

accounts that were not based on names that Google determined to be adequately true and real. Others argue that requiring individuals to use a single identity will create another tragedy of the commons, whereby identities are the overused and underprotected resource. A second line of objection is social rather than economic: that is, individuals who would be threatened in employment or communities for having unpopular views (for example, feminists in Texas, fundamentalist Christians in the San Francisco Bay Area) should be able to speak without their coworkers and employers knowing. This problem was dominant in the cancelation of G+ accounts that lacked true names, with users giving the following reasons for adopting a pseudonym:

- "I am a high school teacher, privacy is of the utmost importance."
- "I publish under my nom de plume, it's printed on my business cards, and all of the thousands of people I know through my social networks know me by my online name."
- "I have used this name/account in a work context, my entire family know this name and my friends know this name. It enables me to participate online without being subject to harassment that at one point in time lead [*sic*] to my employer having to change their number so that calls could get through."
- "I do not feel safe using my real name online as I have had people track me down from my online presence and had coworkers invade my private life."
- "I've been stalked."
- "I'm a rape survivor."
- "I am a government employee that is prohibited from using my IRL [in real life]."[26]

Are global names the price of effective governance? Or can exclusion exist in smaller communities? Certainly, the dual threats of privacy violations and misuse by (authorized or unauthorized) parties have not been adequately addressed by any federated or single-identity proposal. Another challenge requiring an Internet-wide or national solution is that security is a social activity in which the most malicious participants work to be the least visible. Identity thieves will not be hindered by the introduction of a new identity infrastructure any more than foxes would be hindered by an increase in the chicken population.

Exclusion already exists in Internet communities. Just as the Internet is the network of networks, it is the network of communities. Engineering solutions that enable governance in small communities can make a difference in the Internet as a whole, arguably more efficiently and certainly with better privacy. The Internet experience is one of physical disconnectedness and social connection. For every highly verbose commenter who weighs in on a blog post, there are orders of magnitude of more silent readers. The anonymity of the Internet is easily violated, yet users act as if privacy were protected by social contracts communicated via website design.[27] These social contracts could be leveraged with security systems that address the shared costs of security and encourage cooperative governance. Social contracts require exclusion for those constraints to be both binding and enabling.

Social contracts such as those enacted by the users of Facebook existed in the "real" world long before the Internet came into being.[28] In the real world, however, individuals can use the visual, geographical, and tactile information embedded in physical interactions to evaluate the safety, competence, and

trustworthiness of those who control a physical space. For example, merchants offering high-quality products can charge a premium based on reputation and invest their profit in retail spaces that reflect their wealth and standing. These cues are not available online because they have not been integrated into the network – not because they are impossible to engineer. Blacklisting and blocking are usable tools in email, on blogs, chat, and within social networking and recommendation systems. Better engineering that enables self-organization for purposes other than superior advertising targeting is therefore possible; further, it may prove less costly and more effective than a punitive legislative approach.

Social Norms. Certainly, community norms can be altered or enforced by centralized control. In several domains – for example, online communities such as Facebook – information has multiple stakeholders. Acceptable use of knowledge is determined by implicit social and explicit corporate policy norms established through perceived imagined communities.[29] By contrast, current security technology is based on the mental model of the security expert who develops usable security controls for the individual, informed user. This approach limits the effectiveness of such controls by discounting both the social context of use[30] and the mental model of the end user.[31] Considerable research on the creation of norms has examined the development of social conventions in different online communities; material studied ranges from explicit sexual interactions in Second Life (a three-dimensional virtual world where users can socialize) to pattern sharing through Ravelry (a site for knitters and crocheters). Norms must be widely accepted in online communities, but they need not be perfectly enforced. Norms that are violated on rare occasions remain norms, just as the challenges of cybersecurity are manageable without flawless enforcement. Actions that simply are not taken (for instance, there is no norm against nailing one's own foot to the floor) are not suitable for governance by norms.[32] Security controls could enable individuals to set their own norms for interaction. An individual could be empowered to filter and accept different risks automatically – in terms of liability, ownership, rights, and acceptable use – when in different virtual spaces and communities.

Recent work has examined how to nudge users toward positive behaviors through interaction design rather than with inescapable defaults or tedious reminders. Yet even on social networking sites, these efforts target the isolated user sharing his or her own information. In contrast, tools that encourage communal norms and make those norms visible in a user-defined community can encourage members of social networks (as discussed above) to comply with those security and safety norms.

Some of the language used to address computer security may, in fact, discourage compliance. Being a "pirate" may seem desirable by end users because of its appeal in popular culture. "Zombies" are imagined as inherently villainous, but they are also popular (consider the films *Zombieland* and *Shaun of the Dead*). "Phishing" is a purposefully obtuse word, created as an inside joke. Such nomenclature may prompt users to ignore security, or not to take it seriously. Certainly, the language of computer security does not encourage norms of security adoption – nor does the "one size fits all" approach, yet security designs assume that interactions and defaults should be uniform for all people and across the entire browsing or Internet experience.

The term *computer security* is also deeply intertwined with the self-interested protection of copyright holders – to the detriment of many users. The goal of policing the individual directly conflicts with the objective of recruiting the user to participate in securing cyberspace. Disentangling these goals through community involvement, encouragement, and communication can enable security-enhancing norms to emerge. Pursuing these goals through ever-less-functional devices and more expansive definitions of felonies may effectively choke the Internet as an engine of American innovation, without ensuring security for the novice end user. Norms must align with the interest of the community to be adopted. Disentangling very different risks of sharing copyrighted material without permission or hosting a botnet that serves organized crime can encourage norms of security and digital safety.

The Role of Government. Security is in part a good that can be cooperatively produced. This implies recruiting end users into the production of security. The Internet is also a common resource; thus, security is a component that requires shared management. The five characteristics (using the model described above) of shared management are not obviously present on the Internet. Although these five characteristics (resource monitoring, moderate rates of change, social capital, exclusion, and social norms) are not available on the global Internet, they certainly exist in Internet communities.

The challenge of securing cyberspace cannot be met without the cooperation and coordination of the end user, and the conditions to enable this cooperation and coordination exist. While the cybersecurity community, in both technical mechanism design and political discourse, has concentrated on large-scale centralized entities, socially aware security engineering can make a profound difference. Currently, home computer users who seek to remain safe face a patchwork of standards and corporate products. Each user has a unique set of challenges embedded in geography, health, social context, service provider, platforms, and other variables, such as home layout. One size cannot fit all, and one user cannot be expected to face the world alone with no community support.

Similarly, with respect to roadways, water management, and public health, the behavior of nonprofessional participants is critical. Not everyone speeds; littering, smoking, and drinking and driving were all socially appropriate behaviors at one time. These behaviors have been radically reduced by a combination of public education and sporadic enforcement. Changes in norms, more than any other factor, have led to positive change in behaviors.

Tackling more fundamental engineering problems is an important first step. These include the need to build systems that clearly communicate security guidance to users and offer recovery assistance when a computer is compromised. Even in the most optimistic theory or authoritarian regime, individuals cannot be given effective incentives to accomplish tasks beyond their capabilities. Technology to communicate with and empower individuals to protect their own digital assets is sorely needed. Yet the business models for such technologies are uncertain, and the interdisciplinary nature of the research is ill suited for receiving aid from traditional funding agencies.

Scholarship in community production and cooperative research management can be combined with security engineering research to create a socially, and thus technologically, resilient Internet. Re-

search and exploration into which challenges are well suited to community-based production and which require centralized coordination is critical. Engineers and institutions can provide decision-makers with the socially aware technical tools to empower families, individuals, and localities to enhance their cybersecurity. Social engineering has enforced top-down solutions, such as the Trusted Identifiers initiative and other designs for intensive monitoring and control. I am advocating for a different approach, one in which engineering courts community rather than seeking to control it.

Identification of the potential and applicability of community production for cybersecurity at home can make a significant contribution to the total social cost of implementing cybersecurity on a national scale. Strengthening the network will require understanding the potential for citizens to self-govern, with regard to the protection of their own home systems and their personal information and identity.

ENDNOTES

[1] Yochai Benkler, "Sharing Nicely: On Shareable Goods and the Emergence of Sharing as a Modality of Economic Production," *The Yale Law Journal* 114 (2) (2004): 273–359.

[2] Jeff Howe, "The Rise of Crowdsourcing," *Wired*, June 2006.

[3] Ronald H. Coase, "The Nature of the Firm," *Economica* 4 (16) (1937).

[4] Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (London: Methuen and Co., 1776).

[5] Chris DiBona, Sam Ockman, and Mark Stone, eds., *Open Sources: Voices from the Open Source Revolution* (Cambridge, Mass.: O'Reilly, 1999).

[6] Josh Lerner and Jean Tirole, "Some Simple Economics of Open Source," *Journal of Industrial Economics* 50 (2) (2002): 197–234.

[7] Ashish Arora, Rahul Telang, and Hao Xu, "Optimal Policy for Software Vulnerability Disclosure," Third Workshop on the Economics of Information Security, Minneapolis, Minnesota, June 2004.

[8] Edward Balas and Camilo Viecco, "Towards a Third Generation Data Capture Architecture for Honeynets," *Proceedings of the Sixth IEEE Information Assurance Workshop*, West Point, New York, 2005.

[9] Yochai Benkler, "Coase's Penguin, or Linux and the Nature of the Firm," *The Yale Law Journal* 112 (2002).

[10] Zheng Dong and L. Jean Camp, "The Decreasing Marginal Value of Evaluation Network Size," *ACM SIGCAS Computers and Society* (forthcoming).

[11] Niklas Luhmann, "Trust: A Mechanism for the Reduction of Social Complexity," in *Trust; and Power: Two Works* (Chichester, N.Y.: John Wiley and Sons, 1979).

[12] Tyler Moore, Richard Clayton, and Ross Anderson, "The Economics of Online Crime," *Journal of Economic Perspectives* 23 (3) (2009): 3–20.

[13] Thomas Dietz, Elinor Ostrom, and Paul C. Stern, "The Struggle to Govern the Commons," *Science* 302 (5652) (2003): 1907.

[14] David D. Clark, "Window and Acknowledgement Strategy in TCP," RFC 813 (Cambridge, Mass.: MIT Laboratory for Computer Science, July 1982).

[15] Moore, Clayton, and Anderson, "The Economics of Online Crime."

16 Michael van Eeten, Johannes M. Bauer, Hadi Asghari, Shirin Tabatabaie, and Dave Rand, "The Role of Internet Service Providers in Botnet Mitigation: An Empirical Analysis Based on Spam Data," Ninth Workshop on the Economics of Information Security, Cambridge, Massachusetts, June 2010.

17 A recent example is Apple's location information compilation; see Brian X. Chen, "Why and How Apple is Collecting Your Location Data," *Wired* blog, April 21, 2009, http://www.wired.com/gadgetlab/2011/04/apple-iphone-tracking (accessed April 29, 2011).

18 Bruce Schneier, "We Don't Spend Enough on Security," First Workshop on Economics and Information Security, Berkeley, California, May 2002.

19 L. Jean Camp and Catherine Wolfram, "Pricing Security," *Proceedings of the CERT Information Survivability Workshop*, Boston, Massachusetts, October 24–26, 2000, 31–39.

20 Eben Moglen, "Anarchism Triumphant: Free Software and the Death of Copyright," *First Monday* 4 (8) (1999).

21 Hal Varian, "System Reliability and Free Riding," *Proceedings of the Fifth International Conference on Electronic Commerce*, ed. Norman Sadeh (New York: Association for Computing Machinery, 2003).

22 William Yurcik, "Cyberinsurance: A Market Solution to Internet Security Market Failure," Workshop on the Economics of Information Security, Berkeley, California, May 16–17, 2002.

23 Brent Rowe, "ISPs as Cybersecurity Providers," The Ninth Workshop on the Economics of Information Security (WEIS 2010), Harvard University, June 7–8, 2010.

24 Samuel Warren and Louis Brandeis, "The Right to Privacy," *Harvard Law Review* 4 (1890): 193–220.

25 NANOG is the organization of ISPs. Members know each others' corporate habits and practices. Not only network standards but also network norms emerge in groups of network operators and engineers.

26 Kirrily "Skud" Robert, "Preliminary Results of My Survey of Suspended Google+ Accounts," InfoTropism, July 25, 2011, http://infotrope.net/2011/07/25/preliminary-results-of-my-survey-of-suspended-google-accounts/.

27 Jens Riegelsberger and Martina Angela Sasse, "Trustbuilders and Trustbusters," in *Towards the E-Society: E-Commerce, E-Business, and E-Government*, ed. Beat Schmid, Katarina Stanoevska-Slabeva, and Volker Tschammer (Boston: Kluwer, 2001).

28 Russell Hardin, *Trust and Trustworthiness* (New York: Russell Sage, 2002).

29 Alessandro Acquisti and Ralph Gross, "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook," in *Privacy Enhancing Technologies*, Volume 4258 of Lecture Notes in Computer Science, ed. George Danezis and Philippe Golle (Berlin; New York: Springer, 2006), 36–58.

30 Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford, Calif.: Stanford University Press, 2009).

31 L. Jean Camp, "Mental Models of Security," *IEEE Technology and Society Magazine* 28 (3) (2009).

32 Lawrence Lessig, *Code and Other Laws of Cyberspace* (New York: Basic Books, 1999).

Resisting Political Fragmentation on the Internet

R. Kelly Garrett & Paul Resnick

Abstract: Must the Internet promote political fragmentation? Although this is a possible outcome of personalized online news, we argue that other futures are possible and that thoughtful design could promote more socially desirable behavior. Research has shown that individuals crave opinion reinforcement more than they avoid exposure to diverse viewpoints and that, in many situations, hearing the other side is desirable. We suggest that, equipped with this knowledge, software designers ought to create tools that encourage and facilitate consumption of diverse news streams, making users, and society, better off. We propose several techniques to help achieve this goal. One approach focuses on making useful or intriguing opinion-challenges more accessible. The other centers on nudging people toward diversity by creating environments that accentuate its benefits. Advancing research in this area is critical in the face of increasingly partisan news media, and we believe these strategies can help.

R. KELLY GARRETT is an Assistant Professor in the School of Communication at The Ohio State University.

PAUL RESNICK is a Professor in the School of Information at the University of Michigan.

*(*See endnotes for complete contributor biographies.)*

It is fashionable to decry a growing fragmentation of political discourse in America.[1] Gone are the days when Americans of all political stripes relied on a common set of media institutions; liberals and conservatives no longer read the same books or watch the same cable TV talk shows. But the best evidence so far, based on actual reader behavior, suggests that ideological segregation on the Internet is limited. Those who read conservative websites such as RushLimbaugh.com are more likely than the average Internet user to visit *The New York Times* online as well; similarly, visitors to liberal websites such as MoveOn.org are more likely than the average Internet user to visit the Fox News website, too.[2] This phenomenon suggests, perhaps surprisingly, that for now online news consumption is less homogeneous than political dialogue within family or friend networks.

What will happen next? Legal scholar Cass Sunstein, among others, argues that the Internet will inevitably make fragmentation worse over time, as

ever-narrower, more personalized channels allow people to see only the news and opinion stories they want to see.[3] For example, readers can look to the online magazine *Newsmax* for a conservative slant on the news of the day, or to *Slate* for a liberal one. Even narrower channels could increase homogeneity still further. At the extreme, the news aggregator Digg.com, which has long selected front-page articles based on readers' votes, now offers a personalized stream of only the articles a reader's designated friends have voted for. The concern is that if readers choose to follow only like-minded friends, they may never see articles that challenge their preexisting opinions. In this essay, however, we turn the "personalization leads to fragmentation" claim on its head by arguing that personalization could instead be a crucial tool for *resisting* fragmentation.

Individual exposure to opposing viewpoints has several societal benefits. First, it increases tolerance for attitudes and beliefs that differ from one's own. Second, there is a natural tendency for people, particularly those in the minority, to think that their views are shared more broadly than they actually are.[4] Given a better assessment of the true popularity of an opinion, individuals may accept the legitimacy of disagreeable outcomes in the political sphere rather than concoct conspiracy theories to explain how their own will, which they presumed to be in the majority, was thwarted. Third, broader experience with diverse views may prevent polarization. A long history of experiments has shown that deliberation on an issue with like-minded people leads to polarization: that is, everyone tends to end up with more extreme views than they started with.[5] According to one explanation for this finding, people in like-minded groups are exposed to arguments on only one side of the issue.[6] Thus, selective exposure to exclusively attitude-reinforcing news and opinion articles might also lead to opinion shifts to more extreme positions, which may make it harder for society to find political consensus on important issues.

In the past, people could not tailor news according to their individual preferences. Mass-audience broadcast channels offered a mix of information that was not perfectly aligned with any one person's views. According to Sunstein, the social benefit of exposing everyone to some challenging opinions is now at risk in an era of narrowcasting and personalization. In other words, as the Internet increasingly allows people to limit their news consumption to that which reaffirms their own views, the underlying conflict between what is good for society and what individuals will naturally choose becomes apparent.

The risks associated with narrow channels and personalization are undeniable. Even without automated news filtering, millions of Americans have already begun to sort themselves into partisan audiences via their use of cable news networks and ideologically oriented websites. It is a mistake, however, to presume that personalization services make further news market fragmentation inevitable. The technology and how people use it are still malleable; subtle architectural changes could have far-reaching implications for future news consumption patterns.

Our claim is not that technology will miraculously transform people, converting closed-minded ideologues into open-minded deliberators; rather, we argue that it can nudge individuals slightly in the direction of exposure to challenging viewpoints and that most people will prefer news services that provide those nudges to ones that do not. For example, news services could prime norms that promote balanced exposure or could

make readers aware of how popular other viewpoints are. Such strategies could prompt modest increases in people's preferences for exposure to challenging information. Moreover, in many cases technology need not alter people's preferences; instead, it may be sufficient to better serve their existing preferences. As we discuss below, considerable evidence suggests that people gravitate toward confirmation without systematically avoiding challenge. In choosing which news items to view, factors such as informativeness and quality often trump viewpoint. Further, some people prefer to see a mix of perspectives, especially when they anticipate the need to defend their positions. News services that present the right challenging items to the right people and in the right contexts have the potential to be very successful.

Creating automated, diversity-enhancing news services that people want to use, however, will require effort and creativity. Without thoughtful intervention, personalized news aggregation services may not evolve to produce the kind of heterogeneous information streams that people would prefer over homogeneity. For example, researchers have made inroads toward automatically identifying the political slant of news content.[7] If these identification systems were used naively to tailor news consumers' information environments, mechanically screening out political information with which they might disagree, then technology would directly curtail exposure to counterattitudinal information. Thus, considerable research and development may be necessary. There is good reason to be hopeful, though, that such R&D efforts will yield services that win in the marketplace.

In 1944, Columbia University sociologist Paul Lazarsfeld and his colleagues at the Bureau for Social Research published *The People's Choice*, a landmark work based on research conducted in Erie County, Pennsylvania, examining voters' activities and attitudes in the lead-up to a presidential election. In the book, which helped lay the foundations of modern political communication scholarship, the authors observe that "people select their exposure along the line of their political predisposition."[8] This simple claim set the stage for a robust debate that continues to this day.

Selective exposure is premised on what social psychologist Leon Festinger termed *cognitive dissonance*, the negative arousal that individuals experience when they encounter anything suggesting that a prior decision has undesirable implications.[9] For example, voters might experience dissonance upon learning that their preferred candidate in an election has behaved unethically because this knowledge raises questions about their judgment. Given that dissonance is unpleasant, individuals tend to avoid it or mitigate its effects. One strategy for doing so is to discriminate among different types of information based on one's attitudes or opinions, seeking information that confirms prior decisions (confirmation bias) or avoiding disconfirming information (defensive avoidance).

A half-century of research predating the widespread adoption of the Internet, however, suggests that selective exposure has only a modest influence on individuals' political information diet.[10] Political attitudes are only one of many factors that influence news consumption, and that influence is relatively modest. General political interest, issue relevance, and information utility often play bigger roles in shaping media exposure. Furthermore, although individuals frequently exhibit a preference for proattitudinal information, there is very little evidence that they avoid counterattitudinal information.[11] These results suggest that selec-

tive exposure is actually the product of two distinct preferences: an attraction to proattitudinal information paired with a much weaker aversion to counterattitudinal information. This is the context in which concerns about Internet-induced political fragmentation emerged.

Cass Sunstein was among the first to decry the threat the Internet poses to democracy. In *Republic.com*, he presents a compelling vision of how people might use their newfound ability to filter political information online, arguing that fragmentation is the most likely result. To support his claim, Sunstein points to a tendency among political websites to link almost exclusively to other websites that share their political orientation, and he examines the consequence of this behavior through the lens of group polarization. In contrast to prior scholarship on the topic of selective exposure, however, he asserts that, given the opportunity, people *will* systematically screen out information and opinions with which they disagree. In the same year that Sunstein's book was published, political scientists Diana Mutz and Paul Martin released an article offering a similar conclusion.[12] Using cross-national survey data and exploiting exogenous variation in the available news sources in different media markets, the authors demonstrate that the more choice people have in their information environments, the more likely they are to be exposed to proattitudinal instead of counterattitudinal information. Observing that choice abounds online, the authors warn that increasing reliance on the medium could pose a threat to healthy political deliberation.

