THE CONTENT/ENVELOPE DISTINCTION IN INTERNET LAW

MATTHEW J. TOKSON*

ABSTRACT

Whether a component of an Internet communication is classified as "content" or "envelope" information determines in large part the privacy protection it receives under constitutional and statutory law. Courts and Internet law scholars have yet to offer a means of determining the content/envelope status of unique aspects of Internet communications—from email subject lines to website URLs. As a result, data with the potential to expose every website, every Internet file downloaded, and every email sent by an Internet user may be unprotected under current law.

This Article develops a legal framework for distinguishing content from envelope information in unique areas of Internet communications. Drawing on a practical analysis of the structure of the Internet and an evaluation of relevant common law and Fourth Amendment doctrines, the Article proposes that electronic information that can reveal the underlying text or subject matter of an Internet communication must be classified as content. The Article identifies several areas in which application of this principle is necessary to resolve difficult questions about the legal status of an Internet communication, and gives, for the first time, a comprehensive account of the content status of Internet communications, from email body text to website IP addresses. The proposed framework provides a judicially manageable and normatively attractive means for courts to deter-

* Law clerk to the Honorable A. Raymond Randolph of the United States Court of Appeals for the District of Columbia Circuit and 2009 Kauffman Innovation Fellow at the University of Chicago Law School. Thanks to Abigail Abraham, Lisa Bernstein, Stephanos Bibas, Brian Callanan, Robert Helman, Orin Kerr, Judson Littleton, David Pozen, and Arpan Sura for advice and encouragement, and a special thanks to the members of the University of Chicago Legal Scholarship workshop for their helpful critiques of earlier drafts.

mine the legal status of novel communications technologies, present and future. Ultimately, resolving the content/envelope distinction has the potential to clarify other unanswered and controversial questions in Fourth Amendment and statutory privacy law.

TABLE OF CONTENTS

INTRODUCTION

The statutory framework governing the surveillance of electronic communications has been in place since 1986,[1] well before the advent of the World Wide Web and the popularization of the Internet. The relevant constitutional framework was established even earlier, in cases such as *Katz v. United States*[2] and *Smith v. Maryland*,[3] which applied the Fourth Amendment to government agents' interceptions of telephone conversations and dialed phone numbers. Meanwhile, web browsing and web-based email have been widely available to computer owners since 1995,[4] and a large proportion of U.S. households has had Internet access since 1997.[5] Both computer crimes (such as "hacking") and crimes involving the use of computers have increasingly become a major concern of law enforcement agencies.[6] Yet, after a decade of widespread Internet use and several years of prosecutions using Internet-based evidence

---

1. *See* Electronic Communications Privacy Act of 1986, Pub. L. No. 99-508, 100 Stat. 1848 (codified as amended in scattered sections of 18 U.S.C.).

2. 389 U.S. 347, 351-53 (1967).

3. 442 U.S. 735, 741-42 (1979).

4. Dave Crocker, Email History, http://www.livinginternet.com/e/ei.htm (last visited Mar. 11, 2009) (describing how email providers America Online and Delphi began to connect their email programs to the Internet in 1993); Windows History: Internet Explorer History, June 30, 2003, http://www.microsoft.com/windows/WinHistoryIE.mspx (last visited Mar. 11, 2009) (describing the marketing of the Internet Explorer web browser in 1995); *see also* Warren E. Agin & Scott N. Kumis, *A Framework for Understanding Electronic Information Transactions*, 15 ALB. L.J. SCI. & TECH. 277, 281 (2005) (describing how the development of the graphics-based web browser Mosaic in 1993 helped lead to the popularization of the Internet).

5. U.S. CENSUS BUREAU, HOME COMPUTERS AND INTERNET USE IN THE UNITED STATES: 2003, at 1 (2005), *available at* http://www.census.gov/prod/2005pubs/p23-208.pdf (reporting that 18 percent of U.S. homes had Internet access as of 1997, whereas 54.7 percent of homes had access in 2003). Today, the United States has roughly 220 million Internet users. Internet World Stats, Top 20 Countries with the Highest Number of Internet Users, http://www.Internetworldstats.com/top20.htm (last visited Mar. 11, 2009).

6. *See* COMPUTER SEC. INST., 2005 CSI/FBI COMPUTER CRIME AND SECURITY SURVEY (2005), *available at* http://www.cpppe.umd.edu/Bookstore/Documents/2005CSISurvey.pdf; David Finkelhor & Richard Ormrod, *Child Pornography: Patterns from NIBRS*, JUVENILE JUSTICE BULLETIN (U.S. Dep't of Justice, Office of Justice Programs), Dec. 2004, at 6, *available at* http://www.ncjrs.gov/pdffiles1/ojjdp/204911.pdf (reporting that 7 to 13 percent of child pornography crimes committed between 1997 and 2000 involved the use of a computer); Press Release, Fed. Bureau of Investigation, New FBI Computer Crime Survey (Jan. 18, 2006), *available at* http://www.fbi.gov/page2/jan06/computer_crime_survey011806.htm.

within a structure of well-established law, several fundamental questions regarding government surveillance of the Internet remain unanswered. For instance, the Fourth Amendment status of email content remains ambiguous, while the constitutional and statutory status of web surfing data is entirely unresolved.

