**“Any person... a pamphleteer”**
**Internet Anonymity in the Age of Web 2.0**

by

Jonathan R. Mayer

April 7, 2009

A Senior Thesis presented to the Faculty of the Woodrow Wilson School of Public and International Affairs in partial fulfillment of the requirements for the degree of Bachelor of Arts.

In our world, all the sentiments and expressions of humanity, from the debasing to the angelic, are parts of a seamless whole, the global conversation of bits. We cannot separate the air that chokes from the air upon which wings beat.

John Perry Barlow, *A Declaration of the Independence of Cyberspace*

# Contents

# Abstract

Web 2.0, the proliferation of web-based services and applications supporting user generated content, revolutionizes the boundaries of speech: any individual can instantly and costlessly broadcast text, audio, video, or even an interactive experience to a global audience. The application features that support further Web 2.0 innovation come at a cost, however: the very same technologies enable identifying and tracking web users.

Chapter 2 of this work provides a technical overview of the Internet and the identifying information available in its underlying protocols to show that anonymity is technically feasible. After a review and critique of modern anonymizing technologies it proposes two novel deanonymizing techniques and experimental confirmation of their feasibility even against the most robust anonymization tools available. A final section considers the role of browser-based Web 2.0 features in these attacks, and concludes that those seeking anonymity have little influence over future browser developments that render them vulnerable.

In Chapter 3 discussion turns to whether individuals are aware of the identifying information associated with their online activities. A survey of Princeton undergraduates shows even the well-educated and technologically savvy are poorly informed about Internet anonymity and the anonymizing tools available. Qualitative and automated analysis of web search results shows that, while several outstanding resources on Internet anonymity exist, users would face tremendous difficulty locating them. Recent research on the psychology of web search indicates users would instead incorrectly adopt the advice of commercial anonymizing services or out-of-date pages, and leave themselves identifiable despite the perception of anonymity.

Chapter 4 considers this work's apolitical technical and individual findings in a policy context. A case in favor of Internet anonymity shows, with historical examples, its benefits in enhancing the public discourse, national security interests, and privacy. Analysis of legal precedent further suggests action by the U.S. government is constrained by an implicit right to employ Internet anonymity in the First Amendment. Despite these conclusions Internet anonymity does threaten real harms, and a final section proposes policies aimed at mitigating them, including: consumer awareness efforts, support of anonymizing tools, a coherent takedown framework for online content, and separate treatment of commercial interactions.

The benefits of Internet anonymity, by virtue of standardization and software promulgation, extend beyond America's shores, and hold the promise of piercing censorship in all nations. Setting aside the domestic debate over anonymity, the conclusion expounds its unparalleled promise in furthering human rights and national security interests abroad.

# Glossary

| | |
|---|---|
| Client | A mobile, intermittently active host that interacts with a stable server. |
| Gateway | The router on a LAN that links it to other LAN's. |
| Host | A device connected to a network. |
| HTTP | Hypertext Transfer Protocol, the application layer protocol that specifies client-server interactions on the web. TCP provides the reliable connection used by HTTP. |
| IP | Internet Protocol, the network layer protocol underpinning the Internet. |
| LAN | Local Area Network, a local network operating at the link and physical layers. |
| Router | A network layer device that transfers traffic between LAN's. |
| Server | A stationary, always-on host that interacts with unstable clients. |
| TCP | Transmission Control Protocol, a transport layer protocol that establishes reliable bi-directional communication. HTTP runs on top of TCP. |
| UDP | User Datagram Protocol, a transport layer protocol that sends data one-way with integrity but no delivery guarantee. |

# 1 Internet Anonymity and the Web 2.0 Revolution

"Governments of the Industrial World," wrote online rights activist John Perry Barlow in his 1996 *A Declaration of the Independence of Cyberspace*, "you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather."[1] A decade later, Barlow's anarchic utopia of "Internet exceptionalism" has scarcely come to pass.[2] In 1996,[3] 1998,[4] 2000,[5] and again in 2003[6] the U.S. Congress passed legislation restricting certain categories of online speech. The Federal Communications Commission[7] and Federal Trade Commission,[8] among other federal agencies, have both engaged in enforcement activities

1. John Perry Barlow, "A Declaration of the Independence of Cyberspace", `http://homes.eff.org/~barlow/Declaration-Final.html`.
2. On exceptionalism, see Lawrence Lessig, *Code version 2.0* (New York: Basic Books, 2006), 31-37.
3. "Telecommunications Act of 1996", `http://www.fcc.gov/Reports/tcom1996.txt`.
4. "Child Online Protection Act", `http://www4.law.cornell.edu/uscode/html/uscode47/usc_sec_47_00000231----000-.html`.
5. "Childrens' Internet Protection Act", `http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=106_cong_public_laws&docid=f:publ554.106`.
6. "CAN-SPAM Act of 2003", `http://uscode.house.gov/download/pls/15C103.txt`.
7. For example, Federal Communications Commission, "In the Matters of Formal Complaint of Free Press and Public Knowledge Against Comcast Corporation for Secretly Degrading Peer-to-Peer Applications", `http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-08-183A1.pdf`.
8. Federal Trade Commission, "Commission Enforcement Actions Involving the Internet and Online Services", `http://www.ftc.gov/bcp/internet/cases-internet.pdf`.

pertaining to online activity. And an increasing number of countries overseas, meanwhile, have enacted some form of Internet censorship.[9] Regulation of the Internet is here to stay.

Barlow similarly did not imagine the dramatic reshaping of the Internet landscape in the following decade. The explosive growth in users was foreseeable – the Internet user base was already increasing at an exponential rate[10] – but the shift in the very paradigms by which online content is generated and delivered came rapidly and unexpectedly. The Internet of the 1990's offered a limited set of communication tools to those with enough patience and savvy to overcome buggy software, slow data transfer, and the absence of documentation. But the heyday of FTP, newsgroups, chat rooms, and web pages littered with blue underlined links, generally maintained by only organizations or aficionados, has long since passed. Social networking (Facebook, MySpace, and LinkedIn, e.g.), blogging (Blogger and Twitter), photo and video sharing (Flickr and YouTube), file sharing (DropBox), collaborative document editing (Google Docs), knowledge sharing (Wikipedia) and countless other genres of web services and applications falling within the ambit of the "Web 2.0" label enable the average user to broadcast nearly any form of media to a worldwide audience instantaneously. As *Time* elucidated in its citation of "You" as the 2006 "Person of the Year," some "call it Web 2.0, as if it were a new version of some old software. But

9. Ronald J. Deibert et al., eds., *Access Denied: The Practice and Policy of Global Internet Filtering* (MIT Press, 2008), 237-432.

10. International Telecommunication Union, "Free statistics", `http://www.itu.int/ITU-D/ict/statistics/`.

it's really a revolution."[11]

Just as with every prior expansion of the reach of content and ideas,[12] the Web 2.0 revolution raises the quandaries that inevitably accompany free speech: What is permissible? How ought the government respond? Anonymous publication is a prominent facet of these issues. In the history of American public discourse anonymity enjoys a vaunted role; *Common Sense*, the *Federalist Papers*, and "The Sources of Soviet Conduct" were all published under pseudonyms.[13] The same holds true abroad: Dickens, Malthus, Voltaire, and Maréchal, to name but a few, released works without attribution.[14] Despite its popular veneration anonymity is not without its flaws; by rendering accountability impossible, anonymity enables libelous or criminal activity without fear of retribution.[15]

The aim of this work is to examine Internet anonymity, couched in the context of regulation and Web 2.0, from the technological, individual, and policy perspectives. The intended audience is twofold: the computer scientists who advance the Internet's architecture and the public policy practitioners who structure legislation and gov-

11. Lev Grossman, "Time's Person of the Year: You", *Time Magazine* (December 13, 2006).

12. For discussion at length of how the web encourages the development and sharing of new ideas see Yochai Benkler, *The Wealth of Networks* (New Haven: Yale University Press, 2006).

13. Jonathan D. Wallace, "Nameless in Cyberspace: Anonymity on the Internet", *CATO Institute Briefing Papers*, no. 54 (1999): 2-3.

14. Michael H. Spencer, "Anonymous Internet Communication and the First Amendment: A Crack in the Dam of National Sovereignty", *Virginia Journal of Law and Technology* 1, no. 3 (Spring 1998); See also John Mullan, *Anonymity: A Secret History of English Literature* (Princeton: Princeton University Press, 2008).

15. A. Michael Froomkin, "Legal Issues in Anonymity and Pseudonymity", *The Information Society* 15, no. 2 (1999): 113 –127. David Davenport, "Anonymity on the Internet: Why the Price May Be Too High", *Communications of the ACM* 45, no. 4 (2002): 33–35. Gary T. Marx, "What's in a Name? Some Reflections on the Sociology of Anonymity", *The Information Society* 15, no. 2 (1999): 99–112.

ernment action. Only through the cooperation of these two groups will a desirable outcome for Internet anonymity be tenable.

As the framework employed for systematically addressing Internet anonymity, the technological, individual, and policy perspectives form the basis of this work's organization. Chapter 2 examines the technical underpinnings of anonymity online to demonstrate its feasibility, proposes two novel techniques for identifying web users even employing the best anonymizing technology publicly available, and concludes with a discussion of how anonymity is increasingly challenged by Web 2.0 innovations. Having established a technical grounding, Chapter 3 assesses the prevalence of knowledge and availability of information about Internet anonymity from the perspective of the individual user to show most are unaware and unlikely to learn of the anonymizing tools available. In Chapter 4 this work finally considers its deliberately apolitical technological and informational findings in a policy context, presents a case in favor of anonymity motivated by historical examples, builds the legal argument that Internet anonymity is a First Amendment right, acknowledges the harms of anonymity, and finally recommends specific actions for government and non-government actors.

The effects of America's response to Internet anonymity by necessity ripple worldwide, and the stakes of the anonymity debate grow ever higher. In 2008 alone bloggers played key roles in the Zimbabwean election, Egyptian general strike, and Thai protests, among other international incidents.[16] Online writers have been ha-

16. Global Voices, "GlobalVoices Special Coverage", `http://globalvoicesonline.org/specialcoverage/`.

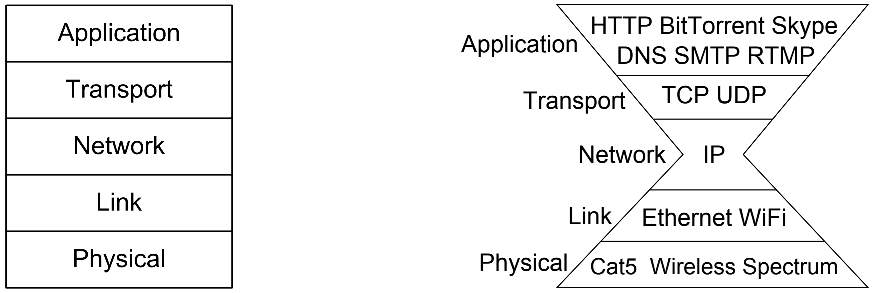
rassed and jailed, both by government and private forces, in a number of countries intolerant of their expositions. A frank discussion of American Internet anonymity policy in the present simultaneously provides real benefits to threatened speech at home and abroad and prepares technologists and policymakers to encounter the future challenges and innovations the Internet will doubtlessly yield.

# 2 The Nuts and Bolts of Internet Anonymity

The Internet was never intended to facilitate anonymity. Originally developed as a means of linking disparate military networks, the protocols underlying the Internet were designed to meet a list of criteria including disruption tolerance, underlying network agnosticism, and low cost.[1] Accountability and identity were among the lowest priorities for the DARPA Internet team, which only envisioned network threats originating from *outside* the Internet. The network architecture that resulted, and is in use on the Internet today, consequently provides only a loose conception of identity and significant opportunities for anonymity. This chapter presents a brief technical overview of the Internet and Internet anonymity, two novel deanonymizing techniques, and a concluding analysis of the strained relationship between Web 2.0 technologies and anonymity.[2]

1. David D. Clark, "The Design Philosophy of the DARPA Internet Protocols", in *SIGCOMM '88* (1988), 106–114.

2. Particularly essential concepts are underlined and defined separately in the glossary. For alternate technical discussions see Ian Goldberg, "Privacy and Anonymity on the Internet", in *Workshop on Vanishing Anonymity, 15th Conference on Computers, Freedom, and Privacy* (2005); Richard Clayton, "Anonymity and traceability in cyberspace", *University of Cambridge Computer Laboratory Technical Reports*, no. 653 (November 2005).



(a) The network layer model.

(b) Common standards at each layer.

Figure 2.1: Layers in a computer network.

## The Internet: A Brief Technical Overview

The prevailing paradigm for modeling computer networks conceptually and graphically separates the functionality of components into vertically distinct "layers," each dependent upon those below it, as shown in Figure 2.1a.[3] At the base is the physical layer, the hardware and media employed to transmit data;[4] common physical platforms include Cat5 Ethernet cable[5] and the wireless spectrum specified by the WiFi[6] standard. The closely related link layer rests atop and consists of a protocol for transmitting data over the physical network.[7] The Ethernet protocol, for example, provides a means of communicating data over any physical network conforming to the Ethernet hardware specifications.

3. James F. Kurose and Keith W. Ross, *Computer Networking: A Top-down Approach Featuring the Internet*, Third Edition (Pearson Education, 2004), 19-30.

4. Ibid., 27.

5. "IEEE 802.3 ETHERNET", `http://www.ieee802.org/3/`; Kurose and Ross, *Computer Networking: A Top-down Approach Featuring the Internet*, 111-120.

6. Also referred to as Wireless Ethernet , "IEEE 802.11, The Working Group Setting the Standards for Wireless LANs", `http://www.ieee802.org/11/`; Kurose and Ross, *Computer Networking: A Top-down Approach Featuring the Internet*, 131-137.

7. Ibid., 27.

The network layer deserves special attention: it allows devices on the network, referred to as hosts, to communicate with a single protocol (dubbed a "thin waist") independent of heterogeneous underlying hardware and link implementations.[8] The Internet Protocol, or IP, was the key innovation made by DARPA[9] and remains the Internet-wide standard at the network layer.[10] Internet traffic is divided into small chunks of data, packets, and directed towards the recipient by specialized hosts known as routers, which pass along a packet until it reaches its destination.[11] IP provides no performance guarantees: packets could be delayed, arrive out of order, be corrupted, or simply disappear.

The transport layer attempts to guarantee properties of data delivery over the unreliable network layer.[12] The majority of Internet traffic employs one of two dominant protocols, the Transmission Control Protocol (TCP) and the User Datagram Protocol (UDP).[13] TCP[14] constructs a reliable two-way connection between a pair of hosts, guaranteeing data will be received intact and in order – or not at all. UDP,[15] on the other hand, provides the more limited guarantee that if data is received it will

8. Kurose and Ross, *Computer Networking: A Top-down Approach Featuring the Internet*, 236-241.

9. Vinton G. Cerf and Robert E. Kahn, "A Protocol for Packet Network Intercommunication", *IEEE Transactions on Communications* 22 (1974): 637–648; Jon Postel, ed., "Internet Protocol", *RFC*, no. 791 (1981).

10. IPv6, an updated version of IP, will provide an increased address space but does not alter the function of the protocol. S. Deering and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", *RFC*, no. 2460 (1998).

11. Kurose and Ross, *Computer Networking: A Top-down Approach Featuring the Internet*, 242-245, 271-294, 299-318.

12. Ibid., 27.

13. Ibid., 374-375.

14. Jon Postel, ed., "Transmission Control Protocol", *RFC*, no. 793 (1981).

15. Jon Postel, ed., "User Datagram Protocol", *RFC*, no. 768 (1980).

not be corrupted.

The application layer is the uppermost layer, and specific to the programs run on each host.[16] The web, streaming media, BitTorrent, Skype, and countless other families of application-specific protocols and standards – the vast majority not ratified by any formal body – all run in this layer. Focusing on the web, the Hypertext Transfer Protocol (HTTP)[17] specifies how hosts interact. In the HTTP paradigm, and in many protocols, one host is the client and the other is the server; servers exist at a stable address and allow ephemeral clients to either request or submit content.[18] Nearly all of Web 2.0 is based on the client/server paradigm; using a web client, whether a browser or custom application, visitors publish content to and view it on a web-based service.

The Internet is not, as is commonly assumed, a global centrally administered network; it is a network of independently operated networks,[19] "autonomous systems" (AS's), which "gossip" with their neighbors about how to reach Internet hosts.[20] Thus, for example, Princeton University tells Patriot Media how to reach hosts on campus, Patriot in turn tells Sprint that it has a connection to Princeton, Sprint informs AT&T of its path to Princeton through Patriot, and finally AT&T tells Yale

16. Kurose and Ross, *Computer Networking: A Top-down Approach Featuring the Internet*, 27.
17. T. Berners-Lee, R. Fielding, and H. Frystyk, "Hypertext Transfer Protocol – HTTP/1.0", *RFC*, no. 1945 (1996); R. Fielding et al., "Hypertext Transfer Protocol – HTTP/1.1", *RFC*, no. 2068 (1997); R. Fielding et al., "Hypertext Transfer Protocol – HTTP/1.1", *RFC*, no. 2616 (1999).
18. Kurose and Ross, *Computer Networking: A Top-down Approach Featuring the Internet*, 16-18.
19. Ibid., 299-301, 316-318.
20. See Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", *RFC*, no. 1771 (2006).

about its path to Princeton through Sprint and Patriot. All of the routers between Princeton and Yale then have sufficient knowledge such that when Yale sends data destined for Princeton to AT&T, the traffic is properly routed all the way to Princeton. Each AS consists of smaller networks in turn; at the lowest level of subdivision is a Local Area Network (LAN), where data is transmitted from host to host at the link layer. LAN's are able to communicate between one another by directing traffic to a local network layer router, the gateway, that connects to other LAN's. The Internet is, in fact, a hierarchy of such LAN's, with countless user networks at the bottom and a small set of "Tier 1" networks at the top.[21]

## Deanonymizing Data in Network Protocols

Before delving into the deanonymizing data available in the aforementioned network layers, one must first define anonymity in the context of Web 2.0. Two general properties are desirable:[22]

1. A service or user cannot gain significant knowledge about another user's identity from the actions they take.[23]

2. A service or user cannot determine whether a set of actions were committed by the same user or group of users.[24]

21. Cooperative Association for Internet Data Analysis, "Visualizing IPv4 Internet Topology at a Macroscopic Scale", 2008, `http://www.caida.org/research/topology/as_core_network/`.

22. The duals of these two properties form the threat model: an adversary could seeks to learn a user's identity or track them.

23. Anonymity is not defined as preventing users from uniquely identifying each other because knowledge of some trait of a user, i.e. that they are a Princeton student, could be sufficient for an adversary to act upon.

24. The "group of users" caveat is included for the same reasoning as above.

The first property follows directly from anonymity requiring a hidden identity, but the second is more nuanced; to be anonymous in a Web 2.0 context a user must not be trackable between interactions *even without knowledge of their identity.* Consider the trivial example of an oppressive regime aspiring to curb critical speech: it need not know the true identity of a critic to silence them, only a systematic way of identifying their content and preventing it from reaching an audience. For the purposes of the following analysis any system that guarantees only the first property is said to provide *weak anonymity*,[25] and any system that ensures both properties provides *strong anonymity.* It should be noted upfront that weak anonymity is fraught with risk – should a user leak their identity to even a single party, their anonymity could be pierced for all past and future interactions.[26] Further, small scraps of independently useless information could be combined across interactions to discover a user's identity.[27]

Data that identifies users on a computer network could potentially exist at each of the network layers discussed earlier.[28] Assuming the absence of explicitly provided identifying information,[29] the primary way to determine identity is through addresses,

25. Note that this conception differs slightly from pseudonymity in that an individual need not have a consistent pseudonym, or even be aware their interactions are trackable.

26. In a sense, then, strong anonymity could be considered a rough parallel of "forward secrecy:" revealing a secret at one point in time does not compromise the secret elsewhere.

27. Arvind Narayanan and Vitaly Shmatikov, "De-anonymizing Social Networks", *IEEE Security and Privacy* (2009).

28. The analysis in this chapter focuses on determining a user's identity by identifying their host. While this is not, of course, always a one-to-one mapping, it is both the best one can hope for without a user explicitly identifying themselves and usually sufficient to, if nothing else, gain significant information about a user's identity.