These claims inspired a new generation of selective exposure research. Network scientists Lada Adamic and Natalie Glance provide a thorough analysis of claims Sunstein made about blog-linking patterns, confirming that bloggers disproportionately link to posts and websites that support their viewpoints.[13] Survey data collected during the 2004 election demonstrate that conservative Republicans and liberal Democrats differ in their media preferences.[14] Conservatives are more likely than liberals to use conservative outlets across a variety of media, including newspapers, radio, television, and the Web. Likewise, liberals show a stronger preference than conservatives for liberal outlets. Experiments confirm that when faced with a choice between proattitudinal and counterattitudinal messages, most individuals choose the former.[15] The results of this recent wave of research can have one of two meanings for selective exposure in the Internet era. One possibility is that changes in the media landscape precipitated by new technology have altered the mechanisms underlying the phenomenon, thereby promoting both confirmation bias and defensive avoidance. Political scientists Lance Bennett and Shanto Iyengar make an argument along these lines, suggesting that changes in the media are producing information "stratamentation," in which politically disinterested individuals simply tune out while the politically involved grow more isolated. As a consequence, the authors predict, few people will ever engage with counterattitudinal information.[16] The alternative explanation, which we advance here, is that the results are driven primarily by confirmation bias.

When individuals choose between proattitudinal and counterattitudinal content, as they have in the studies described above, we cannot know whether they are motivated by confirmation bias, defensive avoidance, or both. Although it seems reasonable to assume that people will avoid content they consider to be dangerous or offensive, recent empirical work indicates that people tend to look

for proattitudinal information *without* systematically screening out other perspectives. Following the work of Sunstein and Adamic and Glance, communications scholar Eszter Hargittai and her colleagues offer a nuanced assessment of cross-ideological discussion on political blogs.[17] Consistent with prior research, they observe that both conservative and liberal bloggers are more likely to link to other like-minded blogs. But they also find that links to blogs of the opposing ideology are pervasive. Survey data collected during the 2004 U.S. presidential election show that greater reliance on online news sources promotes familiarity with proattitudinal information without a corresponding decline in counterattitudinal information.[18] A 2005 study examining consumers' perceptions and use of online political content finds that the more proattitudinal information a news story contains, the more likely the individual is to view it; however, the presence of counterattitudinal information does not have a statistically significant influence on selection.[19]

Perhaps most relevant to the objective of promoting exposure to diverse views is the insight that people differ in their preferences for homogeneous versus diverse streams of news. For example, about one-quarter of participants in an online experiment volunteered, without being asked directly, their preference for ideological heterogeneity in the news; consistent with that claim, when an automated news recommendation system presented them with various combinations of liberal and conservative news items on different days, they reported higher satisfaction with more diverse sets.[20] National survey data reveal that about one-third of partisan online news consumers (those who use political blogs or explicitly ideological news outlets) rely on both supporting and opposing partisan outlets.[21]

Precisely which factors contribute to these preferences remains an open question. Other studies show that when forced to choose between pro- and counterattitudinal information, increasing attitude accessibility, attitude importance, and political interest promote counterattitudinal information exposure.[22] Attitude certainty and defensive confidence – that is, certainty in one's ability to justify and maintain a set of beliefs in the face of counterargument – also increase individuals' willingness to engage with counterattitudinal information.[23] A growing body of research suggests that an individual's ideology may play a role as well. Data from experiments indicate that conservatives tend to explore the information environment less thoroughly than individuals holding other ideologies because they more quickly learn to avoid harmful or costly encounters.[24] Similarly, several studies suggest that conservatives tend to be less tolerant of ambiguity or uncertainty and have a higher need for closure than liberals.[25] This is not to say that all conservatives will engage in strategies of avoidance or that all liberals will seek other perspectives; the results merely show that, on average, those inclinations tend to fall along ideological lines.

Furthermore, exposure preferences are highly contingent on social context. Indeed, several factors promote attention to other perspectives. Notably, decision anxiety can make counterattitudinal information especially desirable when it is expected to be useful, as when one must defend a position.[26] Yet individual anxiety, or threat, can produce the opposite effect within some groups. Individuals holding more authoritarian views have a greater aversion to counterattitudinal information the more threatened they feel.[27] Information scarcity also promotes selective exposure: that is, indi-

viduals are more likely to prioritize proattitudinal over counterattitudinal information when their opportunities to gather information are limited; but proattitudinal preference weakens when information is more abundant.[28]

In sum, several decades of research have shed considerable light on the selective exposure phenomenon. People's attitudes certainly influence their exposure to political information, which has implications for how they use the high levels of choice afforded by the Internet. Political viewpoint, however, is only one of a number of factors that shape exposure decisions, and its influence is modest. Furthermore, it is an error to assert that people consistently prefer homogeneous news streams. Strong evidence shows that confirmation bias is the dominant form of selective exposure; defensive avoidance has relatively little effect despite changes in the media environment. In other words, people have a psychological preference for proattitudinal information without a corresponding aversion to counterattitudinal information. Most important, there are numerous individual and contextual factors that lead people to favor counterattitudinal information in particular settings.

The above discussion brings us back to the question of personalization. If people had access to the news streams they most desire, those streams often would include some information that challenges their preexisting opinions. However, an ideologically segmented news environment – exemplified by the cable TV news market today – encourages consumers to construct relatively homogeneous news streams.[29] In a world of many narrow partisan channels, people must choose between sources offering *either* proattitudinal *or* counterattitudinal information because a source offering both is not an option. Faced with this choice, most will choose the proattitudinal source. Although some individuals will make this choice because they deem the alternatives to be offensive or dangerous, it is more often an unintended consequence of confirmation bias in an environment with limited options. That is, in most cases people exclude opposition channels not out of aversion to other opinions but because they offer less benefit than proattitudinal channels. This choice environment comes at a price for the individuals, who would find certain diverse news streams to be more satisfying, and for society at large, which benefits from a well-informed and tolerant public.

In principle, personalized news aggregators should resolve this problem. Automated personalization services track stories that individuals liked in the past and use that information to identify and recommend "similar" items going forward. People can explicitly mark the items they like, or a program can make automated inferences based on which items they choose to view and for how long. If people truly prefer some challenging articles mixed in with proattitudinal information, the system should "learn" that preference.

Personalization technologies that produce desirably diverse news streams may not emerge naturally. The problem is that unsophisticated interpretations of what constitutes similarity can lead to homogeneous collections. For example, if similarity means that two items cover the same topic, then someone who initially reads a few stories about football might end up with all stories about football and miss the excitement of the Tour de France. If it means that items are "liked" by people whose ratings tend to match one's own, as in the recommender systems of Netflix or Amazon, then liberals might only see stories liked by other liberals. More sophisticated notions of sim-

ilarity are required. For example, if someone reads and likes an article that presents new evidence and argues logically in favor of a particular position, similar articles might include those that present new evidence on any topic, those that are argued logically, those that take a particular position, or any combination thereof. In addition, people may favor collections of items that cannot be reduced to their preferences for individual items. A challenging article may be desirable only if accompanied by a supporting one on the same topic. A reader might find a single challenging item interesting and informative but be annoyed to encounter more than one.

There are three ways that personalized information services could be designed to give people challenging information they would like to have access to but might not otherwise get. One is to provide only high-quality challenging items. A second is to provide challenging information only in the context of specific topics of interest. The third is to reduce the cognitive dissonance associated with challenging information by making it easy for people to access counterarguments that support their views whenever they are exposed to these challenges.[30]

Making intriguing counterattitudinal information more accessible could significantly enhance the level of diversity in people's media exposure choices. Suppose, for example, that an individual's rating of an item depended on two elements: a reinforcement score measuring how well the item matched his or her preexisting opinions and a quality score measuring other attributes such as good writing, humor, and novelty. A reader who generally favors reinforcement may prefer a high-quality, non-reinforcing item to a reinforcing item of much lower quality. Imagine reading the user-contributed comments accompanying a newspaper op-ed. Individuals might enjoy reading the sloganeering responses they agree with as well as the thoughtful ones. If forced to read the sloganeering responses from commenters they disagree with, however, they might be turned off sufficiently to stop reading the comments altogether. But what if they could read only the thoughtful opposing comments? For many people, challenging yet insightful remarks would merit some attention, and might even be more attractive than the less-thoughtful agreeable responses.

This scenario is not as far-fetched as it might seem. The website Slashdot.org already allows users to tag comments as "insightful" or "humorous," for example. Readers can convert those tags into scores and hide comments with scores below certain thresholds. Similarly, Digg.com posts user ratings of individual comments. News services that tracked users' political positions, whether self-identified or automatically estimated based on responses to previous items, might be able to determine whether a particular user would agree or disagree with various comments, setting a higher score threshold for disagreeable comments than agreeable ones. The net effect would be to show each person more agreeable than challenging comments, but to expose everyone to some challenging comments nonetheless. It is at least plausible to assume that individuals would prefer this combination to either a service with only agreeable comments or one with the same quality threshold for agreeable and disagreeable comments. The technique might also be used to filter other types of political content, such as op-ed pages or news analyses.

Determining which challenging items most interest individuals will require considerable experimentation. Selection based on attitude-independent quality metrics, as suggested above, is just one

possibility. In other contexts, it may be interesting to track items that are most popular among the opposition. Yet another intriguing option would be to highlight items that have attracted "strange bedfellows," those that are liked by two clusters of people who do not usually agree with each other. As we develop technologies to select news and opinion items as well as reader comments based on these and similar criteria, the most agreeable method of selecting disagreeable items may become clear.

A second approach is to provide access to challenging information only when people are most curious about and open to it. Research has shown that interest and personal relevance trump ideology in the search for political information. Imagine a service that tracks a user's reading behavior over time. When it detects interest in a topic (for example, reading a news article to completion rather than just scanning the first paragraph), especially a topic the user has not explored recently, the service might suggest topically relevant items representing alternative viewpoints. The technology to cluster news articles by topic already exists. (Google News, for example, offers a single headline and summary story, then lists others sources with more in-depth articles on the same topic.) Various experimental techniques have been developed for automatically clustering items based on political viewpoints. In one intriguing study, Korean researchers found that grouping articles on the same topic into separate opinion clusters led readers to explore more diverse viewpoints.[31]

A third approach is to provide challenging information along with supporting information. Cognitive dissonance theory, the basis for predictions of selective exposure, tells us that viewing counterattitudinal information can produce negative emotions if the exposure leads the viewer to feel badly about a prior decision. The more confident people are in the reasonableness of their opinions, the less threatening counterarguments will appear. To boost reader confidence, information services could make it easier to find proattitudinal information following counterattitudinal exposure. If users could easily navigate from articles containing challenging information to opinion-reinforcing items – ideally, items that respond directly to the arguments in the challenging piece – these reminders of the evidence supporting a prior decision may reassure individuals of their correctness when confronted with counterarguments. Alternatively, it may serve as a face-saving opportunity for individuals who are moved to reconsider their position by demonstrating that they are not alone in their beliefs. Admitting to an error is easier when others have made the same mistake. Either way, this feature would reduce the cognitive dissonance of challenging information so that people feel safer exploring. As described above, another approach would augment a stream of like-minded news items with the periodic inclusion of the "best" of the other side, giving individuals both greater opportunity and fewer disincentives to explore challenges to their opinion. Remembering the justification for one's position creates attitude certainty, which increases people's willingness to engage counterattitudinal information without precluding the possibility of attitude change.

In addition to helping people find the challenging information they want on the occasions they want it, innovative technologies can also provide subtle nudges that encourage people to *seek out* more challenging information. Experimental research has consistently demonstrated that exposure to other viewpoints

is highest when individuals have the most to gain from it. When individuals are warned that they will need to defend or justify their positions, they are more likely to seek out counterattitudinal information, a tendency that becomes more pronounced as decision anxiety increases. The higher the cost of being wrong or uninformed, the more effort people put into verifying the accuracy of their positions (so long as they still have the opportunity to act on the new information). Thus, one way to make other perspectives more attractive to news consumers is to provide information about the prevalence of different opinions on an issue. The realization that one's opinion is not widely shared can increase the value of exploring alternatives because it creates awareness of the need to defend that opinion to others in discussion or to justify it to oneself. This approach could be an effective motivator among those whose views are in the majority as well. Realizing that one's opinion is shared by many should reduce the costs of exploring alternatives: as noted above, attitude certainty makes exposure to counterattitudinal information feel safer. Although it is not clear exactly how a news and opinion aggregation service could integrate polling and reader feedback information into its displays of news articles, this is an area that seems ripe for experimentation.

Another way to nudge people toward consuming challenging information is to accentuate the benefits to self-image that accrue from engaging in counterattitudinal exposure. Most people believe that exposure to a range of political opinions is a good thing. People tune in to political debates and talk shows that highlight opposing perspectives.[32] Individuals at both ends of the political spectrum are unhappy with news media that they perceive to be partisan, and a majority of Americans say they prefer political news sources that do not advocate a particular point of view.[33] Diversity has even been shown to influence perceptions of credibility in some contexts. For example, when people assess the quality of an unfamiliar online information source, they typically rely on cognitive heuristics, mental shortcuts that allow them to decide whether to trust the content. One important heuristic concerns the diversity of views included. A source that offers only one point of view – that is, one whose contributors are all in agreement – is considered less credible, meaning that readers are less likely to trust it.[34] By extension, this finding suggests that political sources explicitly advocating one position without considering others, and without allowing room for dissent, are likely to be viewed skeptically by many.

Normative expectations about diverse exposure apply to consumers as well as producers. In his commencement address at the University of Michigan in 2010, President Obama asked graduates, "How will you keep our democracy going?" Part of the answer, he said, was to "actively seek out information that challenges our assumptions and our beliefs." A quick perusal of online comments from conservative readers of *USA Today* and *Wall Street Journal* articles about the speech confirms the normative status of the message (even though many commentators argued that the messenger was hypocritical for delivering it).

For those who accept that attention to other perspectives is normative, user interfaces can be developed to prime individuals to think about this expectation when they approach the news. For example, we are launching an experiment that adds a simple feedback mechanism to a news aggregator, providing feedback on how many red (conservative) and blue (liberal) articles the user has read recent-

ly.[35] Called "Balance," the tool is designed to encourage exposure to challenging viewpoints: it features a cartoon figure walking a tightrope; if a user's recent reading history is out of balance, the tightrope walker leans precariously to one side.

More research is needed to fully understand how best to prime normative expectations for diverse news gathering, but other intriguing possibilities have been identified. For example, scholars have observed that ideology influences individuals' responses to persuasive messages. Whereas a "benefits" frame effectively motivates liberals, a "loss" frame is more successful among conservatives. Thus, stating that "balance will produce better decisions and is good for society" might move liberals to act, while "if you don't know what the other side is saying, you won't be able to refute their arguments" may be a more effective nudge for conservatives.[36]

Social comparisons can further leverage the desire to conform to the norm of balance. For example, informing people when their viewing histories are less balanced than other users' may trigger a desire to catch up. For some subset of the population, the tracking idea can be taken further and turned into a game in which users compete to accumulate points. Just as Internet-based diet and exercise trackers have turned self-improvement into a competitive game and the mobile application Foursquare has had the same effect on regular attendance at favorite bars and restaurants, a "challenge yourself" application could allow people to earn points for reading challenging information or talking about politics with people they do not know.

We disagree with critics who argue that the Internet inevitably threatens diverse exposure and that society will suffer as a consequence. This outcome is just one of many possibilities. Indeed, we think that these technologies could expose people to a combination of news and opinion pieces that, if selected and presented well, would expand the diversity of the information they receive. The danger will come not from an inherent human desire to filter out other viewpoints; confirmatory information is attractive, but not to the exclusion of everything else. Instead, the threat will come from narrow channels and crude personalization techniques that fail to meet people's true preferences.

We have articulated a number of promising directions for research and development of more sophisticated personalization techniques that could potentially increase the benefits of diverse exposure and help people assemble and access challenging information they will be interested in and receptive to. These strategies include presenting challenging information only if it exceeds a high bar on criteria such as quality and relevance; offering challenging information alongside confirmatory information; providing an opposing view only when people are most open to it; informing people about the prevalence of challenging opinions; and reinforcing the norm of balanced exposure. More sophisticated personalization services based on these approaches could promote more diverse exposure despite the (re)emergence of partisan news media.

ENDNOTES

* Contributor Biographies: R. KELLY GARRETT is an Assistant Professor in the School of Communication at The Ohio State University. His publications include "Politically Motivated Reinforcement Seeking: Reframing the Selective Exposure Debate," *Journal of Communication* (2009); "E-Democracy Writ Small: The Impact of the Internet on Citizen Access to Local Elected Officials" (with Michael Jensen), *Information, Communication & Society* (2011); and "Troubling Consequences of Online Political Rumoring," *Human Communication Research* (2011).

PAUL RESNICK is a Professor in the School of Information at the University of Michigan. He pioneered the development of personalized recommender systems in the early 1990s, work recognized by the 2010 Association for Computing Machinery Software Systems Award. His publications include "Classifying the Political Leaning of News Articles and Users from User Votes" (with Daniel Xiaodan Zhou and Qiaozhu Mei), *Proceedings of the Fifth ICWSM Conference on Weblogs and Social Media* (2011); "Presenting Diverse Political Opinions: How and How Much" (with Sean Munson), *Proceedings of CHI 2010*; and *Building Successful Online Communities: Evidence-Based Social Design* (with Robert Kraut; forthcoming from MIT Press, 2011).

[1] This material is based on work supported by the National Science Foundation under Grant No. IIS-0916099.

[2] Matthew Gentzkow and Jesse M. Shapiro, "Ideological Segregation Online and Offline," *Quarterly Journal of Economics* (forthcoming).

[3] Cass R. Sunstein, *Republic.com* (Princeton, N.J.: Princeton University Press, 2001); see also Eli Pariser, *The Filter Bubble: What the Internet is Hiding from You* (New York: Penguin Press, 2011).

[4] Glenn S. Sanders and Brian Mullen, "Accuracy in Perceptions of Consensus: Differential Tendencies of People with Majority and Minority Positions," *European Journal of Social Psychology* 13 (1983); Lee Ross, David Greene, and Pamela House, "The 'False Consensus Effect': An Egocentric Bias in Social Perception and Attribution Processes," *Journal of Experimental Social Psychology* 13 (1977).

[5] Cass R. Sunstein, *Going to Extremes: How Like Minds Unite and Divide* (New York: Oxford University Press, 2009), 199.

[6] Eugene Burnstein and Amiram Vinokur, "Persuasive Argumentation and Social Comparison as Determinants of Attitude Polarization," *Journal of Experimental Social Psychology* 13 (1977).

[7] Miles Efron, "The Liberal Media and Right-Wing Conspiracies: Using Cocitation Information to Estimate Political Orientation in Web Documents," paper presented at the Association for Computing Machinery Thirteenth Conference on Information and Knowledge Management, Washington, D.C., November 2004; Michael Gamon et al., "BLEWS: Using Blogs to Provide Context for News Articles," paper presented at the Second Association for the Advancement of Artificial Intelligence Conference on Weblogs and Social Media, Seattle, Washington, April 2008.

[8] Paul F. Lazarsfeld, Bernard Berelson, and Hazel Gaudet, *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign* (New York: Duell, Sloan and Pearce, 1944), 164.

[9] Leon Festinger, *A Theory of Cognitive Dissonance* (Stanford, Calif.: Stanford University Press, 1957).

[10] David O. Sears and Jonathan L. Freedman, "Selective Exposure to Information: A Critical Review," *Public Opinion Quarterly* 31 (1967); Dieter Frey, "Recent Research on Selective Exposure to Information," *Advances in Experimental Social Psychology* 19 (1986).

[11] Steven H. Chaffee, Melissa N. Saphir, Joseph Graf, Christian Sandvig, and Kyu S. Hahn, "Attention to Counter-Attitudinal Messages in a State Election Campaign," *Political Communication* 18 (2001).

[12] Diana C. Mutz and Paul S. Martin, "Facilitating Communication across Lines of Political Difference: The Role of Mass Media," *American Political Science Review* 95 (2001).

[13] Lada Adamic and Natalie Glance, "The Political Blogosphere and the 2004 U.S. Election: Divided They Blog," paper presented at the Second Annual Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics, Chiba, Japan, 2005.

[14] Natalie Jomini Stroud, "Media Use and Political Predispositions: Revisiting the Concept of Selective Exposure," *Political Behavior* 30 (2008).