This continuing uncertainty is largely the product of the unique characteristics of current electronic communications law. First, the difficulty of applying the Supreme Court's current Fourth Amendment precedents to modern communications technologies has likely motivated courts to avoid deciding such issues whenever possible.[7] Second, questions involving electronic communications statutes are infrequently litigated in criminal cases because the relevant statutes do not provide an exclusionary remedy for illegal government acquisitions of electronic data.[8] Yet, as both Internet use and the government's surveillance of such use become more pervasive, courts have finally been forced to grapple with some of the difficult questions surrounding the legal protection afforded to Internet communications.[9]

Perhaps the most practically significant of these unresolved questions is whether novel categories of Internet communications data, such as email subject lines,[10] website Uniform Resource Locators (URLs),[11] and website IP addresses,[12] should be protected

---

7. *See* Susan Freiwald, *First Principles of Communications Privacy*, 2007 STAN. TECH. L. REV. 3, ¶ 34, *available at* http://stlr.stanford.edu/pdf/freiwald-first-principles.pdf. For examples of courts avoiding Fourth Amendment questions and resolving cases on alternative, arguably specious grounds, see Warshak v. United States, 532 F.3d 521, 525-26 (6th Cir. 2008) (en banc) (avoiding the question of whether there was a reasonable expectation of privacy in email content by finding that the plaintiff's case was unripe); United States v. Bach, 310 F.3d 1063, 1066 (8th Cir. 2002) (declining to decide whether there was a reasonable expectation of privacy in email content by holding that a warrant served without police presence was reasonably executed); United States v. Ferguson, 508 F. Supp. 2d 7, 9-10 (D.D.C. 2007) (concluding that, because an officer's reliance on the Stored Communications Act (SCA) was reasonable and because the Act did not provide for a suppression remedy, the court "need not consider the constitutionality of the SCA").

8. *See* Orin S. Kerr, *Lifting the "Fog" of Internet Surveillance: How a Suppression Remedy Would Change Computer Crime Law*, 54 HASTINGS L.J. 805, 823-24 (2003).

9. *See, e.g.*, United States v. Forrester, 495 F.3d 1041 (9th Cir. 2007); Warshak v. United States, 490 F.3d 455 (6th Cir. 2007); *In re* Application of the U.S. for an Order Authorizing the Use of a Pen Register, 396 F. Supp. 2d 45 (D. Mass. 2005).

10. *See* discussion *infra* Part II.E.

11. *See* discussion *infra* Part III.

12. *See* discussion *infra* Part III.D.

as the contents of electronic communications, or whether they should be treated as noncontent "envelope" information. The question is a profoundly important one because the legal protection afforded the two types of information is dramatically different in both constitutional and statutory law. Further, due in large part to the confusion over content status, electronic data with the potential to expose the websites visited and email messages sent by a web user may be unprotected from private or government intrusion under current law. Although Internet law scholars and commentators have mentioned the importance and the troublesome complexity of the issue,[13] none have put forth a theory as to which aspects of a communication are content and which are not under constitutional or statutory law.

Obviously, simply identifying the nature of this complex and unresolved area of law is not enough. This Article attempts to develop a conceptual framework of electronic communications content that will allow courts to determine whether even the most novel forms of Internet communications information are content or noncontent as a matter of law. Part I of the Article describes the central importance of the content/noncontent distinction in the constitutional and statutory law of Internet surveillance. Part II examines the content status of relatively well-understood Internet communications information such as email addresses and Internet subscriber information. Part III discusses the unresolved content status of novel forms of communications information such as web surfing data and develops a framework for classifying such information as content or noncontent. Part IV analyzes the implications of the proposed content/noncontent framework for the constitutional and statutory protection afforded to new forms of Internet communications content. It then describes how a clearer conception of content in such communications will leave courts in a better position to confront other complex questions of Internet surveillance law.

13. *See infra* note 82 and accompanying text.

I. THE IMPORTANCE OF THE CONTENT/NONCONTENT DISTINCTION IN INTERNET SURVEILLANCE LAW

For over one hundred years, the Supreme Court has held that the Fourth Amendment protects mailed letters and packages from inspection by postal authorities or other government agents.[14] Yet from the start, the Court has distinguished between the content of a letter and the noncontent information disclosed on its envelope. Whereas noncontent envelope information is exposed and can be examined by anyone, the content of a letter is "as fully guarded from examination and inspection" as it would be if the party mailing the letter had retained it in his or her own home.[15]

The content/noncontent distinction remains important in the constitutional and statutory law governing the inspection of private communications, even as new technologies have dramatically altered the nature of communication itself. Of course, one could question the application of the content/noncontent distinction, which was developed in the unique context of paper mail, to new forms of communications such as the telephone and the Internet. Whether as a normative matter the distinction is insufficiently protective of a privacy interest in the circumstances of one's communications is a question for another day.[16] In any event, the distinction is firmly established in communications surveillance law, and any attempt to dislodge it would likely be quixotic. There is, at least, a good argument that the content/noncontent distinction captures a qualitative difference in the intimacy of different types of communications information: while the lonely traveler might consider the fact that

14. *See, e.g.*, Walter v. United States, 447 U.S. 649, 654 (1980); *Ex parte* Jackson, 96 U.S. 727, 733 (1877). The court has also confirmed that sealed packages given to private carriers are Fourth Amendment "effects" in which the public has a "legitimate expectation of privacy" vis-á-vis the government. United States v. Jacobsen, 466 U.S. 109, 114 (1984). However, "[c]ommon carriers have a common-law right to inspect packages they accept for shipment, based on their duty to refrain from carrying contraband." Illinois v. Andreas, 463 U.S. 765, 769 n.1 (1983). The same principle applies to regular mail. *See Warshak*, 490 F.3d at 474.

15. *Jackson*, 96 U.S. at 733.

16. A forthcoming paper argues in great detail for the normative desirability of the content/noncontent distinction. Orin S. Kerr, *Applying the Fourth Amendment to Internet Communications: A General Approach*, 61 STAN. L. REV. (forthcoming 2009) (manuscript at 4), *available at* http://papers.ssrn.com/so13/papers.cfm?abstract_id=1348322.

he called his wife at midnight from his hotel to be private, surely it is the conversation itself that he considers most intimate and confidential.

Indeed, the content of telephone calls is protected by the Fourth Amendment,[17] whereas phone numbers (which are exposed to the phone company and involve only noncontent information) are not.[18] Statutorily, interception of the content of a telephone call is governed by the Wiretap Act,[19] which sets rigorous standards for court orders permitting government wiretaps and provides for the exclusion of evidence derived from communications intercepted in violation of the Act.[20] By contrast, the interception of noncontent telephone numbers is governed by the Pen Register Act,[21] which mandates court approval of surveillance if the government certifies that the information likely to be obtained is "relevant to an ongoing criminal investigation" and which provides no exclusionary rule.[22] The content/noncontent distinction is equally central to the constitutional and statutory framework that governs state surveillance of Internet communications.