29. It should be noted that user content without explicitly identifying information can be tracked with a variety of techniques, but these generally do not scale to the level of filtering. See, for example, Jiexun Li, Rong Zheng, and Hsinchun Chen, "From Fingerprint to Writeprint", *Communications of*

values that enable hosts to contact one another. While experimental network designs show a network could be operated with a sole address for each host, performance necessitates addresses at multiple layers;[30] as implemented on the Internet, the network stack offers addresses at the link, network, and application layers.

Local uniqueness in addresses is essential at the link layer for hosts to determine which data on the LAN is addressed to them. To enable host mobility among LAN's this local uniqueness requirement is enforced on Ethernet and WiFi networks by ensuring each link layer address is globally unique. Hosts connect to both types of network with a Network Interface Card (NIC), which is imprinted with a globally unique six byte <u>Machine Address Code</u> (MAC) during manufacturing.[31]

Though MAC's are globally unique, they provide little means of identifying an end host on the Internet. As discussed earlier, LAN's are interconnected on the Internet at the network layer by IP. The sender's MAC is associated with data only until it reaches the gateway; the gateway and subsequent routers change the sender and receiver MAC's associated with the data as necessary to forward it along a path towards the recipient – and in many cases the data will transit a network with a different link layer protocol, obliterating any associated MAC's. When the data is finally delivered the receiving host, if using Ethernet or WiFi, only observes its gateway's MAC as the sender and its own MAC as the receiver; the recipient has no

*the ACM* 49, no. 4 (April 2006): 76–82.

30. Matthew Caesar et al., "ROFL: Routing on Flat Labels", in *SIGCOMM '06* (2006), 363–374.

31. The IEEE Registration Authority assigns MAC prefixes to manufacturers, who in turn set the complete, unique MAC for each NIC , "IEEE Registration Authority", `http://standards.ieee.org/regauth/oui/index.shtml`.

way of recovering the sending host's MAC.

The network layer Internet Protocol, on the other hand, provides a fairly reliable means of both identifying and locating hosts on the Internet. Each host is either assigned a globally unique IP address or shares one with a small number of other hosts.[32] Unlike link layer identifiers, sending and receiving IP addresses remain associated throughout a packet's transit of the Internet to facilitate two-way communication. IP addresses can consequently be used to uniquely identify an Internet host or at minimum discover the small group of hosts to which it belongs. Converting an IP address into an identity and location is relatively trivial: IP address blocks are allocated to organizations by Regional Internet Registries (RIR's), who maintain publicly accessible "WHOIS" records of current assignments. The following information, for example, can be instantly gleaned from the North American RIR, ARIN, about a Princeton University IP address:

```
$ whois 128.112.224.200

OrgName:     Princeton University
OrgID:       PRNU
Address:     Office of Information Technology
Address:     87 Prospect Avenue
City:        Princeton
StateProv:   NJ
PostalCode:  08540
Country:     US
...
```

32. Network Address Translation (NAT) maps UDP and TCP traffic to a fixed number of identifiers (ports), limiting the number of hosts a single global IP address can support. K. Egevang and P. Francis, "The IP Network Address Translator (NAT)", *RFC* (1994); IP anycast is an exception to this generalization, but has limited use beyond DNS root servers. C. Partridge, T. Mendez, and W. Milliken, "Host Anycasting Service", *RFC*, no. 1546 (1993); T. Hardie, "Distributing Authoritative Name Servers via Shared Unicast Addresses", *RFC*, no. 3258 (2002).

For the exact identity of a remote host's user an inquiring individual need only approach the address assignee (in this case Princeton University's Office of Information Technology) for its logs of user-address assignments. This very approach has been employed by the recording and movie industries in pursuit of illegal file sharers.[33] In cases where registry information is inaccurate an inquiring host can actively probe the sequence of routers leading to a remote host with the `traceroute` utility.[34] Should both of these avenues fail an Internet host at best can maintain weak anonymity; IP addresses change infrequently if at all, enabling simple correlation of identity across activities on the Internet.

At the application layer, all bets are off. Hosts could transmit data that is completely identifying – a name and address, for example – or data that is wholly non-unique. Even the most deliberate users on occasion make this mistake – in a recent embarrassing incident a number of foreign embassy officials were identified despite using an anonymizing technology below the application layer (Tor, discussed later) because they provided their email user names and passwords in an unencrypted form.[35] Though a gross generalization, by and large the protocols most widely used on

33. Electronic Frontier Foundation, "RIAA v. The People: Five Years Later", (San Francisco, California, United States) (2008), `http://www.eff.org/files/eff-riaa-whitepaper.pdf`.

34. The TTL-based technique employed in `traceroute` is not unique to the program, but it is the most often used implementation , "traceroute", `http://www.freebsd.org/cgi/man.cgi?query=traceroute`.

35. Dan Goodin, "Tor at heart of embassy passwords leak", *The Register* (2007), `http://www.theregister.co.uk/2007/09/10/misuse_of_tor_led_to_embassy_password_breach/`; Kim Zetter, "Embassy E-mail Account Vulnerability Exposes Passport Data and Official Business Matters", *Wired Threat Level* (August 31, 2007), `http://blog.wired.com/27bstroke6/2007/08/embassy-e-mail-.html`; Kim Zetter, "Rogue Nodes Turn Tor Anonymizer Into Eavesdropper's Paradise", *Wired Threat Level* (September 10, 2007), `http://www.wired.com/politics/security/news/2007/09/embassy_hacks`.

the Internet and for Web 2.0 interactions contain a minimum of identifying information unless users add it. The following is an example HTTP request generated by the Mozilla Firefox 3.0.5 browser on Mac OS X 10.5 for `http://www.princeton.edu/`:

```
GET / HTTP/1.1
Host: www.princeton.edu
User-Agent: Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10.5; en-US; rv:1.9.0.5) Gecko/2008120121 Firefox/3.0.5
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive
```

While a requester's operating system, web browser, and default language are provided, no particularly unique attributes can be gleaned.

HTTP cookies, short text strings a server assigns to clients, present a more successful web-based identification method. Once a client has a cookie stored for a site it includes the cookie in every subsequent request to the site. Embedding content from the same source consequently allows cross-site tracking with a single cookie; behavior-based advertising[36] and cross-site request forgery attacks[37] rely on this very mechanism. Adobe's Flash plug-in can also be manipulated to store cookie-like data,[38] as can the browser's cache.[39] All three of these approaches allow tracking

36. Miguel Helft, "Google to Offer Ads Based on Interests", *The New York Times* (March 11, 2009).

37. Adam Barth, Collin Jackson, and John C. Mitchell, "Robust Defenses for Cross-Site Request Forgery", in *15th ACM Conference on Computer and Communications Security (CCS 2008)* (2008).

38. Corey Benninger, "AJAX Storage: A Look at Flash Cookies and Internet Explorer Persistence", *Foundstone White Papers* (2006).

39. Martin Poole, "meantime: non-consensual http user tracking using caches" (2000), `http://sourcefrog.net/projects/meantime/`.

| Anonymizing mechanism | Web access | Weak anonymity | Strong anonymity |
|---|---|---|---|
| IP Spoofing | No$*$ | Maybe$\circ$ | Maybe$\circ$ |
| IP Proxy | Yes | Maybe† ‡ $\bullet$ | Maybe‡$\bullet$ |
| Web Proxy | Yes | Maybe† ‡ $\bullet$ | Maybe‡$\bullet$ |
| "Private Browsing" | Yes | No$\star$ | No$\star$ |
| Tor and Torbutton | Yes | Yes | Yes |

$*$ No return traffic. † Need sufficient traffic. ‡ Need trustworthy proxy. $\star$ IP visible, often errors in implementation. $\circ$ If no monitoring of the LAN. $\bullet$ Traffic to proxy must be encrypted.

Table 2.1: Analysis of anonymizing techniques for a web client.

a user from site to site, violating the second anonymity principle and allowing weak anonymity at best.

## Existing Techniques for Anonymity

Of the threats to online anonymity discussed above, IP addresses and cookies pose the two gravest risks. This section addresses the techniques used to minimize the impact of each in turn, also summarized in Table 2.1.

Before continuing on to the IP address anonymizing techniques available to more honorable Internet users, it should be acknowledged that anonymity can be attained by compromising another individual's computer and routing encrypted traffic through it. Botnets use this very approach on a large scale to anonymously send spam email, for example. The relatively frequent discovery of vulnerabilities in end host software[40] suggests this will remain a viable tactic in future for those willing to commit criminal acts.

40. United States Computer Emergency Readiness Team, "US-CERT Vulnerability Notes", `http://www.kb.cert.org/vuls/`.

The intuitive solution to the IP address problem would suggest simply using alternate computers for anonymous activity as might be available at a public library or cybercafe. While on face the concept is appealing, it is rife with possibilities for identity disclosure. Shared computers largely conform to one of two models: either they give free reign over software, or they restrict activity to little beyond web access. In the former case, while anonymizing tools could be employed, there is also no technical barrier to another user installing (purposefully or inadvertently) spyware, keyloggers, or other potentially identity compromising software. The latter option, on the other hand, ensures IP addresses and cookies will be available to a party seeking to discover the user's identity. Moreover, in either scenario a nation could impose mandatory real-time monitoring of user activity, eliminating any potential for anonymity.[41] As for the physical anonymity provided by a public computer – that even if an adversary were to discover a specific computer was used for certain online activity, no person would be implicated – several nations including South Korea, China, and Italy have begun requiring identification and even a photo before access to a shared computer.[42] Payment by credit card and video surveillance provide additional means for cybercafes and prying states to monitor shared computer use.[43]

Another approach an end user might adopt to anonymize their IP address is

41. Deibert et al., *Access Denied: The Practice and Policy of Global Internet Filtering*, 109.

42. "Cameras Draw Closer to Beijing's Internet Cafes", *The Wall Street Journal China Journal Blog* (October 17, 2008), `http://blogs.wsj.com/chinajournal/2008/10/17/cameras-draw-closer-to-beijings-internet-cafes/`; Sofia Celeste, "Want to check your e-mail in Italy? Bring your passport.", *The Christian Science Monitor* (October 4, 2005); Deibert et al., *Access Denied: The Practice and Policy of Global Internet Filtering*, 83, 65.

43. Ibid., 65.

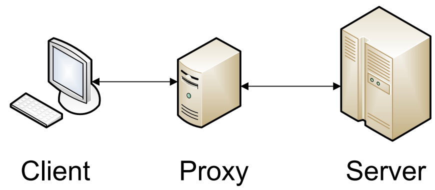



Figure 2.2: Routing traffic through a proxy.

lying: without much special configuration a computer can be set to use any address. While this technique is viable for one-way UDP traffic, all reverse direction traffic – essential for any interactive or TCP-based application, including HTTP and Web 2.0 services – will not be delivered; the receiving host would send packets destined for the fake IP, but they would be dropped either at the host actually assigned that IP or by the first router to recognize no path exists to the IP. Moreover, should an adversary be able to monitor the sender's LAN, they will be able to identify the sending host by the MAC address on the spoofed packets.

An alternative approach is the adoption of IP-level proxies or Virtual Private Networks (VPN's), often operated either as a public free[44] or private commercial[45] service. All of a client's IP traffic is routed through the proxy, leaving remote hosts with knowledge of only the proxy's IP address (Figure 2.2). In fact, even on an Internet that associated identity with every packet (as has been proposed on several occasions[46]) routing through a proxy would remain a feasible anonymizing technique

44. For example, Public Proxy Servers, "Public Proxy Servers", `http://www.publicproxyservers.com/`.
45. For example, Anonymizer, Inc., "Anonymous Surfing", `http://www.anonymizer.com/consumer/products/anonymous_surfing/`.
46. i.e. Declan McCullagh, "U.N. agency eyes curbs on Internet anonymity", *CNET News* (September 12, 2008), `http://news.cnet.com/8301-13578_3-10040152-38.html`.

– all sub-application layer identifiers are erased upon transiting the proxy. Three flaws significantly impede this approach's success, however. First, a proxy must be trustworthy; it knows the true IP address of the client, and if compromised (legally or otherwise) would provide no anonymity. Second, enough clients must route their traffic through the proxy such that a given user's data will be sufficiently indistinguishable from other traffic passing through the proxy.[47] In the worst case, suppose only one user routed their traffic through a proxy; while remote sites would not know the user's true IP address, they could still easily track activity from site to site, thereby guaranteeing only weak anonymity. Third, without encryption between the client and proxy an adversary need only monitor the connection between the two to determine what traffic belongs to the client.

Web proxies offer similar anonymizing properties. Like IP proxies these systems protect identity by redirecting traffic – in this case HTTP – through an intermediary. In addition to the weaknesses of IP redirection discussed above, web proxies depend upon forcing web browsers to load content through them. For proxies compliant with the HTTP proxy standard[48] this support is built into the web browser; a user need only configure their browser to route all traffic through the proxy. Some proxies[49] avoid any user configuration by rewriting web content to preserve anonymity. Both approaches disclose a user's IP address in non-HTTP traffic (i.e. certain types

47. Roger Dingledine and Nick Mathewson, "Anonymity Loves Company: Usability and the Network Effect", in *Proceedings of the Fifth Workshop on the Economics of Information Security (WEIS 2006)* (2006).
48. Fielding et al., "Hypertext Transfer Protocol – HTTP/1.1".
49. For example , "The Cloak", `http://www.the-cloak.com`.

of streaming video), however, and the latter requires careful modification of all web content to ensure no requests originate from the client. Though some sites attempt to cleverly provide a safety net from inadvertently requesting content without the proxy by maintaining a secure connection and setting the user's browser to issue a warning upon leaving a secure website, this mechanism is easily defeated by directing the user to another secure site.

Turning to cookies and the cookie-like techniques discussed earlier, solutions are similarly problematic. Though most web browsers provide a setting for disabling cookies, Flash and cache "cookies" remain effective tracking mechanisms. Many newer browsers now offer some form of "private browsing" mode that purports to disable all cookie-like tracking, but a December 2008 study found weaknesses in the implementations in all four major web browsers[50] – and one browser[51] where private browsing appeared to have no effect on cookies![52] The study also concluded that clearing Flash cookies is too complicated for a lay user, and disabling them completely requires navigating a convoluted process on, counterintuitively, Adobe's website.

The Tor Project presents the most thorough and widely-adopted solution to the IP address and cookie problems yet.[53] Unlike traditional IP proxies, which rely on

50. Microsoft Internet Explorer, Mozilla Firefox, Apple Safari, and Google Chrome.

51. Apple Safari on Windows.

52. Katherine McKinley, "Cleaning Up After Cookies", *iSEC Partners White Papers* (2008), http://www.isecpartners.com/files/iSEC_Cleaning_Up_After_Cookies.pdf.

53. Roger Dingledine, Nick Mathewson, and Paul Syverson, "Tor: The Second-generation Onion Router", in *Proceedings of the 13th USENIX Security Symposium* (2004), 303–320; "Tor: Overview", http://www.torproject.org/overview.html.en; other projects, including Java Anon Proxy, I2P, and Mixminion (for email) offer similar anonymizing properties below the application layer, but this analysis is limited to Tor because of its popularity , "Mixes for Privacy and Anonymity in the Internet", http://anon.inf.tu-dresden.de/develop/doc/mix_short/; "Introducing I2P",

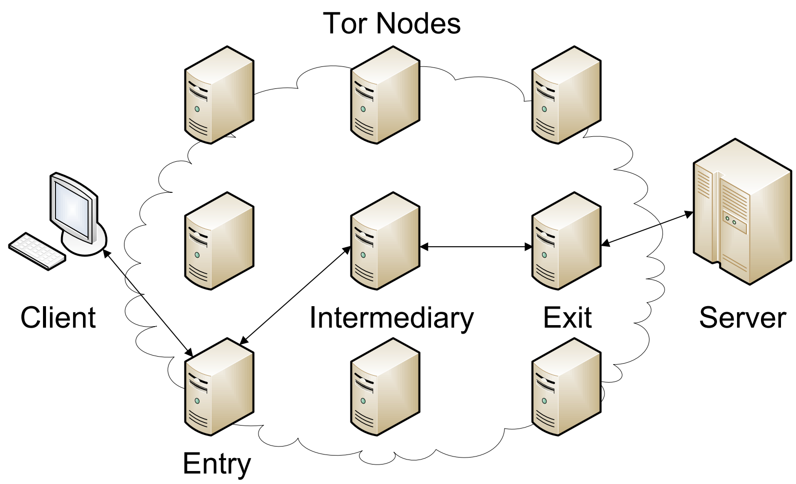



Figure 2.3: Tor routing in the client-server paradigm. Only the entry node knows the client's IP address, and only the exit node and server can observe the client's traffic.

a single trusted intermediary, Tor builds short-duration paths ("virtual circuits") of three shared *untrusted* intermediary nodes (Figure 2.3). In a process dubbed "onion routing" the client layers encryption[54] on their data, which is stripped off (analogous to layers of an onion) by each of the three intermediaries in succession; only the first intermediary knows the client's IP address, and only the last intermediary can decode the data and send it to the receiving host. Below the application layer, then, Tor provides strong anonymity: a user's apparent IP address changes on a short timescale, and the use of shared nodes ensures a user's traffic is sufficiently mixed with other traffic to be indistinguishable. Moreover, Tor's reliance on untrusted intermediaries allows it to expand through node contributions from altruistic parties without fear of breaching user anonymity. At the application layer the Torbutton[55] Firefox plug-in

`http://www.i2p2.de/techintro.html`; George Danezis et al., "Mixminion: Design of a Type III Anonymous Remailer Protocol", in *Proceedings of the 2003 IEEE Symposium on Security and Privacy* (2003), 2–15.

54. Ian Goldberg, "On the Security of the Tor Authentication Protocol", in *Proceedings of the Sixth Workshop on Privacy Enhancing Technologies (PET 2006)* (2006).

55. "Torbutton FAQ", `https://www.torproject.org/torbutton/faq.html.en`.

solves the cookie problem by disabling cookies, caching, and plug-ins known to store tracking data (including Flash) with a single click.

Tor was first introduced in 2004, as of 2006 was estimated to have roughly 250,000 users,[56] and at present consists of roughly 1,250 nodes.[57] A variety of attacks have been levied against Tor, including externally identifying the nodes participating in a virtual circuit,[58] monitoring traffic pattern correlation at malicious entry and exit nodes to determine a client's IP address,[59] injecting data at the exit node to make the prior attack more feasible,[60] and applying the traffic pattern attack to virtual circuit construction.[61] While these attacks have been moderately effective in small test networks, they require an adversary to gain control of a significant proportion of nodes; given the size of the current Tor network, it seems unlikely this could be accomplished without arousing significant suspicion. Tor clients also limit the number of entry nodes they choose from, minimizing the likelihood of deanonymization even if such attacks are feasible.[62] Similarly, though an adversary could perform traffic analysis with logs of a user's sent traffic and a server's received traffic to deanonymize Tor (or any other proxy-based system) users, the magnitude of data collection and

56. Goldberg, "On the Security of the Tor Authentication Protocol".

57. "TorStatus - Tor Network Status", `http://torstatus.kgprog.com/`.

58. Steven J. Murdoch and George Danezis, "Low-Cost Traffic Analysis of Tor", in *SP '05: Proceedings of the 2005 IEEE Symposium on Security and Privacy* (2005), 183–195.

59. Lasse Øverlier and Paul Syverson, "Locating Hidden Servers", in *SP '06: Proceedings of the 2006 IEEE Symposium on Security and Privacy* (2006), 100–114.

60. Timothy G. Abbott et al., "Browser-Based Attacks on Tor", in *Privacy Enhancing Technologies* (2007).

61. Kevin Bauer et al., "Low-resource Routing Attacks Against Tor", in *WPES '07: Proceedings of the 2007 ACM Workshop on Privacy in Electronic Society* (2007), 11–20.

62. With high probability the client will select non-malicious entry nodes, mitigating nearly all threats of this sort , "TheOnionRouter/TorFAQ", `http://wiki.noreply.org/noreply/TheOnionRouter/TorFAQ`.

analysis required would likely prohibit any sort of large-scale implementation. Combined with Torbutton, then, Tor has consequently been assumed to provide strong anonymity to web clients. The novel deanonymizing techniques presented in the following sections threaten this premise, and suggests more drastic measures yet are required to provide strong anonymity.