[15] Charles S. Taber and Milton Lodge, "Motivated Skepticism in the Evaluation of Political Beliefs," *American Journal of Political Science* 50 (2006); Silvia Knobloch-Westerwick and Jingbo Meng, "Looking the Other Way: Selective Exposure to Attitude-Consistent and Counterattitudinal Political Information," *Communication Research* 36 (2009).

[16] W. Lance Bennett and Shanto Iyengar, "A New Era of Minimal Effects? The Changing Foundations of Political Communication," *Journal of Communication* 58 (2008).

[17] Eszter Hargittai, Jason Gallo, and Matthew Kane, "Cross-Ideological Discussions among Conservative and Liberal Bloggers," *Public Choice* 134 (2008).

[18] R. Kelly Garrett, "Politically Motivated Reinforcement Seeking: Reframing the Selective Exposure Debate," *Journal of Communication* 59 (2009).

[19] R. Kelly Garrett, "Echo Chambers Online?: Politically Motivated Selective Exposure among Internet News Users," *Journal of Computer-Mediated Communication* 14 (2009).

[20] Sean A. Munson and Paul Resnick, paper presented at CHI 2010: Association for Computing Machinery Conference on Human Factors in Computing Systems, Atlanta, Georgia, April 10–15, 2010.

[21] The data are unpublished. For survey details, see R. Kelly Garrett, "Troubling Consequences of Online Election Rumoring," *Human Communication Research* 37 (2011).

[22] Knobloch-Westerwick and Meng, "Looking the Other Way."

[23] Dieter Frey, "Recent Research on Selective Exposure to Information"; Dolores Albarracín and Amy Mitchell, "The Role of Defensive Confidence in Preference for Proattitudinal Information: How Believing That One Is Strong Can Sometimes Be a Defensive Weakness," *Personality and Social Psychology Bulletin* 30 (2004).

[24] Natalie J. Shook and Russell H. Fazio, "Political Ideology, Exploration of Novel Stimuli, and Attitude Formation," *Journal of Experimental Social Psychology* 45 (2009).

[25] David M. Amodio, John T. Jost, Sarah L. Master, and Cindy M. Yee, "Neurocognitive Correlates of Liberalism and Conservatism," *Nature Neuroscience* 10 (2007); John T. Jost, Jack Glaser, Arie W. Kruglanski, and Frank J. Sulloway, "Political Conservatism as Motivated Social Cognition," *Psychological Bulletin* 129 (2003).

[26] Nicholas A. Valentino, Antoine J. Banks, Vincent L. Hutchings, and Anne K. Davis, "Selective Exposure in the Internet Age: The Interaction between Anxiety and Information Utility," *Political Psychology* 30 (2009).

[27] Howard Lavine, Milton Lodge, and Kate Freitas, "Threat, Authoritarianism, and Selective Exposure to Information," *Political Psychology* 26 (2005).

[28] Peter Fischer, Eva Jonas, Dieter Frey, and Stefan Schulz-Hardt, "Selective Exposure to Information: The Impact of Information Limits," *European Journal of Social Psychology* 35 (2005).

[29] The prospect of a partisan press is not altogether new. To the contrary, sociologist Michael Schudson notes that the norm of journalistic objectivity is a relatively recent invention, dating back to the early twentieth century. The intervening years may be the exception, making a return to more propagandistic journalism likely. See Michael Schudson, "The Objectivity Norm in American Journalism," *Journalism* 2 (2001).

[30] These strategies are focused on media exposure and do not address fragmentation due to homophilous interpersonal networks. Although there may be analogous ways of addressing homophily, they will not be identical because the underlying motivations are distinct. Most notably, many people *do* have a strong aversion to interpersonal disagreement.

[31] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song, *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, Boston, Massachusetts, 2009.

[32] Michael X. Delli Carpini, Fay Lomax Cook, and Lawrence R. Jacobs, "Public Deliberation, Discursive Participation, and Citizen Engagement: A Review of the Empirical Literature," *Annual Review of Political Science* 7 (2004).

[33] Andrew Kohut, Carroll Doherty, Michael Dimock, and Scott Keeter, *Ideological News Sources: Who Watches and Why* (Washington, D.C.: The Pew Research Center for the People & the Press, 2010).

[34] Miriam J. Metzger, Andrew J. Flanagin, and Ryan B. Medders, "Social and Heuristic Approaches to Credibility Evaluation Online," *Journal of Communication* 60 (2010).

[35] Items are automatically classified based on the political leanings of the people who voted to promote the article to the website's front page.

[36] Howard Lavine et al., "Threat, Authoritarianism, and Voting: An Investigation of Personality and Persuasion," *Personality and Social Psychology Bulletin* 25 (1999).

Who Speaks? Citizen Political Voice on the Internet Commons

Kay Lehman Schlozman, Sidney Verba & Henry E. Brady

Abstract: Using an August 2008 representative survey of Americans conducted by the Pew Internet & American Life Project, we investigate the consequences of Internet-based political activity for long-standing patterns of participatory inequality. There is little evidence of change in the extent to which political participation is stratified by socioeconomic status, even when we account for the fact that the well educated and affluent are more likely to be Internet users. However, because young adults are much more likely than their elders to be comfortable with electronic technologies and to use the Internet, the Web has ameliorated the well-known participatory deficit among those who have recently joined the electorate. Still, among Internet users, the young are not especially politically active. How these trends play out in the future depends on what happens to the current Web-savvy younger generation and the cohorts that follow as well as on the rapidly developing political capacities of the Web.

KAY LEHMAN SCHLOZMAN, a Fellow of the American Academy since 2003, is the J. Joseph Moakley Endowed Professor of Political Science at Boston College.

SIDNEY VERBA, a Fellow of the American Academy since 1968, is the Carl H. Pforzheimer University Professor Emeritus and a Research Professor of Government at Harvard University.

HENRY E. BRADY, a Fellow of the American Academy since 2003, is Dean of the Goldman School of Public Policy and Class of 1941 Monroe Deutsch Professor of Political Science and Public Policy at the University of California, Berkeley.

*(*See endnotes for complete contributor biographies.)*

From the Greek agora to the Habermasian public sphere, the public commons is a space, open to all citizens, where political discourse and contestation take place; where citizens gather to discuss and possibly influence public policy; where they inform each other about relevant facts and share and debate their preferences. In the ideal commons, discussion is open and civil and essential to democracy. The public commons takes many forms, ranging from small-scale gatherings in a town meeting to national election campaigns that engage millions. The Internet has added a new commons, a virtual space for citizen communication. The novel properties of the Internet raise many questions: Is the political information on the Internet accurate? Does the Internet encourage understanding among those with different views? Does it create community? In short, what does it mean for democracy?

We are concerned not so much with the nature and quality of the discourse taking place in the Internet commons as with who participates in that discourse. In its earliest incarnation, the Greek agora,

the commons was open only to a limited set of Athenians. Similarly, for much of our history, full participation in the American political commons was denied to many – in particular, women and African Americans. Although access is more nearly universal now, many are excluded by their youth, incarceration, or immigrant status, and still others take little or no part. More specifically, our question is whether – compared to traditional, pre-Internet modes of expression of citizen political voice – the virtual commons on the Internet makes opportunities for public discourse more egalitarian in terms of who takes part.

The Internet has generated a great deal of discussion about its consequences for democratic equality. Observers have claimed that "[t]he Internet changes everything,"[1] that it functions as "the great equalizer"[2] and as "our last, best chance to rekindle the great American dream."[3] According to some commentators, the Internet permits ordinary citizens to short-circuit political elites and deal directly with one another and with public officials; to encourage deliberation, enhance trust, and create community[4]; and – of special concern to us – to facilitate political participation.

The following example, one of many that could be culled from the press, illustrates the Internet's promise in creating networks for organized political action. The Help America Vote Act, passed in response to the irregularities associated with the 2000 election, resulted in the replacement of old-fashioned punch card and lever voting systems with optical scan and Direct Record Electronic (DRE) systems. Then, beginning in 2003, an Internet-based movement among computer scientists led to questions about the security of electronic voting systems and potential for electronic corruption of DREs. Skeptics established websites about the issue and then moved into more traditional forms of advocacy in opposition to paperless electronic systems. By 2007, twenty-seven states had adopted provisions mandating a paper trail.[5] This story, a textbook example of a jointly concerned group of citizens working together to have an impact on government, has been used as evidence of the positive consequences of the Internet for democracy.

Most political scientists who study the impact of the Internet on politics have been cautious in their assessments of its implications. According to one such perspective, "[F]ar from revolutionizing the conduct of politics and civic affairs in the real world . . . the Internet tends to reflect and reinforce the patterns of behavior of that world" and constitutes "politics as usual conducted mostly by familiar parties, candidates, interest groups, and news media."[6] In fact, what is known about political participation renders the political success of what began as an Internet-based movement among computer professionals as not fully unexpected. While computer nerds have hardly been the most active group in American politics, they have characteristics – in particular, high levels of education – that predispose them to take part in politics should the occasion arise. Not all citizens bring such advantages to political participation on the Internet.

Our interest in studying the role of the Internet in the democratic functioning of the commons grows out of our long-term concern with the issue of political equality: the ideal – though not the reality – of equal voice for each citizen in political matters.[7] Citizens in American democracy who wish to have an impact on politics can choose from a variety of options for exercising political voice; they can act on their own, with others, or in formal organizations. Working individually or collectively, they can communicate their con-

cerns and opinions to policy-makers in order to have a direct effect on public policy, or they can attempt to affect policy indirectly by influencing electoral outcomes. They can donate their time or their money. They can use conventional techniques or protest tactics. They can work locally or nationally. They can even have political input as the unintended by-product when, for reasons entirely outside politics, they affiliate with an organization or institution that is politically active. Through their political activity, citizens communicate information to public officials about their political opinions and priorities and generate pressure on public officials to pay attention.

One of the basic principles of democracy is the equal consideration by the government of the preferences, concerns, and needs of all citizens. The ideal of equal political voice embodied in the principle of "one person, one vote" is never fully achieved in any democracy, but the deviation from political equality is larger in the United States than in other developed democracies.[8]

Political voice is stratified on many bases, including income and education, race or ethnicity, age, and gender. Our focus is on socioeconomic status (SES), the advantage or disadvantage (based on a combination of income and education) that underlies many forms of inequality in American society. In our investigations, we have been struck by the power and durability of SES-based inequalities in political voice.[9] Not only are participatory inequalities deeply rooted in American institutions and practices but they are persistent, dating back at least the half-century for which we have systematic evidence and presumably longer.

Our past work has demonstrated the multiple ways that SES is associated with various kinds of political activity.[10] Those who are affluent and, especially, well educated are in many ways more likely to be motivated and able to take part in politics. They are, for example, more likely to be politically interested and informed; to think that they can make a difference if they take part; to have the kinds of jobs that develop the communications and organizational skills that facilitate activity in politics; to be actively engaged in religious institutions and nonpolitical organizations and, thus, to have further opportunities to develop such civic skills and greater exposure to a variety of political cues; to have the financial wherewithal to make contributions to campaigns and other political causes; to be located in the social networks through which requests for political activity are mediated; and to be asked by others to take part. In sum, all the factors that foster political participation have roots in socioeconomic circumstances.

We were, thus, concerned as to whether the Internet – which has been so transformative in many other ways – might have the effect of promoting equal citizen voice in politics. In this essay, we draw on a large-scale study of political behavior to assess the impact that opportunities for online political participation have on the stratification of political voice.[11] The survey, which was conducted during the presidential campaign in August 2008, provides a unique opportunity to consider whether online political activity – including newer forms of online activity on blogs and social networking sites – has the possibility of remedying the inequalities of political voice so characteristic of traditional, offline participation. Does the Internet bring new people into politics? Even if the Internet is effective in generating political participation, do the activists simply duplicate the participatory inequalities among offline participants? Or is the Internet bringing new kinds of people into political activity?

We must caution at the outset that, in common with all studies of the impact of the Internet on some aspect of democratic politics, we are dealing with a phenomenon that is very much a moving target – a technology that is, according to political scientist Matthew Hindman, "in its adolescence."[12] In view of our focus on inequalities in political participation, a second possible objection – that the survey on which we rely was conducted at a particular point during the campaign, in August 2008, after the parties had selected their presidential candidates but before the candidates had been officially nominated and before the campaign was in its final autumn sprint – might, in fact, not be cause for concern. The 2008 presidential campaign had unusual characteristics, including significant activity by younger adults and a candidate who made self-conscious efforts to incorporate the Internet into his campaign. However, Obama's experience as a community organizer and his obvious appeal to the young and to persons of color could imply that this August 2008 survey more likely understates rather than overstates the extent of class- and age-based participatory inequalities.

In many ways, the Internet makes it easier to be active in politics. The Internet contains a wealth of political information: from the press, government offices, public officials, and interest organizations. Many kinds of political activity are faster and more efficient online. Making political contributions requires a credit card but neither envelope nor stamp. It is possible to contact large numbers of people quickly and cheaply with a persuasive message or a request to sign a petition, attend a protest, or take some political action. These capacities of the Internet have the potential to increase the numbers of political activists.[13] Our concern, however, is not with the consequence of the Internet on the *amount* of citizen activity, but with the *equality* of citizen voice. Even if it were unambiguous that Internet use increases political participation, a higher level of political participation does not necessarily imply a less unequal distribution of political activity. While we often associate the use of the Internet as a tool of citizen activation with emergent groups and underdog candidates operating on a shoestring, such use is now common among established as well as emergent interests. In short, if any increase in political participation derives from the same people, or the same kinds of people, who are already active, then a possible consequence of the process is to replicate or even exacerbate existing political inequalities.[14]

For more than a decade, social observers have been concerned that the "digital divide" is leaving behind a substantial portion of the public – with implications for equal opportunity in economic life and equal voice in political life. Although the metaphor of the digital divide originally referred to lack of hardware access and suggested a chasm separating cyber haves from the cyber have-nots, it is more appropriate to think of a continuum ranging from, at one end, those who have no Internet access or experience to those, at the other, who have broadband access at home, use the Internet frequently, and are comfortable with a variety of online techniques.[15] Using the Internet to learn about politics and to be politically active requires not simply access to hardware but an array of skills: the capacity both to operate the computer and to seek and understand political information on the Web.[16] But what is critical for our concern with participatory inequalities is not simply that some Americans have been left out of the technological advances of recent decades but that the contours of the

digital divide hew so closely to the socioeconomic stratification that is characteristic of political activity in the United States.[17]

Data from the 2008 Pew Internet & American Life survey that provide the basis for our analysis confirm the unevenness in access to the Internet. Reflecting patterns that have emerged from earlier studies, these data show that the attributes associated with access to hardware are in many ways familiar ones that, in important respects, track the socioeconomic class stratification that has such powerful implications for equal political participation.[18] Roughly half of those in the lowest income category (family incomes below $20,000 in 2007) are online; that is, they send or receive email or otherwise use the Internet at least occasionally. In contrast, at least occasional Internet or email use is nearly universal among those in the highest income category (family incomes of $150,000 or more in 2007). Similarly, only 38 percent of those who did not graduate from high school, compared to 95 percent of those with at least some graduate education, are online.

In terms of the Internet's political capacities for providing opportunities for participation, access to information, and requests for activity, there is a difference between having Internet access at home and elsewhere – say, at work or the local library. In addition, even for those with Internet at home, there is a difference between dial-up and broadband access. The Pew data indicate that in 2008, three-quarters of those who were online – or 56.5 percent of all respondents – had high-speed Internet at home. Once again, there is a sharp socioeconomic gradient: 30 percent of those in households with annual incomes below $20,000, compared to 88 percent in households with annual incomes above $150,000, reported having high-speed Internet access at home; the analogous figures for education are 22 percent for respondents who did not finish high school, as opposed to 81 percent for those with education beyond college. If we assume, not unreasonably, that a high-speed connection at home is an important resource for political engagement, it is interesting to note that those in the top quintile of SES are four times as likely as those in the bottom quintile to have such a connection.[19]

Beyond access to and skillful use of the Internet is the inclination to use it for political purposes. The overwhelming share of Internet use is for nonpolitical activities that range from finding directions to viewing pornography to keeping up with others via a social networking site. Studies of political participation make clear that the predisposition to devote leisure time – that is, time not spoken for by obligations at home, school, or work – to political activity is structured by both age and SES. We were suspicious that, beyond the demographic bias in access to hardware, online political participation might not function to redefine the kinds of people who are active politically but might instead reproduce the widely acknowledged stratification in offline participation.[20]

The 2008 Pew survey provides a unique opportunity to investigate whether political participation on the Internet overcomes the representational biases that have long been observed as characterizing offline political activity. The survey asked about a series of political activities. Five of these can be performed either online or offline: contacting a national, state, or local government official; signing a petition; sending a "letter to the editor" to a newspaper or magazine; communicating with fellow members of a political or community group; and making a political contribution. We constructed two ac-

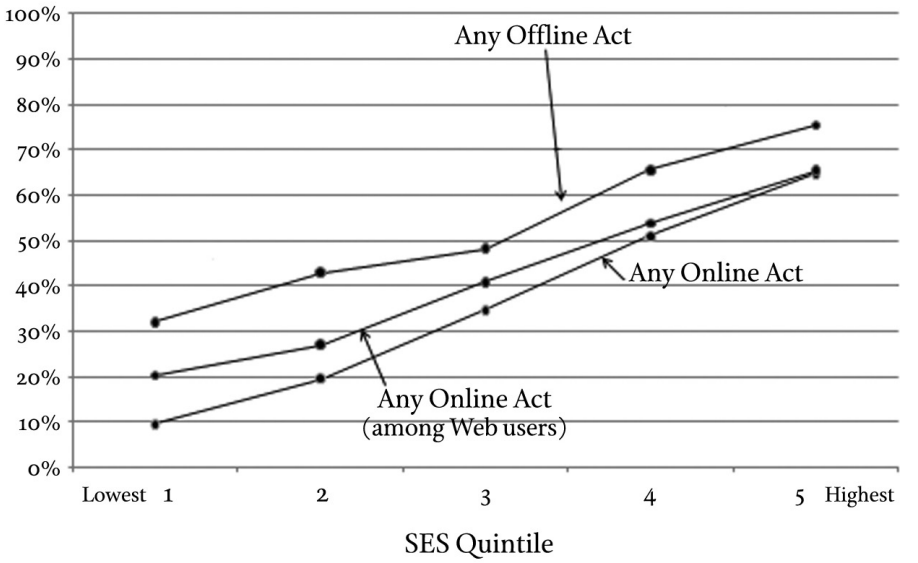
tivity scales measuring either online or offline activity in the counterpart acts.[21]

Figure 1 presents data about the percentage of citizens, divided into five groups based on SES, who engage in at least one participatory act offline and the percentage who engage in at least one act online. The top line shows the proportion who undertake at least one of the five activities *offline*, the bottom line the proportion who undertake at least one *online*. Across the five SES quintiles, offline activity is more frequent than online activity. But more relevant to our concern is the fact that there is a steep slope upward regardless of whether the activity is offline or online. What is more, the gap between the bottom and top socioeconomic categories is greater for online than for offline paricipation. Clearly, whether political activity is traditional or Internet-based, it rises sharply with SES.

The middle line on Figure 1 shows the proportion engaging in at least one of the five online political activities *among Web users*, that is, those who use the Internet or email at least occasionally. The upward slope of the line for online activity among those with Internet access makes clear an important point: lack of access is only part of the story of the SES structuring of online political activity. Even omitting those who are not online and considering only those who use the Internet or email, we see a strong association between political participation and SES. Note also that the difference in activity between Internet users and those not connected is visible at the bottom of the SES scale and disappears at the top, indicating the double barrier to those with low SES. While lack of access to the Internet obviously makes their online political activity impossible, those who lack Internet access would not necessarily use it for political activity if they were to get connected. Still, the digital divide presumably depresses levels of online political activity for those at the lower end of the SES ladder. In contrast, at the upper end, where Internet use is nearly universal, the level of online activity is not affected by lack of access to hardware. Thus, far from acting as a great equalizer, the possibility of political activity on the Internet reproduces long-standing patterns of SES stratification not only because the digital divide has an SES component but because the SES-disadvantaged among those online are not using the Internet for political participation.

Because making political contributions is the form of political activity most obviously dependent on access to financial resources – which are distributed unequally across citizens – and because a great deal of attention has been paid to the success of some candidates in raising large numbers of small donations over the Web, we were particularly interested to look more carefully at political giving. The Pew data contain helpful items about political giving that allow us to ascertain not only whether but also how much respondents gave in political contributions, both offline and on the Web.[22] These data show that Internet contributions are less common than offline donations: 6 percent of respondents made an online contribution, with 15 percent making an offline one. They also suggest that behind the widely discussed success of Internet-based fundraising in collecting political money in smaller amounts is a more complex pattern. On one hand, the average offline contribution is larger than the average online contribution. On the other hand, the percentage of contributions that were $50 or less – 38 percent for online and 39 percent for offline – is virtually identical as is the proportion of contributions that were between $51 and $100: 28 percent for online and 29 percent for offline.