A. The Content/Noncontent Distinction in the Fourth Amendment Law of Internet Communications

The Supreme Court has determined that both tangible things (papers, effects) and intangible things (such as the sounds of a telephone conversation) can be searched within the meaning of the Fourth Amendment.[23] A search occurs wherever the government

---

17. Katz v. United States, 389 U.S. 347, 351-53 (1967).
18. Smith v. Maryland, 442 U.S. 735, 741-42 (1979).
19. 18 U.S.C. §§ 2510-2522 (2006).
20. *Id.* § 2518.
21. 18 U.S.C. §§ 3121-3127 (2006).
22. *Id.* § 3123(a), (c).
23. Katz v. United States, 389 U.S. 347, 351-52 (1967). The Fourth Amendment states that:

> The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

U.S. CONST. amend. IV.

infringes upon a person's reasonable "expectation of privacy."[24] In order for the Fourth Amendment to apply, a two-pronged test must be met: (1) the person must have an actual, subjective expectation of privacy, and (2) the expectation must be one that society is prepared to recognize as objectively reasonable.[25]

Whether Internet users have a reasonable expectation of privacy in their emails and web surfing data is largely unresolved. Unlike traditional letters, emails and web surfing communications are often copied in transit by Internet Service Providers (ISPs) and are (in theory) easily accessed by ISP employees.[26] Because emails and other forms of Internet communications are arguably exposed to third parties during transmission, it remains controversial whether the Fourth Amendment protections that apply to the contents of letters and telephone calls can apply to them.[27] Yet, an emerging body of case law suggests that the content/noncontent distinction is crucial in determining whether Internet communications are protected by the Fourth Amendment.

Most cases decided since the popularization of the Internet have dealt with the unique situations of public chat groups or employees

---

24. *Katz*, 389 U.S. at 361 (Harlan, J., concurring); *see, e.g.*, Terry v. Ohio, 392 U.S. 1, 9 (1968) ("[W]herever an individual may harbor a reasonable 'expectation of privacy' ... he is entitled to be free from unreasonable governmental intrusion.").

25. *Katz*, 389 U.S. at 361; *see, e.g.*, California v. Greenwood, 486 U.S. 35, 40 (1988). Note that the second prong is generally the controlling factor, and Fourth Amendment protection might be found even where an individual lacks a subjective expectation of privacy. *See* Hudson v. Palmer, 468 U.S. 517, 525 n.7 (1984). For example, when an individual's subjective expectation is eliminated by the announcement of a normatively unacceptable means of government intrusion (such as a policy of warrantless home searches), a normative inquiry based solely on the second prong would be proper. Smith v. Maryland, 442 U.S. 735, 740 n.5 (1979).

26. *See In re* DoubleClick Privacy Litig., 154 F. Supp. 2d 497, 503-04 (S.D.N.Y. 2001) (describing how a website recorded URL data); PRESTON GRALLA, HOW THE INTERNET WORKS 87 (2001) (describing how emails are transmitted); Deirdre K. Mulligan, *Reasonable Expectations in Electronic Communications: A Critical Perspective on the Electronic Communications Privacy Act*, 72 GEO. WASH. L. REV. 1557, 1563 (2004) (describing email transmission and noting that unencrypted email messages could in theory be read by ISP employees); Ryan Singel, *Which ISPs Are Spying on You?*, WIRED, May 5, 2007, *available at* http://www.wired.com/politics/onlinerights/news/2007/05/isp_privacy (suggesting that many ISPs record URL data); *see also* Kerr, *supra* note 8, at 812-16 (suggesting that the transmission process may eliminate any Fourth Amendment expectation of privacy in email). Also, emails are often scanned by anti-spam software. Warshak v. United States, 490 F.3d 455, 474 (6th Cir. 2007).

27. *See* discussion *infra* Part IV.

using at-work computer systems that were regularly monitored by their employers.[28] Finally, in the 2007 case *Warshak v. United States*, a panel of the Sixth Circuit Court of Appeals held that the Fourth Amendment applies to the contents of emails.[29] In so holding, it placed a great deal of weight on the content/noncontent distinction, stating that the Supreme Court's "combined precedents of *Katz* and *Smith*" applying the Fourth Amendment to telephone calls "recognize a heightened protection for the *content* of the communications."[30] The court found a strong "content-based privacy interest" in a user's emails, one that cannot be eliminated in the absence of total access by a third party to their contents.[31] However, the Sixth Circuit subsequently granted a petition for rehearing en banc in *Warshak* and ultimately avoided the constitutional issue, vacating the decision on the ground that the case was unripe.[32] Yet, despite the en banc judgment, the arguments put forth in the panel opinion are likely to surface again as other circuits are forced to rule on the Fourth Amendment status of Internet communications, either in response to facial challenges or to as-applied challenges arising in criminal cases or *Bivens* suits.[33] The *Warshak* panel's holding suggests a likelihood that at least some circuits will find a

28. *See* United States v. Simons, 206 F.3d 392, 398 (4th Cir. 2000) (holding that an employee had no reasonable expectation of privacy in records of his Internet use when his employer's policy was to monitor computer use); United States. v. Charbonneau, 979 F. Supp. 1177, 1184 (S.D. Ohio 1997) (finding no expectation of privacy in conversations posted to public chat rooms); United States v. Monroe, 52 M.J. 326, 330 (C.A.A.F. 2000) (finding no expectation of privacy in email messages stored on a monitored, government-owned computer system).

Prior to 2007, one court found that the Fourth Amendment protected the content of email sent via the private ISP America Online. *See* United States v. Maxwell, 45 M.J. 406, 417 (C.A.A.F. 1996). The court relied largely upon America Online's policy of not reading or disclosing users' email. *Id.* Its value as precedent is probably limited due to its status as a military case. *See* United States v. Hambrick, 55 F. Supp. 2d 504, 508 (W.D. Va. 1999).