## Deanonymizing Web Clients with Quirkiness

As end users, we love to customize our computers. We select operating systems, displays, web browsers, plug-ins, add-ons, media viewers, document editors, and a variety of other features to best meet our needs. The major web browsers have in turn developed simple interfaces for extending their built-in capabilities and enriching the user experience, forming a virtuous cycle of add-on demand and integration. The question naturally arises, then: just as no two users are alike, with all the customizations now available ("quirkiness") are any two web browsing environments identical? And if not, is there any way a web server or other user could exploit this fact to identify users?

A web client's quirkiness stems from its underlying operating system and hardware in addition to display settings, browser settings, plug-ins, and add-ons. Though a variety of values will be wholly unique to the system – MAC addresses, the processor serial number, and the operating system license key are just a few that come to mind – web servers are constrained to gathering only that data which web clients are willing to volunteer and are able to recover from their limited "sandbox" access to underlying hardware and operating system functionality.

Three popular browser-based code environments have potentially sufficient access to hardware and settings to gather quirkiness:[63] Java Applets, Flash ActionScript, and JavaScript. Java has the greatest access to underlying hardware, but the SecurityManager class restricts unsigned applets from accessing nearly all identifying data.[64] Flash faces a similar issue with its own security controls,[65] leaving JavaScript, a language both largely interpreted and purposefully implemented in a manner independent of hardware, giving it unpredictable performance characteristics[66] and nearly no knowledge of hardware quirkiness. Nonetheless, it is privy to a wide range of browser settings.

Early in JavaScript's development the web browser firm Netscape recognized websites would benefit from the ability to ask of a visitor's browser, "Do you accept cookies?", "What is your screen resolution?", and other questions informing content presentation. It consequently implemented a series of standard objects, `navigator`, `screen`, `Plugin`, and `MimeType` (Table 2.2), that provide programmatic access to

---

63. With increasing adoption Microsoft Silverlight will, in future, present a fourth viable option. Given its ability to run Common Language Runtime code, it could provide significant hardware access.

64. Sun Microsystems, "JDK 6 Security-related APIs & Developer Guides", `http://java.sun.com/javase/6/docs/technotes/guides/security/index.html`.

65. Adobe, Inc., "Adobe Flash Player 9 Security", `http://www.adobe.com/devnet/flashplayer/articles/flash\_player\_9\_security.pdf`.

66. Anecdotal experience suggests the same script could vary in execution time by several orders of magnitude.

| `navigator` | `screen` | `Plugin` | `MimeType` |
|---|---|---|---|
| appCodeName | availHeight | name | type |
| appMinorVersion | availWidth | filename | description |
| appVersion | colorDepth | description | suffixes |
| cookieEnabled | pixelDepth | length | enabledPlugin |
| language | height | | |
| mimeTypes (array of `MimeType`'s) | width | | |
| opsProfile | | | |
| platform | | | |
| plugins (array of `Plugin`'s) | | | |
| systemLanguage | | | |
| userAgent | | | |
| userLanguage | | | |
| userProfile | | | |
| javaEnabled() | | | |
| taintEnabled() | | | |

Table 2.2: Built-in JavaScript objects.

a browser's options, display settings, installed plug-ins, and supported file formats respectively. Modern Mozilla- and WebKit-based browsers including Mozilla Firefox, Apple Safari, and Google Chrome continue to fully support the Netscape legacy objects, while Microsoft's Internet Explorer (IE) provides access to identical information albeit through a query-based interface. Testing for the Adobe Acrobat Reader plug-in, for example, only requires:

```
try
{
   new ActiveXObject('AcroPDF.PDF');
}
catch(e)
{
   // Acrobat Reader is not installed
}
```

IE plug-in objects themselves offer largely the same information available from the

| Plug-in | Version Accessor |
|---|---|
| AcroPDF.PDF | GetVersions() |
| ShockwaveFlash.ShockwaveFlash | getVariable("$version") |
| Quicktime.Quicktime | QuickTimeVersion |
| RealPlayer | GetVersionInfo() |
| SWCtl.SWCtl | ShockwaveVersion("") |
| WMPlayer.OCX | versionInfo |

Table 2.3: Proprietary version accessors in Internet Explorer plug-in objects.

standard `Plugin` object, though authors specify the non-standard fields available,[67] as demonstrated in Table 2.3.

This seemingly trivial difference in interfaces has staggering repercussions: by providing plug-ins and file types in a list, Mozilla- and WebKit-based browsers inadvertently risk adding extra quirkiness through the ordering of the list! In practice it appears both code bases provide the list of plug-ins from a hash set that uses file modification timestamps as an element of the hash, as shown in the snippet from the WebKit source[68] below.

```
unsigned hashCodes[3] = {
        m_description.impl()->hash(),
        m_lastModified.dwLowDateTime,
        m_lastModified.dwHighDateTime
};
```

The subtle issue that arises is so long as two web clients possess alternate plug-in order of installation or file modification the ordering of their `navigator.plugins` and `navigator.mimeTypes` lists could differ. Thus, for two Mozilla- or WebKit-

67. Matthew Ratzloff, "Detecting plugins in Internet Explorer (and a few hints for all the others)", April 26, 2007, `http://www.builtfromsource.com/2007/06/26/detecting-plugins-in-internet-explorer-and-a-few-hints-for-all-the-others/`.
68. "WebKit", `http://trac.webkit.org/browser`.

based web clients to have guaranteed matching quirkiness they must not only share the same plug-ins and settings, but also have installed or last modified their plug-ins *at exactly the same time*. This property was easily demonstrated in Firefox 3.0.8 (the latest version at the time of writing) by uninstalling and reinstalling the Adobe Flash plug-in on Mac OS X 10.5 and Ubuntu Linux 8.04, and the Apple Quicktime plug-in on Windows XP, all resulting in an altered plug-in ordering.[69] The chance of two web clients with identical settings and plug-ins randomly sharing the same ordering of plug-ins is incredibly small: assuming 15 plug-ins, there are $15! \approx 10^{12} \approx 2^{40}$ possible combinations, several orders of magnitude more than there are web clients in existence. Moreover, unless computers are cloned from an identical image, the install time of the web browser itself should be sufficient to induce a unique ordering of plug-ins and file types on first use – setting aside the changes that arise from adding or updating plug-ins!

User behaviors also contribute to non-obvious quirkiness in plug-ins and file types. The tendency to not regularly update one's browser[70] or plug-ins[71] adds uniqueness, as does the selection of which plug-in to use for playing back a specific type of media.

---

69. Further experimentation is required to determine why uninstalling and reinstalling Adobe Flash on Windows XP did not induce an ordering change.

70. Stefan Frei, Thomas Duebendorfer, and Bernhard Plattner, "Firefox (In)Security Update Dynamics Exposed", *ACM SIGCOMM Computer Communications Review* 39, no. 1 (2009): 16–22.

71. For example, only roughly half of Adobe Flash users had updated to the latest version as of December 2008. Adobe, Inc., "Flash Player Version Penetration", 2008, `http://www.adobe.com/products/player_census/flashplayer/version_penetration.html`.

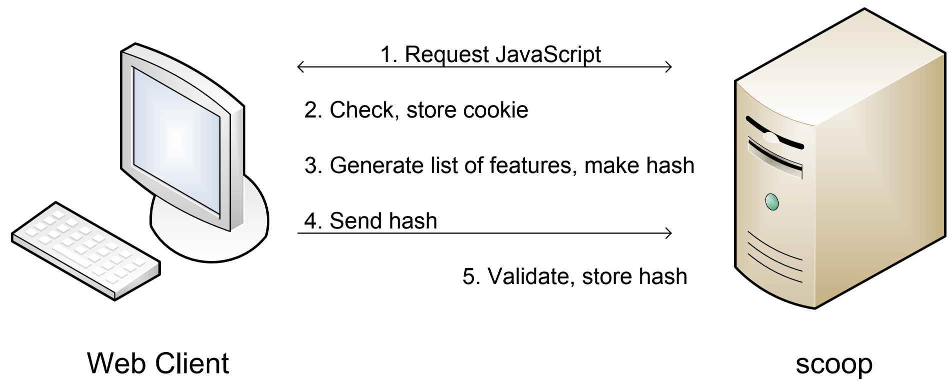



Figure 2.4: Experiment design.

## Experimentally Measuring Quirkiness

### Goals

Quirkiness-based identification rests upon hosts having enough semi-unique traits to combine into a unique identifier. While qualitative experience suggests users customize their browsing experience to the point of uniqueness, more quantitative proof is desirable before evaluating quirkiness' potential for upsetting web anonymity. The following experiment is designed to show that not only does sufficient quirkiness exist in web clients for individual identification, but also that it can be retrieved through JavaScript enabling a web server or user with the ability to insert JavaScript into a page to uniquely identify a client.

### Design

A simple experiment for achieving the above goals could directly compare web browsers' JavaScript-accessible features. Transmitting and storing complete

JavaScript objects could prove quite inefficient, however, and potentially compromise features participants would prefer remain undisclosed. This basic design was therefore slightly modified such that clients create and transmit a one-way cryptographic hash of their features[72] as opposed to a list of the traits themselves, providing the twofold benefits of eliminating transmission and storage overhead, as well as protecting participant privacy.

Participants were solicited by a posting at the popular Princeton Center for Information Technology blog `freedomtotinker.com` for a two week period (February 3, 2009 to February 17, 2009) to visit `scoop.princeton.edu` with their preferred web browser. Upon loading the site a visitor's browser would execute a small JavaScript snippet that first checks for a cookie from the site. If none is present, a cookie is set and the contents of the `navigator`, `screen`, `navigator.plugins`, and `navigator.mimeTypes` objects are efficiently concatenated[73] and hashed using the MD5 algorithm,[74] resulting in a unique 128-bit identifier. This value is then transmitted with an XMLHttpRequest to a separate PHP script on `scoop.princeton.edu`, which validates and stores it in a MySQL database of hashes and hash frequencies

72. A cryptographic hash is completely determined by an input value, but appears random without knowledge of that value. The experiment applies a hash to generate a short value that (nearly) uniquely represents each participant's quirkiness but cannot be meaningfully reversed into a list of a participant's settings and plug-ins.

73. Objects are added to a list, which is only concatenated once to reduce the computational burden on clients and thereby decrease user wait time; in practice the entire process was nearly instantaneous.

74. Though MD5 is now known to have certain cryptographic failings, it is still viable as a tool for comparing uniqueness. Xiaoyun Wang and Hongbo Yu, "How to break MD5 and other hash functions", in *EuroCrypt* (2005); R. Rivest, "The MD5 Message-Digest Algorithm", *RFC*, no. 1321 (1992).

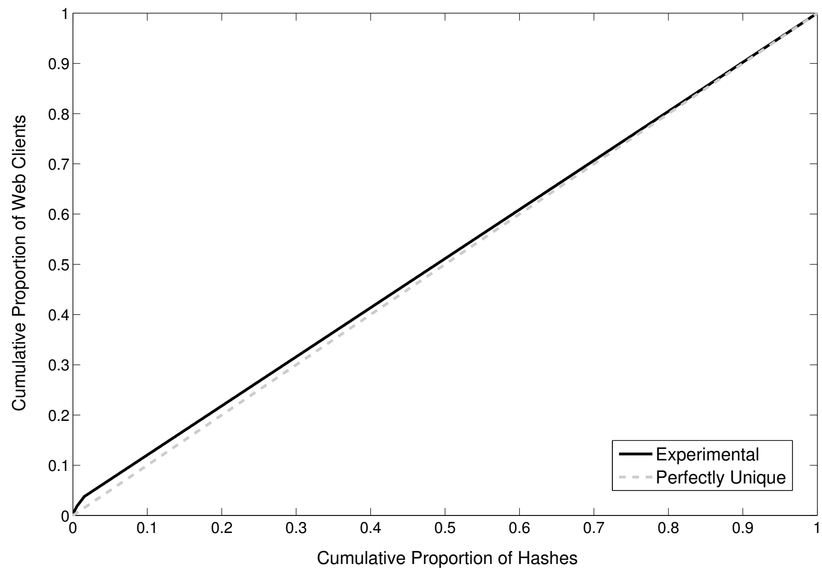



Figure 2.5: Experimental results plotted against perfect uniqueness.

| **Web Clients Per Hash** | **Hash Count** |
|---|---|
| 1 (unique) | 1278 |
| 2 | 13 |
| 3 | 5 |
| 4 | 1 |
| 5 | 1 |
| **Total Web Clients** | 1328 |
| **Total Hashes** | 1298 |

Table 2.4: Experimental results.

for later analysis.[75] The entire process is depicted in Figure 2.4, and the JavaScript and PHP used in the experiment can be located in Appendix C.

## Results and Statistical Inferences

Over the two week period $N = 1328$ web clients participated in the experiment. As shown in Table 2.4 1278 of the visitors (96.23%) could be uniquely identified

75. The astute reader will note this design is subject to cross-site request forgery attacks and data poisoning, but given the absence of harm to the client the former was discounted and the open nature of the experiment unfortunately implies the latter will always be a possibility.

even with the limited set of traits explored. Guesswork led to discovery of one of the hash collisions: Apple's iPhone, which offers few browser customization options. Presumably this property generalizes to all unmodifiable or identical browsers as might be found on other mobile devices or imaged computers. That said, the restrictions inherent to web clients of this sort often extend beyond browser settings; many have a fixed IP address (or range of IP addresses) and not easily cleared cookies and caches, like the public library and cybercafe scenarios discussed earlier. More traditional identification techniques should therefore be largely sufficient in these cases where deanonymization through quirkiness fails.

While it would be desirable to extrapolate from the experimental results that a significant proportion of web clients could be uniquely identified by their quirkiness, or at minimum reduced into a small "anonymity set" of web clients with identical quirkiness, statistical testing shows that a dataset several orders of magnitude larger is required to approach any reasonable degree of confidence in drawing such conclusions.

An initial attempt at statistical inference could model the proportion of web clients that can be uniquely identified, $p$. Through the simplified, approximate proportion test,[76] a confidence interval for $p$ would be $\left[\hat{p} - z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$ where $n$ is the experimental sample size, $\hat{p}$ is the experimental proportion of uniquely identifiable clients, and $z_{1-\alpha}$ is the normal cumulative distribution function z-value for a given level of statistical significance $1-\alpha$.[77] The experimental data would therefore

76. Applicable here because of large $n$.

77. Jay Devore, *Probability and Statistics for Engineering and the Sciences*, 7th ed. (Pacific Grove:

show a 95% ($\alpha = .05$) confidence interval $p \in [.952, .973]$.

This analysis is flawed, however, in that anonymity sets, represented by hashes, do not devolve into either a unique or non-unique population; sets that appear unique could well be non-unique, but a second instance simply never appeared in the sample. Consider the extreme case where there are precisely two users on the Internet in each anonymity set. The likelihood a sample of size $n$ would appear entirely unique is $\frac{(M)(M-2)...(M-2n+2)}{(M)(M-1)...(M-n+1)}$ where $M$ is the size of the user population, roughly 1.4 billion according to the latest International Telecommunication Union statistics.[78] For $n =$ 1000, then, there is a 99.96% chance the entire sample would appear unique... but in reality the population is 0% unique!

A more accurate model of the data recognizes that the hash values do not fall into identifiable and unidentifiable populations, but rather a group of anonymity sets, at maximum $M$,[79] each with a positive number of members. A claim of uniqueness is equivalent to a claim that a given set has a sole member, and the proportion of the overall population of web clients in a given set $i$, $p_i$, is equal to $1/M$. Adopting this view allows development of a maximum likelihood model for the values of $p_i$ given the experimental result; Appendix B gives a rigorous proof that this model holds, intuitively, the global population proportion in each anonymity set equal to the sample population proportion in each set.

Duxbury Press, 2007), 265-266.

78. International Telecommunication Union, "Free statistics".

79. In the case where each web client is unique.

For the purposes of the experiment this result is quite unfortunate. What would make the sample most likely is if those anonymity sets that appeared unique actually contained $1/1328 = .075\%$ of the Internet population, or roughly $M/1328 = 1,054,217$ web clients! Thus, while on an intuitive level the pressures for unique quirkiness discussed earlier should exist globally, statistical results from the experiment unfortunately provide no firm support. In fact, to approach any statistical confidence in uniqueness, the sample size would have to near the global population! The variance analysis in Appendix B shows that, to not have evidence at 95% confidence that a given anonymity set has more than one member, $n \approx 312.3$ million participants are required!

## Experimental Flaws

Setting aside the extent of evidence offered by the experiment in favor of unique identification through quirkiness, it suffers from several endemic flaws:

1. Sample Bias: The population that visits `freedomtotinker.com` is technologically savvy and likely employs more customized web browsers than the average individual, skewing results towards uniqueness. Furthering this effect, few mobile web clients are likely to visit the site.

2. Internet Explorer Plug-in Support: The experiment's JavaScript code did not employ Internet Explorer's plug-in querying architecture, understating the quirkiness available in Internet Explorer and skewing the results away from uniqueness. That said, the audience of `freedomtotinker.com` likely uses alternative browsers more popular with advanced users, minimizing the degree of this effect.

3. Scale: As shown above, a much larger experiment is necessary to provide statistical confidence in the uniqueness of quirkiness.

4. Contributions to Quirkiness: The hash-based experiment design masks which properties contribute most to a web client's quirkiness.

5. Change Over Time: Users change their browser settings over time, but the experiment only takes a single snapshot of quirkiness.

A more robust followup experiment could be run on a large Internet ad network to resolve the sample bias and scale issues. The contribution and change problems could be easily addressed through hashes of specific properties and hash storage in cookies.

## Deanonymization with Web-based Applications

The plethora of new functionality being integrated into modern web browsers to support web-based applications, most notably in the draft HTML 5 standard,[80] presents a new vector for deanonymizing attacks.[81] Two features appear readily exploitable: the ability to register handlers and local application storage.

In the traditional usage model users install plug-ins capable of handling certain types of media and select which plug-in to associate with each type. HTML 5 attempts to allow web services to fulfill the content handling role previously the domain of plug-ins: sites are able to register, through JavaScript, specific protocol (i.e. FTP) or content (i.e. MP3 audio) handlers that activate when the browser encounters a reference to the specified protocol or content on the site. Instead of handing the reference to a plug-in, the browser instead forwards it to the page specified by

80. Web Hypertext Application Technology Working Group, "HTML 5 Draft Recommendation", April 4, 2009, `http://www.whatwg.org/specs/web-apps/current-work/`.

81. Google Gears and Native Client could offer similar vulnerabilities, but given their lack of universal acceptance and unclear futures are excluded from this analysis. Google Inc., "Gears API", `http://code.google.com/intl/en/apis/gears/design.html`; Bennet Yee et al., "Native Client: A Sandbox for Portable, Untrusted x86 Native Code", in *2009 IEEE Symposium on Security and Privacy* (2009).

the handler. By manipulating the registered page to be unique for each user, a malicious site could easily employ this mechanism for tracking: upon receiving instruction to load any resource from the malicious site that matches the specified protocol or content – which does not even require JavaScript! – the browser will automatically issue a request to the handler page and thereby uniquely identify itself. In a brief test this attack appeared feasible on Firefox 3.0.8 with the latest Tor and Torbutton (1.2.0); the user need only be convinced to accept an innocuous-looking prompt to register the handler. Worse yet, revoking a registered handler appears to be impossible with Firefox's user-facing preferences – a user must navigate Firefox's internal configuration to remove it.

Local application storage, which has yet to be fully implemented, presents the same risk as the cache "cookies" discussed earlier. A malicious site could store a unique application, and every subsequent activation would uniquely identify the browser.

## Consequences for Anonymity Online

Identification through quirkiness poses a twofold risk to all web anonymizing technologies, including Tor. First, by providing what is, in effect, an indelible cookie, users attain weak anonymity at best. More dangerously, though, browsing at any time without anonymization risks associating quirkiness with an individual's identity and eradicating anonymity for all past and future interactions. Tor appears the only anonymizing tool to have recognized the risk of quirkiness and attempts through Torbutton to replace the objects discussed earlier with a generic set.[82] At present the original `screen` object is recoverable,[83] however, and a similar attack could possibly reveal portions of the `navigator` object.