Figure 1
Percent Engaged in Political Activity, by Socioeconomic Status (SES)



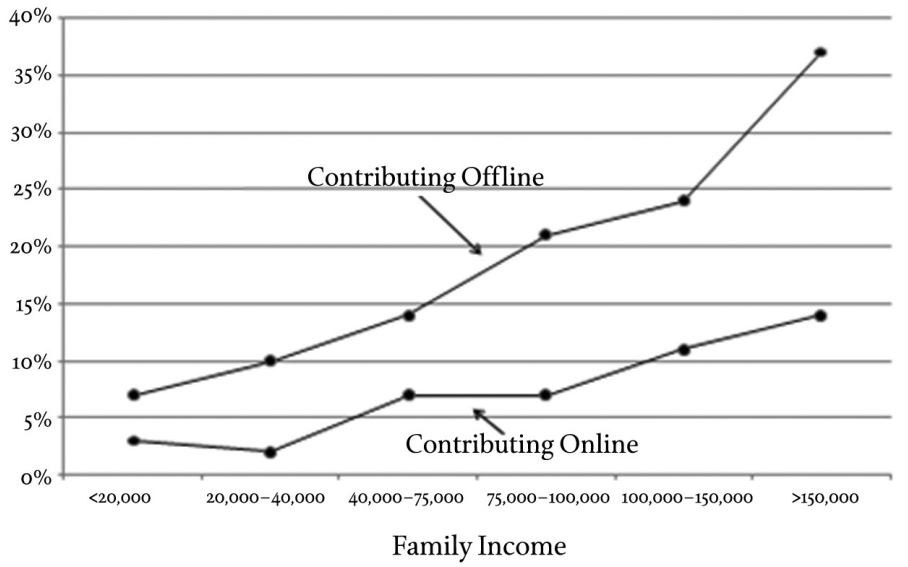
Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project. "Web users" are defined here as those who use the Internet or email at least occasionally.

The very large donations that figure so importantly in campaign war chests are much less likely to come via the Web: less than 1 percent of the online contributions – as opposed to nearly 5 percent of the offline contributions – were for amounts greater than $1,000. We are not certain why donors who give large gifts are less likely to use the Internet. Perhaps, out of security concerns, they are reluctant to enter a credit card number attached to a large donation on the Web. Or perhaps major donors like to be invited to events where they can rub elbows with politicos and celebrities, or they like to contribute in such a way as to allow a friend or political ally to get credit for the donation.[23]

But what about the contributors? Does the Internet encourage donations from less affluent donors? Figure 2, which presents data about the proportion of respondents in various family income groups who make political contributions, shows a familiar pattern.[24] Regardless of whether we consider offline or online political donations, the share of respondents who contribute rises sharply with family income and is more than five times greater in the highest family income group than in the lowest.

Many analysts of campaign finance emphasize expanding the ranks of small donors as the solution to the conundrum of money in democratic politics. Because small donations are unlikely to arrive with a set of policy instructions attached and can exercise limited leverage even when they do, small donations seem to ameliorate the possibilities for compromise of political equality in a campaign finance system that relies heavily on contributions from individuals. Hence, it is noteworthy that even those who made what would seem to be very small donations of $50 or less in 2008 were relatively unlikely to be drawn from the lower rungs of

Figure 2
Percent Making a Campaign Contribution, by Family Income



Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project.

the income ladder, a regularity that characterizes online as well as offline donors.[25] If anything, online donors were somewhat better off financially: 18 percent of all the respondents in the Pew survey – compared to 21 percent of those who made small contributions offline and 29 percent of those who made such contributions online – reported family incomes over $100,000. Thus, at least in the time period covered by the Pew data, the Internet seems to have brought in more small donors but not to have brought in a less affluent set of small donors.

Social movements fascinate precisely because they are not politics as usual and because they hold out the promise of mobilizing outsiders who would not otherwise take part in politics. Processes of mobilization are indeed potent for generating political participation, and social movements often bring into politics previously quiescent publics, thus diminishing inequalities of political voice. Still, much more common than mobilization through social movements are the processes through which neighbors, workmates, and fellow organization and church members ask one another to take part in politics. A great deal of political activity occurs in response to such ordinary processes of recruitment. However, those who seek to get others involved in politics act as rational prospectors, directing their requests at people with characteristics that make it likely that they will assent when asked and that they will be effective when they take part.[26] The result of rational prospecting is to exaggerate existing participatory biases – including the class stratification of political activity – rather than to ameliorate them. Those who take part in response to requests from others are even better educated and more affluent than those who participate at their own initiative.

What happens if recruitment is via the Internet? The Internet provides a number of modalities – of which email and social networking sites are, at present, especially prominent – that make it nearly costless to multiply the number of specially crafted messages to selected publics. In fact, the level of Internet-based political recruitment has already expanded to nearly the same level as offline recruitment: 29 percent of our respondents indicated that they receive an email and 35 percent that they receive a phone call at least once a month asking them to get involved politically.

Figure 3, which reports requests for political activity that come by phone or by email, allows us to compare offline and online recruitment with respect to the extent to which it is structured by SES. Regardless of whether the request arrives by phone or email, the probability that a respondent reports a request for political activity rises steadily with SES. In fact, the curve is much steeper for email requests for political participation. The pattern shows that political recruitment exacerbates the class-based inequality in political activity; inequality is even more pronounced when requests arrive over the Internet.

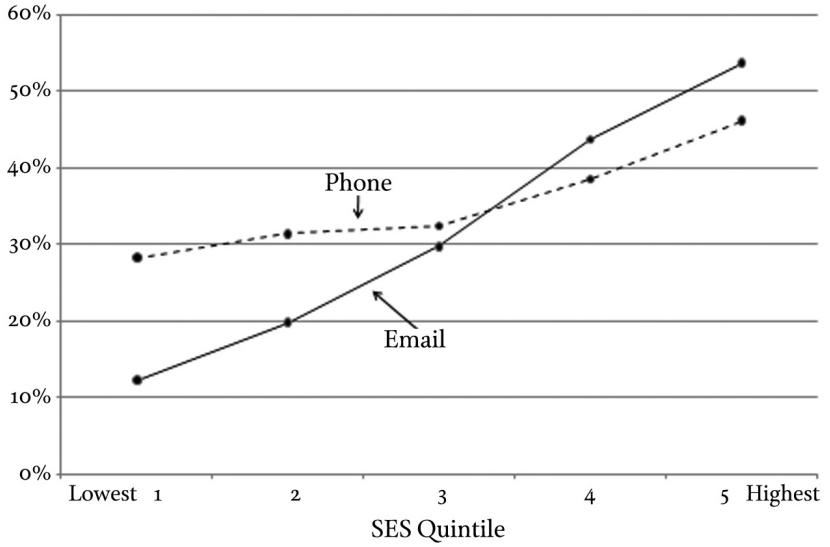
We have seen no evidence that class-based inequalities of political voice are reduced when political participation is online rather than offline. However, we find a very different pattern when it comes to age. The young are more likely than their elders to use the Internet. Every study of Internet access and use, no matter what the measure, shows a steady, sharp decline with age. In the 2008 Pew survey, 88 percent of those who were between the ages of eighteen and twenty-four reported using the Internet, and 70 percent reported having a high-speed connection at home. In contrast, the figures for those in their fifties are 76 percent and 56 percent and, for those over sixty, 44 percent and 29 percent, respectively.

The generational component to Internet use suggests that, unlike the case with SES, the age profile of political activists will not be the same for offline and online activity. Figure 4, which shows, for each of seven age groups, the percentage who undertake at least one participatory act online and at least one offline, confirms that suspicion. Consider the top line, which shows, for offline political participation, a pattern long known to characterize the relationship between age and political activity: a roughly curvilinear trajectory over the life cycle. Political participation starts at a relatively low level, rises with age, peaks among those in their fifties, and falls off among the sixty-somethings and those over seventy. Still, age is much less powerful in structuring political activity than is SES: the gap in participation between the most and least active of the seven age groups is much smaller than the gap between the lowest and highest of the SES quintiles.

When it comes to online activity – shown in the bottom line for all respondents, regardless of whether they are Internet users – the difference between the youngest group and the middle-aged is relatively small, much smaller than for offline activity. For those under sixty, there is little relationship between age and online political activity. However, those who are over sixty are considerably less likey than those in any of the younger age groups to undertake any political activity online. In contrast to what we observed for offline political activity, the absence of online activity among the elderly represents, we assume, not a fall-off from previous Internet-based participation, but instead a "never was."

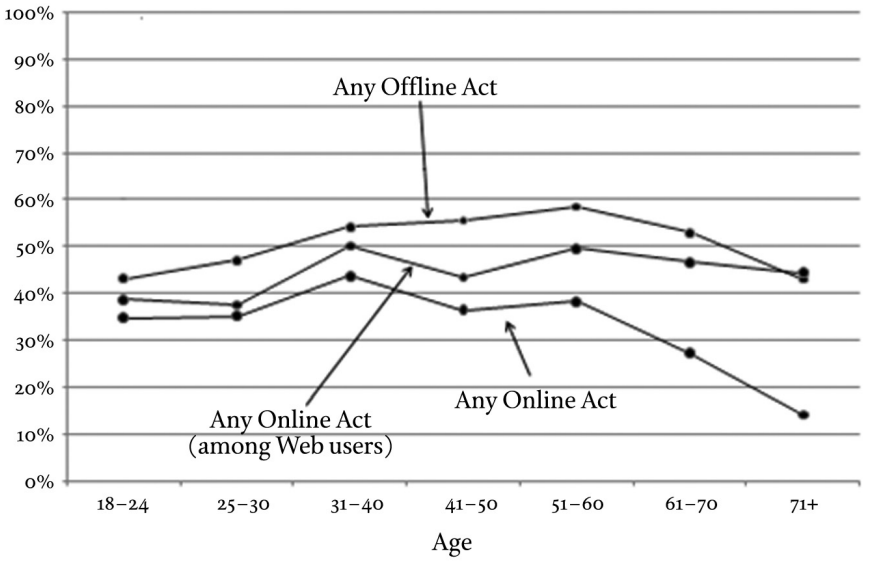
The middle line on Figure 4, which shows the frequency of online political

Figure 3
Percent of Requests for Political Activity that Came by Phone or Email



Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project.

Figure 4
Percent Engaged in Political Activity, by Age



Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project. "Web users" are defined here as those who use the Internet or email at least occasionally.

activity among Internet users only, is striking in showing no pattern at all. There is relatively little difference among age groups in the share who are active online. Among those who use the Internet and email, the oldest groups are not especially inactive, and the young are the least likely to be politically active online. Thus, the digital divide has its greatest impact among older respondents. The small number of Web users among older respondents – a group that surely is not a random selection – are quite politically active on the Internet.

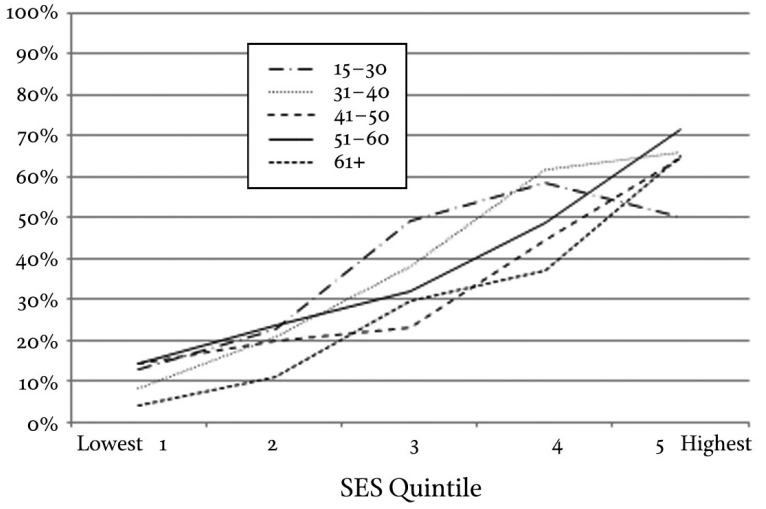
Figure 5 reaffirms the centrality of SES for online political participation. The lines on Figure 5 report for the various age groups the percent undertaking at least one participatory act online across SES quintiles.[27] The overall pattern shows the impact of SES and the comparative irrelevance of age. The five lines are bunched quite closely; they rise in tandem with SES. Each age group shows the expected association between SES and political activity. Within any SES quintile, there is much less variation among age groups and little consistent pattern as to which age group is the most active. The data reinforce our understanding of the strength of the relationship between SES and political activity.

The activities we have just considered are political acts that existed before the advent of the Internet – which is what allows us to compare them in their offline and online manifestations. Certain modes of Internet-based engagement have no direct offline counterpart, including posting comments on blogs (whether one's own or someone else's) and using social networking sites like MySpace, Facebook, or LinkedIn. Most people who post to blogs or, especially, join social networking sites do so for reasons having nothing to do with politics. Figure 6 gives information about the proportion of respondents in each age group who reported blogging or using social networking sites, whether or not for politics. Figure 6 makes clear that the young are much more likely to exploit these relatively recent and rapidly developing Internet capabilities. Especially striking are the data for social networking, which show that the overwhelming majority of respondents under twenty-five are social networkers, a proportion that has undoubtedly grown since the survey was conducted.

Both modes of Internet engagement can also be used for political purposes. The Pew study asked explicitly about political blogging: that is, writing about a political or social issue on a blog, either one's own or, more frequently, someone else's. The survey also asked about political use of social networking, namely, doing any of the following on a social networking site: getting campaign or candidate information; starting or joining a political group or group supporting a cause; signing up as a "friend" of any candidates; or posting political news for friends or others to read.[28]

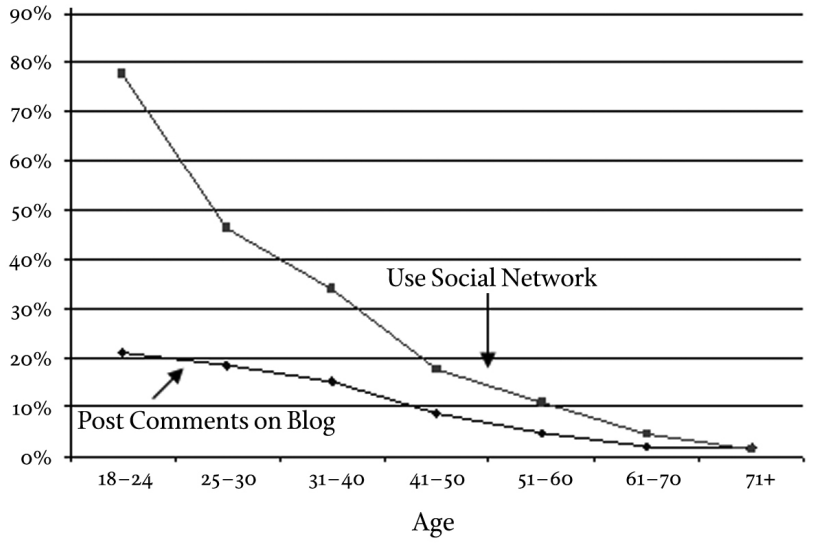
The forms of political engagement in these venues about which the Pew survey asked do not fall squarely under the definition of political participation as "activity that has the intent or effect of influencing government action – either directly by affecting the making or implementation of public policy or indirectly by influencing the selection of people who make those policies."[29] The items from the Pew survey focus on the ways that a social networking site is more a forum for political talk than for organized political effort; even the political groups formed are more about affinity than concerted political action. "Friending" a candidate is not the same as working in a campaign. In many ways, these modes of political involvement reflect some of the distinctive

Figure 5
Percent Engaged in Any Online Political Activity, by Socioeconomic Status (SES) and Age



Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project.

Figure 6
Percent Engaged in Any Blog or Social Network Use, by Age



Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project.

civic tastes of post-Boomer cohorts: their preference for participatory forms that are anchored in nonhierarchical and informal networks and that eschew such traditional political intermediaries as campaigns, parties, and interest groups.[30]

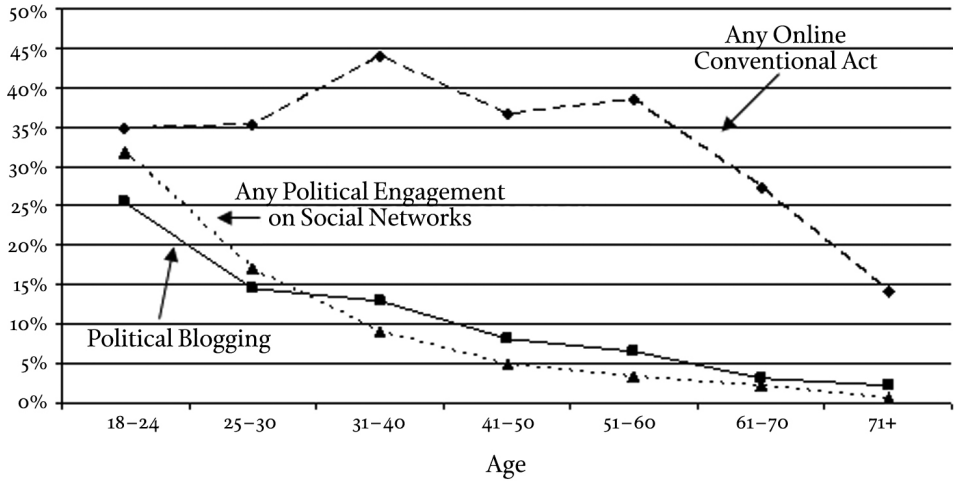
Thus, as Figure 7 makes clear, blogging about political and social issues and political social networking are closely connected to age. The lower two lines – which show the percentage in various age groups reporting that, in the past year, they have posted comments about a political or social issue on a website or blog and the percentage reporting that they have undertaken at least one of the four political activities on a social networking site – fall sharply from the level for those under twenty-five. Figure 7 also repeats the data from Figure 4 about the proportion of respondents who engage in the online version of "conventional" political activities, such as making a political contribution or getting in touch with a public official. As before, the pattern is quite different. Although there is a steep drop-off among those over sixty, the youngest groups are not especially active in conventional online political activity.

Although these possibilities for political engagement through social networking sites do not simply reproduce participation as we have always known it, they may nevertheless lead to forms of online and offline political participation as conventionally understood. Besides, in the period since the Pew survey was conducted, these modes of involvement have become less exclusively the province of the young and have continued to evolve. There is a well-known pattern such that new technologies initially look a lot like the older technologies they eventually replace before their unique capacities are developed. For example, before the power of visual images was refined, early campaign ads on television used talking heads with wordy messages suitable for radio. In certain ways, as increasing numbers of politicians move from maintining websites to establishing a presence on Facebook, what is happening is almost the opposite. More conventional forms of political discourse and advocacy have established a beachhead in this brave new world.

Do these new types of activity hold the promise of diminished inequality of political voice when it comes to social class? We are reluctant to draw conclusions from these August 2008 data about the extent to which these Web 2.0 phenomena have the potential to overcome the structuring of political participation by SES. Figure 8 shows data analogous to Figure 7, but in this instance plots the data based on SES quintile rather than age. As revealed in Figure 1, the relationship of more traditional political activity carried out on the Internet slopes sharply upward with SES. The lines for political social networking and blogging about political and social issues also rise with SES, but the increase is much less pronounced.

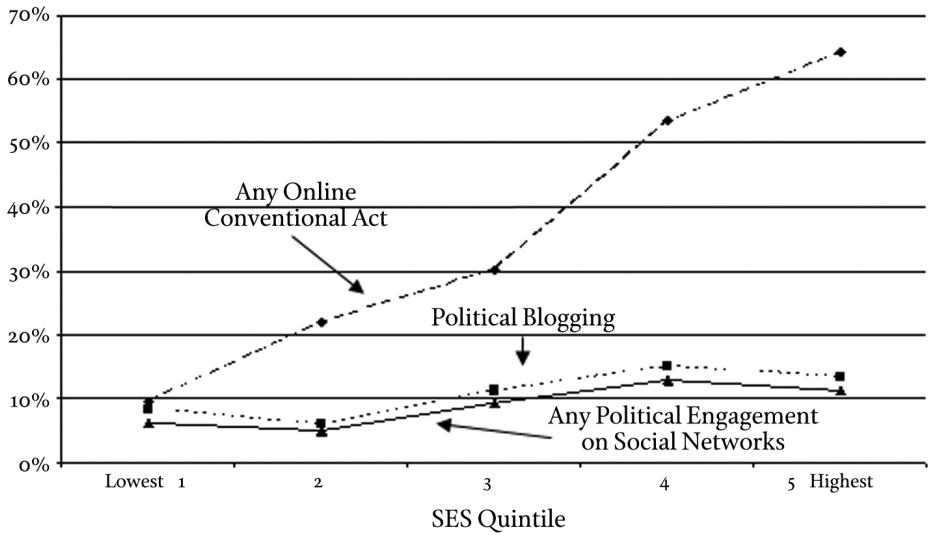
Before we conclude prematurely that new forms of political engagement on the Web might break the long-standing association between social class and political participation, let us go one step further. Most of the political bloggers and political social networkers are twenty-somethings. It is difficult to measure SES for young adults. Forty-two percent of the respondents in the Pew survey who are between eighteen and thirty reported still being in school either full or part time. This group, which has not yet achieved its full educational attainment, includes many respondents whose measured incomes are artificially depressed by their student status but whose incomes will, in the future, rise more sharply than those in their cohort who left school earlier.