29. 490 F.3d 455, 470-71 (6th Cir. 2007), *vacated en banc*, 532 F.3d 521 (6th Cir. 2008).

30. *Id.* at 471.

31. *Id.* at 474; *see also id.* at 476 (requiring that the government show "complete access" to email content in order to demonstrate the lack of a reasonable expectation of privacy).

32. Warshak v. United States, 532 F.3d 521, 526, 529-30 (6th Cir. 2008) (en banc) (finding that Warshak, who had sought an injunction against the government, was unlikely to be the target of future email searches and therefore did not present an Article III case or controversy).

33. *See id.* at 531-32 (describing how Warshak could have obtained a ruling on the constitutional issue through a *Bivens* action or potentially through a motion to suppress evidence in his criminal trial).

constitutional expectation of privacy in email content. In fact, in a recent case dealing primarily with text messages, the Ninth Circuit indicated that email content should receive the same constitutional protections as the content of letters.[34]

By contrast, the few cases that have dealt with noncontent information related to Internet communications strongly suggest that there is no Fourth Amendment protection for such information. For instance, in *United States v. Hambrick*, the Fourth Circuit held that there was no Fourth Amendment protection for the subscriber information that a user submitted to an ISP in order to set up an email account, relying on the fact that noncontent information is not protected under *Smith*.[35] Further, in the recent case *United States v. Forrester*, the Ninth Circuit directly held that there was no Fourth Amendment interest in the to/from addresses of emails, Internet Protocol (IP) addresses of websites, or the total volume of file transfers associated with an Internet user's account.[36] The *Forrester* court also based its judgment on the significant distinction in *Smith* between the content and noncontent information of communications.[37] According to the court, surveillance techniques that capture IP addresses relating to Internet communications were constitutionally indistinguishable from the pen registers approved in *Smith*, because the IP addresses obtained were no more private or intimate than phone numbers.[38] The opinion reflects the current

34. Quon v. Arch Wireless Operating Co., 529 F.3d 892, 905 (9th Cir. 2008) (finding a reasonable expectation of privacy in the content of text messages stored electronically by an electronic communication service, and analogizing such messages to email content).

35. No. 99-4793, 2000 WL 1062039, at **3-4 (4th Cir. Aug. 3, 2000); *see also* United States v. Kennedy, 81 F. Supp. 2d 1103, 1110 (D. Kan. 2000) (holding that the defendant had no reasonable expectation of privacy in the subscriber information associated with his IP address, which was already known to the government).

36. 495 F.3d 1041, 1048-49 (9th Cir. 2007).

37. *Id.*

38. *Id.*; *see also In re* Application of the U.S. for an Order Authorizing the Use of a Pen Register, 396 F. Supp. 2d 45, 48 (D. Mass. 2005) (stating that pen register surveillance of website IP address communications, if carried out properly, would be "no problem" under electronic communications statutes).

The court acknowledged in a footnote that surveillance techniques that reveal not only website IP addresses but also Uniform Resource Locators (URLs) of the specific pages visited "might be more constitutionally problematic," though it did not need to decide the issue. *Forrester*, 495 F.3d at 1049 n.6. URLs are the website addresses that direct users to a specific page on a website. *Id. See generally* Marshall Brain, How Web Servers Work, http://computer.howstuffworks.com/web-server.htm (last visited Mar. 11, 2009).

lack of constitutional protection for anything but clearly defined Internet content.

Looking forward, it remains difficult to predict whether the content/noncontent distinction will remain the central determinant of constitutional protection for email and website communications. As discussed below in Part IV.A, the content/noncontent distinction was just one of several rationales given in *Smith* for finding no reasonable expectation of privacy in telephone number information. *Smith* also analogizes phone number information disclosed to the phone company and appearing on monthly bills to information disclosed to a third party (such as a government informant).[39] This "third party doctrine," if applied to email content and/or URLs, may dictate that there is no reasonable expectation of privacy even in the content of emails, which may be considered "disclosed" to third party ISPs. However, despite the persistent ambiguity in the law of Internet communications, the few courts that have thus far examined the constitutional status of email and web surfing activity have all relied heavily and often exclusively on the content/noncontent distinction in reaching their conclusions.

B. The Content/Noncontent Distinction in the Electronic Communications Privacy Act

The distinction between content and noncontent information is critical in determining the level of statutory protection provided to Internet communications.[40] The Electronic Communications Privacy Act of 1986 (ECPA), which provides the statutory framework for government surveillance of Internet communications, offers far less protection to noncontent information than content information in a variety of surveillance contexts.[41] The Act includes three sections—the Wiretap Act,[42] which governs the interception of

39. Smith v. Maryland, 442 U.S. 735, 743-44 (1979).

40. *See, e.g.*, Susan Freiwald, *Online Surveillance: Remembering the Lessons of the Wiretap Act*, 56 ALA. L. REV. 9, 48 (2004) (stating that the statutory protection for noncontent information "could not differ more" from that provided for content information).

41. Pub. L. No. 99-508, 100 Stat. 1848 (codified as amended in scattered sections of 18 U.S.C.).

42. 18 U.S.C. §§ 2510-2522 (2006).

communications content in transit; the Pen Register Act,[43] which governs the interception of noncontent communications attributes such as phone numbers and email to/from information; and the Stored Communications Act,[44] which regulates retrospective access to communications held in electronic storage. The following overview of the ECPA is not intended to be comprehensive;[45] instead, it outlines the general function of the relevant statutes and highlights the significance of the content/noncontent distinction in determining the level of privacy protection provided by the Act.