The web-based application deanonymizing techniques pose the same risks and are suggestive of a broader problem: the interests of browser vendors and web services – developing and making available Web 2.0 content and application platforms – are often at odds with those of anonymity-seeking individuals. The latter parties are not represented in the development of web browsers and standards, however, while the former are the major sponsors.[84] The browsers and standards that have resulted consequently offer features like the JavaScript objects, handler registration capabilities, and local application storage discussed above, which pose grave risks to anonymity. Adopting the verbiage coined by Internet pioneer David Clark, web clients in the age of Web 2.0 are not designed to support an anonymity "tussle";[85] those desiring anonymity have no means of easily modifying browsers and standards to meet their needs, and must instead play a perpetual game of cat and mouse with new browser features and standards.

82. "TheOnionRouter/TorFAQ".
83. Krishna E. Bera, "Torbutton Bug Report: Unmask Screen", March 24, 2009; Gregory Fleischer, "Unmask Screen - Iframe JavaScript", `http://pseudo-flaw.net/tor/torbutton/unmask-screen-iframe-javascript.html`.
84. Including Google and Apple in the HTML 5 effort.
85. David D. Clark et al., "Tussle in Cyberspace: Defining Tomorrow's Internet", *SIGCOMM '02* (2002).

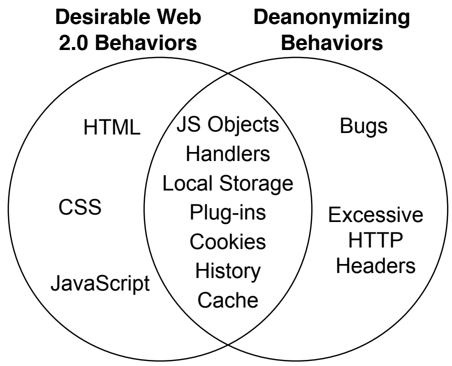



Figure 2.6: Desired Web 2.0 behaviors and deanonymizing behaviors; there is significant overlap, consistently arbitrated in favor of Web 2.0.

Recognizing there is an intersection of behaviors that are desirable for Web 2.0 functionality but not from an anonymizing standpoint (Figure 2.6), and that anonymity will not have a major voice in setting such behavior, anonymizing tools should reconsider the approach of relying on major web browsers. They would be well served to, instead, develop a new browser designed strictly to ensure anonymity. One promising option is a browser contained in a Java Applet: using Java's built-in signing mechanism it could be securely delivered from even the most untrustworthy source, and the host employing it need only have a Java-capable browser installed – not even the privileges to install a program. Compared with the "quite complex"[86] installation of existing tools like Tor, a web-based solution could be far more usable.

In the meantime, Tor and similar technologies do provide a high degree of anonymity, and are indicative that intellectual capital is motivated to pursue anonymizing technologies. That said, if users are not knowledgeable about the challenges facing

86. Hal Roberts, Ethan Zuckerman, and John Palfrey, "2007 Circumvention Landscape Report: Methods, Uses, and Tools", March 2009, `http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/2007_Circumvention_Landscape.pdf`.

Internet anonymity and tools available, such projects will be for naught. The following chapter examines this proposition by exploring the level of knowledge possessed and resources for learning about anonymity online.

# 3 Individual Perceptions of Internet Anonymity

Even the most technically rigorous of the anonymizing tools discussed in Chapter 2 will have minimal impact if users remain unaware of its existence. In this chapter qualitative and quantitative evidence provides support for the view that not only are users woefully ignorant of the tools at their disposal, but few impartial, accurate, easily discovered resources exist to inform users of both threats to anonymity and available anonymizing tools. Justification of the former claim stems from a survey of Princeton undergraduates which suggests even the well-educated possess limited knowledge of Internet anonymity. Subjective experience and automated analysis of web search results further show that, though several comprehensive resources on Internet anonymity exist, an individual's independent discovery of them appears unlikely. A final section examines this dilemma in the context of recent research on the psychology of web search, and advances experimental findings as evidence an inquiring user would not settle on a trustworthy source about Internet anonymity.

## Survey Methodology

In anecdotal experience detailed knowledge of Internet anonymity appears limited to those technologically inclined – and even then, significant confusion exists over the import of IP addresses, MAC addresses, cookies, and other risks to anonymity. While suggestive of a trend, occasional personal conversations are no firm basis for policy analysis; in association with this work a short online survey of Princeton undergraduates was consequently conducted to quantitatively evaluate these anecdotal conclusions against a broader population. This survey suffers from far too many biases to accurately model the Princeton undergraduate or global population, including:

1. Princeton undergraduates are far more likely to study a technical field than the average individual.
2. Even those students not directly involved in technical studies are required to employ computers and the Internet on a daily basis.
3. Being of a younger generation, undergraduates have grown up accustomed to using the Internet.
4. Respondents by necessity were sufficiently knowledgeable to open an email and navigate a web page to complete the survey.
5. Many students who responded were likely interested in the survey's topic, implying a degree of technical curiosity.

Noting that all these biases would favor increased awareness of Internet anonymity, the survey should instead be construed as a loose upper bound on the average individual's knowledge.[1] Full text of the survey can be found in Appendix A; questions probed students' perceived Internet competency and degree of anonymity to fellow users

1. Given these issues no variance-based analysis is presented in this chapter.

and websites, knowledge of anonymizing tools, and research methods to learn about anonymous access. An initial email solicitation was dispatched to 800 randomly selected undergraduates on February 27, 2009, a reminder email was sent on March 19, and the survey was closed shortly thereafter. To encourage participation students were enticed with a drawing for a 2GB iPod shuffle, valued at \$50; $N = 190$ recipients had completed the survey at close. The following analysis treats qualitative ordinal responses as ranging from 1-4 (1 lowest, 4 highest) and employs a category coding of free responses distilled from the responses themselves.

## Survey Results

The survey's findings are largely congruent with experience. Respondents correctly recognized they are far less anonymous to the sites they visit, which have limitless control over the content presented, than to other users who must turn to less direct channels[2] (mean of 1.76 vs. 2.32). Gaps in knowledge quickly grew apparent when pressed on achieving anonymity, however; 50.8% of respondents indicated with average confidence (mean of 2.59) they did not believe anonymity is attainable! Those respondents who indicated anonymity is within reach were further ill-informed of techniques for achieving it. Only 23.9% of this population explicated the importance of not sharing personally identifiable information, 19.6% recommended the use of a proxy, and 14% stated they would either turn off or clear cookies. The most popular

2. For example, submitting a comment that includes JavaScript.

response, at 28.3%, recommended the use of another computer – but as discussed in Chapter 2, this approach is highly problematic. The more nuanced threats to anonymity received even less attention from this population; only 2.2% recommended disabling plug-ins and another 2.2% pointed out the importance of clearing cache, both essential steps for negating cookie-like deanonymizing mechanisms. As for Tor, a scant 6.5% of the population explicitly referenced the system. Nearly as many, 5.4%, suggested "private browsing," with most pointing to Google Chrome's Incognito mode. Even more worryingly, 13.0% of the population indicated firewalls and 4.4% antivirus as essential for anonymity; though these security products are beneficial in preventing malicious software from executing, that they contribute to concealing identifying information is a significant misperception. Many free responses expressed a high level of technical knowledge but absence of understanding. One participant recommended using "one of those fake IP addresses you can get online," while another advocated the unnecessary step of "a program that can change and randomize your MAC address." Less informed responses included "turn off IP address," "get spyware on my computer," and "unregister your computer."

A followup question probed which resources respondents would turn to for information on Internet anonymity. An online search was the unambiguous favorite, advocated by 44.3% of respondents; 25.7% explicitly referenced Google. The next most frequent response, at 30.4%, was approaching a trusted individual for advice. This merely pushes the burden of knowledge onto the second party, however, and

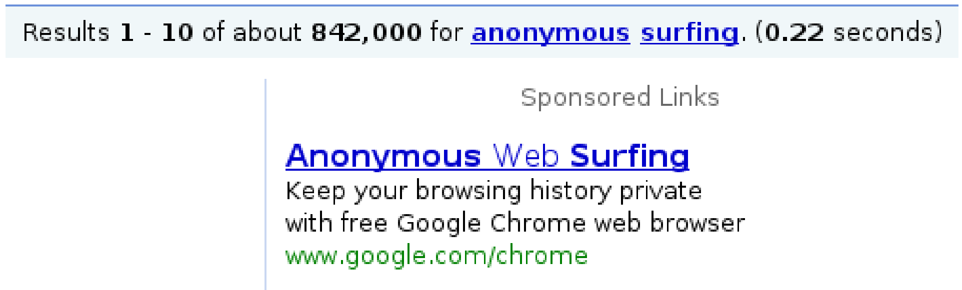



Figure 3.1: Google paid results for "anonymous surfing" showing a top listing for the Chrome browser.

given the minimal level of knowledge possessed among even the tech-savvy – only one of seven computer science concentrators was aware of Tor! – they in turn would likely rely on a search themselves. A final popular recommendation, at 9% incidence, was to look for a book on anonymity. This approach also seems likely to result in a search, as even the most authoritative texts in related fields[3] contain little on how to achieve anonymity.

## Resources on Internet Anonymity

Given the dearth of knowledge on anonymity and likelihood individuals would either directly or indirectly derive information on the topic from a web search, two questions naturally arise: What resources are available online? And how accessible are they? This section addresses each question in turn.

Experience suggests the vast majority of online resources on anonymity are

3. For example, Deibert et al., *Access Denied: The Practice and Policy of Global Internet Filtering*; Michael Y. Dartnell, *Insurgency Online: Web Activism and Global Conflict* (Toronto: University of Toronto Press, 2006); Shanthi Kalathil and Taylor C. Boas, *Open Networks, Closed Regimes: The Impact of the Internet on Authoritarian Rule* (Washington, D.C.: Carnegie Endowment for International Peace, 2003).

of a commercial nature, and offer scant technical details or links to alternatives. Anonymizer, Inc.,[4] for example, appears in the first page of many Internet anonymity-related queries. The firm's site champions the benefits of its "Anonymous Surfing" software product, available for an annual subscription of $30, and claims it offers "one-click privacy" through "IP hiding" (a proxy)... but there is no discussion of cookies, caching, plug-ins, or any of the myriad other threats to anonymity! A competitor, Tenebril GhostSurf,[5] provides a short analysis of the impact of IP addresses and cookies on anonymity, but only deletes cookies and clears cache at a user's request and takes no action against plug-ins. Web proxies are largely similar; the-cloak, for example, makes the same omissions as Tenebril in its documentation, and provides uncertain protection against Flash and other plug-ins.[6] Even well-known, trusted companies provide poor documentation of the limitations of their anonymizing products. Google, for example, advertises its Chrome browser as providing "Anonymous Web Surfing" (Figure 3.1) when it only offers the limited benefits of "private browsing" as discussed in the prior chapter.[7] The notices provided by the major web browsers themselves, meanwhile, are technically oriented and of little use to the lay user (Figure 3.2). Only Google Chrome provides an indication that a user remains

4. Anonymizer, Inc., "Anonymous Surfing".

5. Tenebril, Inc., "GhostSurf", `http://www.tenebril.com/consumer/ghostsurf/ghostsurf_standard.php`.

6. "The Cloak".

7. Microsoft Inc., "Internet Explorer 8: More secure, private, and reliable", `http://www.microsoft.com/windows/internet-explorer/beta/features/browse-privately.aspx`; Apple Inc., "Apple - Safari", `http://www.apple.com/safari/features.html`; Google Inc., "Explore Google Chrome Features: Incognito Mode", `http://www.google.com/support/chrome/bin/answer.py?answer=95464&hl=en`.

trackable and a link for further information, but the linked page gives few additional details and no directions on how to attain anonymity.

Another broad class of site that appears with frequency is outdated discussion of technical means for attaining anonymity. Sources range from defunct projects[8] to old articles,[9] but all provide no indication of where to properly turn at present.

So much for web search results. Another venue individuals might explore is Wikipedia, a popular wiki-based encyclopedia with nearly three million articles by latest count.[10] Here, too, actionable information is hard to come by: the articles on "Anonymity,"[11] "Anonymous web browsing,"[12] and "Anonymous web proxy"[13] all offer no practical advice on how to use the Internet anonymously. The "Tor" article[14] is quite accurate about the project's history and technical properties, but provides little information on how to use the tool.

Several outstanding resources do exist, however. The Tor Project maintains immaculate documentation of Tor's design and use, at the level of both simple instructions and academic papers.[15] Blogging activism supporter Global Voices goes a step further and offers details on both installing Tor and setting up an anonymous blog

8. e.g. Jacob Palme and Mikael Berglund, "Anonymity on the Internet", http://people.dsv.su.se/~jpalme/society/anonymity.html.

9. e.g. Thomas C Greene, "Do-it-yourself Internet anonymity", *The Register* (November 14, 2001), http://www.theregister.co.uk/2001/11/14/doityourself_internet_anonymity/.

10. "Statistics", http://en.wikipedia.org/wiki/Special:Statistics.

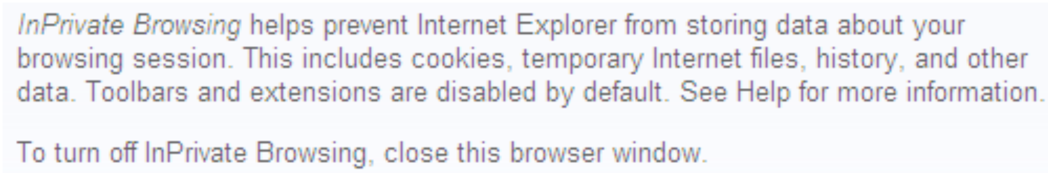
11. "Anonymity", http://en.wikipedia.org/wiki/Anonymity.

12. "Anonymous Web Browsing", http://en.wikipedia.org/wiki/Anonymous_web_browsing.

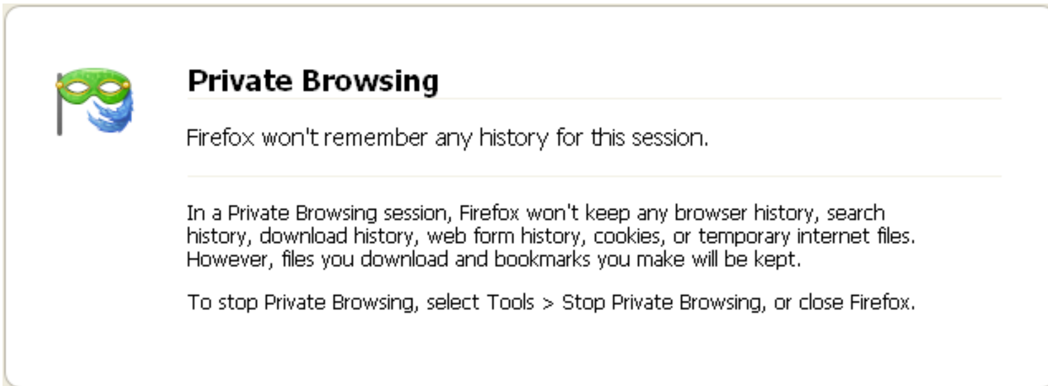
13. "Anonymous Web Proxy", http://en.wikipedia.org/wiki/Anonymous_web_proxy.

14. "Tor (anonymity network)", http://en.wikipedia.org/wiki/Tor_(anonymity_network).

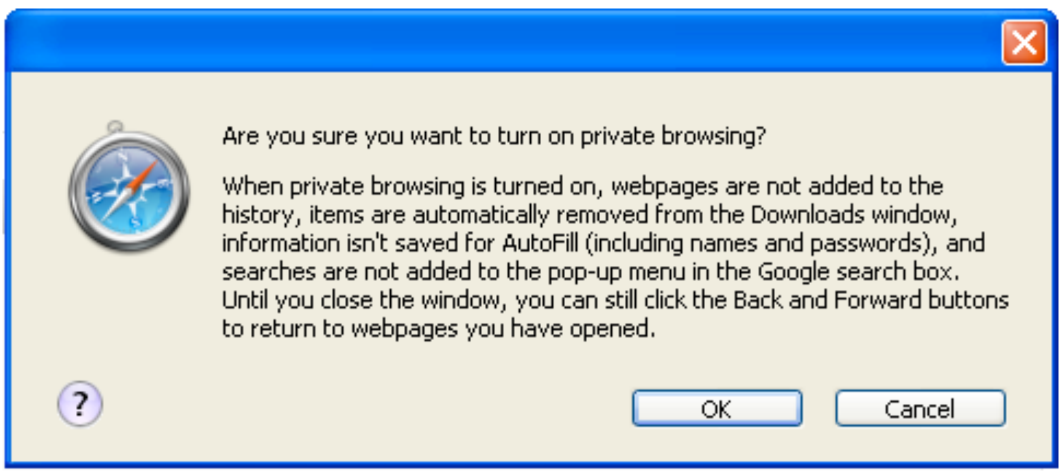
15. "Tor: Overview".



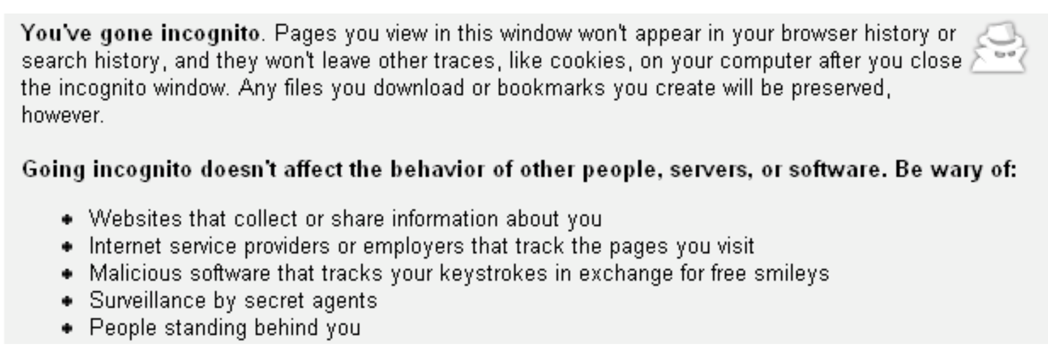

(a) Microsoft Internet Explorer 8




(b) Mozilla Firefox 3.1 Beta




(c) Apple Safari 4 Beta




(d) Google Chrome Beta

Figure 3.2: Private browsing notices provided by the major browsers in Windows XP.

| Anonymity words | Action words | Object words |
|---|---|---|
| anonymous | browse | internet |
| anonymously | browsing | online |
| anonymity | surf | web |
| | surfing | |

Table 3.1: Query terms expected to be employed by a web searcher attempting to find information on Internet anonymity.

| Result | Number of Queries | Percentage of Queries |
|---|---|---|
| Did not appear | 147 | 56.3% |
| First page | 43 | 16.5% |
| Second page | 30 | 11.5% |
| Third page | 25 | 9.6% |
| Fourth page | 8 | 3.1% |
| Fifth page | 8 | 3.1% |
| **Total queries** | 261 | 100% |

Table 3.2: The appearance of the Tor Project's site in response to generated Google search queries.

with WordPress.[16] The Electronic Frontier Foundation, meanwhile, offers a high-level whitepaper that flags Tor and several other tools for exploration.[17]

Unfortunately, these sites routinely fail to appear in Google search results, and the Wikipedia article "Tor" is only discovered through a "Related Pages" link at the bottom of the "Anonymous web browsing" article. To empirically demonstrate this subjective analysis a script was developed to perform automated evaluation of Google search results. Queries followed a simple three-word template intended to roughly model what an individual might construct:[18] some permutation of an anonymity-

16. Global Voices Advocacy, "Anonymous Blogging with Wordpress and Tor", `http://advocacy.globalvoicesonline.org/projects/guide/`.

17. Electronic Frontier Foundation, "How to Blog Safely (About Work or Anything Else)", `http://www.eff.org/wp/blog-safely`.

18. Historical search patterns show users are most likely to submit a three word or less query. Nadine Hochstotter and Martina Koch, "Standard parameters for searching behaviour in search engines and their empirical evaluation", *Journal of Information Science* 35, no. 1 (2009): 45–65.

related term and neither, either, or both of an action related to using the Internet and an object akin to the Internet (Table 3.1). The first five pages of results for each query were tested for the Tor Project site as, in experience, it seemed the only of the reliable resources discussed likely to appear in search results. The outcome of this experiment is shown in Table 3.2; the Tor Project appeared on the first page of results in a scant 16.5% of queries!