Figure 7
Percent Involved in Online Conventional Political Activity, Political Blogging, and Political Engagement on Social Networks, by Age



Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project.

Figure 8
Percent Involved in Online Conventional Political Activity, Political Blogging, and Political Engagement on Social Networks, by Socioeconomic Status (SES)



Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project.

Table 1
Personal and Political Use of the Internet among Respondents under Thirty (among Web Users)

	Personal Use of Internet	Political Use of Social Network	N
Educational Status			
Current Student	54%	43%	94
Highest Education Level Achieved			
High School or Less	60%	18%	129
Some College	58%	25%	62
College Graduate	62%	36%	49

Personal use of the Internet includes seeking information about someone one knows or would like to get to know or finding dates; political use of a social networking site includes seeking political information on a social networking site, joining a political group, signing up as a "friend" of a candidate, or posting political messages. Source: Data from the August 2008 survey conducted by the Pew Internet & American Life Project.

Table 1 allows us to look more closely at the social networking behavior of these younger respondents. In addition to asking about their use of social networks to engage in political activities, the Pew study queried respondents about their use of the Internet for personal reasons (to learn more about people they knew or hoped to get to know or to find dates). Instead of focusing on SES, we compare groups based on educational attainment: those who are still full-time students; and, among those no longer in school, those with no education beyond high school, those with some college, and college graduates. Even in an election year that witnessed an upsurge of activity by younger citizens, those under thirty were considerably more likely to use the Internet for personal objectives – to find information about people or to find dates – rather than to use social networking sites for political ones. In addition, when it comes to personal use of the Internet, there is no association with current student status or, for non-students, with educational attainment. The pattern for use of social networking sites for political purposes is quite different. Those who are still students are the most active and, among non-students, the higher the education level, the more likely someone is to take political actions on the Internet. This finding is especially germane to our concern with class-based inequalities of political voice and suggests that even these new forms of Internet-based political involvement may not act as the circuit breaker interrupting the long-standing connection between SES and citizen political activity.

Because the Internet continues to create new possibilities for communication and the dissemination of information with astonishing rapidity, we are reluctant to make predictions about its future conse-

quences for inequalities of political voice. The opportunities for online political engagement continue to proliferate both in ways that mimic older forms of political participation and in ways that were not imagined even a few years ago. At present, political engagement on blogs and social networking sites clearly overcomes the historical underrepresentation of younger citizens with respect to political activity, but its impact on the socioeconomic stratification of participation is less certain. As older cohorts quickly register on social networking sites, the extent to which the young dominate these venues is also less certain. Moreover, we cannot know whether the current techno-savvy generation will be trumped by their successors who are now in elementary school. We consider it premature to conclude, as others have suggested, that interactive forms of online political participation hold the key to unlocking the association between political participation and SES. The links between social class and political participation have proved to be powerful and enduring. We are not ready to bet our lives, our fortunes, and our sacred honor that the Internet will sunder them.

ENDNOTES

* Contributor Biographies: KAY LEHMAN SCHLOZMAN, a Fellow of the American Academy since 2003, is the J. Joseph Moakley Endowed Professor of Political Science at Boston College. Her publications include *The Private Roots of Public Action: Gender, Equality, and Political Participation* (with Nancy Burns and Sidney Verba, 2001), *Voice and Equality: Civic Voluntarism in American Politics* (with Sidney Verba and Henry E. Brady, 1995), and *Organized Interests and American Democracy* (with John T. Tierney, 1986).

SIDNEY VERBA, a Fellow of the American Academy since 1968, is the Carl H. Pforzheimer University Professor Emeritus and a Research Professor of Government at Harvard University. His publications include *The Private Roots of Public Action: Gender, Equality, and Political Participation* (with Nancy Burns and Kay Lehman Schlozman, 2001), *Voice and Equality: Civic Voluntarism in American Politics* (with Kay Lehman Schlozman and Henry E. Brady, 1995), and *Designing Social Inquiry: Scientific Inference in Qualitative Research* (with Gary King and Robert O. Keohane, 1994).

HENRY E. BRADY, a Fellow of the American Academy since 2003, is Dean of the Goldman School of Public Policy and Class of 1941 Monroe Deutsch Professor of Political Science and Public Policy at the University of California, Berkeley. His publications include *Rethinking Social Inquiry: Diverse Tools, Shared Standards* (with David Collier; 2nd ed., 2010), *Capturing Campaign Effects* (edited with Richard Johnston, 2006), and *Voice and Equality: Civic Voluntarism in American Politics* (with Sidney Verba and Kay Lehman Schlozman, 1995).

[1] The title of an article by Stephen M. Johnson, "The Internet Changes Everything: Revolutionizing Public Participation and Access to Government Information through the Internet," *Administrative Law Review* 50 (1998): 277–337.

[2] Howard Rheingold, "The Great Equalizer," *Whole Earth Review* (Summer 1991): 6, quoted in Bruce Bimber, "The Internet and Political Transformation: Populism, Community, and Accelerated Pluralism," *Polity* 31 (1998): 138.

[3] William Wresch, *Disconnected: Haves and Have-nots in the Information Age* (New Brunswick, N.J.: Rutgers University Press, 1996), 237; and Daniel Burstein and David Kline, *Road Warriors: Dreams and Nightmares along the Information Highway* (New York: Dutton, 1995), 360, quoted in Richard Davis, *Politics Online: Blogs, Chatrooms, and Discussion Groups in American Democracy* (New York: Routledge, 2005), x.

4 A strong statement of this theme is contained in Steve Davis, Larry Elin, and Grant Reeher, *Click on Democracy: The Internet's Power to Change Political Apathy into Civic Action* (Boulder, Colo.: Westview Press, 2002).

5 For a brief version of this story, see Paul Herrnson, Richard G. Niemi, Michael J. Hanmer, Benjamin B. Bederson, Frederick C. Conrad, and Michael W. Traugott, *Voting Technology* (Washington, D.C.: Brookings Institution, 2008), 11–12.

6 Michael Margolis and David Resnick, *Politics as Usual: The Cyberspace "Revolution"* (Thousand Oaks, Calif.: Sage Publications, 2000), vii.

7 Our major work on the subject is Sidney Verba, Kay Lehman Schlozman, and Henry E. Brady, *Voice and Equality: Civic Voluntarism in American Politics* (Cambridge, Mass.: Harvard University Press, 1995). See also, Henry E. Brady, Kay Lehman Schlozman, and Sidney Verba, "Prospecting for Participants: Rational Expectations and the Recruitment of Political Activists," *American Political Science Review* 93 (1999): 153–168; Sidney Verba, Kay Lehman Schlozman, and Henry E. Brady, "Rational Action and Political Participation," *Journal of Theoretical Politics* 12 (2000): 243–268; Henry E. Brady, Kay Lehman Schlozman, Sidney Verba, and Laurel Elms, "Who Bowls?: The (Un)Changing Stratification of Participation," in *Understanding Public Opinion*, ed. Barbara Norrander and Clyde Wilcox (Washington, D.C.: CQ Press, 2002); and Sidney Verba, "Would the Dream of Political Equality Turn out to Be a Nightmare?" *Perspectives on Politics* 1 (2003): 663–679. In addition, see Nancy E. Burns, Kay Lehman Schlozman, and Sidney Verba, *The Private Roots of Public Action: Gender, Equality, and Political Participation* (Cambridge, Mass.: Harvard University Press, 2001).

8 See, for example, Sidney Verba, Norman Nie, and Jae-on Kim, *Participation and Political Equality: A Seven Nation Comparison* (New York: Cambridge University Press, 1979); G. Bingham Powell, Jr., "Political Representation in Comparative Politics," *Annual Review of Political Science* (Palo Alto, Calif.: Annual Reviews, 2004), 273–296; and Russell J. Dalton, *Citizen Politics: Public Opinion and Political Parties in Advanced Industrial Democracies*, 4th ed. (Washington, D.C.: CQ Press, 2006).

9 On these themes, see our forthcoming book, tentatively titled *The (Un)heavenly Chorus: Unequal Political Voice in American Democracy*.

10 Verba, Schlozman, and Brady, *Voice and Equality*, chap. 15.

11 During Summer 2008, we collaborated with Lee Rainie and Scott Keeter of the Pew Internet & American Life Project to design a survey to collect information about Internet use and about political activity both off and on the Internet. We are very grateful to them for having responded to our suggestion about the importance of collecting systematic national data comparing online and offline participation, for allowing us to be partners in the design of the questionnaire, and for making those data available to us.

12 Matthew Hindman, *The Myth of Digital Democracy* (Princeton, N.J.: Princeton University Press, 2009), 129.

13 A succinct and sober estimate of the participation-enhancing capacities of the Internet is contained in Richard Davis, *The Web of Politics* (New York: Oxford University Press, 1999), 20–27.

14 On this issue, see Pippa Norris, *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide* (Cambridge: Cambridge University Press, 2001), 230–231.

15 Anthony G. Wilhelm, *Democracy in the Digital Age* (New York: Routledge, 2000), 67ff. See also, Anthony G. Wilhelm, "Civic Participation and Technological Inequality: The 'Killer Application' Is Education," in *The Civic Web*, ed. David M. Anderson and Michael Cornfield (Lanham, Md.: Rowman and Littlefield, 2002).

16 Karen Mossberger, Caroline Tolbert, and Mary Stansbury call these capacities, respectively, "technical skills" and "information literacy"; see Karen Mossberger, Caroline J. Tolbert, and Mary Stansbury, *Virtual Inequality: Beyond the Digital Divide* (Washington, D.C.: George-

town University Press, 2003), 40–50. In an interesting study that parallels what we find here, Samuel Best and Brian Krueger demonstrate that online skills (measured as the sum of whether the respondent has designed a Web page, sent an attachment via email, posted a file to the Internet, or downloaded a program from the Internet) function in predicting Internet-based political activity in just the same way that organizational and communications civic skills (using the measure in Verba, Schlozman, and Brady, *Voice and Equality*) do in predicting offline activity; see Samuel Best and Brian Krueger, "Analyzing the Representativeness of Internet Political Participation," *Political Behavior* 27 (2005): 183–216.

17 For discussion of inequalities in access to and use of the Internet and citations to the literature, see Paul DiMaggio, Eszter Hargittai, Coral Celeste, and Steven Shafer, "Digital Inequality: From Unequal Access to Differentiated Use," in *Social Inequality*, ed. Kathryn M. Neckerman (New York: Russell Sage Foundation, 2004), chap. 9; and Karen Mossberger, Caroline Tolbert, and Ramona McNeal, *Digital Citizenship: The Internet, Society, and Participation* (Cambridge, Mass.: MIT Press, 2008), chap. 1.

18 For a general discussion, see Michael Alvarez and Thad E. Hall, *Point, Click, and Vote* (Washington, D.C.: Brookings Institution, 2004), 44–53. Other data sets show similar patterns to those presented here. See the October 2003 Current Population Survey contained in National Telecommunications and Information Administration 2004 (Table A-1); and John B. Horrigan and Aaron Smith, "Home Broadband Adoption 2007," Pew Internet & American Life Project, http://www.pewinternet.org/Reports/2007/Home-Broadband-Adoption-2007.aspx?r=1 (accessed May 18, 2010).

19 We generated a scale based on education and family income and divided respondents into five roughly equal groups or quintiles. Although there is very little missing data on educational attainment, we lack information about family income for 19 percent of respondents. While the respondents for whom family income is missing are distributed fairly evenly along the educational hierarchy, they are somewhat less active politically – especially with respect to online political activity – than are those who reported family income.

20 There are so many different paths by which the Internet might influence political activity that we have no reason to expect that studies focusing on different participatory acts or focusing on Internet mobilization as opposed to online participation would find identical results. Nevertheless, all studies of particular political acts find that online participants are not representative of the public as a whole. See, for example, Bruce Bimber, "The Internet and Citizen Communication with Government: Does the Medium Matter?" *Political Communication* 16 (1999): 409–428; Michael Alvarez and Jonathan Nagler, "The Likely Consequences of Internet Voting for Political Representation," *Loyola of Los Angeles Law Review* 34 (2001): 1115–1154; John Clayton Thomas and Gregory Streib, "The New Face of Government: Citizen-Initiated Contacts in the Era of E-Government," *Journal of Public Administration Research and Theory* 13 (2003): 83–102; and Davis Schlosberg, Stephen Zavestoski, and Stuart W. Schulman, "Democracy and E-Rulemaking: Web-Based Technologies, Participation, and the Potential for Deliberation," *Journal of Information Technology and Politics* 4 (2007): 37–55.

21 With regard to any particular form of participation, the survey recorded whether respondents were active offline, online, or both. The questionnaire can be found at http://pewinternet.org/Shared-Content/Data-Sets/2008/August-2008--Civic-Engagement.aspx.

22 We should remind the reader that these data were collected in August 2008, before Obama's Web-based September fundraising blitz. Thus, one must be cautious in generalizing from them to the situation closer to Election Day. However, the Pew survey is the first large-scale survey to collect data about the size of political contributions since our 1990 Citizen Participation Study and contains valuable data comparing online and offline giving. We must also mention that the two-stage design of the Citizen Participation Study in 1990 permitted the oversampling of those who made large contributions, thus facilitating the analysis of political activity in which the input is money rather than time. With very few large donors in the

Pew survey, we do not feel comfortable drawing conclusions about those who make very large contributions.

23 We thank Michael Malbin for the first suggestion and Daniel Schlozman for the second.

24 Because the size of political contributions has been shown to be a function of family income rather than education, we substitute categories based on family income for SES quintiles.

25 As expected, donors of large amounts are almost exclusively affluent. As mentioned above, the Pew data contain too few cases of those who make very large campaign contributions to justify drawing conclusions. However, the fact that the Pew respondents who indicated having made campaign contributions of more than $2,500 are drawn almost uniformly from the highest income category is consistent with earlier studies.

26 Brady, Schlozman, and Verba, "Prospecting for Participants."

27 In order to facilitate the graphic presentation, we have reduced the number of age groups from seven to five.

28 Respondents to the Pew survey were asked about "signing up as a 'friend' of any candidates on a social networking site." However, supporting a political figure on Facebook is not the same as "friending" someone.

29 Verba, Schlozman, and Brady, *Voice and Equality*, 38.

30 See, for example, Cliff Zukin, Scott Keeter, Molly Andolina, Krista Jenkins, and Michael X. Delli Carpini, *A New Engagement? : Political Participation, Civic Life, and the Changing American Citizen* (Oxford and New York: Oxford University Press, 2006), chap. 4.

Prosocial Behavior on the Net

Lee Sproull

Abstract: Volunteers and charitable organizations contribute significantly to community welfare through their prosocial behavior: that is, discretionary behavior such as assisting, comforting, sharing, and cooperating intended to help worthy beneficiaries. This essay focuses on prosocial behavior on the Internet. It describes how offline charitable organizations are using the Net to become more efficient and effective. It also considers entirely new models of Net-based volunteer behavior directed at creating socially beneficial information goods and services. After exploring the scope and diversity of online prosocial behavior, the essay focuses on ways to encourage this kind of behavior through appropriate task and social structures, motivational signals, and trust indicators. It concludes by asking how local offline communities ultimately could be diminished or strengthened as prosocial behavior increases online.

LEE SPROULL is the Leonard N. Stern School Professor Emerita in the Stern School of Business at New York University. Her publications include *Organizational Learning* (edited with Michael D. Cohen, 1996); *Connections: New Ways of Working in the Networked Organization* (with Sara Kiesler, 1991); and *Technology and Organizations* (with Paul S. Goodman, 1990).

In 2009, 63.4 million Americans contributed more than 8 billion hours of volunteer service in the offline world. With the exception of short-term service and relief programs like Habitat for Humanity, most volunteering benefits a volunteer's local community and consists of fundraising/selling items to raise money (26.6 percent of volunteers); collecting, preparing, distributing, or serving food (23.5 percent); engaging in general labor or providing transportation (20.5 percent); and tutoring or teaching (19 percent).[1] Beneficiaries are designated by the community as needy and worthy; they are typically children, the poor, the aged and infirm, and community resources such as libraries, schools, and parks. Volunteers and volunteer organizations enhance local community welfare by improving the situation of direct beneficiaries and by increasing community members' social capital: the bonds of trust and reciprocity created in social networks.

Volunteering is a form of what psychologists call *prosocial behavior*, that is, discretionary behavior such as assisting, comforting, sharing, and cooperating intended to help people other than oneself. Organized prosocial behavior is planned, relatively long-

term, non-obligated by family or friendship, and situated within an organizational context. Organized prosocial behavior in the form of charitable donation of time and money represents an enormously important social resource in the offline world.

Prosocial behavior can also be found on the Internet, although most discretionary behavior on the Internet today is devoted to asocial behaviors – shopping or looking for information and entertainment – or social behaviors such as connecting with family and friends.[2] While prosocial behavior represents only a small fraction of discretionary Internet behavior, it is significant for at least four reasons. First, it can reduce transaction costs for offline volunteer organizations and activities, thereby allowing them to operate more efficiently. Second, it can extend the reach and impact of offline volunteer organizations by engaging volunteers and beneficiaries for whom offline opportunities are inconvenient or inaccessible. Third, it can offer new ways for people to strengthen important social institutions.[3] Fourth, it can produce socially beneficial information goods and services that are undersupplied by the market. This essay sketches some of the scope and diversity of prosocial behavior on the Net today; describes the building blocks for producing online prosocial behavior; and suggests what can promote online prosocial behavior in the future. It concludes with questions about the relationship between prosocial behavior on the Net and in the offline world.

All major types of online prosocial projects share a small number of attributes derived from the underlying network technology and communications applications. One is that the people or projects that need help and the volunteers willing to help are able to find one another and interact independent of geographic or social location. Another is that they are able to interact asynchronously. A third shared attribute is that volunteers are able to participate in brief segments of time at any hour of the day or night.

Opportunities for online prosocial behavior are evolving in a pattern seen in other sectors of human behavior on the Internet, such as electronic commerce. In this process of evolution, early efforts designed to make previously offline work more efficient give way to later efforts that create substantially new models of Net-based interaction. Like other types of activity on the Internet, some prosocial sites simply make offline prosocial behavior more efficient through reducing transaction costs. Others involve people in new models of helping. Tables 1 and 2 provide several examples of online prosocial behavior.

Websites that support and modify familiar offline models of volunteering can generally be divided among charitable giving, service projects, and online health-support groups.

Charitable Giving. Net-based charitable giving is increasing as more people use the Net regularly and are comfortable using it for financial transactions. In 2009, Americans made $15.4 billion in online charitable donations (compared with $300,000 in 1997).[4] Early donation sites, connected to specific causes or organizations such as the Hurricane Katrina relief fund or the American Red Cross, provided the convenience of online donation but left all decision authority with the sponsoring organization. By 2007, individuals could create their own fundraising campaigns on the Net, using social networking sites and crowdsourcing to encourage friends (and friends of friends) to donate to causes or projects of one's own choosing or design. The newest of these sites – for instance, www.crowdrise

Table 1
Online Prosocial Projects that Support or Modify Familiar Types of Volunteering

Website	Year Founded	Purpose	Scope/Impact
Charitable Giving			
www.firstgiving.com	2003	Support charitable giving campaigns	>8,000 nonprofits; >13 million online donors; >$1 billion raised online
www.causes.com	2007	Support user-created advocacy groups	>$27 million raised online
www.crowdrise.com	2010	Crowdsourcing for volunteers and charities	N/A
Service Projects			
www.volunteermatch.org	1998	Searchable directory of offline volunteer opportunities	4.9 million referrals since 1998
www.volunteerspot.com	2009	Tools for mobilizing and coordinating community volunteers	>450,000 volunteers, mostly parents in schools
www.mentornet.org	1998	Online mentoring for college-level women and underrepresented minorities in science and engineering	>27,000 matched pairs of mentors and protégés since 1998; >90% retention rate in science and engineering fields for protégés
www.icouldbe.org	2000	School-based curriculum supported by online mentors	>10,000 students served
www.onlinevolunteering.org	2000	Projects to support organizations addressing UN millennium development goals	>36,000 online volunteer assignments completed since 2000
Health Support			
www.dailystrength.org	2006	Health discussion boards and expert advice	>0.5 million members in >500 online groups
www.mdjunction.com	2006	Online health-support groups	>700 online groups
www.patientslikeme.com	2004	Support groups and tools for sharing medical/health information	>69,000 public patient profiles; >4,500 treatment reports; >3,500 symptom reports

Source: Table compiled by author.