1. The Wiretap Act—Interception of Communication Content

The Wiretap Act within the ECPA evolved from previous legislation passed in 1968 regulating government and private wiretapping or bugging of "wire" or "oral" communications.[46] In 1986, the Wiretap Act was amended to include "electronic communications," defined as any "transfer of signs, signals, writing, images, sounds, data, or intelligence of any nature transmitted in whole or in part by a wire, radio, electromagnetic, photoelectronic or photooptical system" which is not a wire or oral communication.[47] The Wiretap Act imposes significant criminal penalties on anyone who intercepts such communications (a minimum ten thousand dollar fine per violation and up to five years imprisonment),[48] though it provides no exclusionary rule for evidence derived from electronic communications. It also requires the government to obtain a "super-warrant" court order, including a showing that other investigative procedures have failed or are likely to fail, a complete statement under oath of facts and circumstances sufficient to justify a belief that the order

43. 18 U.S.C. §§ 3121-3127 (2006).

44. 18 U.S.C. §§ 2701-2711 (2006).

45. For a detailed overview of the statutes governing electronic communications surveillance, see for example, Patricia L. Bellia, *Surveillance Law Through Cyberlaw's Lens*, 72 GEO. WASH. L. REV. 1375 (2004); Mulligan, *supra* note 26; Daniel J. Solove, *Digital Dossiers and the Dissipation of Fourth Amendment Privacy*, 75 S. CAL. L. REV. 1083 (2002).

46. *See* Omnibus Crime Control and Safe Streets Act of 1968, Pub. L. No. 90-351, 82 Stat. 197, 211-25 (codified as amended at 18 U.S.C. §§ 2510-2522). Wire and oral communications are telephone conversations or private face-to-face conversations that contain the human voice. *See* 18 U.S.C. § 2510(1)-(2).

47. 18 U.S.C. § 2510(12).

48. *See id.* §§ 2511, 2520.

should issue, a particular description of the details of the suspected crime, and a particularized description of the communications sought.[49]

The content/noncontent distinction in the Wiretap Act is found in the definition of "intercept," which refers to the "acquisition of the *contents* of any ... electronic ... communication."[50] "Contents" appears to be defined fairly broadly in the ECPA,[51] as it consists of "any information concerning the substance, purport, or meaning of [a] communication."[52]

The body of an email is considered content under the Wiretap Act, and thus capable of interception.[53] However, email content, which is generally copied by ISP servers in the course of transmission to the recipient's ISP,[54] is highly unlikely to be intercepted under the Wiretap Act. Any "intercept" under the ECPA must occur contemporaneously with transmission,[55] and courts applying the Wiretap Act to the acquisition of emails have concluded that even email stored temporarily on an ISP's servers is in storage and not in transmission within the meaning of the Act.[56] As a result of this interpretation, the government has little motive to capture emails during the fraction of a second when they are transmitted to or from the ISP. By waiting until they are in storage on the ISP server, the government can acquire the contents under the less stringent standards of the Stored Communications Act.[57]

---

49. *Id.* § 2518(1)(b)-(c); *see, e.g.*, Orin S. Kerr, *Internet Surveillance Law After the USA Patriot Act: The Big Brother That Isn't*, 97 NW. U. L. REV. 607, 630 (2003) (discussing "super-warrant" court orders).

50. 18 U.S.C. § 2510(4) (emphasis added).

51. *See* discussion *infra* Part IV.B.

52. 18 U.S.C. § 2510(8).

53. *See, e.g.*, Steve Jackson Games, Inc. v. U.S. Secret Serv., 36 F.3d 457, 460 (5th Cir. 1994).

54. *See* GRALLA, *supra* note 26, at 89; *see also* H.R. REP. NO. 99-647, at 22 (1986) (discussing ISP copying of users' emails).

55. *See* United States v. Turk, 526 F.2d 654, 658 (5th Cir. 1976).

56. *See, e.g.*, Fraser v. Nationwide Mut. Ins. Co., 352 F.3d 107, 113 (3d Cir. 2003); *Steve Jackson Games*, 36 F.3d at 461-62.

57. However, if URLs or website IP addresses were deemed to be "content" for the purposes of the ECPA, a device set up to capture them as they are transmitted would actually be intercepting them under the Wiretap Act. *See* IIT RESEARCH INST., INDEPENDENT TECHNICAL REVIEW OF THE CARNIVORE SYSTEM: FINAL REPORT, at C-5 (2000), *available at* http://epic.org/privacy/carnivore/carniv_final.pdf [hereinafter IITRI REPORT] (describing how the government's Carnivore pen register surveillance software captures IP addresses and is

2. The Pen Register Act—Interception of Noncontent Communication Attributes

Created in 1986, the Pen Register Act initially applied to devices that recorded outgoing and incoming telephone numbers on a particular telephone line. A "pen register" records outgoing numbers, whereas a "trap and trace device" records the numbers of persons calling the monitored line.[58] The USA PATRIOT Act amended the definition of pen registers and trap and trace devices to include any "device or process" that records "dialing, routing, addressing, or signaling information" (DRAS information), other than content information, associated with an electronic communication.[59]

The Pen Register Act provides no exclusionary rule, and provides for less significant criminal penalties than those in the Wiretap Act.[60] Instead of the Wiretap Act's heightened requirements, the Pen Register Act mandates that courts "shall authorize" pen registers if the government applicant certifies that the information likely to be obtained is "relevant to an ongoing ... investigation."[61] Courts generally do not challenge or investigate the relevance certification, and thus judicial review of pen register applicants is virtually nonexistent in practice.[62] The difference in the standards for court approval of content-capturing wiretaps and noncontent-capturing pen registers is dramatic—content information is protected by a "super-warrant," noncontent information by a rubber stamp.

---

capable of capturing URLs); Kerr, *supra* note 49, at 645 n.180 (describing how Carnivore's successor, DCS-1000, is also capable of capturing URLs). The interception would be subject to the stringent regulations that apply to wiretaps, 18 U.S.C. §§ 2510-2522 (2006), rather than the relatively lax standards governing the acquisition of noncontent information under the Pen Register.

58. *See* 18 U.S.C. § 3127(3)-(4) (2006).

59. *Id.* ("[Captured information] shall not include the contents of any communication.").

60. *Id.* § 3121(d) (mandating fines "under this title" or imprisonment for not more than one year for violations).