## The Psychology of Web Search: A Bleak Picture

Recent research results suggest this paucity of relevant search results is particularly pernicious: in conducting an "informational"[19] search with limited a priori knowledge, individuals have no means of judging a source's quality. They will, consequently, tend to "satisfice" their quest for knowledge with sub-par resources. A 2005 eyetracking study confirmed this phenomenon by re-ordering Google search results and tasking subjects with quantifying each site's quality; users exhibited a "trust bias" towards higher ranked sites (from the search engine's historical tendency to provide high quality results) as well as a "quality bias" of judging sites relative to neighboring results.[20] A 2008 experiment with similar methodology reached the same conclusions, and further found that, in general, users will favor the first two to three

19. Andrei Broder, "A Taxonomy of Web Search", *SIGIR Forum* 36, no. 2 (2002): 3–10.

20. Thorsten Joachims et al., "Accurately Interpreting Clickthrough Data as Implicit Feedback", in *ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)* (2005), 154–161.

results far disproportionately to the rest – independent of quality.[21]

Combining these observations with the prior analysis of online resources on anonymity, a bleak picture forms. Though as discussed in Chapter 2 there are a variety of tools available for attaining a high degree of anonymity, few reliable resources exist to either advocate or provide information about their use, individuals have little means of discovering them, and will likely adopt the advice of the poor resources they do find. Given this state of affairs a policy intervention is essential to both rectify the absence of consumer awareness and determine if and how Internet anonymity should be supported by the government. The subsequent section addresses this topic and provides a framework for government response.

21. Mark T. Keane, Maeva O'Brien, and Barry Smyth, "Are People Biased in Their Use of Search Engines?", *Communications of the ACM* 51, no. 2 (February 2008): 49–52.

# 4 Internet Anonymity Policy

The foregoing chapters were deliberately apolitical in the interest of establishing a factual basis for policy analysis. In Chapter 2 this work examined purely technical aspects of Internet anonymity, concluding that it is feasible though current schemes have weaknesses and future Web 2.0 developments threaten to add more. Chapter 3 quantified knowledge about and resources on Internet anonymity, finding both that individuals are largely unaware of how to browse anonymously and would face significant difficulty in learning how to do so. Having established a firm factual basis, this final chapter shines Internet anonymity through a policy lens. After reviewing the benefits and legal status in the United States of anonymity online, it acknowledges the real harms associated and posits policy recommendations that maximizes benefits and minimizes the associated downsides.

The past decade has seen a dearth of scholarship on Internet anonymity policy. A conference convened by the American Association for the Advancement of Science[1] and a variety of law journals considered the issues raised by anonymity online until

1. Al Teich et al., "Anonymous Communication Policies for the Internet: Results and Recommendations of the AAAS Conference", *The Information Society* 15, no. 2 (1999): 71–77.

roughly the turn of the millennium; work since has primarily been either technical or conflated with the novel, growing problem of Internet censorship. In the intervening years, however, the technological and legal landscapes of the Internet have been vastly transformed. This chapter consequently provides an updated – and actionable – policy analysis of Internet anonymity in the age of Web 2.0.

## The Case for Anonymity Online

Proponents of anonymity often turn to lofty arguments about individual liberty and self-efficacy, adopting the language of human rights.[2] Others enumerate potential benefits of anonymity writ large without providing analytical or anecdotal depth as to how they would accrue from anonymity online.[3] Finding such techniques largely unpersuasive in a policy context, the case for anonymity presented in this section instead argues advances in the more familiar quantities public discourse, national security, and privacy, with evidence drawn from historical examples. Readers should note this section is intended to be a sufficiently persuasive, but by no means exhaustive, account of anonymity's upsides.

The most immediate benefit to Internet anonymity lies in enhancing the public discourse through encouraging free speech. The first significant means by which it accomplishes this is the elimination of repercussions; in traditional discourse those

2. Teich et al., "Anonymous Communication Policies for the Internet: Results and Recommendations of the AAAS Conference", 73.
3. Marx, "What's in a Name? Some Reflections on the Sociology of Anonymity", 102-104.

critical of empowered individuals, organizations, or government entities often face real penalties, whether in the form of social stigma, economic cost, or physical harm threatened or carried out.[4] Anonymity is a shield against all such threats; it "levels the playing field"[5] on which lone individuals and colossal powers interact.

Numerous political dissidents and whistleblowers throughout history have provided scathing opinion and incriminating information anonymously for just this reason. Thomas Paine, for example, published his 1776 inflammatory, pro-independence pamphlet *Common Sense* anonymously for fear of treason prosecution.[6] In more recent memory, ex-RAND employee Daniel Ellsberg requested *The New York Times* reporter Neil Sheehan conceal his role in leaking the series of classified Department of Defense documents in 1971 that would come to be known, infamously, as "The Pentagon Papers."[7] Their contents, and the federal government's attempts to prevent publication – culminating in the *New York Times Co. v. United States* Supreme Court decision in the press' favor[8] – are credited for bridging disenchantment with the conduct of the Vietnam War into the domestic mainstream.[9] Similarly, reporter Bob Woodward of *The Washington Post* closely guarded the identity of his source on the Watergate scandal;[10] only three decades after President Richard Nixon's re-

4. Marx, "What's in a Name? Some Reflections on the Sociology of Anonymity", 102; Peter Wayner, "Technology for Anonymity: Names by Other Nyms", *The Information Society* 15, no. 2 (1999): 92; Froomkin, "Legal Issues in Anonymity and Pseudonymity", 115.

5. Ibid., 115.

6. Wallace, "Nameless in Cyberspace: Anonymity on the Internet", 2.

7. Nicholas Lemann, "Paper Tiger", *The New Yorker*, November 4, 2002.

8. "New York Times Co. v. United States", `http://www.law.cornell.edu/supct/html/historics/USSC_CR_0403_0713_ZO.html`.

9. Lemann, "Paper Tiger".

10. Bob Woodward, *The Secret Man: The Story of Watergate's Deep Throat* (New York, NY:

sulting resignation, and nearing his death from congestive heart failure, did former FBI Deputy Director W. Mark Felt identify himself as "Deep Throat."[11] A scant five years ago Department of Justice attorney Thomas Tamm anonymously tipped off journalists at *The New York Times* about the National Security Agency's Terrorist Surveillance Program involving warrantless wiretaps, resulting in a Pulitzer Prize-winning exposé.[12] The subsequent leak hunt, late 2007 ransacking of his home, and threats of federal prosecution confirm that Tamm had much to fear from revealing his identity alongside the government's illegal acts.[13]

The second mechanism by which anonymity encourages free speech is through minimizing the effects of the author's identity on perception of the speech and vice versa. An author's identity may connote specific societal, ethical, or political viewpoints, and cause recipients to discount the speech or perceive it in a way that diminishes its value. Alternatively, the inclusion of identity could lead recipients to focus on identity itself as the salient feature of the speech and ignore its contents. Anonymity simultaneously forces listeners to focus solely on the content of speech and judge its merits on that basis alone.

Much of the discussion surrounding the ratification of the United States Constitution was conducted by pseudonym for these very reasons. Prominent opponents included "Cato" and "Brutus,"[14] while Alexander Hamilton, John Jay, and James

---

Simon & Schuster, 2005), 4.

11. John D. O'Connor, "I'm the Guy They Called Deep Throat", *Vanity Fair* (July 2005).
12. Michael Isikoff, "The Fed Who Blew the Whistle", *Newsweek* (December 13, 2008).
13. Ibid.
14. Richard C. Box, *Public Administration and Society* (Armonk: M.E. Sharpe, 2003), 70.

Madison famously responded with a series of essays, the *Federalist Papers*, penned by "Publius" between 1787 and 1788 in support of ratification. Such tactics "forced readers to focus on arguments rather than authors;" otherwise, "they, rather than their arguments, would have become part of the debate over the Constitution."[15] Withholding identity also prevented altering perceptions of the authors, allowing "politicians to develop ideas free from public pressures, change their minds during deliberations, and explore differences until conclusions were reached."[16] When State Department Director of Policy Planning George F. Kennan published the article "The Sources of Soviet Conduct" in the July 1947 issue of *Foreign Affairs*, he adopted the pseudonym "X," almost assuredly to both guarantee the article's reception would not be affected by his position and prevent perception of the article as official U.S. foreign policy.[17]

A third, closely related means by which anonymity encourages free speech is through reducing the need for follow-up on the author's part. Proffering examples is more difficult here as authors would be unlikely to acknowledge such a self-centered motivation, but two candidates come to mind. Famed astronomer Carl Sagan's decision to publish a 1971 essay in favor of marijuana use under the pseudonym "Mr. X" could in part be construed as a desire to focus on his scientific and educational work instead of social activism.[18] In a more clear-cut case, the 1996 novel *Primary Colors*, a roman à clef of President Bill Clinton's first term in office penned by journalist

15. Box, *Public Administration and Society*, 70.
16. Ibid., 70.
17. Wallace, "Nameless in Cyberspace: Anonymity on the Internet", 2.
18. David A. Hollinger, "Star Power", *The New York Times* (November 28, 1999).

Joe Klein, was (and still is) published under the pseudonym "Anonymous;" among other reasons Klein hoped to cover the upcoming presidential election without the distraction of fielding questions about his literary work.[19]

As a final mechanism for enhancing the public discourse, anonymity reduces the psychological burden of sharing compromising or embarrassing information. Support groups frequently premise their programming on anonymity to encourage open conversation about usually private issues; the well known Alcoholics Anonymous, for example, cites anonymity as "the spiritual foundation of all our traditions, ever reminding us to place principles before personalities."[20] This effect is particularly pronounced online: researchers at the University of Toronto found interest in anonymous online counseling quickly surged far beyond traditional in-person and phone venues.[21]

Anonymous Internet access not only promotes free speech in the above ways with unparalleled effectiveness, whether in the form of text, audio, video, or even interactive content, but also affords the unprecedented ability to instantly and costlessly broadcast that speech to a worldwide audience. Moreover, as traditional print media,[22] mail,[23] and the landline telephone[24] continue their slow decline and replacement by blogs, e-mail, and VoIP respectively, the Internet will increasingly become

19. Doreen Carvajal, "Columnist's Mea Culpa: I'm Anonymous", *The New York Times* (July 18, 1996).
20. *A Brief Guide to Alcoholics Anonymous* (New York: Alcoholics Anonymous World Services, Inc., 1972), 14.
21. Jill Mahoney, "Troubled youth find an open ear on-line", *The Globe and Mail* (August 1, 2005).
22. David Carr, "Mourning Old Media's Decline", *The New York Times*, October 28, 2008.
23. Anick Jesdanun, "Postal agencies respond to mail decline", *Associated Press*, February 4, 2008.
24. Daniel Gross, "Phones Without Homes", *Newsweek*, July 28, 2008.

the only viable platform for sharing and publishing speech anonymously.[25]

A second, independent benefit to Internet anonymity is its utility for national security purposes. Anonymity allows domestic intelligence agencies to scour the web for "open source" intelligence on foreign powers and non-state actors without revealing which resources have been tapped. As one Central Intelligence Agency (CIA) official bluntly explained, "We want to operate anywhere on the Internet in a way that no one knows the CIA is looking at them."[26] Anonymity similarly provides a means of probing the security of other nations' online infrastructure and disabling or disrupting online services abroad without revealing the U.S. government's role. The veil of anonymity additionally extends protection to intelligence assets overseas, who are able to transmit reports without fear of revealing their ties to the U.S. government. For these very reasons the U.S. Navy's Office of Naval Research funded the initial research on onion routing, including Tor, and only discontinued primary support in 2004 once the Tor Project had secured a new home at the Electronic Frontier Foundation.[27] The CIA made its own strategic investment in anonymizing technology through its venture capital firm, In-Q-Tel; in 2001 it purchased a share of SafeWeb, and required the (ultimately unsuccessful[28]) company to implement support for proprietary CIA

25. While mail and in-person interactions (and to a lesser extent books) will remain venues for anonymous speech, none offer rapid publication to a wide audience and only mail offers a rigorous anonymity guarantee.

26. Neil King Jr., "Small Start-Up Helps CIA Mask Its Moves on Web", *Wall Street Journal* (February 12, 2001).

27. U.S. Naval Research Laboratory, "Onion Routing: History", `http://www.onion-router.net/History.html`.

28. Significant deanonymizing bugs led to the product's failure at market. David Martin and Andrew Schulman, "Deanonymizing Users of the SafeWeb Anonymizing Service", February 11, 2002, `http://www.cs.bu.edu/techreports/pdf/2002-003-deanonymizing-safeweb.pdf`.

encryption standards in its product.[29]

A final independent benefit of Internet anonymity is its relation to privacy; "anonymity ensures privacy."[30] Though the early Internet held the promise of few prying eyes, commercial and government interests quickly learned the wealth of information that could surreptitiously be gleaned from browsing and shopping habits.[31] Amidst pervasive tracking cookies and commercial data mining, anonymity provides the greatest privacy guarantee; so long as all other parties are unaware of an individual's identity they have no means of, adopting a generic construal of privacy, learning some information about the individual they would rather have kept secret.

## The Legal Status of Anonymity in the United States

On numerous occasions the U.S. Congress and Supreme Court have recognized the benefits of anonymity, and a long history of legislation and jurisprudence suggests Internet anonymity is a First Amendment right.[32] Four broad threads inform this finding: the right to circulate speech, the right to congregate anonymously, the right to publish and circulate anonymously, and the finding that online speech is afforded

29. King Jr., "Small Start-Up Helps CIA Mask Its Moves on Web"; In-Q-Tel, "In-Q-Tel Commissions SafeWeb for Internet Privacy Technology", February 14, 2001, `http://www.iqt.org/news-and-press/press-releases/2001/Safeweb_02-14-01.html`.

30. L. Jean Camp, "Web Security and Privacy: An American Perspective", *The Information Society* 15, no. 1 (1999): 249–256.

31. Robert O'Harrow, *No Place to Hide* (New York: Free Press, 2006), 34-73, 214-246; Hal Abelson, Ken Ledeen, and Harry Lewis, *Blown to Bits: Your Life, Liberty, and Happiness After the Digital Explosion* (Upper Saddle River: Addison-Wesley, 2008), 19-72.

32. The First Amendment reads: "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances."

the same protections as printed speech.[33]

The 1877 *Ex parte Jackson* Supreme Court decision, striking down a statute restricting mailings about legal lotteries, first established the right to not only publish, but also circulate speech: "liberty of circulating is as essential to that freedom [of the press] as liberty of publishing."[34] Issues of circulation arose again in the 1938 *Lovell v. Griffin*, in which a Jehovah's Witness was arrested for distributing pamphlets in violation of a Griffin, Georgia statute requiring prior written permission.[35] The Supreme Court extended the right to circulate to individuals in a unanimous decision, finding the statute unconstitutional on its face as a violation of the First Amendment.

Anonymous congregation developed as a point of contention during the civil rights movement, when conservative states sought to combat the effects of growing pro-civil rights organizations. In the 1958 case *NAACP v. Alabama* the state of Alabama attempted to compel the local National Association for the Advancement of Colored People (NAACP) chapter with a contempt citation to provide a list of its members.[36] The Supreme Court recognized in another unanimous decision "the vital relationship between freedom to associate and privacy in one's associations," and that "inviolability of privacy in group association may in many circumstances be indispensable to preservation of freedom of association, particularly where a group

33. Please note: the author of this text is not a lawyer (though he does hope to have a JD in three years); **none of the following should be considered qualified legal advice**.
34. "Ex parte Jackson", `http://supreme.justia.com/us/96/727/case.html`.
35. "Lovell v. Griffin", `http://supreme.justia.com/us/303/444/case.html`.
36. "NAACP v. Alabama", `http://supreme.justia.com/us/357/449/case.html`.

espouses dissident beliefs."[37] Two years later the Supreme Court broadened the right to anonymous association in *Bates v. Little Rock*, rejecting a municipal regulation that required charities to provide their membership in exchange for tax exemption.[38] Thus, outside of particularly compelling interests – far beyond those presented by the Southern states and municipalities – government can neither require nor even incentivize breaches of anonymous association.

Anonymous publishing first entered major jurisprudence in 1913 with the case *Lewis Publishing Co. v. Morgan*.[39] The Supreme Court upheld a Congressional statute that "provided lower postal rates to newspapers and magazines" if they furnished "information regarding ownership, managerial and editorial personnel, and circulation."[40] Congress could offer such a discount, the Court found, only because the Postal Service already provided reasonable service to periodicals which did not wish to comply. In so doing, the Supreme Court created the germ of an implicit right to anonymously publish. The landmark 1960 case *Talley v. California*, another offspring of the civil rights movement, solidified anonymity's protection under the First Amendment.[41] Overturning a Los Angeles ordinance requiring an author's name and address on all handbills as overly restrictive, Justice Hugo Black wrote for the Court:

> Anonymous pamphlets, leaflets, brochures and even books have played an important role in the progress of mankind. Persecuted groups and sects

37. "NAACP v. Alabama".
38. "Bates v. Little Rock", `http://supreme.justia.com/us/361/516/case.html`.
39. "Lewis Publishing Co. v. Morgan", `http://supreme.justia.com/us/229/288/case.html`.
40. David M. Rabban, *Free Speech in its Forgotten Years, 1870-1920* (Cambridge: Cambridge University Press, 1999), 151.
41. "Talley v. California", `http://supreme.justia.com/us/362/60/case.html`.

> from time to time throughout history have been able to criticize oppressive practices and laws either anonymously or not at all.

While *Talley* set a high bar for restrictions on anonymous speech, state governments continued attempts to carve exceptions. The Supreme Court finally clarified the legal status of anonymous publication in the 1995 case *McIntyre v. Ohio*, challenging an Ohio law that required the same author and address information on all campaign-related literature.[42] The Court held the decision to publish anonymously falls within the ambit of a speaker's control over their speech's content, and consequently is subject to the same level of scrutiny as the speech itself; "an author's decision to remain anonymous, like other decisions concerning omissions or additions to the content of a publication, is an aspect of the freedom of speech protected by the First Amendment."[43] Justice John Paul Stevens summarized the Court's view of anonymous speech in the opinion:

> Under our Constitution, anonymous pamphleteering is not a pernicious, fraudulent practice, but an honorable tradition of advocacy and dissent. Anonymity is a shield from the tyranny of the majority... It thus exemplifies the purpose behind the Bill of Rights, and of the First Amendment in particular: to protect unpopular individuals from retaliation–and their ideas from suppression–at the hand of an intolerant society.

Though *Talley* and *McIntyre* leave the door open to narrowly tailored regulation of anonymous speech that is either indecent, obscene, or otherwise particularly harmful, in the vast majority of cases speakers consequently have a constitutional right to remain anonymous. In practice the high regulatory standard imposed on legislators

42. "McIntyre v. Ohio", `http://supct.law.cornell.edu/supct/html/93-986.ZO.html`.
43. Ibid.

is nearly unattainable; in *Hiibel v. Nevada* the Supreme Court barely upheld on a 5-4 decision a Nevada "stop and identify" law requiring individuals to identify themselves to police officers "under circumstances which reasonably indicate that the person has committed, is committing or is about to commit a crime."[44]

The final trend in legislation and jurisprudence that informs the legal status of Internet anonymity is a series of laws and decisions examining whether the Internet as a medium should be afforded the same protections as verbal communication, print, and other traditional "free speech zones." The 1997 *Reno v. ACLU* decision examined the constitutionality of the 1996 Communications Decency Act (CDA), which criminalized "the knowing transmission of [or making available] obscene or indecent messages to any recipient under 18 years of age" and effectually required sites serving indecent content to, at not insignificant cost, utilize credit card authentication to check each visitor's age.[45] The federal government argued the Internet should be subject to the same restrictions as broadcast media, and that in prior cases the Court had upheld various regulations on indecent speech; both lines of argument were rejected. The Court found that the twin factors which justify the regulation of broadcast media, scarcity – that only a certain number of publishers can co-exist in a medium – and invasiveness – that broadcast "invades" the home and cannot be avoided – "are not present in cyberspace," and concluded that the Internet is entitled to the highest level of free speech protection:

44. "Hiibel v. Nevada", `http://supct.law.cornell.edu/supct/html/03-5554.ZO.html`.
45. "Reno v. ACLU", `http://www.law.cornell.edu/supct/html/96-511.ZO.html`.

> Through the use of chat rooms, any person with a phone line can become a town crier with a voice that resonates farther than it could from any soapbox. Through the use of Web pages, mail exploders, and newsgroups, the same individual can become a pamphleteer...We agree with its [the District Court's] conclusion that our cases provide no basis for qualifying the level of First Amendment scrutiny that should be applied to this medium.