.com – incorporate competitive elements taken from online gaming, such as leader boards and prizes, to encourage not just one-time donation but ongoing commitment to volunteer activity.

Service Projects. In the simplest case, an Internet site can provide a way for projects located in the offline world to describe their needs for volunteers and for people to see offline volunteer opportunities displayed by zip code. The actual volunteering and management of volunteers happens offline; the sites merely offer a way for people and projects to find one another efficiently. For example, VolunteerMatch lets users search for oppor-

Table 2
Online Prosocial Projects that Support New Types of Volunteering

Website	Year Founded	Purpose	Scope/Impact
Creating Information Goods			
www.apache.org	1995	Collaborative volunteer software projects	Nearly 100 projects; >2,500 committers; Apache Web server powers >65% of world's websites
www.iFixit.com	2003	Collaboratively written consumer-product repair manuals	>2,500 repair manuals; >800 volunteer repair technicians
www.wikipedia.org	2001	Collaboratively written encyclopedia project	>16 million articles in >270 languages; >91,000 active volunteers
www.pgdp.net	2000	Support the digitization of public domain books	>19,000 books digitized; 2,500 active volunteers per month
Citizen Science			
http://boinc.berkeley.edu	2002	Donate idle cycles to generate computing power for scientific research	>50 projects; daily average computing power of about 4 PetaFLOPs
http://implicit.harvard.edu	2003	Participate in Web-based social psychology research projects	Volunteers have completed >3.5 million tasks
http://ebird.org/content/ebird/	2002	Contribute to global bird observation database	>35,000 volunteers have submitted >21 million bird records
www.gwap.com	2008	Tag words, music, clips, or photographs to improve search engine performance	>1 million images labeled
www.galaxyzoo.org	2007	Categorize telescopic images of galaxies	>250,000 volunteers
http://fold.it/portal	2008	Manipulate renderings of proteins to predict stable structures	>100,000 volunteers

Source: Table compiled by author.

tunities to volunteer by zip code, city, or state in order to "make it easier for good people and good causes to connect."[5] VolunteerSpot provides tools to mobilize and coordinate local volunteers for community projects like PTA fairs and youth sports leagues.

Some sites support projects that need online volunteers rather than offline ones. These sites typically offer more features than ones that simply list offline opportunities. They provide guidance to listing organizations about the kinds of projects that are suitable for online volunteers; guidance to potential volunteers about what kinds of information to include in a volunteer application; and templates to support various parts of the process. Projects in developing countries, in particular, often need volunteers whose skill sets are in short supply on the ground. If those skills can be delivered via computer – by creating a website, writing a grant proposal, building a database, or translating

a document, among other acts – geographic distance is no barrier to volunteering. The United Nations Online Volunteering Service, for example, connects online volunteers with organizations working for sustainable human development. Its website proclaims: "Everyone can make a difference. Share your skills, knowledge and ideas – from a computer anywhere in the world."[6]

In contrast with sites that serve as intermediaries between volunteers and projects, some sites run their own volunteer programs. Mentoring, a familiar offline model, matches young people or novices with more experienced or older mentors who offer them guidance and support. In its online form, mentoring differs in that protégés usually have little or no opportunity for offline interaction with potential mentors; mentors have little or no time or opportunity to engage in face-to-face mentoring meetings. Online mentoring programs operate independent of geographic and time constraints.

Online Health-Support Groups. About six million U.S. adults report having used an online health-support group in 2009.[7] In addition to offering access to personal experience and advice about medications, procedures, symptoms, and so on, these groups also reduce the social and emotional isolation of people who find little or no support in their offline communities. The following comment is representative of the thousands of comments posted to these groups over the years: "i live in a remote rural community and had no support and little therapy options due to location. the support i have received [from this group] in not feeling alone has made a tremendous difference in my life and gives me strength."[8]

Websites offering new types of volunteering often fall into the categories of information goods and citizen science.

Information Goods. The Net allows volunteers to collaborate to create and freely distribute high-quality socially beneficial information goods such as public domain e-books, repair manuals, software, and reference articles that function as public goods.[9] Each person's contribution to an information product is accessible to and modifiable by other volunteers with the consequence that product quality can improve over time. Observers note that these projects can produce high-quality information goods more rapidly with much lower overhead than similar projects organized and distributed by corporations and markets.[10] Thousands of volunteers proofread scanned pages of books in the public domain to convert them into e-books for free distribution over the Net. Hundreds of thousands of programmers voluntarily contribute code, bug reports, and patches to freely downloadable open-source software projects. Although many of these are one-person vanity projects, some have attracted thousands of volunteer programmers and have become large enough and reliable enough to substitute for commercial software. The online reference site Wikipedia displays more than 16 million articles written by volunteers in more than 270 languages.[11] Anyone with Web access can write or edit an article on any topic; articles evolve because they are visible to everyone and can be freely edited by anyone. Over time, mistakes and lacunae are corrected; and through this process, at least some articles become equivalent in quality to those found in professionally produced publications.[12]

Citizen Science. A number of sites on the Net encourage interested amateurs to volunteer to help scientists and scientific projects. Opportunities for volunteer engagement range in level of effort from simply donating idle PC cycles to contributing data to categorizing and analyzing scien-

tific data. A relatively low-effort type of project lets people donate computing cycles from an otherwise idle PC to a project that is analyzing massive amounts of scientific or mathematical data. An early example, SETI@home, was launched in 1999 to detect narrow-bandwidth radio signals from space. More than three million people have donated idle cycles to this project.[13] By 2010, volunteers had donated idle PC time to more than fifty projects in biology and medicine, earth sciences, mathematics, astronomy, physics, and chemistry.

Volunteers contribute data that they have personally collected about the physical world to sites for ongoing scientific projects that maintain databases for research and public distribution. In most cases, the volunteers are hobbyists who have been collecting data for their own pleasure. Submissions are verified, aggregated, and published to the broader scientific and citizen community. For example, bird, insect, or plant watchers upload field observations. People with backyard weather stations upload data to a server that makes them available to the National Oceanic and Atmospheric Administration (NOAA) and other weather services.

Volunteers also contribute data about themselves. The PatientsLikeMe site allows members to join support groups organized by condition or disease that serve the functions of the online support groups described above. But the site also supports collaborations with medical and pharmaceutical researchers. Support group members can report data about such factors as their health condition, symptoms, and medications; data are then anonymized and aggregated for sharing with medical researchers. Researchers can describe clinical trials to groups whose members may be eligible for enrollment. On the Implicit site, volunteers can participate in Web-based social psychology experiments that allow researchers to access a large, relatively heterogeneous (compared to a pool of all college sophomores) research population at minimal cost. Volunteers have completed more than 3.5 million comparison tasks producing "the largest database on implicit attitudes and knowledge currently available."[14]

Individuals can contribute their mental effort to help categorize existing knowledge or to create new knowledge. With a small amount of effort, volunteers tag music clips or photographs found on the Web in order to improve the performance of search engines. With substantially more effort, more than 250,000 volunteers categorize millions of telescopic images of galaxies to create a database of detailed galaxy shape information. More than 100,000 volunteers manipulate renderings of proteins to predict stable structures.[15]

There are no comprehensive surveys of online volunteering to provide an estimate of its overall magnitude and impact. Scattered evidence suggests that online prosocial projects do produce beneficial consequences. Studies of specific online health-support groups have shown that participation provides both informational and emotional benefits to group members.[16] Broad studies of Internet health information, not just information in online support groups, suggest that its impact is more positive than negative.[17] Online mentoring projects report significant retention rates. Millions of people use reference material sites like Wikipedia and depend on software produced in open-source projects. Scholars have drawn on citizen-science databases in the production of hundreds of scholarly papers.

Every context for online behavior is a symbolically differentiated place on the Net. Different people seek out different

places, and the same people behave differently in different places. Each context for online prosocial behavior can be understood by characteristics of its task and social structures, by what motivates its participants, and by its trust dynamics. The shared features of the Net described above are influential across contexts, but most of the factors that influence online volunteer behavior are situational and specific to particular contexts.[18] It is important to focus on the details of specific contexts because factors encouraging prosocial behavior in one context may impede it in another. (For example, leader boards highlighting top contributors might be useful for charitable donation sites but inappropriate for health support groups. Peer ratings might be useful for information product sites but inappropriate for citizen science sites.)

Task and Social Structures. A website's task structure embodies the purpose and goals of the group and focuses volunteers' attention on particular activities. Task structures can be more or less modular or decomposable with more or less fine-grained tasks and more or less complex aggregating mechanisms. Modularity and granularity determine the level of effort a volunteer must expend to participate. For example, on an online mentoring site, the task structure is relatively undifferentiated; each protégé represents one task unit. A volunteer's task – helping a protégé – entails multiple different kinds of interconnected actions and a relatively long time commitment. In an information product site, volunteers have their choice of different types of relatively circumscribed tasks (such as posting a bug report or a feature request, writing a review, editing an article, tagging an image, checking references, or providing a quality rating); each task type is relatively fine-grained and homogeneous. The greater the degree of modularity and granularity, the easier it is for a volunteer to make a small contribution.[19]

A website's social structure, which includes explicit roles and governance rules, and implicit norms, represents how volunteers are organized to accomplish tasks. Sites with a relatively modular task structure often use volunteers' performance history to define roles and manage role differentiation. (For example, roles for Wikipedia include editor, featured article editor, administrator, bureaucrat, and steward.[20]) Performance history may be assessed objectively by measuring the number of contributions or contributions with particular characteristics, subjectively through peer ratings of contributions, or some combination of both. Volunteers who make more contributions of greater value can be recognized with higher status roles. Higher status roles may serve as rewards for past performance; they may also confer additional task rights and responsibilities on those achieving such status. Higher status roles are also associated with more authority. Role differentiation helps volunteers understand why some members have more authority than others and how to attain more authority for themselves.

Motivating Participants. Analysts and commentators are fascinated by prosocial behavior on the Net that benefits unknown (and unknowable) others, such as contributing to an online health-support group or to an information product. Underlying the studies of volunteer behavior on support group sites is the question, why would people help others whom they do not know and with whom they could not have a close personal relationship? Early analysts argued that donating high-quality advice or help to unknown and unknowable others was a classic public goods problem and could not be sustained.[21] These early analyses were flawed by framing the problem as "contributing to databases of

information" rather than as "helping people." Researchers investigating why people offer help online found that volunteers clearly developed relationships with specific other online people and with the generalized "other" represented by the online group. Shared physical location (and all the information associated with proximity, such as physical indicators of participants' worthiness) was unnecessary to create these relationships. Researchers typically found a mix of motivations for contributing to support groups, a mix dominated by altruism and generalized reciprocity: volunteers reported wanting to help others because "it is the right thing to do," because they had been helped in the past or anticipated they might need help in the future, or because they wanted to spare other people the pain that they had experienced.[22]

In terms of volunteer projects to create information products, the question is a matter of why people would do this work for free when they could be paid to do it. Studies of open-source programmers found that some programmers *were* being paid by their employers for their work on open-source projects, and some hoped to translate reputational benefit from their open-source work into economic benefit by securing better jobs. But more programmers reported contributing because it was fun, educational, improved their own programming environment, or benefited the open-source "cause."[23]

Social scientists distinguish between behavior and motivation and note that instances of the same behavior can be motivated by different reasons. Thus, prosocial behavior can be motivated by altruism (desire to benefit others with no concern for self), egoism (desire to benefit the self), or a combination of the two. Prosocial behavior can be motivated by extrinsic forces such as the expectation of positive regard or reward from others; intrinsic forces such as compassion, curiosity, or the desire for mastery; or a combination of the two.[24] It is a mistake to expect all contributors to prosocial projects to be motivated exclusively, or even primarily, by altruism.

It is a truism in the offline world that volunteers receive more than they give.[25] Online volunteers follow the same pattern. The rewards are not only expressive and emotional; they can also be utilitarian. Volunteers in any domain, not just open-source programming, may develop skills leading to career advancement. For example, consider this statement from someone working with UNVolunteering on a technical documentation project to turn pig waste into energy in West Africa: "I hold an Advanced Technical Diploma in Electrical Engineering, and I wanted to learn more about this form of renewable energy.... I have learned a lot on all levels, strengthening my managerial, training and development project management skills, amongst others."[26]

Some citizen-science projects have been designed to appeal explicitly to hedonic motivations: that is, to the desire to be entertained and have fun and to do so through participating in competitive games.[27] In some games people compete or cooperate with one another to label images or guess commonplace words; their behavior contributes to the scientific goals of improving search or semantic understanding. In contrast to such simple online games, Foldit, a protein-folding game, offers much more complex challenges. Players use joystick-like motions to jiggle, twist, and pull protein elements in order to find the lowest energy arrangement most likely to exist in nature. More experienced players are assigned more complex proteins. Players can compete against one another, solo, or in teams. A recent competition offered down-to-the-wire excitement as the lead changed

hands frequently in the waning minutes. The ultimate winner, a reminder that geographic and social distance are not barriers to volunteer participation on the Net, was a thirteen-year-old boy from Virginia with the screen name Cheese.[28]

Instead of using video games as an organizing model, the creators of Galaxyzoo, where volunteers label attributes of galaxies displayed from space telescope images, rely on the intrinsic beauty of the images themselves to motivate volunteers: "gorgeous imagery of hundreds of thousands of galaxies . . . more detailed and beautiful than ever."[29] In response to a forum question about why volunteers take part in the project, many volunteers commented on the beauty of the images and the joy of volunteering. One said:

> This stuff is beautiful, I could stare at it all day.…I am overwhelmed – honored – overjoyed (language fails me) with the opportunity. Imagine, pissant little old *me* doing original research in Astrophysics! It doesn't matter that it's at kindergarten level – that's all I'm capable of. It doesn't matter that I may be (probably am) often wrong – I'm doing my best, and learning more each minute. I am making a contribution! And I don't care that I'm not gonna get a PhD for it – you guys (some of you, anyway) will. All I can say is "Wow!"[30]

The symbolically differentiated place on the Web called Galaxyzoo offers a particular combination of task structure, social structure, and motivational prompts that wowed this volunteer. Millions of volunteers are equally wowed by other sites that allow them to make their own online prosocial contribution a few minutes at a time, at any time, independent of geographic and social distance.

Trust. The production of prosocial behavior, whatever its goals and motivations, depends on bilateral trust. A volunteer must trust that a beneficiary's need is legitimate and that the beneficiary will make use of the donation as represented. A beneficiary must trust that a donor's gift is free to be given and that it is honestly represented. Sociologists describe three sources of trust production in social settings: personal characteristics of the participants; the record of their past performance; and social institutions or intermediaries that vouchsafe participants and performance.[31] Offline contexts for prosocial behavior afford many overlapping indicators of trustworthiness for both donors and recipients. Personal characteristics of participants, including physical appearance and demeanor, are visible. Records of attendance and contribution are kept and reported. Membership and performance insignia such as badges, uniforms, or logos from social institutions or intermediaries are conferred and visible. Untrustworthy situations can certainly be found offline, including scam charities, bogus beneficiaries, and malicious or predatory "donors." But the diversity and abundance of trust indicators in offline prosocial contexts signify and reassure potential donors and recipients that participation is likely to be worthwhile.

Online contexts for prosocial behavior also must provide indicators of trustworthiness in order to attract and retain participants.[32] Information about some personal characteristics may be captured and displayed through registration and profile creation processes. Some types of indicators are more difficult to convey online than offline, such as indicators of participants' demeanor and sensitivity. Some indicators of competence are relatively easy to convey. In the case of software projects, a first-order indicator of contributor competence and contribution quality is an affirmative answer to the question, does the code run? On many sites, information about a volunteer's performance history is available. In the case of text-based

information product sites like Wikipedia, some indicators can be produced because a complete contribution history is available. Page-related indicators of past history – number of edits over time and personal characteristics of editors, for example – have been shown to influence readers' perceptions of article trustworthiness.[33] When the task in citizen-science projects is tagging, labeling, or classifying, contribution quality can be assessed via comparing multiple completions of the same task. Multiple contributors are presented with the same stimulus; contributions are compared and the extent to which they agree is an indicator of quality.

Many online contexts for prosocial behavior use human intermediaries to provide additional indicators of trustworthiness. These intermediaries are typically identified and recognized through a context's role structure. Support-group moderators, where they are used, signal that replies have been vetted. People designated as module maintainers who accept code contributions in open-source projects confirm not only that the code runs but also that it does not break other contributors' code. The featured articles editor for Wikipedia oversees an elaborate process to identify the highest quality articles, which are indicated by a bronze star – a visible institutional insignia of quality.

In addition to gauging the trustworthiness of volunteers and their contributions, volunteers also need assurance that the beneficiaries of their contributions are trustworthy. Acting in accordance with the norms of the group is an indicator that people seeking help belong to the group and are legitimate and worthy of help. For example, questions that conform to a support group's norms are more likely to be answered than those that do not.[34] Open-source licenses assure potential contributors to information products that their freely given contribution will not be appropriated for private gain. Institutional sponsorship indicates to citizen scientists that data they contribute will be treated according to the norms of science. It also indicates that, if they are asked to download code to their PCs, the code is benign.

Encouraging online prosocial behavior requires attracting, retaining, and motivating arms-length volunteers. Social scientists believe that if people agree to do a small prosocial act, they will be more likely to agree to do a related, larger one in the future.[35] The Net offers numerous ways for people to perform small prosocial acts that can be aggregated across many people for substantial social good. Donating idle PC cycles to scientific projects is one example. Tagging an image, answering a question, posting a bug report, or providing a small amount of data are others. Social scientists also observe that people like to do what their friends like to do. Recent projects using social networking sites and friendship dynamics to promote prosocial behavior build on this observation.

As the number of volunteers and the amount of online prosocial behavior increase, infrastructure sites that provide tools to create and manage prosocial projects are emerging. In the early days of the Internet, anyone who wanted to create or maintain a volunteer project such as an online health-support group had to spend significant time managing technical issues like mailing lists, archives, and software upgrades.[36] Today, sites like DailyStrength and MDJunction make it relatively easy for anyone to create and manage an online health-support group. BOINC gives scientific researchers the tools to create and manage projects using donated PC cycles. Sites like SourceForge provide tools for managing open-source software

projects. Sites like FirstGiving and Causes make it relatively easy to create and manage a charitable donation project. As infrastructure sites make it simpler to create and manage opportunities for online prosocial behavior, these opportunities should increase.

All these sites recognize the importance of factors that help manage task and social structures and that increase trust and commitment on the part of participants. Among these factors are persistent identity, aggregation and coordination mechanisms, motivational prompts, and quality control. Tools and processes supporting these factors vary substantially across different types of sites, but their presence is evident in all of them. And evidence is accumulating about ways that are more and less effective to implement them. Scholars from several universities have organized a multidisciplinary, multiyear effort to understand, predict, and improve contribution behavior across all types of online communities. Their work has resulted in a number of scholarly papers and a forthcoming book.[37]

Many of the examples described above are less than ten years old. Efforts to understand how to create, manage, nurture, and sustain them are in their infancy. The scope and diversity of online prosocial behavior should continue to grow and expand as new applications are developed, new volunteers are recruited, and new social connections are forged via the Net. As this growth continues, it will become more important to understand the contours of prosocial behavior on the Net. One mechanism for developing this understanding will be comparative research that simultaneously investigates different types of online contexts to understand how participant characteristics, motivations, and behaviors differ across them.[38] Another mechanism will be national sample surveys that specifically include questions about online prosocial behavior in comparison with other behavior. For example, national surveys of offline volunteering, such as those produced by the Corporation for National and Community Service, should be augmented to include questions about online volunteering. Surveys about the behavior of Internet users, such as those produced by the Pew Research Center, should include questions about online prosocial behavior.[39]

What impact will online prosocial behavior have on offline volunteers – those sixty-three million Americans who contributed more than eight billion hours of offline service in 2009 – and local community well-being in the offline world? It is possible that local volunteer organizations and activities will be diminished as online prosocial behavior expands, just as some local bricks-and-mortar businesses have been weakened by online commerce. Local fundraising campaigns may be weakened as potential donors find it just as easy and perhaps more compelling to donate to more exotic causes that are featured on the global Net. Local volunteer organizations may be diminished as potential local volunteers turn to more convenient Net-based endeavors. New models of online prosocial behavior may redirect the attention of people who previously would have been local volunteers. When someone living in Fort Wayne, Indiana, can proofread a science lesson for a school in Africa over the Net, while sitting pajama-clad in front of her computer, how likely is it that she will do that instead of (or in addition to?) running a bake sale at her local elementary school? Alternatively, Net-based tools may allow local volunteer programs to enlist new volunteers and to operate more efficiently and effectively. Moreover, new models of online prosocial behavior may

attract people who never would have engaged in conventional offline volunteer activity, thereby increasing the total volunteer pool.