61. *Id.* § 3123(a), (c).

62. Freiwald, *supra* note 40, at 62.

3. The Stored Communications Act—Retrospective Surveillance of Content and Noncontent Information

The Stored Communications Act (SCA) protects emails and other electronic communications that are not in the process of transmission.[63] The SCA also provides for the disclosure of "record[s] or other information," not including content, pertaining to a customer of computing services.[64] The SCA has no exclusionary rule, and its penalties are less severe than those of the Wiretap or Pen Register Acts—minimum fines of one thousand dollars and up to six months imprisonment.[65] The standard for court orders differs according to the duration of electronic storage and whether the information obtained is content or noncontent. At the highest level, ordering an ISP to turn over the contents of electronic communications stored for 180 days or less requires a standard search warrant.[66] Communications contents stored for more than 180 days can be obtained either with a standard warrant or a subpoena (or a § 2703(d) court order) that must be coupled in most cases with prior notice to the ISP subscriber.[67] A subscriber who has been notified that his personal information has been subpoenaed would likely have the opportunity to challenge the subpoena on the grounds of irrelevance, improper purpose, or procedural flaws.[68] Numerous courts have recently held that a privacy interest inherent in many personal records (such as credit card or employment personnel records) allows the subject of the records to challenge subpoenas issued to third parties.[69]

---

63. *See supra* text accompanying notes 52-56; *see also* 18 U.S.C. §§ 2701-2712 (2006).

64. *Id.* § 2703(c).

65. *Id.* § 2701(b).

66. *Id.* § 2703(a).

67. *Id.* § 2703(b); *see id.* § 2705 (providing that notice may be delayed for ninety days in order to avoid "endangering the life or physical safety of an individual; ... flight from prosecution; ... destruction of or tampering with evidence; ... intimidation of potential witnesses; or ... otherwise seriously jeopardizing an investigation or unduly delaying a trial").

68. *See* United States v. Powell, 379 U.S. 48, 57-58 (1964) (describing the criteria that determine the validity of a subpoena).

69. *See* 9A CHARLES ALAN WRIGHT ET AL., FEDERAL PRACTICE AND PROCEDURE § 2459 n.7 (3d ed. 2008) (listing cases that recognize the right to challenge such subpoenas); *see also* Reserve Solutions, Inc. v. Vernaglia, No. 05 Civ. 8622 VM RLE, 2006 WL 1788299, at *1 (S.D.N.Y. June 26, 2006) (finding a personal privacy right in credit card records). *See generally* Reisman v. Caplin, 375 U.S. 440, 445 (1964) (stating that interested third parties may attack

At the lowest level, only a subpoena is required to compel an ISP to disclose basic noncontent subscriber information, including name, address, records of session times, length and type of subscription, telephone number or network address, and source of payment including credit card number.[70] A somewhat higher level of protection is granted to "other" noncontent records pertaining to the subscriber, a category that generally covers all transactional information (such as phone usage records or records of email headers) other than basic subscriber information.[71] For these records, the government must generally obtain a § 2730(d) court order, which can be issued only if the government applicant provides "specific and articulable facts" demonstrating "reasonable grounds to believe that the [records] are relevant and material to an ongoing criminal investigation."[72] Ironically, this standard is significantly higher than the standard governing Pen Register Act intercept orders, which generally provides no judicial review and requires no showing of specific and articulable facts.[73] The "reasonable grounds for relevance" standard is lower than probable cause (and is akin to the general relevance standard for subpoenas),[74] but does provide some degree of judicial scrutiny for noncontent record requests.[75] Still, the standards for obtaining content in the ECPA, even content stored for more than 180 days, are substantially higher than those for noncontent. The *lowest* standard for obtaining stored content still requires notice to the subscriber (and the corresponding opportunity to challenge the surveillance); on the other hand, the *highest* standard for noncontent information does not require notice to the

summonses before a district court and applying that rule to administrative hearings). Note also that when a subscriber is notified that his or her records have been subpoenaed by the government, the possibility of negative publicity for the ISP holding the records may incentivize the ISP to bring its own challenge to the subpoena.

70. *See* 18 U.S.C. § 2703(c)(2).

71. *See* COMPUTER CRIME & INTELLECTUAL PROP. SECTION, U.S. DEP'T OF JUSTICE, SEARCHING AND SEIZING COMPUTER AND OBTAINING ELECTRONIC EVIDENCE IN CRIMINAL INVESTIGATIONS 90-91 (2002), *available at* http://www.usdoj.gov/criminal/cybercrime/s&smanual2002.htm [hereinafter CCIPS MANUAL]; Mulligan, *supra* note 26, at 1567.

72. 18 U.S.C. § 2703(d).

73. *See supra* notes 61-62 and accompanying text.

74. *See* United States v. Powell, 379 U.S. 48, 57-58 (1964).

75. *See* Anthony E. Orr, Note, *Marking Carnivore's Territory: Rethinking Pen Registers on the Internet*, 8 MICH. TELECOMM. & TECH. L. REV. 219, 236 (2002).

subscriber whose records are being observed.[76] The statutory protection afforded to electronic communication information depends in large part on whether the information is classified as content or noncontent information under the ECPA.

II. CONTENT AND NONCONTENT IN INTERNET COMMUNICATIONS

As discussed in Part I, whether communications information is considered content or noncontent (attribute) information is perhaps the most important determinant of the constitutional and statutory protection which that information receives. Yet, whether Internet communications should be classified as content or noncontent remains, outside of certain well-defined categories (for instance, the body of an email), a source of controversy and confusion among legal scholars. The greatest of the content controversies has arisen over the content status of web surfing information—the IP addresses of websites and the URL addresses of the individual pages viewed. Advocates of broad privacy protection have expressed concern that URLs may reveal intimate personal information about web users,[77] but have largely focused on critiquing *Smith v. Maryland*'s content/noncontent distinction[78] and the "third party doctrine" in general.[79] To them, the fact that URL data might be easily obtained by the government is another piece of evidence that cases like *Smith* are insufficiently protective of citizens' privacy interests.[80] Yet by

76. *See supra* notes 66-70 and accompanying text.