As for the CDA's restrictions on indecent content, the Court found they failed the strict scrutiny standard for First Amendment compliance by being overly broad and burdensome for website operators. Revisiting the issue of online speech again in 2004 with *Ashcroft v. ACLU*, the Supreme Court struck down the Child Online Protection Act (COPA) because, though narrower in scope than the CDA by only extending its provisions to commercial content, it too failed the strict scrutiny standard; content filters on web clients are more effective and less intrusive than requiring website operators to take action.[46] Though the 2003 *United States v. ALA* did allow content filtering on federally funded library computers, the court's finding was contingent upon libraries deactivating their filters on the request of an adult.[47] *Reno* and *Ashcroft* are a strong signal that the Internet is a protected "free speech zone" where all legislation and jurisprudence on free speech applies.

Weaving these four threads together, individuals have the right to publish and circulate speech, the right to congregate anonymously, the right to publish anonymously, and all of these rights apply on the Internet. While no legislation or Supreme

46. "Ashcroft v. ACLU", `http://supct.law.cornell.edu/supct/html/03-218.ZO.html`.

47. "United States v. American Library Association", `http://supct.law.cornell.edu/supct/html/02-361.ZO.html`.

Court jurisprudence specifically protects Internet anonymity, little inference is required to find it implicitly guaranteed in the aforementioned case law. Lending credence to this theory, in the 1997 *ACLU v. Miller* a District Court followed similar reasoning in overturning a Georgia statute banning anonymous or pseudonymous online communications.[48] The U.S. policy response to Internet anonymity is, consequently, constrained from directly opposing it.

## The Case Against Anonymity Online

Even the most fervent supporter of anonymity must admit the significant harms that result, however. As with the case in favor of anonymity online, arguments against can devolve into unwarranted claims about "loss of trust," "a general deterioration of morals," and that "bravery, honesty, and openness should be encouraged."[49] Some critics go so far as to claim Internet anonymity would result in apocalypse: "by allowing anonymous communication we actually risk an incremental breakdown of the fabric of our society."[50] Given the relative degree of anonymity already present on the Internet, however, such arguments have little traction.

A second weak line of attack claims Internet anonymity would be ineffective at best; "messages sent anonymously are... unlikely to have much impact on their own," and, at any rate, "the very notion of free speech under law means protecting the

48. "ACLU v. Miller", `http://www.aclu.org/privacy/speech/15521lgl19970620.html`.
49. Davenport, "Anonymity on the Internet: Why the Price May Be Too High", 34.
50. Ibid., 33.

speaker from prosecution and persecution, thus the speaker's identity is known."[51] The former claim is handily rebutted by the wealth of historical evidence introduced earlier, and the latter mistakes that free speech protections prevent private individuals or organizations from heaping repercussions upon a speaker.

Turning to the more valid claims against Internet anonymity, one of the strongest is its utility for libelous activity. By eliminating the accountability normally associated with speech, anonymity allows malicious individuals to spread libelous content without fear of reprisal. This concern is by no means academic; in a notable 1995 case a victim's name and telephone number were posted alongside materials glorifying the Oklahoma City Federal Building bombing to an America Online (AOL) message board.[52] Threatening phone calls poured into the victim's home at a rate of roughly one every two minutes, and the individual was ultimately assigned a protective police detail until the calls subsided weeks later. A lawsuit holding AOL liable for delaying in removal of the offensive postings, *Zeran v. AOL*,[53] was unsuccessful due to Section 230 of the CDA which provides immunity ("safe harbor") to online services from claims arising from user submitted content: "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."[54] In a more recent

51. Davenport, "Anonymity on the Internet: Why the Price May Be Too High", 34.
52. Abelson, Ledeen, and Lewis, *Blown to Bits: Your Life, Liberty, and Happiness After the Digital Explosion*, 242-245.
53. "Zeran v. America Online, Inc.", `http://www.law.emory.edu/4circuit/nov97/971523.p.html`.
54. "Telecommunications Act of 1996".

case, a pair of Yale Law School students were repeatedly defamed on the forum at `AutoAdmit.com`.[55] Having no remedy against the site itself or Google, which indexed the site's contents, under Section 230, the students had no choice but to engage in the arduous and costly process of filing individual libel claims against each of the posters.[56]

Criminal activity enabled by anonymity is another harm, and a reality on the Internet today. Advance-fee scams, most notably operated from Nigeria, promise an individual future riches in exchange for an upfront payment.[57] By hiding behind a cloak of anonymity even once the confidence trick has been recognized the perpetrators offer victims little prospect for recovering their loss. Phishing, the practice of directing users to phony websites where they reveal login information, similarly relies on anonymity; tracing down those responsible is a daunting technical task. The burgeoning field of auction fraud also often makes use of anonymity and forged mailing addresses to prevent the aggrieved party or law enforcement from tracking down the thief. In 2007 alone the Federal Bureau of Investigation received notice of $239 million in individual American losses to online fraud,[58] and a recent survey estimated online merchants lost nearly $4 billion in 2008.[59] Even the inventor of the web, Sir

55. David Margolick, "Slimed Online", *Portfolio* (March 2009).
56. Ibid.
57. Abraham McLaughlin, "Nigeria cracks down on e-mail scams", *The Christian Science Monitor* (December 15, 2005).
58. Internet Crime Complaint Center, "2007 Internet Crime Report", `http://www.ic3.gov/media/annualreport/2007_IC3Report.pdf`.
59. CyberSource, "Online Fraud Report", 2009, `http://forms.cybersource.com/forms/FraudReport2009NACYBSwww020309`, 4.

Tim Berners-Lee, recently fell prey to a scam.[60] Terrorist organizations also benefit from access to Internet anonymity. Combined with encryption software, which Al-Qaeda and other groups are known to utilize,[61] operatives can communicate over the Internet without revealing either that they are talking to one another or the contents of their discussion.

A third downside to Internet anonymity is the possibility of its use to blamelessly broadcast speech that threatens to incite criminal activity, formalized in the United States as the "imminent lawless action" test from *Brandenburg v. Ohio*[62] or more popularly known from the since overturned *Schenck v. United States*[63] as posing a "clear and present danger" akin to "falsely shouting fire in a theater." Legal scholar Cass Sunstein posits that online speech is unusually risky as a medium for inciting criminal activity owing to the immense size of its audience:[64]

> Suppose that an incendiary speech, expressly advocating illegal violence, is not likely to produce lawlessness in any particular listener or viewer. But suppose too that it is believed that of the millions of listeners, one or two, or ten, may well be provoked to act, and perhaps to imminent, illegal violence. Might the government ban advocacy of criminal violence in mass communications when it is reasonable to think that one person, or a few, will take action? *Brandenburg* offers a reasonable approach to the somewhat vague speech in question in that case, which was made in a setting where relatively few people were in earshot. But the case offers unclear guidance on the express advocacy of criminal violence via

60. Dan Goodin, "Web scam hoodwinks web founding father", *The Register* (March 16, 2009), http://www.theregister.co.uk/2009/03/16/berners_lee_burned/.
61. Ellen Messmer, "Al-Qaeda group's encryption software stronger, security firm confirms", *Network World* (February 1, 2008), http://www.networkworld.com/news/2008/020108-al-qaeda-encryption.html.
62. "Brandenburg v. Ohio", http://www.law.cornell.edu/supct/html/historics/USSC_CR_0395_0444_ZO.html.
63. "Schenck v. United States", http://www.law.cornell.edu/supct/html/historics/USSC_CR_0249_0047_ZO.html.
64. Cass Sunstein, "Constitutional Caution", *University of Chicago Legal Forum* 1996 (1996): 370.

the airwaves or the Internet.

While some might consider the prevalence of content a boon – corrupting influences would have to compete for attention[65] – one must recall that the ability to rapidly consume content online is a strong countervailing force: where before one might attend a white supremacist rally at a farm hours away on occasion, the Internet user could view one every half hour. Though the absence of a physical crowd could also decrease the likelihood of violence,[66] online video significantly restores the sensation of presence. Moreover, this observation suggests a further risk: real-world rallies either tend to be at remote locations or well-monitored by local law enforcement, and are attended by other individuals who could both actively temper a would-be criminal's ire (perhaps for fear of prosecution themselves) and passively set an example of non-violence. A lone viewer would have neither of these checks. At present the risk of anonymous online speech sparking illegal activity appears contained to the Internet, such as with the recent incitation of Russian hackers to deface and disable Estonian online services,[67] but real-world violence may be a realistic threat in future; there is no technical barrier on the Internet to, for example, severe hate speech or instructions on converting an assault rifle to fully automatic.

Obscene and harmful-to-minors content poses a fourth downside. With no potential for punishment individuals have little incentive to refrain from sharing child

65. Discussion with Professor Felten, March 20, 2009.
66. Discussion with Professor Felten, March 27, 2009.
67. "Estonia and Russia: A Cyber-riot", *The Economist* (May 10, 2007).

pornography and other materials with, in the view of many, no socially redeeming value and significant harms. Again, such content is not widespread on the Internet at present, but could be going forwards.

Copyright infringement and other intellectual property violations form a fifth issue. Whether or not one considers the Recording Industry Association of America's lawsuits against file sharers effective,[68] individuals using anonymizing technologies successful evade any degree of disincentive provided. This practice appears to already be occurring: observation of a Tor exit node in late 2007 found that 40% of traffic was the popular file sharing protocol BitTorrent,[69] and it does not take much stretch of the imagination to believe the vast majority of such traffic was in violation of intellectual property rights.

A sixth and final concern is that other nations or non-state actors could use Internet anonymity, like the U.S. and its allies, to gather intelligence and carry out attacks with impunity. Again the trouble is not a mere hypothetical: in repeated incidents the computer systems behind presidential campaigns,[70] defense departments,[71] and in numerous other sensitive contexts have been breached by unknown attackers.[72] China in particular is known to have dedicated significant resources towards

68. Sarah McBride and Ethan Smith, "Music Industry to Abandon Mass Suits", *Wall Street Journal* (December 19, 2008).

69. Damon McCoy et al., "Shining Light in Dark Places: Understanding the Tor Network", in *Privacy Enhancing Technologies* (2008).

70. Lee Glendinning, "Obama, McCain computers 'hacked' during election campaign", *Guardian* (November 7, 2008).

71. "Several countries trying to hack into US military system: Pentagon", *AFP* (September 3, 2007).

72. PBS, "frontline: cyber war!: the warnings?", `http://www.pbs.org/wgbh/pages/frontline/shows/cyberwar/warnings/`.

developing a cyberattack capacity,[73] while a variety of other nations were revealed to be employing anonymizing technology for routine business in a 2007 incident where an individual monitored a Tor exit node's unencrypted traffic.[74] Penetration testing, meanwhile, suggests private American infrastructure is severely at risk from cyberattack; in one recent simulation Department of Homeland Security attackers were able to disable a power plant without any physical intervention.[75]

## A Mature Policy for a Mature Internet

The Internet has changed significantly in the past decade, and policy must catch up. This concluding section proposes a coherent set of policies that would retain the congruent benefits and minimize the harms of Internet anonymity, taking into account the technological realities imposed by Web 2.0.

The status quo is far from ideal: libelous speakers, criminals, and foreign attackers exploit the Internet with relative impunity, while whistleblowers and dissenters are harassed and detained. That said, reactions calling for identifying information to be associated with Internet traffic (often referred to as "traceback") are overbroad and ignore the numerous benefits of Internet anonymity. Such calls have emanated from foreign governments attempting to enforce censorship through international in-

73. Tim Reid, "China's cyber army is preparing to march on America, says Pentagon", *The Times* (September 8, 2007).

74. Zetter, "Embassy E-mail Account Vulnerability Exposes Passport Data and Official Business Matters"; Zetter, "Rogue Nodes Turn Tor Anonymizer Into Eavesdropper's Paradise".

75. Jeanne Meserve, "Sources: Staged cyber attack reveals vulnerability in power grid", *CNN* (September 26, 2007), `http://www.cnn.com/2007/US/09/26/power.at.risk/index.html`.

stitutions[76] as well as domestically from law enforcement officials,[77] politicians,[78] and policy projects[79] single-mindedly focused on the security risks posed by the Internet. In a recent report detailing a cybersecurity policy framework for the incoming Obama administration, for example, the Center for Strategic and International Studies concluded, "Creating the ability to know reliably what person or device is sending a particular data stream in cyberspace must be part of an effective cybersecurity strategy,"[80] and the recently proposed Cybersecurity Act of 2009 sets a research agenda for exploring this recommendation.[81] Given the international nature of the Internet, however, a mechanism of this sort – if even worthwhile given the technical objections in Chapter 2 – would only arise through cooperation with the very regimes who would doubtlessly employ it to censor and quash dissidence. Moreover, historical experience with the National Security Agency's Project SHAMROCK[82] and Terrorist Surveillance Program[83] suggest inappropriate and illegal uses of the system would quickly abound, and a 2004-2005 compromise of a Greek cellphone wiretap system

76. McCullagh, "U.N. agency eyes curbs on Internet anonymity".

77. "Hearing of the Commerce, Justice, State, and the Judiciary Subcommittee of the Senate Appropriations Committee", March 10, 1998, `http://w2.eff.org/Censorship/Internet_censorship_bills/1998/19980310_freeh_allen_sen_cjs_app.testimony`.

78. Abelson, Ledeen, and Lewis, *Blown to Bits: Your Life, Liberty, and Happiness After the Digital Explosion*, 161-165.

79. Center for Strategic and International Studies Commission on Cybersecurity for the 44th Presidency, "Securing Cyberspace for the 44th Presidency", December 2008, `http://www.csis.org/media/csis/pubs/081208_securingcyberspace_44.pdf`.

80. Ibid., 62.

81. "Cybersecurity Act of 2009", `http://cdt.org/security/CYBERSEC4.pdf`.

82. United States Senate Select Committee to Study Governmental Operations With Respect to Intelligence Activities, "Supplementary Detailed Staff Reports On Intelligence Activities and the Rights of Americans, Book III", 1976, `http://www.icdc.com/~paulwolf/cointelpro/churchfinalreportIIIj.htm`.

83. James Risen and Eric Lichtblau, "Bush Lets U.S. Spy on Callers Without Courts", *The New York Times* (December 16, 2005).

reminds that security vulnerabilities exist in monitoring mechanisms as well, allowing adversaries to employ them against the same governments who developed them.[84] And, as discussed in Chapter 2, end host vulnerabilities will provide anonymity to the technically savvy and legally unencumbered even with traceback.

The U.S. government ought instead embrace the reality of Internet anonymity and adopt policies that simultaneously minimize its harms and magnify its benefits. The first step, essential even without adopting a stance on anonymity, is an overhaul of the mechanisms for addressing objectionable content online, whether libelous, intellectual property infringing, or criminal. The Communications Decency Act of 1996 (CDA)[85] provides near-immunity to web services for user submitted defamatory content, while the Digital Millennium Copyright Act of 1998 (DMCA)[86] provides a means for rights-holders to issue "takedown" notices to websites, initiating a back-and-forth between the rights-holder, the site, and the poster; failure to comply renders the site liable for contributory copyright infringement. While the origin of this incongruity is clear – rights-holding organizations lobby, but aggrieved individuals do not – it makes little sense from the policy perspective. Moreover, the CDA was enacted at a time when the Internet was (at least perceived to be) in a nascent, fragile state. The slow ossification of protocols and online conventions, as well as the spread of Internet restrictions abroad, suggest it is now resilient to increased regulation.

84. Steven M. Bellovin et al., "Risking Communications Security: Potential Hazards of the Protect America Act", *IEEE Security and Privacy* 6, no. 1 (2008): 30.

85. "Telecommunications Act of 1996".

86. "Digital Millenium Copyright Act", `http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=105_cong_public_laws&docid=f:publ304.105`.

A variety of proposals have attempted to rectify the situation. Tim O'Reilly of the manual writing and technology content producing O'Reilly Media initiated a collaborative process to develop a voluntary "Blogger's Code of Conduct" regulating content and submissions in 2007.[87] The subsequent lack of adoption suggests such voluntary standards lack traction, however. An alternative approach proposed by cyberlaw attorney Mike Godwin would condition CDA Section 230 on a "right of reply" to post content neighboring the alleged libel. The offending materials would remain online causing harm, however, and the presence of a response could even add legitimacy. At any rate, the solution does not generalize beyond libel, and provides little guidance on how to approach other issues of offending speech. One of the Yale Law School students libeled on `AutoAdmit.com` has proposed a third resolution, conditioning Section 230 immunity on adoption of a universal rating system.[88] Under her regime sites would prominently disclose whether they allow anonymous comments, the degree of "offensive content" tolerated, the policy towards rectifying incorrect information, and restrictions on submitters. Setting aside the immense standardization issues inherent in the proposal, it too would leave the offensive content available online and provide no resolution for the other forms of harmful speech.

While victims of libel no doubt seek justice of their persecutors and the government aims to pursue those sharing obscene materials, immediately eliminating the

87. Brad Stone, "A Call for Manners in the World of Nasty Blogs", *The New York Times* (April 9, 2007). "Blogging:Blogger's Code of Conduct", `http://blogging.wikia.com/wiki/Blogger%27s_Code_of_Conduct`.

88. Caitlin Hall, "A Regulatory Proposal for Digital Defamation: Condition § 230 Safe Harbor on the Provision of a Site 'Rating'", *Stanford Technology Law Review*, no. 1 (2008).

content causing harm is a more pressing concern in both cases. Recognizing this, legislators should revisit the CDA and DMCA and develop a coherent takedown framework that flexibly addresses all forms of offending content, and provides remedies parallel to contributory copyright infringement against web services for other harms.[89] A decade of interactions under the CDA, DMCA, and current libel law must inform new regulation, however; takedown abuses and strategic lawsuits against public participation (SLAPP's) run rampant,[90] and failure to assign attorney fees results in a perverse incentive to sue even immunized web services.[91] New law must impose real penalties for issuing invalid takedown notices, provide anti-SLAPP and/or counterclaim "SLAPPback" provisions,[92] and burden the unsuccessful party in all actions with the full cost of litigation. Lawmakers must also revisit the question of default behavior after a takedown notification; at present the default is to remove the offending content only to reinstate it upon the poster's rebuttal, but perhaps retaining the content until receiving a reply would be a more appropriate standard for libelous content. Having developed such a scheme, the U.S. should seek to promulgate it abroad, allowing U.S. citizens to issue takedowns to foreign sites and vice versa. Critics will quickly, and rightly, point out the immense burden and responsibility such a proposal places on

89. Applying the takedown framework to libel is discussed at length in Bradley A. Areheart, "Regulating Cyberbullies Through Notice-Based Liability", *The Yale Law Journal Pocket Part* 117, no. 41 (2007): 41–47.

90. Electronic Frontier Foundation, "Unintended Consequences: Ten Years under the DMCA", October 2008, `http://www.eff.org/wp/unintended-consequences-ten-years-under-dmca`; Electronic Frontier Foundation, "CyberSLAPP", `http://www.eff.org/issues/cyberslapp`.

91. Anthony Ciolli, "Chilling Effects: The Communications Decency Act and the Online Marketplace of Ideas", *University of Miami Law Review* 63, no. 1 (2008): 137–268.

92. See, for example, California's anti-SLAPP legislation: California Anti-SLAPP Project, "California Statutes", `http://www.casp.net/statutes/calstats.html`.

web services. But at the point at which the DMCA already requires the capacity to arbitrate takedown claims, retooling mechanisms to address a broader array of content issues seems well within reason. A second claim, that "such a regime would be ineffective, because by the time a victim realizes the problem, notifies the website operator, and has the material removed, it may have spread to other sites, becoming effectively impossible to contain,"[93] is not borne out, at least in the case of libel, by experience – most incidents appear contained to a single site even when materials are taken down at the site's discretion.

Over time a coherent takedown system of this sort should reduce the likelihood of libelous content being shared in the first place. In a rough parallel of "deterrence by denial" from international relations,[94] should an individual know their false statements or illegal materials will enjoy only a short lifespan they may be disincentivized to share them at all.

Fraud and other commercially-related anonymity ills are easily dispatched by either disallowing or issuing strong warnings against anonymous payment systems. As many ethical systems, and certainly U.S. law, have recognized on many occasions, commercial speech does not enjoy the same protections as political or other "core" speech.[95] Shunning anonymous transactions neither stymies ideas in the marketplace nor has broader chilling effects.