Encouraging prosocial behavior, whether online or offline, is a worthy societal endeavor. Understanding interactions between the two and how each contributes to broader social welfare will be important in our continuing effort to understand and build the Internet as a public commons.

ENDNOTES

1 *Volunteering in America 2010 : National, State, and City Information* (Washington, D.C. : Office of Research and Policy Development, Corporation for National and Community Service, June 2010), http ://www.volunteeringinamerica.gov/.

2 http ://www.pewinternet.org/Static-Pages/Trend-Data/Online-Activities-Daily.aspx.

3 For analyses of how the Net can broaden access to social support, see, among others, Jonathon Cummings, Lee Sproull, and Sara Kiesler, "Beyond Hearing : Where Real World and Online Support Meet," *Group Dynamics : Theory, Research, and Practice* 6 (2002) : 78 – 88 ; Katlyn McKenna and John Bargh, "Coming Out in the Age of the Internet : Identity Demarginalization through Virtual Group Participation," *Journal of Personality and Social Psychology* 75 (1998) : 681 – 694 ; Andrea Meier et al., "How Cancer Survivors Provide Support on Cancer-Related Internet Mailing Lists," *Journal of Medical Internet Research* 9 (2) (2007), http ://www.jmir.org/2007/2/e12/. For how volunteers on the Net can create nonmarket resources with economic value, see, among others, Yochai Benkler, *The Wealth of Networks : How Social Production Transforms Markets and Freedom* (New Haven, Conn. : Yale University Press, 2006). Yochai Benkler and Helen Nissenbaum suggest that an increase in participation in virtuous behavior on the Net could lead to an increase in public virtue more generally ; see Yochai Benkler and Helen Nissenbaum, "Commons-Based Peer Production and Virtue," *The Journal of Political Philosophy* 14 (4) (2006) : 394 – 419.

4 http ://mashable.com/2010/09/18/social-good-infographic/.

5 http ://www.volunteermatch.org/about/.

6 http ://www.onlinevolunteering.org/en/vol/index.html.

7 Robin A. Cohen and Barbara Stussman, "Health Information Technology Use among Men and Women Aged 18 – 64 : Early Release of Estimates from the National Health Interview Survey, January – June 2009," Health E-Stats, National Center for Health Statistics, February 2010.

8 http ://www.dailystrength.org/support-community-testimonials ; capitalization as in the original.

9 Not all freely produced online information goods have an overtly prosocial orientation. Hobby sites may be personally satisfying for members, but they do little to address broader social goals. Some information sites are explicitly antisocial or malevolent in their orientation.

10 Benkler, *The Wealth of Networks*.

11 See "About Wikipedia," at http ://en.wikipedia.org/wiki/Wikipedia :About (accessed August 24, 2010).

12 Jim Giles, "Internet Encyclopedias Go Head to Head," *Nature* 438 (7070) (December 15, 2005) : 900 – 901. Lara Devgan et al., "Wiki-Surgery ? Internal Validity of Wikipedia as a Medical and Surgical Reference," *Journal of the American College of Surgeons* 205 (3) (September 2007) : S76 – S77.

13 http ://boincstats.com/.

[14] http://www.projectimplicit.net/about.php.

[15] See http://www.gwap.com/gwap/ for tagging; http://www.galaxyzoo.org for galaxy shape information; and http://fold.it/portal/ for protein structures.

[16] For example, McKenna and Bargh, "Coming Out in the Age of the Internet"; Cummings et al., "Beyond Hearing"; and Meier et al., "How Cancer Survivors Provide Support on Cancer-Related Internet Mailing Lists."

[17] In 2008, the Pew Internet & American Life Project reported a benefit to cost ratio of ten to one. Thirty-one percent of e-patients say they or someone they know has been significantly helped by following medical advice or health information found on the Internet. Three percent say they or someone they know has been seriously harmed.

[18] Although some theorists postulate the existence of an altruistic personality type, most scholars believe that helping others is usually motivated through the interaction between personality factors and situational factors. Daniel Batson and Adam A. Powell, "Altruism and Prosocial Behavior," in *Handbook of Psychology*, vol. 5: *Personality and Social Psychology*, ed. Theodore Millon and Melvin J. Lerner (Hoboken, N.J.: Wiley, 2003), 463–484. Jane Allyn Piliavin and Hong-Wen Charng, "Altruism: A Review of Recent Theory and Research," *Annual Review of Sociology* 16 (1990): 27–65. Louis A. Penner, "Volunteerism and Social Problems: Making Things Better or Worse," *Journal of Social Issues* 60 (2004): 645–666.

[19] Task structures may be supported by more or less complex software to support coordination and aggregation. Mentoring sites may have complex matching algorithms to create mentor/protégé pairs, but then rely upon email and a simple website for information exchange. Scientific analysis sites rely upon more complex software to present images to volunteers, record their contributions, perform quality checks, upload data to the database, produce displays of results, and so on.

[20] http://en.wikipedia.org/wiki/Wikipedia:About#Editorial_administration.2C_oversight.2C_and_management.

[21] For example, Brian K. Thorn and Terry Connolly, "Discretionary Data Bases: A Theory and Some Experimental Findings," *Communication Research* 14 (1987): 512–528.

[22] For example, Cummings et al., "Beyond Hearing"; McKenna and Bargh, "Coming Out in the Age of the Internet"; and Meier et al., "How Cancer Survivors Provide Support on Cancer-Related Internet Mailing Lists."

[23] Karim Lakhani and Robert Wolf, "Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects," in *Perspectives on Free and Open Source Software*, ed. Joseph Feller et al. (Cambridge, Mass.: MIT Press, 2005), 3–22. Jeffrey Roberts, Il-Horn Hann, and Sandra A. Slaughter, "Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects," *Management Science* 52 (7) (July 2006): 984–999.

[24] E. Gil Clary et al., "Understanding and Assessing the Motivations of Volunteers: A Functional Approach," *Journal of Personality and Social Psychology* 74 (6) (1998): 1516–1530. John Wilson, "Volunteering," *Annual Review of Sociology* 26 (2000): 215–240. Batson and Powell, "Altruism and Prosocial Behavior." Piliavin and Charng, "Altruism."

[25] For example, see Wilson, "Volunteering," 231–233.

[26] http://www.onlinevolunteering.org/en/vol/stories/2009_africavenir.html.

[27] In 2008, 23 percent of U.S. adults played games online, and the majority of them played at least a few times a week. Americans watch 200 billion hours of television each year. Clay Shirky points out that if even a tiny fraction of those hours was converted to entertainment with a purpose, substantial social benefit could accrue; see Clay Shirky, *Cognitive Surplus: Creativity and Generosity in a Connected Age* (New York: Penguin, 2010), 10.

[28] John Bohannon, "Gamers Unravel the Secret Life of Protein," *Wired Magazine*, April 2009, http://www.wired.com/medtech/genetics/magazine/17-05/ff_protein?currentPage=all.

29 http://www.galaxyzoo.org/.

30 http://www.galaxyzooforum.org/index.php?PHPSESSID=81257edabee1bbf433ca69c4b4acae5a&topic=68.0.

31 Lynne Zucker, "Production of Trust: Institutional Sources of Economic Structure," *Research in Organizational Behavior* 8 (1986): 53–111.

32 The topic of online trust has been pursued most extensively within the context of e-commerce. For example, see Ye Diana Wang and Henry H. Emurian, "An Overview of Online Trust: Concepts, Elements, and Implications," *Computers in Human Behavior* 21 (2005): 105–125.

33 Aniket Kittur, Bongwon Suh, and Ed H. Chi, "Can You Ever Trust a Wiki?" *CSCW 2008 Proceedings* (2008): 477–480.

34 Jolene Galegher, Lee Sproull, and Sara Kiesler, "Legitimacy, Authority, and Community in Electronic Support Groups," *Written Communication* 15 (1998): 493–530. Karl Sassenberg, "Common Bond and Common Identity Groups on the Internet: Attachment and Normative Behavior in On-Topic and Off-Topic Chats," *Group Dynamics* 6 (2002): 27–31.

35 Jerry M. Burger, "The Foot-in-the-Door Compliance Procedure: A Multiple-Process Analysis and Review," *Personality and Social Psychology Review* 3 (1999): 303–325.

36 Brian Butler, Lee Sproull, Sara Kiesler, and Robert Kraut, "Community Effort in Online Groups: Who Does the Work and Why?" in *Leadership at a Distance*, ed. Suzanne Wiesband (New York: Lawrence Erlbaum Associates, 2008), 171–193.

37 Robert E. Kraut and Paul Resnick, *Evidence-Based Social Design: Mining the Social Sciences to Build Online Communities* (Cambridge, Mass.: MIT Press, forthcoming).

38 For a recent example of comparative research, see Shaul Oreg and Oded Nov, "Exploring Motivations for Contributing to Open Source Initiatives: The Roles of Contribution Context and Personal Values," *Computers in Human Behavior* 24 (2008): 2055–2073.

39 A recent Pew Research Center survey found that 80 percent of U.S. adult Internet users, compared with 56 percent of non-Internet users, participate in one or more different kinds of groups including "charitable or volunteer organizations." Their use of the Internet was reported to be primarily as a way of keeping up with group news and organizing activities. Fewer than 25 percent of Internet users reported that the Net had a major impact on their ability to contribute time to their group. It did not ask explicitly about Internet-based volunteering. See Lee Rainie, Kristen Purcell, and Aaron Smith, *The Social Side of the Internet* (Washington, D.C.: Pew Research Center, January 2011), http://pewinternet.org/Reports/2011/The-Social-Side-of-the-Internet.aspx.

WikiLeaks and the PROTECT-IP Act: A New Public-Private Threat to the Internet Commons

Yochai Benkler

Abstract: The WikiLeaks affair and proposed copyright bills introduced in the Senate are evidence of a new, extralegal path of attack aimed at preventing access and disrupting the payment systems and advertising of targeted sites. In this model, the attacker may be a government agency seeking to circumvent constitutional constraints on its power or a private company trying to enforce its interests beyond those afforded by procedural or substantive safeguards in the law. The vector of attack runs through the targeted site's critical service providers, disrupting technical services, such as Domain Name System service, cloud storage, or search capabilities; and business-related services, such as payment systems or advertising. The characteristics that make this type of attack new are that it targets an entire site, rather than aiming for removal or exclusion of specific offending materials; operates through denial of business and financial systems, in addition to targeting technical systems; and systematically harnesses extralegal pressure to achieve results beyond what law would provide or even permit.

YOCHAI BENKLER is the Berkman Professor of Entrepreneurial Legal Studies at Harvard University, where he also serves as Faculty Co-director of the Berkman Center for Internet and Society. His publications include "The Commons as a Neglected Factor of Information Policy," Telecommunications Policy Research Conference (1998); "From Consumers to Users: Shifting the Deeper Structures of Regulation Toward Sustainable Commons and User Access," *Federal Communications Law Journal* (2000); and *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (2006).

In December 2010, a website that the Pentagon had described in 2008 as dedicated "to expos[ing] unethical practices, illegal behavior, and wrongdoing within corrupt corporations and oppressive regimes in Asia, the former Soviet bloc, Sub-Saharan Africa, and the Middle East," and that in 2009 had received the Amnesty International New Media Award for reporting on extrajudicial killings in Kenya, came under a multisystem denial-of-service attack intended to prevent it from disseminating information. The attacks combined a large-scale technical distributed-denial-of-service (DDoS) attack with new patterns of attack aimed to deny Domain Name System (DNS) service and cloud-storage facilities, disrupt payment systems services, and disable an iPhone app designed to display the site's content.

The site was WikiLeaks. The attackers ranged from unidentified DDoS attackers to Senator Joseph Lieberman and, more opaquely, the Obama admin-

istration. The latter attack is of particular interest here, having entailed an extralegal public-private partnership between politicians gunning to limit access to the site, functioning in a state constrained by the First Amendment, and private firms offering critical functionalities to the site – DNS, cloud storage, and payments, in particular – that were not similarly constrained by law from denying service to the offending site. The mechanism coupled a legally insufficient but publicly salient insinuation of illegality and dangerousness with a legal void. By publicly stating or implying that WikiLeaks had acted unlawfully, the attackers pressured firms skittish about their public image to cut off their services to WikiLeaks. The inapplicability of constitutional constraints to nonstate actors created the legal void, permitting firms to deny services to WikiLeaks. This, in turn, allowed them to obtain results (for the state) that the state is prohibited by law from pursuing directly. The range of systems affected by the attack was also new: in addition to disrupting technical service providers – which had been familiar targets since efforts to control the Net began in the 1990s – the attack expanded to include payment systems.

This pattern of attack is not an aberration. One need only observe its similarities to current efforts by the copyright industries to shut down sites that challenge their business models. This objective was laid out most explicitly in the first draft of the Combating Online Infringements and Counterfeits Act (COICA)[1] that was introduced in September 2010, and a powerful version of it remains in the present version of the bill, the Preventing Real Online Threats to Economic Creativity and Theft of Intellectual Property Act (PROTECT-IP Act) of 2011.[2] The COICA/PROTECT-IP approach, which replicates the dynamics of the WikiLeaks attack, endeavors to create a relatively procedure-free context for designating sites as legally suspect actors, while making critical service providers immune from responsibility for any action they take by denying technical, payment, and business process systems to targeted sites. Together, these elements form the basis for extralegal attacks on critical services, thereby creating a shortcut to shutting down allegedly offending sites. The insinuation of illegality creates the basis for public pressure on the service providers to deny service; immunity replicates the legal void that allows service-provider action well beyond anything a court would have ordered.

Combining denial-of-payment systems with the use of extrajudicial mechanisms and private party enforcement appears to extend basic techniques developed in the war on terrorism into the civilian domain. It represents a new threat not only to the networked commons, but to the very foundations of the rule of law in the United States.

On November 28, 2010, WikiLeaks, in cooperation with *The New York Times*, the *Guardian*, *Der Spiegel*, *Le Monde*, and *El País*, began to release a set of leaked U.S. embassy cables. The following is a condensed version of a detailed and fully documented event study of the response to that disclosure.[3] WikiLeaks, a site dedicated to making materials leaked by whistleblowers public, had published a series of items from the Pentagon and the State Department between April and November 2010. The first release, a video showing American helicopters shooting a Reuters photographer and his driver, exposed previously hidden collateral damage incurred in the pursuit of insurgents. The video was followed by the release of thousands of war logs in which field commanders described conditions on the ground in Afghanistan and Iraq. The disclosures were initially described by the administration as highly

damaging to the security of troops and human rights workers, but as time passed, formal Pentagon assessments sent to Congress suggested that no such harm had occurred.[4] On November 28, WikiLeaks and its traditional-media partners began to release documents selected from a cache of about 250,000 classified cables that U.S. embassies around the world had sent to the State Department. In late November and December they published, in redacted form, a few hundred of these cables. WikiLeaks's decision to publish the materials, including when and how they were published, was protected by First Amendment law. Indeed, precedents established at least as far back as the *Pentagon Papers* case support the proposition that a U.S. court would not have ordered removal or suppression of the documents, nor would it have accepted a criminal prosecution of WikiLeaks or any of its editors and writers.[5]

Despite the constitutional privilege that allowed WikiLeaks to publish the leaked documents, American political figures widely denounced the disclosures. Moreover, critics appeared to blame only WikiLeaks, even though traditional outlets such as *The New York Times* were providing access to the same cables, and in the same form. The most effective critic, Chairman of the Senate Homeland Security Committee Senator Joseph Lieberman, urged companies providing services to WikiLeaks to cease doing so. Senator Lieberman issued his call on December 1, 2010, following a well-crafted letter from the State Department to WikiLeaks sent November 27, 2010. That letter did not take the legally indefensible position that WikiLeaks itself had broken the law. Instead, it correctly asserted that the law had been broken (by someone), insinuating that WikiLeaks was the offending party. Not surprisingly, implicated service providers were among those who misread the letter. In a critical move, PayPal discontinued its service to WikiLeaks; a vice president of the firm, commenting publicly, pointed to the November 27 letter, not to Senator Lieberman's call, as the reason that PayPal believed WikiLeaks had broken the law, thus triggering the firm's decision to stop payment service to WikiLeaks.[6] The State Department letter was complemented by a series of public statements that tried to frame WikiLeaks's embassy cable release as international terrorism. Secretary of State Hillary Clinton called the release of the cables "an attack on the international community." Vice President Joseph Biden explicitly stated that Julian Assange, the founder of WikiLeaks, was "more like a high-tech terrorist than the Pentagon Papers." Senator Dianne Feinstein wrote a *Wall Street Journal* editorial calling for Assange's prosecution under the Espionage Act. Some right-wing politicians simply called for his assassination on the model of U.S. targeted killings against Taliban and Al Qaeda leaders.[7]

Against the backdrop of this massive public campaign against WikiLeaks, Senator Lieberman's December 1 public appeal was immediately followed by a series of service denials:

- December 1: *Storage*. Amazon removes WikiLeaks materials from its cloud-storage facility.
  - –*Countermeasure:* WikiLeaks moves storage to OVH in France.
- December 2: *DNS*. EveryDNS, the DNS registrar serving the WikiLeaks.org domain, stops pointing the domain name to WikiLeaks's server.
  - –*Countermeasure:* WikiLeaks uses numeric IP addresses updated through Twitter and begins to rely more heavily on WikiLeaks.ch DNS as well as on mirroring by various volunteers throughout the Net.

- December 3, 5: *Storage*. French Minister of Industry Eric Besson calls on OVH to cease providing storage; by December 5, OVH removes WikiLeaks content.
  - *Countermeasure:* WikiLeaks moves again, to Sweden, initially to the servers of the Pirate Party, a Swedish political party, and later to a Swedish storage provider.
- December 4: *Payment systems*. PayPal stops processing donations for WikiLeaks, cutting off a major source of funding. A vice president of PayPal points to the State Department's November 27 letter to WikiLeaks as the reason PayPal concluded that WikiLeaks was acting illegally and terminated service.
  - *Countermeasure:* No effective response. WikiLeaks loses substantial revenue as PayPal ceases to process donations. Loss of revenue continues with the credit card stoppages that follow.
- December 6: *Payment systems*. MasterCard stops servicing WikiLeaks. The Swiss Postal Bank closes Julian Assange's personal account with the Swiss bank for his failure to provide an adequate address.
- December 7: *Payment systems*. Visa joins MasterCard. Bank of America discontinues services ten days later.
- December 20: *App store*. Apple removes a third-party app created to allow iPhone users to access and search WikiLeaks embassy cables.
  - *Countermeasure:* WikiLeaks has no possible recourse. However, apps for the Android smartphone were not removed.

None of these companies was compelled by legal order to deny services to WikiLeaks. Indeed, under First Amendment law, it would have been impossible for the government or anyone else to obtain such an order. That aspect of U.S. constitutional law justifies describing this set of events as an *attack* on WikiLeaks. Put differently, the service denials to WikiLeaks were the result of an effort by the government to shut down the site irrespective of the fact that the law prohibited the government from doing so. In private conversations, individuals within and close to the administration emphatically denied any backchannel communications threatening or cajoling the companies. These claims seem plausible, and for purposes of analysis here, I consider them to be true. My claim, however, is based not on intent or the likelihood of conspiracy, but on effect. A public media campaign against WikiLeaks, led by top administration figures and some of the most senior politicians in the president's party, triggered vigilante actions by corporations that, unfettered by the laws constraining public-sector responses, likely saw themselves as acting in the national interest as they degraded the site's capabilities. Regardless of how its actions were perceived, WikiLeaks was engaged in classic fourth-estate functions at the core of freedom-of-the-press protections. In order to guard against similar outcomes in the future, it is important to understand and correctly characterize the events against the site as an attack on an important practice in the networked commons.

From a technical perspective, the attack was largely unsuccessful. The site proved enormously robust, using the core modes of networked resilience, namely, redundancy and decentralized cooperation. When WikiLeaks.org was denied DNS service, the site used a range of numeric IP addresses circulated on blogs and Twitter. It moved through a series of non-U.S. domains, the most important of which was the Swiss domain name WikiLeaks.ch. The Swiss DNS service provider, Switch, refused to capitulate to pressures to cease service to WikiLeaks. When cloud storage was denied in the United States, the site

moved first to France, where service was again denied under pressure from the French government, and then to Sweden. Moreover, thousands of mirror sites sprang up to permit access to the documents that had been released up to that point. However, where the system was not Internet based, as in the case of the iPhone app, it was impossible to replace. Nonetheless, the relative insignificance of the app, as long as an open Internet alternative existed, minimized the importance of that pathway. However, the fact that the WikiLeaks app was not easily replaceable provides an important indication of how vulnerable information is when available only over an iPhone or iPad-accessed network; the open Internet, by contrast, is robust.