77. *See* Daniel J. Solove, *Reconstructing Electronic Surveillance Law*, 72 GEO. WASH. L. REV. 1264, 1287 (2004).

78. *See, e.g.*, *id.* at 1286-88.

79. For the seminal third party doctrine case, also frequently criticized by advocates of broader privacy protection, see United States v. Miller, 425 U.S. 435, 442 (1976) (holding that a defendant had no reasonable expectation of privacy in banking records because they had already been disclosed to the bank's employees in the ordinary course of business).

80. *See* Matthew D. Lawless, *The Third Party Doctrine Redux: Internet Search Records and the Case for a "Crazy Quilt" of Fourth Amendment Protection*, 2007 UCLA J.L. & TECH. 2, ¶ 21; Gavin Skok, *Establishing a Legitimate Expectation of Privacy in Clickstream Data*, 6 MICH. TELECOMM. TECH. & L. REV. 61, 82-83 (2000); Solove, *supra* note 77, at 1288; Solove, *supra* note 45, at 1137-38, 1156; Jayni Foley, Note, *Are Google Searches Private? An Originalist Interpretation of the Fourth Amendment in Online Communication Cases*, 22 BERKELEY TECH. L.J. 447, 468-70 (2007); Peter J. Georgiton, Note, *The FBI's Carnivore: How Federal Agents May Be Viewing Your Personal Email and Why There Is Nothing You Can Do About It*, 62 OHIO ST. L.J. 1831, 1845-46 (2001).

jumping to this normative conclusion, these scholars are inadvertently weakening their case for stronger Internet privacy protection by failing to recognize the breadth of Internet communications data that should be classified as content under constitutional and statutory law. Meanwhile, supporters of a bright line definition of content (and also many neutral observers) largely have not addressed the content status of web surfing information in any detail; often, they simply assume that URLs and IP addresses should be treated as noncontent data or assume that per *Smith* they are obviously not protected by the Fourth Amendment, no matter what their content status.[81] Probably as a result of these normative biases, we lack a robust conceptual framework for determining whether new forms of communications information, such as web surfing data, should be classified as content or noncontent. Or perhaps it is simply because determining whether web surfing "communications" are content or not—and sorting out what that would mean in terms of the Fourth Amendment and the ECPA—presents a complex legal and technical question.[82]

Determining whether different types of Internet communications information are content requires decoupling the question of content/noncontent status from the question of whether the information is protected under *Smith*. As discussed below, *Smith* may foreclose a reasonable expectation of privacy in even intimate content if the

---

81. *See* Christopher Slogobin, *Transaction Surveillance by the Government*, 75 MISS. L.J. 139, 153 (2005) (stating that courts are likely to find no reasonable expectation of privacy in web surfing data); Rich Haglund, Note, *Applying Pen Register and Trap and Trace Devices to Internet Communications: As Technology Changes, Is Congress or the Supreme Court Best-Suited To Protect Fourth Amendment Expectations of Privacy?*, 5 VAND. J. ENT. L. & PRAC. 137, 141-42 (2003) (arguing that government software programs that collect web surfing data do not violate any reasonable expectation of privacy); Christian David Hammel Schultz, Note, *Unrestricted Federal Agent: "Carnivore" and the Need To Revise the Pen Register Statute*, 76 NOTRE DAME L. REV. 1215, 1241-42 (2001) (arguing that *Smith* dictates that there is no expectation of privacy in web surfing data disclosed to a third party ISP); Brian I. Simon, Note, *The Tangled Web We Weave: The Internet and Standing Under the Fourth Amendment*, 21 NOVA L. REV. 941, 967 (1997) (analogizing web surfing to traveling public streets and arguing that there is no expectation of privacy in Internet browsing); *see also* Brief for Professor Orin S. Kerr as Amicus Curiae Supporting Appellant at 6, United States v. Bach, 310 F.3d 1063 (8th Cir. 2002) (No. 02-2138) (making a similar argument that email content, like a postcard, is not protected by the Fourth Amendment).

82. *See, e.g.*, Kerr, *supra* note 49, at 645; David McPhie, *Almost Private: Pen Registers, Packet Sniffers, and Privacy at the Margin*, 2005 STAN. TECH. L. REV. 1, ¶ 26; Solove, *supra* note 77, at 1287.

user is found to have disclosed the content to the carrier of the communication (the ISP).[83] But conflating *Smith*'s analysis of the content/noncontent distinction in telephone calls with its analysis of a reasonable expectation of privacy in such calls risks obscuring the question of what "content" actually is.

The following discussion seeks to develop a conceptual framework capable of determining which types of Internet communications information are content and which are noncontent. The inquiry focuses on the semantic and/or common law definition of "content," presumably the relevant definition for the word as it is used in *Smith*[84] and the lower court Fourth Amendment cases that focus on the content/noncontent distinction in Internet communications.[85] Note that although developing this framework will help determine whether a type of Internet communication represents "contents" under the ECPA,[86] the statutory inquiry is technically separate because "contents" and other relevant terms are separately defined in the Act.

As a general semantic matter, the definition of "content" (and "contents") has remained largely the same since before *Smith* was decided. Indeed, according to the Oxford English Dictionary, it has remained largely the same since the early sixteenth century.[87] Commonly, the first meaning of "content" is "that which is contained" in something; the second meaning is the "subject-matter" of a speech or piece of writing; and the third meaning is the "sum or substance of what is contained in a document; tenor, purport."[88] As we can see, content is generally defined as not only the actual written words of a document, but also the general subject matter of

83. *See infra* Part IV.A.

84. Smith v. Maryland, 442 U.S. 735, 741 (1979) ("Yet a pen register differs significantly from the listening device employed in *Katz*, for pen registers do not acquire the *contents* of communications.").