93. Danielle Keats Citron, "Cyber Civil Rights", *Boston University Law Review* 89 (2009): 123.
94. For a detailed explanation of "deterrence by denial" see David S. Yost, "Debating security strategies", *NATO Review*, no. 4 (2003).
95. Froomkin, "Legal Issues in Anonymity and Pseudonymity", 118.

As for the threat of foreign intelligence and cyberattack, as discussed earlier anonymity will always be available to those willing to compromise end hosts. Only a robust cybersecurity policy, which falls well outside the scope of this work, can adequately address such issues.[96] Forms of strong online identification may be a component of such policy, but should only be required when absolutely necessary – for example, for access to sensitive government resources.

Policy should aim to support anonymizing technology as a non-excludable, non-rival public good under-provided by the market. Delivering funding will require gingerly navigating the widespread distrust of government online; the U.S. may find it best, counterintuitively, to not immediately involve itself in anonymizing technologies, thereby lending them credence. Funding should first be attempted through direct support and in the guise of academic endeavor, though.[97] All possible opportunities for partnership with the private sector should similarly be tentatively explored.

At the same time, the Federal Trade Commission should aim to ensure the level of knowledge and quality of resources about Internet anonymity are significantly improved from the dire snapshot in Chapter 3, framing the issue as a consumer awareness problem. Its efforts should begin with direct information resources along the lines of its existing "OnGuard Online" site, which provides materials on broad-

---

96. See, for example, Center for Strategic and International Studies Commission on Cybersecurity for the 44th Presidency, "Securing Cyberspace for the 44th Presidency"; "The National Strategy to Secure Cyberspace", `http://www.dhs.gov/xlibrary/assets/National_Cyberspace_Strategy.pdf`.

97. The Tor Project currently receives funding from the Broadcasting Board of Governors, parent of Voice of America, but the amount is unclear and the association with propaganda is less than desirable , "Tor: Sponsors", `http://www.torproject.org/sponsors.html.en`.

band, spam, and a variety of other Internet-related topics that impact American consumers.[98] Partnerships with industry and academia should aim to develop similar informational sites linked to products, academic projects, and IT. A secondary focus should be complaints and, if necessary, enforcement action against the anonymizing technology firms and browser vendors for not adequately representing their products' capabilities to users; browsers and anonymizing services should provide clear notice to users of the extent of anonymization offered by their software. If possible the FTC, or perhaps a Congressional committee acting in a consumer-oriented capacity, should encourage browser vendors to increase the level of anonymity offered by their products to be on par with the best systems available. An ideal outcome would be if Torbutton were built into Firefox, for example; not only would end users have more ready access to anonymity, but the influx of users would further cloak identity.[99] As discussed in Chapter 2, without such direct market intervention browser developers will have little incentive to focus on the privacy traits of their products.[100]

A final element of U.S. anonymity policy must be the pursuit of anonymity-friendly outcomes in Internet standardization and governance bodies;[101] policymakers and scientists must recognize that technical design decisions often bear not just scien-

98. "OnGuard Online", `http://www.onguardonline.gov/`.

99. Dingledine and Mathewson, "Anonymity Loves Company: Usability and the Network Effect".

100. This problem is not particularly unique; markets consistently fail to provide security and privacy features in software products. See Mark F. Grady and Francesco Parisi, eds., *The Law and Economics of Cybersecurity* (Cambridge: Cambridge University Press, 2006).

101. For discussion of international Internet governance see John Mathiason, *Internet Governance* (New York: Routledge, 2009); also David G. Post, *In Search of Jefferson's Moose: Notes on the State of Cyberspace* (Oxford: Oxford University Press, 2009), 126-186.

tific, but also political ramifications. Traceback proposals have surfaced on numerous occasions in the Internet's history, most recently in the United Nations' International Telecommunications Union,[102] and a watchful eye will be required to ensure none pass in future.

102. McCullagh, "U.N. agency eyes curbs on Internet anonymity".

# 5 The "Virus of Liberty"

In the years ahead the locus of power in Internet protocol standardization and software development will doubtlessly shift away from the United States. But until that time, U.S. government policy and private industry practice set the global Internet agenda. Ensuring access to anonymity domestically will, consequently, guarantee all populations with Internet access have the ability to go nameless subject to minimal technical and legal assumptions – an individual need only be able to run arbitrary software and access hosts in the United States with impunity. Given the criticality of software and the Internet for success in the globalized economy, the set of nations meeting this baseline will only grow with time. Adopting the verbiage made famous by cyberlaw scholar Lawrence Lessig, the influence of America's "East Coast Code," law, is bounded by its shores, but "West Coast Code," software, reaches into even the most oppressive of regimes.[1]

The United Nations Declaration of Human Rights, adopted by the General Assembly on December 10, 1948, envisioned a world in which

> Everyone has the right to freedom of thought, conscience and religion; this

1. Lessig, *Code version 2.0*, 72-74.

> right includes freedom to change his religion or belief, and freedom, either alone or in community with others and in public or private, to manifest his religion or belief in teaching, practice, worship and observance.[2]

and

> Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.[3]

While great strides have been made in human rights over the past half century, a number of recalcitrant regimes regulate domestic media, remain unaccountable to their citizens, and continue to persecute and prosecute individuals on the basis of their speech.[4] Internet anonymity is a shield against the club of censorship, and combined with the innovations of Web 2.0 both guarantees free speech to a global audience and, through the abundance of user generated content, enables novel forms of constructive dissent. An episode from 2007 foreshadows the future of political criticism: an individual gathered photographic evidence of the Tunisian president's plane visiting popular vacation and shopping destinations throughout Europe from airplane affinity sites and posted them in a video on a pseudonymous blog.[5] And a recent study concluded that even in China, a nation that has made significant efforts

2. United Nations General Assembly, "The Universal Declaration of Human Rights", December 10, 1948, `http://www.un.org/Overview/rights.html`, Article 18.
3. Ibid., Article 19.
4. Freedom House, "Freedom in the World", 2008, `http://www.freedomhouse.org/template.cfm?page=15`.
5. Astrubal (pseudonym), "Tunisie: Qui utilise l'avion de la présidence de la République Tunisienne?" (August 29, 2007), `http://astrubal.nawaat.org/2007/08/29/tunisie-avion-presidentiel/`.

to censor the Internet,[6] venues remain for posting critical content.[7] Thus, not only does Internet anonymity bestow all the benefits to society discussed earlier, but it ensures they accrue to those in most dire need of them.

Some readers will find this outcome more persuasive in a national security context; proponents of Democratic Peace Theory may note that free speech could pressure for democratic reform, which in turn increases the stability of the international order.[8] Whatever the justification, Internet anonymity bears the promise of Barlow's dream fulfilled: a "virus of liberty" "creating a world where anyone, anywhere may express his or her beliefs."[9]

---

6. Deibert et al., *Access Denied: The Practice and Policy of Global Internet Filtering*, 263-271; For a technical discussion, see Richard Clayton, Steven J. Murdoch, and Robert N. M. Watson, "Ignoring the Great Firewall of China", in *Privacy Enhancing Technologies* (2006).

7. Rebecca MacKinnon, "Studying Chinese blog censorship" (November 29, 2008), `http://rconversation.blogs.com/rconversation/2008/11/studying-chines.html`.

8. For a discussion of Democratic Peace Theory see Michael E. Brown, Sean M. Lynn-Jones, and Steven E. Miller, eds., *Debating the Democratic Peace* (Cambridge: MIT Press, 1996).

9. Barlow, "A Declaration of the Independence of Cyberspace".

# Appendix A: Survey

**1.** What is your major?
(free response)

**2.** How would you rate your level of competency with computers and the Internet?
1: Novice user. Inexperienced with using a computer.
2: Common user. Regularly use a computer for word processing, email, and web surfing. Very limited knowledge of the technical workings of the Internet.
3: Advanced user. Often install and use new programs or web services, some knowledge of the technical workings of the Internet.
4: Expert user. Some expertise in the technical workings of computers and the Internet.

**3.** How anonymous do you believe you are to the web sites you visit?
1: Not at all anonymous. Your actions can be completely tracked.
2: Fairly anonymous. The site could track your actions online if it so desired.
3: Very anonymous. A determined site could track your actions online with difficulty.
4: Completely anonymous. Even a determined site would be unable to track your actions online.

**4.** Which of the following do you believe a web site you visit could easily determine?
( ) Your web browser
( ) Your browsing history
( ) Your location
( ) Your name

**5.** How anonymous do you believe you are to other users on the Internet?
1: Not at all anonymous. Your actions can be completely tracked.
2: Fairly anonymous. A user could track your actions online if it so desired.
3: Very anonymous. A determined user could track your actions online with difficulty.
4: Completely anonymous. Even a determined user would be unable to track your actions online.

**6.** Which of the following do you believe a user on the Internet could easily deter-

mine?
( ) Your web browser
( ) Your browsing history
( ) Your location
( ) Your name

**7a.** Do you believe you have the ability to browse the web anonymously if necessary?
( ) Yes
( ) No

**7b.** How confident are you in the above response?
1: Not very confident.
2: Somewhat confident.
3: Confident.
4: Very confident.
5: Certain.

**8.** Please briefly (3-4 sentences or bullets) describe the steps you would take to browse the web as anonymously as possible.
(free response)

**9.** Please list the resources you would use to learn how to browse the web anonymously.
(free response)

# Appendix B: Proofs

## Maximum Likelihood Model

In a two anonymity set (binomial) case, developing the maximum likelihood model is trivial. The probability of a particular sample is

$$\mathrm{P}(\mathrm{N}_0 = n_0, \mathrm{N}_1 = n_1) = {p_0}^{n_0}(1-p_0)^{M-n_0}\binom{M}{n_0}$$

where $N_i$ is a random variable valued at the number of occurances of the $i$th observed anonymity set in the sample, $n_i$ is a specific assignment, and $M$ is the sample size. Taking the derivative of the log likelihood (an acceptable move because logarithms are monotonic functions) and setting to 0 gives

$$\begin{aligned}
\ln\mathrm{P} = n_0 \ln p_0 + (M-n_0)\ln(1-p_0) + \ln\left[\binom{M}{n_0}\right] \\
\frac{d\ln\mathrm{P}}{dp_0} = 0 = \frac{n_0}{p_0} + -1\frac{M-n_0}{1-p_0} \\
\frac{M-n_0}{1-p_0} = \frac{n_0}{p_0} \\
p_0 = \frac{n_0}{M}.
\end{aligned}$$

As intution might suggest, then, the maximum likelihood model of the global population holds the global population proportion in each anonymity set equal to the sample population proportion in each anonymity set. Developing a proof of this intuition in the multinomial case of more than two anonymity sets, as resulted from the experiment, is less straightforward. Though taking the same approach as the binomial case except with $L$ anonymity sets seems the intuitive move, it's a trap:[2]

$$\begin{aligned}
\mathrm{P}(\forall i : \mathrm{N}_i = n_i) &= \prod_{i=0}^{L-1} \left[ {p_i}^{n_i} \binom{M - \sum_{j=0}^{i-1} n_j}{n_i} \right] \\
\ln\mathrm{P} &= \sum_{i=0}^{L-1} n_i \ln p_i + \sum_{i=0}^{L-1} \ln \left[ \binom{M - \sum_{j=0}^{i-1} n_j}{n_i} \right] \\
\frac{\partial \ln\mathrm{P}}{\partial p_i} &= 0 = \frac{n_i}{p_i} \\
0 &= \frac{n_i}{p_i},
\end{aligned}$$

which gives no satisfying set of assignments at all! The proof above errs in underconstraining the multinomial; given $L - 1$ anonymity set sample proportions, the final set sample proportion is already determined. Formulating this realization by re-expressing the final set in terms of the prior sets gives

2. Admiral Ackbar, "Return of the Jedi", *Star Wars* 6 (1983).

$$\mathrm{P}(\forall i : \mathrm{N}_i = n_i) = \prod_{i=0}^{L-2} \left[ {p_i}^{n_i} \binom{M - \sum_{j=0}^{i-1} n_j}{n_i} \right] p_{L-1}^{n_{L-1}}$$

$$\mathrm{P} = \prod_{i=0}^{L-2} \left[ {p_i}^{n_i} \binom{M - \sum_{j=0}^{i-1} n_j}{n_i} \right] (1 - \sum_{i=0}^{L-2} p_i)^{L - \sum_{i=0}^{M-2} n_i}$$

$$\ln\mathrm{P} = \sum_{i=0}^{L-2} n_i \ln p_i + (M - \sum_{i=0}^{L-2} n_i)\ln(1 - \sum_{i=0}^{L-2} p_i) + \sum_{i=0}^{L-2} \ln \left[ \binom{M - \sum_{j=0}^{i-1} n_j}{n_i} \right]$$

$$\frac{\partial \ln\mathrm{P}}{\partial p_i} = 0 = \frac{n_i}{p_i} + -1 \frac{M - \sum_{i=0}^{L-2} n_i}{1 - \sum_{i=0}^{L-2} p_i}$$

$$\frac{n_i}{n_{L-2}} = \frac{p_i}{p_{L-2}}$$

$$\frac{p_i}{n_i} = \frac{p_{L-2}}{n_{L-2}}.$$

From this last step the ratio of population proportion to sample incidence must always be the same,

$$\forall i : \frac{p_i}{n_i} = \frac{p_{L-2}}{n_{L-2}}$$

$$\forall i : p_i = \frac{n_i p_{L-2}}{n_{L-2}},$$

and summing over $i$ with the knowledge that $\sum_{i=0}^{L-1} p_i = 1$ in the likelihood maximizing case ($\sum_{i=0}^{L-1} p_i < 1$ would assuredly result in a lower likelihood) finally gives the intuitive result:

$$
\begin{aligned}
\sum_{i=0}^{L-1} p_i &= \frac{p_{L-2}}{n_{L-2}} \sum_{i=0}^{L-1} n_i \\
1 &= \frac{p_{L-2} M}{n_{L-2}} \\
\frac{p_{L-2}}{n_{L-2}} &= \frac{1}{M} \\
\forall i : p_i &= \frac{n_i}{M}.
\end{aligned}
$$

## Variance Analysis

Where $M$ is now the number of web clients worldwide and $n$ is the sample size, the one-sided statistical test against complete uniqueness for a particular anonymity set is

$$
z_{1-\alpha} = \frac{1/n - 1/M}{\sqrt{\frac{(1/M)(1-1/M)}{n}}}.
$$

Solving for $n$ through a proof akin to the quadratic formula gives

$$
\begin{aligned}
c = z_{1-\alpha}\sqrt{(1/M)(1-1/M)} &= \frac{1}{\sqrt{n}} - \frac{\sqrt{n}}{M} \\
n + cM\sqrt{n} &= M \\
\left(\sqrt{n} + \frac{cM}{2}\right)^2 &= M + \frac{c^2 M^2}{4} \\
n &= \left(\sqrt{M + \frac{c^2 M^2}{4}} - \frac{cM}{2}\right)^2,
\end{aligned}
$$

where $c$ is an intermediary variable, as the sample size required to not have confidence at the $1-\alpha$ level that a certain anonymity set has more than one member.

# Appendix C: Source Code

Redacted. Please contact the author for source code.

# Bibliography


*A Brief Guide to Alcoholics Anonymous*. New York: Alcoholics Anonymous World Services, Inc., 1972.

Abbott, Timothy G. et al. "Browser-Based Attacks on Tor". In *Privacy Enhancing Technologies*. 2007.

Abelson, Hal, Ken Ledeen, and Harry Lewis. *Blown to Bits: Your Life, Liberty, and Happiness After the Digital Explosion*. Upper Saddle River: Addison-Wesley, 2008.

"ACLU v. Miller". `http://www.aclu.org/privacy/speech/15521lgl19970620.html`.

Admiral Ackbar. "Return of the Jedi". *Star Wars* 6 (1983).

Adobe, Inc. "Adobe Flash Player 9 Security". `http://www.adobe.com/devnet/flashplayer/articles/flash\_player\_9\_security.pdf`.

———. "Flash Player Version Penetration". 2008. `http://www.adobe.com/products/player_census/flashplayer/version_penetration.html`.

"Anonymity". `http://en.wikipedia.org/wiki/Anonymity`.

Anonymizer, Inc. "Anonymous Surfing". `http://www.anonymizer.com/consumer/products/anonymous_surfing/`.

"Anonymous Web Browsing". `http://en.wikipedia.org/wiki/Anonymous_web_browsing`.

"Anonymous Web Proxy". `http://en.wikipedia.org/wiki/Anonymous_web_proxy`.

Apple Inc. "Apple - Safari". `http://www.apple.com/safari/features.html`.

Areheart, Bradley A. "Regulating Cyberbullies Through Notice-Based Liability". *The Yale Law Journal Pocket Part* 117, no. 41 (2007): 41–47.

"Ashcroft v. ACLU". `http://supct.law.cornell.edu/supct/html/03-218.ZO.html`.

Astrubal (pseudonym). “Tunisie: Qui utilise l’avion de la présidence de la République Tunisienne?” (August 29, 2007). http://astrubal.nawaat.org/2007/08/29/tunisie-avion-presidentiel/.

Barlow, John Perry. “A Declaration of the Independence of Cyberspace”. http://homes.eff.org/~barlow/Declaration-Final.html.

Barth, Adam, Collin Jackson, and John C. Mitchell. “Robust Defenses for Cross-Site Request Forgery”. In *15th ACM Conference on Computer and Communications Security (CCS 2008)*. 2008.

“Bates v. Little Rock”. http://supreme.justia.com/us/361/516/case.html.

Bauer, Kevin et al. “Low-resource Routing Attacks Against Tor”. In *WPES ’07: Proceedings of the 2007 ACM Workshop on Privacy in Electronic Society*, 11–20. 2007.

Bellovin, Steven M. et al. “Risking Communications Security: Potential Hazards of the Protect America Act”. *IEEE Security and Privacy* 6, no. 1 (2008): 24–33.

Benkler, Yochai. *The Wealth of Networks*. New Haven: Yale University Press, 2006.

Benninger, Corey. “AJAX Storage: A Look at Flash Cookies and Internet Explorer Persistence”. *Foundstone White Papers* (2006).

Bera, Krishna E. “Torbutton Bug Report: Unmask Screen”. March 24, 2009.

Berners-Lee, T., R. Fielding, and H. Frystyk. “Hypertext Transfer Protocol – HTTP/1.0”. *RFC*, no. 1945 (1996).

“Blogging:Blogger’s Code of Conduct”. http://blogging.wikia.com/wiki/Blogger%27s_Code_of_Conduct.

Box, Richard C. *Public Administration and Society*. Armonk: M.E. Sharpe, 2003.

“Brandenburg v. Ohio”. http://www.law.cornell.edu/supct/html/historics/USSC_CR_0395_0444_ZO.html.

Broder, Andrei. “A Taxonomy of Web Search”. *SIGIR Forum* 36, no. 2 (2002): 3–10.

Brown, Michael E., Sean M. Lynn-Jones, and Steven E. Miller, eds. *Debating the Democratic Peace*. Cambridge: MIT Press, 1996.

Caesar, Matthew et al. “ROFL: Routing on Flat Labels”. In *SIGCOMM ’06*, 363–374. 2006.

California Anti-SLAPP Project. “California Statutes”. http://www.casp.net/statutes/calstats.html.

“Cameras Draw Closer to Beijing’s Internet Cafes”. *The Wall Street Journal China Journal Blog* (October 17, 2008). http://blogs.wsj.com/chinajournal/2008/10/17/cameras-draw-closer-to-beijings-internet-cafes/.

Camp, L. Jean. "Web Security and Privacy: An American Perspective". *The Information Society* 15, no. 1 (1999): 249–256.

"CAN-SPAM Act of 2003". `http://uscode.house.gov/download/pls/15C103.txt`.

Carr, David. "Mourning Old Media's Decline". *The New York Times*, October 28, 2008.

Carvajal, Doreen. "Columnist's Mea Culpa: I'm Anonymous". *The New York Times* (July 18, 1996).

Celeste, Sofia. "Want to check your e-mail in Italy? Bring your passport." *The Christian Science Monitor* (October 4, 2005).

Center for Strategic and International Studies Commission on Cybersecurity for the 44th Presidency. "Securing Cyberspace for the 44th Presidency". December 2008. `http://www.csis.org/media/csis/pubs/081208_securingcyberspace_44.pdf`.