Targeting WikiLeaks's business systems proved much more successful as a line of attack. WikiLeaks, which depends on donations from supporters to fund its operations, apparently lost 80 to 90 percent of its revenue stream in the first two months of the attack, and only gradually was able to create a set of proxies for receiving donations.[8] As was the case with the iPhone app, in the absence of a competitive market to offer significant redundant pathways for payment systems, persuading two or three companies to deny service was sufficient to severely hamper the site's payment operations. Whether a targeted site is a nonprofit dependent on donations or a for-profit or low-profit enterprise funded by transactions or advertising, an attack on the business systems a site depends on for financing appears harder to avert. This particular attack on payment systems seems to derive from the war-on-terror rhetoric applied to WikiLeaks as well as from a decade-old program established to compel payment and financial services firms to shut off funds flowing to terrorist organizations.[9]

The attack on WikiLeaks largely failed to achieve its goals. If it was aimed to prevent people around the world from accessing the leaked materials, it failed. The materials were made available on both distributed mirror sites and the sites of traditional media partners, whose public visibility seems to have made them invulnerable to the kind of informal, extralegal pressure that worked to deny service to WikiLeaks. If it was aimed to discredit the reports, it clearly failed here because WikiLeaks's partnership with traditional media helped raise visibility and add credibility to the documents. The technical aspect of the attack failed almost entirely: redundancy and the ability to move from one country to another allowed for robust storage, and the creation of thousands of mirror sites by individuals around the world made DDoS and DNS attacks ineffective.

Moreover, not all firms folded as easily as Amazon, PayPal, MasterCard, and Visa. Refusing to follow the U.S.-based EveryDNS, the Swiss DNS registrar continued to point to WikiLeaks.ch. Twitter declined to respond to document requests until subject to subpoena. Google did not remove related apps from the Android system or drop WikiLeaks results from its search engine. The success of an attack that relies on public pressure and a legal void in which to act depends on service providers' concern about being perceived as helping the targeted site; this concern must outweigh the providers' interest in maintaining their image as providers of robust, incorruptible services to the Internet-using public. Thus, the new form of informal, extralegal attack can be only partially effective if not all service providers are on board. Nonetheless, the denial of payment systems greatly affected WikiLeaks's cash flow and was likely the most effective and dangerous aspect of the attack.

This new pattern of attack (a) targeted an entire site; (b) was carried out through

denial of service by commercial service providers of critical technical and business capabilities; and (c) circumvented constitutional protections by creating an extralegal public-private partnership for censorship, using the inapplicability of constitutional limitations to private companies together with the relatively loose regulation of the standard-form contracts that govern the relations between service providers and their customers.

The WikiLeaks affair might properly have been dismissed as a one-off set of events if not for a similarly structured attack at the center of copyright legislation introduced in the Senate since late 2010. The PROTECT-IP Act is the most recent iteration of the U.S. copyright industries' seventeen-year-long drive to enlist various intermediaries and service providers of networked facilities to enforce their rights through law and public policy.[10] Beginning in the Clinton White House with a 1995 white paper[11] and culminating with the Digital Millennium Copyright Act (DMCA) of 1998,[12] the industries sought to create a set of liabilities that would lead Internet service providers (ISPs) and Web-hosting companies to remove infringing materials. The safe harbor notice and takedown procedures adopted in the DMCA represented the settlement of the first half-decade of policymaking in this field. Under these provisions, pure telecommunications carriers were excluded from the requirements of policing content. Providers of caching, Web-hosting, and search engines and Web directories were required to have a procedure in place for receiving notices regarding specific offending materials, and for taking down those materials; but they were not required to search out such content themselves or to block entire sites.

The following decade witnessed a legislative stalemate. On the one hand, the content industries hoped to expand control over materials on the Net in order to preserve and increase their revenues. On the other hand, a coalition of computer, software, and communications businesses that profited from the free flow of information and cultural goods online, together with civil society organizations aiming to preserve a space for a cultural commons, was concerned that efforts to impose controls would hamper the open, creative, participatory structure of the networked environment. While Republicans seemed less responsive to pressures from Hollywood, since 2006, Democrats controlling the Senate have pushed through a slate of laws designed to implement the Motion Picture Association of America's long-standing agenda. Most pertinent are the Prioritizing Resources and Organization for Intellectual Property Act (PRO-IP Act) of 2008, which created an IP czar in the White House and funded additional resources for criminal copyright enforcement,[13] and provisions in the Higher Education Opportunity Act of 2008[14] that required colleges to redesign their networks and develop offerings to protect the interests of Hollywood and the recording industry against their students. These laws include the two main elements of the bills currently under consideration: that is, they expanded the involvement of criminal enforcement authorities in what was traditionally an area of private commercial law, and they used state leverage to harness private platform providers to enforce the interests of the copyright industries.

Unlike the settlement of the 1990s, the most recent set of bills targets not offending content, but offending sites. While the DMCA focused on specific documents that violated copyright, new legislation – in the same vein as the WikiLeaks case – seeks to take out entire sites, specifically those defined as primarily dedicated to unauthorized distribution of copyrighted ma-

terials. It also substantially expands the set of addressees who are enlisted to aid the content industries. In addition to carriers, caching providers, and Web-hosting companies (which, in today's incarnation, cover cloud-storage facilities), the new bills cover DNS providers, advertising providers, and payment systems such as PayPal or credit card companies. From a procedural standpoint, the newest bills combine elaborate procedures that would allow a court order against sites or domain names not subject to U.S. jurisdiction, with subtle efforts to harness and formalize the extralegal public-private partnership exhibited in the WikiLeaks affair.

Introduced in September 2010 as the first bill in this series, COICA clearly identified its target as sites that have "no demonstrably commercially significant purpose" other than providing access through downloading, streaming, or linking to unauthorized materials. The breadth of the definition, however, captures much more, including "providing access to any goods or services in violation of the Copyright Act" or enabling a violation. The more tightly defined target is only an example of this broader set. For instance, the broader definition would include a creative site dedicated to anime music videos that provides the underlying songs, as is so often the case with the genre, in full or in substantial part – even though the work is transformative. The breadth of coverage becomes clearer when considering the blacklist described below; developed by a copyright industry firm in June 2011, the list included Archive.org and distribution of basic technical tools such as BitTorrent. Here, my point is not to challenge the definition, but to outline the method of attack on sites targeted under the proposed law. COICA empowers the Attorney General – the same government division that the 2008 legislation bolstered – to enforce copyright through criminal law. If the Department of Justice determines that a given domain name is associated with a site that falls under COICA's definition of unlawful behavior, it can petition for a court order that would obligate DNS providers in the United States to stop resolving the domain; or, if the domain is registered with a DNS provider used by U.S. customers but not subject to U.S. jurisdiction, any U.S. service provider, ISPs in particular, is required to take reasonable measures to prevent the domain name from resolving to the offending site. Moreover, "financial transaction providers" are required to cease servicing the site and enforce their copyrights to prevent the site from using their logos. Finally, contextual advertising providers are required to stop serving ads to the site. The innovations embedded in COICA, relative to prior legislation, are (a) the introduction of a broad-based attack at the site level, rather than removal of discrete documents, and (b) the harnessing of payment systems and advertising to deny economic viability to the site. In this sense, COICA presaged the attack on WikiLeaks through the payment system.

Another element of the original COICA was its particularly crisp platform for extrajudicial enforcement. Although the original has since been abandoned in favor of more subtle versions, the original form crystallizes the intent of the later versions. In its initial form, COICA required the Attorney General to "maintain a public listing of domain names that, upon information and reasonable belief, the Department of Justice determines are dedicated to infringing activities but for which the Attorney General has not filed an action under this section." The threshold for designation as an offending site is extremely low: the Department of Justice simply must allege "upon information and reasonable belief" that a site is dedi-

cated to infringing activities. This provision invokes standard language used in litigation to indicate the minimum level of knowledge required for plaintiffs to sustain a complaint without subjecting themselves to sanctions; it suggests a generalized suspicion more than a real investigation. Once a site is blacklisted, DNS service providers, ISPs, payment system providers, and advertising providers are immunized from liability if they deny service to the site listed as offending.

Note that the technique employed here is similar to the one utilized in the attack on WikiLeaks. The evidentiary threshold for state designation of a "bad actor" is well below what would be necessary to obtain judicial approval of that actor's "badness." For this reason, the statute cannot demand that private third parties comply with the enforcement efforts. Nonetheless, this substandard designation of bad-actor status can be used to pressure private service providers into acquiescence. By combining the extrajudicial designation with immunity for firms that discontinue service to the targeted sites, the state increases the likelihood that private parties will comply. The promise of immunity both expresses the state's expectation that cooperative private providers will, in fact, act against the designated entities and minimizes the risk and cost of doing so. The immunity creates the legal void necessary for vigilante enforcement and shows that such actions are desirable to the state. By contrast, the targeted site owner's defense becomes expensive. The procedure proposed would not create a legal black hole: the Attorney General was required to create mechanisms for allowing site owners to challenge their blacklisting and to appeal an unfavorable decision to a reviewing court. But the process reverses the normal presumptions of innocence. The "bad actors" blacklist, coupled with immunity, allows the state to place substantial pressure on sites deemed offending without obtaining a judicial determination prior to triggering the attack.

What makes this form of attack so worrisome? Ultimately, cases will be subject to judicial review, and if the court rules that the closure is unjustified, it will be lifted. The problem is that this procedure allows for effective elimination of revenues and technical access for lengthy periods pending review. Because there is no specific order or process prior to blacklisting, a site can find itself technically inaccessible and unable to use payment systems or advertising. Unless a site can immediately reestablish a backup presence – that is, use the redundancy of multiple sites – it will likely be economically dead by the time it can challenge the listing.

In combination, COICA expands the vectors of attack to include payment systems and advertising networks and provides an extralegal avenue of attack without prior judicial approval that can be sustained for an unspecified period while administrative and judicial appeals are pending. These elements largely, though not completely, enable the state to circumvent or severely curtail the requirements of legality and the protections of procedure.

The Senate abandoned this explicit entanglement of the state in extralegal enforcement. The procedure was replaced by an immunity provision that created space for private enforcement of the multisystem attack. In the revised bill, the provision simply states: "No domain name registry, domain name registrar, financial transaction provider, or service that provides advertisements to Internet sites shall be liable to any person on account of any action described in this section voluntarily taken if the entity reasonably believes the Internet site is dedicated to infringing activities." The promise of

immunity creates a legal space for informal pressures on advertising and financial services firms to deny services to potentially offending sites. It effectively invites private entities to create blacklists of their own. Similar to the reasonable belief envisioned in the original COICA bill, those lists could provide the justification for blocking targeted sites.

The current draft of the PROTECT-IP Act replicates this latter approach. It expands on COICA by (a) creating a private right of action, which gives the copyright industries the power to initiate and enforce the attacks and (b) making the immunity provision applicable with regard to any site accused, rather than only non-U.S. sites, as was the case in COICA. Section 5 of the PROTECT-IP Act immunizes any service provider that in "good faith and based on credible evidence has a reasonable belief that the Internet site is an Internet site dedicated to infringing activities." This weak standard encourages the creation of industry-maintained blacklists to implicate sites allegedly engaged in offending activities. In turn, the legal immunity creates the perfect context for putting pressure on private infrastructure, payment systems, and advertising providers to deny service to the blacklisted sites. Not surprisingly, in June 2011, less than a month after publication of the most recent iteration of this type of immunity, the advertising firm GroupM, whose clients include Universal Music, Paramount, and Warner Bros., developed a blacklist of more than two thousand sites to which it would not serve ads.[15] The list reportedly includes sites that indeed appear to provide primarily illegal downloads as well as sites whose practices are clearly non-offending, such as Archive.org and a broad range of basic technology sites that could, in principle, be used for file sharing.[16] Reliance on such a list is unlikely to fail the "good faith and based on credible evidence" test of "reasonable belief" set out in the PROTECT-IP Act. This makes the blacklist, however imperfect, a base from which to launch an extrajudicial attack on payment systems, contextual advertising, DNS, and other technical services of these sites, entirely circumventing the procedural and substantive protections embedded in the Copyright Act and the federal rules of civil procedure.

The years 2010 and 2011 have witnessed the introduction of a new pattern of attack on controversial websites, one that involves both the state and major private actors in a public-private partnership formed to suppress offending content. WikiLeaks publishes content that is of primary concern to the state; the suppression of such content is prohibited by the First Amendment. The attack on the site sought to circumvent constitutional protections by applying informal pressure (which is not reviewable under the Constitution) to private actors (who were not subject to constitutional constraints) to further the state's objective of suppressing the publication of the materials in question. PROTECT-IP represents the inverse of this public-private partnership for censorship. Here, the interests are those of certain segments of the business community – the copyright industries – seeking to use the state to help harness other private actors to enforce their interests.

The elements common to both methods of attack are the denial of business and technical systems and the use of extralegal or very weakly legally constrained forms to designate the target of attack and to define the pattern of denial of service. The effect is to dispense with, or at least limit, the procedural and substantive protections afforded to targeted sites, and to degrade, if not completely prevent, the operations of the organizations that use the site. All this is achieved with

practically no need for judicial approval before the action, and with only relatively expensive and slow judicial review while the attack is ongoing.

The features of the attack are eerily familiar. They are the common characteristics of what was described as early as September 24, 2001, as "the financial front in the Global War on Terrorism."[17] The COICA model for designating bad actors to be blocked by private parties replicates the model developed in 2001 that allowed the Treasury Department to designate "blocked persons," a label that triggers obligations by banks and others to freeze assets and deny further use of payment systems. Administrative designation without need for judicial order, or weak-to-nonexistent procedural protection for targets, combined with the use of private business systems providers to execute the goals of the state is rooted in the model developed for the "war on terror" of the first decade of the 2000s. This model now appears to be introducing two new elements into much more mundane areas of social policy and organization. The first is the use of extrajudicial models for designating targets for attack. The second is harnessing private actors, in particular business and financial systems providers, to choke off fund flows to suspected organizations. Setting aside debates over whether those elements can be justified when the targets are suspected terrorist organizations, observing them metastasize to the civilian part of normal political and economic life in a democratic, networked society is extremely troubling and should be resisted – politically, legally, and technically.

ENDNOTES

1 Combating Online Infringements and Counterfeits Act of 2010, 111th Cong., 2nd. sess., September 20, 2010, S. 3804.

2 Preventing Real Online Threats to Economic Creativity and Theft of Intellectual Property Act of 2011, 112th Cong., 1st. sess., May 12, 2011, S. 968.

3 The description is drawn from the extensively documented study, Yochai Benkler, "A Free Irresponsible Press: WikiLeaks and the Battle Over the Soul of the Networked Fourth Estate," *Harvard Civil Rights-Civil Liberties Law Review* 46 (2011). The study is also available at http://www.benkler.org/Benkler_Wikileaks_current.pdf.

4 Adam Levine, "Gates: WikiLeaks Don't Reveal Key Intel but Risks Remain," CNN.com, October 16, 2010, http://articles.cnn.com/2010-10-16/us/wikileaks.assessment_1_julian-assange-wikileaks-documents?_s=PM:US.

5 Benkler, "A Free Irresponsible Press," Part III.

6 According to a PayPal executive, "What happened is that on November 27th [the day before WikiLeaks began releasing cables] the State Department, the US government basically, wrote a letter saying that the WikiLeaks activities were deemed illegal in the United States. And so our policy group had to take a decision to suspend the account.... It was straightforward from our point of view"; Benkler, "A Free Irresponsible Press," n.146 – 148.

7 Ibid., Part II.A.

8 Ibid. This information is based on statements by Julian Assange in comments on an early draft of the article.

9 Patrick D. Buckley and Michael J. Meese, "The Financial Front in the Global War on Terrorism," U.S. Army War College Strategic Studies Institute, 2001, http://www.au.af.mil/au/awc/awcgate/army/usma_terrorists_finances.pdf.

[10] There are many histories of this long battle. See James Boyle, *The Public Domain: Enclosing the Commons of the Mind* (New Haven, Conn.: Yale University Press, 2008); Jessica Litman, *Digital Copyright: Protecting Intellectual Property on the Internet* (Amherst, N.Y.: Prometheus Books, 2001); and Yochai Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (New Haven, Conn.: Yale University Press, 2006).

[11] *Intellectual Property and the National Information Infrastructure: The Report of the Working Group on Intellectual Property Rights* (Washington, D.C.: U.S. Department of Commerce, 1995).

[12] Digital Millennium Copyright Act of 1998, Public Law 105-304, 105th Cong., 2nd. sess., October 28, 1998.

[13] Prioritizing Resources and Organization for Intellectual Property Act of 2008, Public Law 110-403, 110th Cong., October 13, 2008.

[14] The Higher Education Opportunity Act of 2008, Public Law 110-315, 110th Cong., August 14, 2008.

[15] Mark Sweney, "WPP Blacklists More than 2,000 US Websites," *Guardian*, June 8, 2011, http://www.guardian.co.uk/media/2011/jun/08/wpp-groupm-sir-martin-sorrell.

[16] "BitTorrent.com and Archive.org Blacklisted as Pirate Sites by Major Advertiser," TorrentFreak, October 6, 2011, http://torrentfreak.com/bittorrent-com-and-archive-org-blacklisted-as-pirate-sites-110610/.

[17] Buckley and Meese, "The Financial Front in the Global War on Terrorism."

Poems by Michael Longley

Puff-Ball

When we picked mushrooms at midnight
Among intersecting fairy rings,
You said moonlight had ripened them.

Later I found the moon's image –
The full moon's – a giant puff-ball
Taking shape as in a low cloud.

Notebook

Why did I never keep a notebook
That filled up with reed buntings
And blackcaps and chiffchaffs, their
Songs a subsong between the lines?

Early April. I am seventeen.
Under an overhanging whin bush
I have spotted linnets building.
A robin has laid her first egg.

Firewood

Out of the darkness and
Up the spiral staircase
I am carrying logs,
An armful every day,
Firewood for winter when
I shall not be here – wild
Fig perhaps – white sap
For curing warts, scrotum-
Concealing leaves – blackthorn,
Chestnut – for all I know –
From the skinny waterfall,
Antique olive branches,
Sycamore, mulberry –
At the back of the wood pile
Underneath the *casa*
Logs that will never burn
Disintegrating year
By year, forgetfulness,
Woodlouse, scorpion.

Tongue Orchid

I pass the first dilapidated
Chestnut that holds in its leaves
The waterfall's hurlygush,
When you call me back
Through tangles of paradise
Lily, bastard balm,
Nightshade, vetch to our very
First wasp-seducing
Tongue orchid, brownish red
Napkins neatly folded
As for a love-feast:
Why can't we find a name
For purple candelabra
And dusk-stars like signals
To amorous fireflies, yet
So white in their thicket
They mark the path ending
And things coming to an end?

Michael Longley, a Foreign Honorary Member of the American Academy since 2009, is one of Ireland's most prominent contemporary poets. His recent poetry collections include "The Weather in Japan" (2000), "Snow Water" (2004), and "Collected Poems" (2007). His newest collection, "A Hundred Doors," was published by Jonathan Cape in 2011. "Puff-Ball," "Notebook," "Firewood," and "Tongue Orchid,"

coming up in Dædalus:

On the American Narrative	Denis Donoghue, Rolena Adorno, Gish Jen, E. L. Doctorow, David Levering Lewis, Jay Parini, Michael Wood, William Chafe, Philip Fisher, Craig Calhoun, Larry Tribe, Peter Brooks, David A. Hollinger, William Ferris, Linda Kerber, and others
The Alternative Energy Future	Robert Fri, Stephen Ansolabehere, Steven Koonin, Michael Graetz, Pamela Matson & Rosina Bierbaum, Mohamed El Ashry, James Sweeney, Ernest Moniz, Daniel Schrag, Michael Greenstone, Jon Krosnick, Naomi Oreskes, Kelly Sims Gallagher, Thomas Dietz, Paul Stern & Elke Weber, Roger Kasperson & Bonnie Ram, Robert Stavins, Michael Dworkin, Holly Doremus & Michael Hanemann, Ann Carlson, Robert Keohane & David Victor, and others
Science in the 21st Century	Jerrold Meinwald, May Berenbaum, Jim Bell, Shri Kulkarni, Paul McEuen, Daniel Nocera, Terence Tao, M. Christina White, Bonnie Bassler, Neil Shubin, Joseph DeRisi, Gregory Petsko, G. David Tilman, Chris Somerville, and others
Public Opinion	Lee Epstein, Jamie Druckman, Robert Erikson, Linda Greenhouse, Diana Mutz, Kevin Quinn & Jim Greiner, Gary Segura, Jim Stimson, Kathy Cramer Walsh, and others

plus The Common Good, Immigration & the Future of America &c.

AMERICAN ACADEMY
OF ARTS & SCIENCES
Cherishing Knowledge · Shaping the Future

U.S. $13; www.amacad.org