Cerf, Vinton G. and Robert E. Kahn. "A Protocol for Packet Network Intercommunication". *IEEE Transactions on Communications* 22 (1974): 637–648.

"Child Online Protection Act". `http://www4.law.cornell.edu/uscode/html/uscode47/usc_sec_47_00000231----000-.html`.

"Childrens' Internet Protection Act". `http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=106_cong_public_laws&docid=f:publ554.106`.

Ciolli, Anthony. "Chilling Effects: The Communications Decency Act and the Online Marketplace of Ideas". *University of Miami Law Review* 63, no. 1 (2008): 137–268.

Citron, Danielle Keats. "Cyber Civil Rights". *Boston University Law Review* 89 (2009): 61–125.

Clark, David D. "The Design Philosophy of the DARPA Internet Protocols". In *SIGCOMM '88*, 106–114. 1988.

Clark, David D. et al. "Tussle in Cyberspace: Defining Tomorrow's Internet". *SIGCOMM '02* (2002).

Clayton, Richard. "Anonymity and traceability in cyberspace". *University of Cambridge Computer Laboratory Technical Reports*, no. 653 (November 2005).

Clayton, Richard, Steven J. Murdoch, and Robert N. M. Watson. "Ignoring the Great Firewall of China". In *Privacy Enhancing Technologies*. 2006.

Cooperative Association for Internet Data Analysis. "Visualizing IPv4 Internet Topology at a Macroscopic Scale". 2008. `http://www.caida.org/research/topology/as_core_network/`.

"Cybersecurity Act of 2009". `http://cdt.org/security/CYBERSEC4.pdf`.

CyberSource. “Online Fraud Report”. 2009. http://forms.cybersource.com/forms/FraudReport2009NACYBSwww020309.

Danezis, George et al. “Mixminion: Design of a Type III Anonymous Remailer Protocol”. In *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, 2–15. 2003.

Dartnell, Michael Y. *Insurgency Online: Web Activism and Global Conflict*. Toronto: University of Toronto Press, 2006.

Davenport, David. “Anonymity on the Internet: Why the Price May Be Too High”. *Communications of the ACM* 45, no. 4 (2002): 33–35.

Deering, S. and R. Hinden. “Internet Protocol, Version 6 (IPv6) Specification”. *RFC*, no. 2460 (1998).

Deibert, Ronald J. et al., eds. *Access Denied: The Practice and Policy of Global Internet Filtering*. MIT Press, 2008.

Devore, Jay. *Probability and Statistics for Engineering and the Sciences*. 7th ed. Pacific Grove: Duxbury Press, 2007.

“Digital Millenium Copyright Act”. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=105_cong_public_laws&docid=f:publ304.105.

Dingledine, Roger and Nick Mathewson. “Anonymity Loves Company: Usability and the Network Effect”. In *Proceedings of the Fifth Workshop on the Economics of Information Security (WEIS 2006)*. 2006.

Dingledine, Roger, Nick Mathewson, and Paul Syverson. “Tor: The Second-generation Onion Router”. In *Proceedings of the 13th USENIX Security Symposium*, 303–320. 2004.

Egevang, K. and P. Francis. “The IP Network Address Translator (NAT)”. *RFC* (1994).

Electronic Frontier Foundation. “CyberSLAPP”. http://www.eff.org/issues/cyberslapp.

———. “How to Blog Safely (About Work or Anything Else)”. http://www.eff.org/wp/blog-safely.

———. “RIAA v. The People: Five Years Later”. (San Francisco, California, United States) (2008). http://www.eff.org/files/eff-riaa-whitepaper.pdf.

———. “Unintended Consequences: Ten Years under the DMCA”. October 2008. http://www.eff.org/wp/unintended-consequences-ten-years-under-dmca.

“Estonia and Russia: A Cyber-riot”. *The Economist* (May 10, 2007).

“Ex parte Jackson”. http://supreme.justia.com/us/96/727/case.html.

Federal Communications Commission. "In the Matters of Formal Complaint of Free Press and Public Knowledge Against Comcast Corporation for Secretly Degrading Peer-to-Peer Applications". http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-08-183A1.pdf.

Federal Trade Commission. "Commission Enforcement Actions Involving the Internet and Online Services". http://www.ftc.gov/bcp/internet/cases-internet.pdf.

Fielding, R. et al. "Hypertext Transfer Protocol – HTTP/1.1". *RFC*, no. 2068 (1997).

———. "Hypertext Transfer Protocol – HTTP/1.1". *RFC*, no. 2616 (1999).

Fleischer, Gregory. "Unmask Screen - Iframe JavaScript". http://pseudo-flaw.net/tor/torbutton/unmask-screen-iframe-javascript.html.

Freedom House. "Freedom in the World". 2008. http://www.freedomhouse.org/template.cfm?page=15.

Frei, Stefan, Thomas Duebendorfer, and Bernhard Plattner. "Firefox (In)Security Update Dynamics Exposed". *ACM SIGCOMM Computer Communications Review* 39, no. 1 (2009): 16–22.

Froomkin, A. Michael. "Legal Issues in Anonymity and Pseudonymity". *The Information Society* 15, no. 2 (1999): 113 –127.

Glendinning, Lee. "Obama, McCain computers 'hacked' during election campaign". *Guardian* (November 7, 2008).

Global Voices Advocacy. "Anonymous Blogging with Wordpress and Tor". http://advocacy.globalvoicesonline.org/projects/guide/.

Global Voices. "GlobalVoices Special Coverage". http://globalvoicesonline.org/specialcoverage/.

Goldberg, Ian. "On the Security of the Tor Authentication Protocol". In *Proceedings of the Sixth Workshop on Privacy Enhancing Technologies (PET 2006)*. 2006.

———. "Privacy and Anonymity on the Internet". In *Workshop on Vanishing Anonymity, 15th Conference on Computers, Freedom, and Privacy*. 2005.

Goodin, Dan. "Tor at heart of embassy passwords leak". *The Register* (2007). http://www.theregister.co.uk/2007/09/10/misuse_of_tor_led_to_embassy_password_breach/.

———. "Web scam hoodwinks web founding father". *The Register* (March 16, 2009). http://www.theregister.co.uk/2009/03/16/berners_lee_burned/.

Google Inc. "Explore Google Chrome Features: Incognito Mode". http://www.google.com/support/chrome/bin/answer.py?answer=95464&hl=en.

———. "Gears API". http://code.google.com/intl/en/apis/gears/design.html.

Grady, Mark F. and Francesco Parisi, eds. *The Law and Economics of Cybersecurity.* Cambridge: Cambridge University Press, 2006.

Greene, Thomas C. "Do-it-yourself Internet anonymity". *The Register* (November 14, 2001). http://www.theregister.co.uk/2001/11/14/doityourself_internet_anonymity/.

Gross, Daniel. "Phones Without Homes". *Newsweek*, July 28, 2008.

Grossman, Lev. "Time's Person of the Year: You". *Time Magazine* (December 13, 2006).

Hall, Caitlin. "A Regulatory Proposal for Digital Defamation: Condition § 230 Safe Harbor on the Provision of a Site 'Rating'". *Stanford Technology Law Review*, no. 1 (2008).

Hardie, T. "Distributing Authoritative Name Servers via Shared Unicast Addresses". *RFC*, no. 3258 (2002).

"Hearing of the Commerce, Justice, State, and the Judiciary Subcommittee of the Senate Appropriations Committee". March 10, 1998. http://w2.eff.org/Censorship/Internet_censorship_bills/1998/19980310_freeh_allen_sen_cjs_app.testimony.

Helft, Miguel. "Google to Offer Ads Based on Interests". *The New York Times* (March 11, 2009).

"Hiibel v. Nevada". http://supct.law.cornell.edu/supct/html/03-5554.ZO.html.

Hochstotter, Nadine and Martina Koch. "Standard parameters for searching behaviour in search engines and their empirical evaluation". *Journal of Information Science* 35, no. 1 (2009): 45–65.

Hollinger, David A. "Star Power". *The New York Times* (November 28, 1999).

"IEEE 802.11, The Working Group Setting the Standards for Wireless LANs". http://www.ieee802.org/11/.

"IEEE 802.3 ETHERNET". http://www.ieee802.org/3/.

"IEEE Registration Authority". http://standards.ieee.org/regauth/oui/index.shtml.

In-Q-Tel. "In-Q-Tel Commissions SafeWeb for Internet Privacy Technology". February 14, 2001. http://www.iqt.org/news-and-press/press-releases/2001/Safeweb_02-14-01.html.

International Telecommunication Union. "Free statistics". http://www.itu.int/ITU-D/ict/statistics/.

Internet Crime Complaint Center. "2007 Internet Crime Report". http://www.ic3.gov/media/annualreport/2007_IC3Report.pdf.

"Introducing I2P". http://www.i2p2.de/techintro.html.

Isikoff, Michael. "The Fed Who Blew the Whistle". *Newsweek* (December 13, 2008).

Jesdanun, Anick. "Postal agencies respond to mail decline". *Associated Press*, February 4, 2008.

Joachims, Thorsten et al. "Accurately Interpreting Clickthrough Data as Implicit Feedback". In *ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 154–161. 2005.

Kalathil, Shanthi and Taylor C. Boas. *Open Networks, Closed Regimes: The Impact of the Internet on Authoritarian Rule*. Washington, D.C.: Carnegie Endowment for International Peace, 2003.

Keane, Mark T., Maeva O'Brien, and Barry Smyth. "Are People Biased in Their Use of Search Engines?" *Communications of the ACM* 51, no. 2 (February 2008): 49–52.

King Jr., Neil. "Small Start-Up Helps CIA Mask Its Moves on Web". *Wall Street Journal* (February 12, 2001).

Kurose, James F. and Keith W. Ross. *Computer Networking: A Top-down Approach Featuring the Internet*. Third Edition. Pearson Education, 2004.

Lemann, Nicholas. "Paper Tiger". *The New Yorker*, November 4, 2002.

Lessig, Lawrence. *Code version 2.0*. New York: Basic Books, 2006.

"Lewis Publishing Co. v. Morgan". http://supreme.justia.com/us/229/288/case.html.

Li, Jiexun, Rong Zheng, and Hsinchun Chen. "From Fingerprint to Writeprint". *Communications of the ACM* 49, no. 4 (April 2006): 76–82.

"Lovell v. Griffin". http://supreme.justia.com/us/303/444/case.html.

MacKinnon, Rebecca. "Studying Chinese blog censorship" (November 29, 2008). http://rconversation.blogs.com/rconversation/2008/11/studying-chines.html.

Mahoney, Jill. "Troubled youth find an open ear on-line". *The Globe and Mail* (August 1, 2005).

Margolick, David. "Slimed Online". *Portfolio* (March 2009).

Martin, David and Andrew Schulman. "Deanonymizing Users of the SafeWeb Anonymizing Service". February 11, 2002. http://www.cs.bu.edu/techreports/pdf/2002-003-deanonymizing-safeweb.pdf.

Marx, Gary T. "What's in a Name? Some Reflections on the Sociology of Anonymity". *The Information Society* 15, no. 2 (1999): 99–112.

Mathiason, John. *Internet Governance*. New York: Routledge, 2009.

McBride, Sarah and Ethan Smith. "Music Industry to Abandon Mass Suits". *Wall Street Journal* (December 19, 2008).

McCoy, Damon et al. "Shining Light in Dark Places: Understanding the Tor Network". In *Privacy Enhancing Technologies*. 2008.

McCullagh, Declan. "U.N. agency eyes curbs on Internet anonymity". *CNET News* (September 12, 2008). `http://news.cnet.com/8301-13578_3-10040152-38.html`.

"McIntyre v. Ohio". `http://supct.law.cornell.edu/supct/html/93-986.ZO.html`.

McKinley, Katherine. "Cleaning Up After Cookies". *iSEC Partners White Papers* (2008). `http://www.isecpartners.com/files/iSEC_Cleaning_Up_After_Cookies.pdf`.

McLaughlin, Abraham. "Nigeria cracks down on e-mail scams". *The Christian Science Monitor* (December 15, 2005).

Meserve, Jeanne. "Sources: Staged cyber attack reveals vulnerability in power grid". *CNN* (September 26, 2007). `http://www.cnn.com/2007/US/09/26/power.at.risk/index.html`.

Messmer, Ellen. "Al-Qaeda group's encryption software stronger, security firm confirms". *Network World* (February 1, 2008). `http://www.networkworld.com/news/2008/020108-al-qaeda-encryption.html`.

Microsoft Inc. "Internet Explorer 8: More secure, private, and reliable". `http://www.microsoft.com/windows/internet-explorer/beta/features/browse-privately.aspx`.

"Mixes for Privacy and Anonymity in the Internet". `http://anon.inf.tu-dresden.de/develop/doc/mix_short/`.

Mullan, John. *Anonymity: A Secret History of English Literature*. Princeton: Princeton University Press, 2008.

Murdoch, Steven J. and George Danezis. "Low-Cost Traffic Analysis of Tor". In *SP '05: Proceedings of the 2005 IEEE Symposium on Security and Privacy*, 183–195. 2005.

"NAACP v. Alabama". `http://supreme.justia.com/us/357/449/case.html`.

Narayanan, Arvind and Vitaly Shmatikov. "De-anonymizing Social Networks". *IEEE Security and Privacy* (2009).

"New York Times Co. v. United States". `http://www.law.cornell.edu/supct/html/historics/USSC_CR_0403_0713_ZO.html`.

O'Connor, John D. "I'm the Guy They Called Deep Throat". *Vanity Fair* (July 2005).

O'Harrow, Robert. *No Place to Hide*. New York: Free Press, 2006.

"OnGuard Online". http://www.onguardonline.gov/.

Palme, Jacob and Mikael Berglund. "Anonymity on the Internet". http://people.dsv.su.se/~jpalme/society/anonymity.html.

Partridge, C., T. Mendez, and W. Milliken. "Host Anycasting Service". *RFC*, no. 1546 (1993).

PBS. "frontline: cyber war!: the warnings?" http://www.pbs.org/wgbh/pages/frontline/shows/cyberwar/warnings/.

Poole, Martin. "meantime: non-consensual http user tracking using caches" (2000). http://sourcefrog.net/projects/meantime/.

Post, David G. *In Search of Jefferson's Moose: Notes on the State of Cyberspace*. Oxford: Oxford University Press, 2009.

Postel, Jon, ed. "Internet Protocol". *RFC*, no. 791 (1981).

———, ed. "Transmission Control Protocol". *RFC*, no. 793 (1981).

———, ed. "User Datagram Protocol". *RFC*, no. 768 (1980).

Public Proxy Servers. "Public Proxy Servers". http://www.publicproxyservers.com/.

Rabban, David M. *Free Speech in its Forgotten Years, 1870-1920*. Cambridge: Cambridge University Press, 1999.

Ratzloff, Matthew. "Detecting plugins in Internet Explorer (and a few hints for all the others)". April 26, 2007. http://www.builtfromsource.com/2007/06/26/detecting-plugins-in-internet-explorer-and-a-few-hints-for-all-the-others/.

Reid, Tim. "China's cyber army is preparing to march on America, says Pentagon". *The Times* (September 8, 2007).

Rekhter, Y., T. Li, and S. Hares. "A Border Gateway Protocol 4 (BGP-4)". *RFC*, no. 1771 (2006).

"Reno v. ACLU". http://www.law.cornell.edu/supct/html/96-511.ZO.html.

Risen, James and Eric Lichtblau. "Bush Lets U.S. Spy on Callers Without Courts". *The New York Times* (December 16, 2005).

Rivest, R. "The MD5 Message-Digest Algorithm". *RFC*, no. 1321 (1992).

Roberts, Hal, Ethan Zuckerman, and John Palfrey. "2007 Circumvention Landscape Report: Methods, Uses, and Tools". March 2009. http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/2007_Circumvention_Landscape.pdf.

"Schenck v. United States". http://www.law.cornell.edu/supct/html/historics/USSC_CR_0249_0047_ZO.html.

"Several countries trying to hack into US military system: Pentagon". *AFP* (September 3, 2007).

Spencer, Michael H. "Anonymous Internet Communication and the First Amendment: A Crack in the Dam of National Sovereignty". *Virginia Journal of Law and Technology* 1, no. 3 (Spring 1998).

"Statistics". http://en.wikipedia.org/wiki/Special:Statistics.

Stone, Brad. "A Call for Manners in the World of Nasty Blogs". *The New York Times* (April 9, 2007).

Sun Microsystems. "JDK 6 Security-related APIs & Developer Guides". http://java.sun.com/javase/6/docs/technotes/guides/security/index.html.

Sunstein, Cass. "Constitutional Caution". *University of Chicago Legal Forum* 1996 (1996).

"Talley v. California". http://supreme.justia.com/us/362/60/case.html.

Teich, Al et al. "Anonymous Communication Policies for the Internet: Results and Recommendations of the AAAS Conference". *The Information Society* 15, no. 2 (1999): 71–77.

"Telecommunications Act of 1996". http://www.fcc.gov/Reports/tcom1996.txt.

Tenebril, Inc. "GhostSurf". http://www.tenebril.com/consumer/ghostsurf/ghostsurf_standard.php.

"The Cloak". http://www.the-cloak.com.

"The National Strategy to Secure Cyberspace". http://www.dhs.gov/xlibrary/assets/National_Cyberspace_Strategy.pdf.

"TheOnionRouter/TorFAQ". http://wiki.noreply.org/noreply/TheOnionRouter/TorFAQ.

"Tor (anonymity network)". http://en.wikipedia.org/wiki/Tor_(anonymity_network).

"Tor: Overview". http://www.torproject.org/overview.html.en.

"Tor: Sponsors". http://www.torproject.org/sponsors.html.en.

"Torbutton FAQ". https://www.torproject.org/torbutton/faq.html.en.

"TorStatus - Tor Network Status". http://torstatus.kgprog.com/.

"traceroute". http://www.freebsd.org/cgi/man.cgi?query=traceroute.

United Nations General Assembly. "The Universal Declaration of Human Rights". December 10, 1948. http://www.un.org/Overview/rights.html.

United States Computer Emergency Readiness Team. "US-CERT Vulnerability Notes". http://www.kb.cert.org/vuls/.

United States Senate Select Committee to Study Governmental Operations With Respect to Intelligence Activities. "Supplementary Detailed Staff Reports On Intelligence Activities and the Rights of Americans, Book III". 1976. http://www.icdc.com/~paulwolf/cointelpro/churchfinalreportIIIj.htm.

"United States v. American Library Association". http://supct.law.cornell.edu/supct/html/02-361.ZO.html.

U.S. Naval Research Laboratory. "Onion Routing: History". http://www.onion-router.net/History.html.

Wallace, Jonathan D. "Nameless in Cyberspace: Anonymity on the Internet". *CATO Institute Briefing Papers*, no. 54 (1999).

Wang, Xiaoyun and Hongbo Yu. "How to break MD5 and other hash functions". In *EuroCrypt*. 2005.

Wayner, Peter. "Technology for Anonymity: Names by Other Nyms". *The Information Society* 15, no. 2 (1999): 91–97.

Web Hypertext Application Technology Working Group. "HTML 5 Draft Recommendation". April 4, 2009. http://www.whatwg.org/specs/web-apps/current-work/.

"WebKit". http://trac.webkit.org/browser.

Woodward, Bob. *The Secret Man: The Story of Watergate's Deep Throat*. New York, NY: Simon & Schuster, 2005.

Yee, Bennet et al. "Native Client: A Sandbox for Portable, Untrusted x86 Native Code". In *2009 IEEE Symposium on Security and Privacy*. 2009.

Yost, David S. "Debating security strategies". *NATO Review*, no. 4 (2003).

"Zeran v. America Online, Inc." http://www.law.emory.edu/4circuit/nov97/971523.p.html.

Zetter, Kim. "Embassy E-mail Account Vulnerability Exposes Passport Data and Official Business Matters". *Wired Threat Level* (August 31, 2007). http://blog.wired.com/27bstroke6/2007/08/embassy-e-mail-.html.

———. "Rogue Nodes Turn Tor Anonymizer Into Eavesdropper's Paradise". *Wired Threat Level* (September 10, 2007). http://www.wired.com/politics/security/news/2007/09/embassy_hacks.

Øverlier, Lasse and Paul Syverson. "Locating Hidden Servers". In *SP '06: Proceedings of the 2006 IEEE Symposium on Security and Privacy*, 100–114. 2006.