85. *See, e.g.*, Warshak v. United States, 490 F.3d 455, 471 (6th Cir. 2007); United States v. Hambrick, No. 99-4793, 2000 WL 1062039, at **3-4 (4th Cir. Aug. 3, 2000).

86. *See* discussion *infra* Part IV.B.

87. III OXFORD ENGLISH DICTIONARY 815 (J.A. Simpson & E.S.C. Weiner eds., 2d ed. 1989).

88. *Id.* The meanings are the same in dictionaries published at the time of the *Smith* decision. S*ee, e.g.*, I SHORTER OXFORD ENGLISH DICTIONARY 411 (William Little et al. eds., 3d ed. 1977); WEBSTER'S NEW WORLD DICTIONARY OF AMERICAN LANGUAGE 307 (David B. Guralnik ed., 2d College ed. 1970). They remain the same today. S*ee, e.g.*, ENCARTA WEBSTER'S DICTIONARY OF THE ENGLISH LANGUAGE 413 (2d ed. 2004).

the document and the purport of its message. Applied to the email context, a very narrow definition of "contents" might refer to only the actual letters contained in the body of the email; a more expansive definition—like the one used in the ECPA[89]—would include the overall gist of the message contained, or even the general subject matter discussed. Although the distinction between these narrow and broad definitions of content may be useful in determining the status of web surfing information, it is ultimately not central to determining whether most other forms of Internet communication information are content or noncontent as a matter of law. In any event, it is very likely that the term "content" as used in *Smith* refers to the broader definition of "content" adopted by the Wiretap Act; *Smith* approvingly quotes a 1977 Supreme Court decision discussing how pen registers do not reveal the purport of telephone communications.[90]

A. The Body of Email Messages

Emails are transmitted in "packets" of digitized information that travel through the Internet's infrastructure, and these packets may contain both email address ("header") and body text information.[91] When the email reaches its destination server, the server reassembles the information into the header and the body of the email.[92] Yet, nothing about this transmission process implicates the content status of the body itself in semantics or in constitutional law (or under the ECPA). Setting aside questions of whether certain email body information is protected by the Fourth Amendment,[93] the "message in the body of the email itself," like a letter or a phone conversation,[94] clearly constitutes the content of a communication.

89. 18 U.S.C. § 2510(8) (2006).

90. Smith v. Maryland, 442 U.S. 735, 741 (1979) (quoting United States v. N.Y. Tel. Co., 434 U.S. 159 (1977)).

91. Kerr, *supra* note 49, at 614-15.

92. *Id.* at 615.

93. *See* Warshak v. United States, 490 F.3d 455, 469 (6th Cir. 2007), *vacated en banc*, 532 F.3d 521 (6th Cir. 2008) (vacating the suit on ripeness grounds).

94. *See* Kerr, *supra* note 49, at 612.

B. Email To/From Information

Email to/from routing information is contained in the header portion of the email packet, and generally consists of email addresses, IP addresses, packet information, and information about servers that are transmitting the email message.[95] Though this sort of information is used exclusively for directing emails and seems to be directly analogous to noncontent phone number information, some writers and privacy advocates have argued that it reveals more personal information than do phone numbers, and therefore should receive greater legal protection.[96] Email addresses may, for instance, indicate the identity of the communicators with more specificity than phone numbers, which might be used by anyone in a household.[97] Depending on the domain name of the address, email addresses may also reveal the communicator's workplace or university affiliation. Certainly, the marginal increase in specificity and the disclosure of a communicator's work or school affiliation might represent a tiny increase in the invasiveness of email pen registers relative to telephone pen registers.[98] But none of the information disclosed by an email address concerns the actual content of any emails.[99]

---

95. For a sample email header containing such information, see *id.* at 612-13.

96. *See* Paul Taylor, *Issues Raised by the Application of the Pen Register Statutes To Authorize Government Collection of Information on Packet-Switched Networks*, 6 VA. J.L. & TECH. 4, ¶ 4 (2001); Robert Ditzion, Note, *Electronic Surveillance in the Internet Age: The Strange Case of Pen Registers*, 41 AM. CRIM. L. REV. 1321, 1341 (2004); Georgiton, *supra* note 80, at 1845-46; Kevin Butler, *Is Big Brother Surfing the Internet? FBI's 'Carnivore' Raises Privacy Issue*, INVESTOR'S BUS. DAILY, Aug. 9, 2000, at A22, *available at* http://www.investors.com/editorialcontent.asp?secid=1501&status=article&id=160428&secure=7406 (quoting privacy advocates' arguments that *Smith* does not apply to email headers).

97. *See* Georgiton, *supra* note 80, at 1846. *But see* Kerr, *supra* note 49, at 643.

98. Of course, anyone with access to a person's computer could send an email from their address, assuming the person uses Microsoft Outlook or another email service that activates without a password. Neither phone numbers nor email addresses identify a user with 100 percent certainty, yet both will identify a user in the great majority of cases. The differences are slight.

99. Note that one could make a "hyper-literal" argument that under the narrowest definition of contents, only email and website IP addresses are truly envelope information, because every other byte of information is contained in the packets transmitted to these addresses. *See* Kerr, *supra* note 49, at 614-15. Under this conception of content versus envelope, email addresses and URLs would be considered content because they are "contained" in packets. Though this might provide robust protection for Internet communi-

To be sure, government investigators might be able to make educated guesses as to the purpose of an email (or a phone call) between two parties,[100] for example, a known drug addict and a known drug dealer. But one cannot argue that such to/from information actually *reveals* the contents of communications. Of course, observing the addressing information very likely reveals the identities of the communicators. But even that information does not reveal the details (or even the purpose or subject) of the communication itself—two people might talk about any subject. As even most proponents of additional privacy protection acknowledge, email to/from information should not be considered content.[101]

C. Subscriber Information

Internet users often provide personal and billing information to their ISP in the course of subscribing for Internet services.[102] This

---

cations data, it would be contrary to the current constitutional approach to content questions. *See, e.g.*, United States v. Forrester, 495 F.3d 1041, 1049-50 (9th Cir. 2007); Warshak v. United States, 490 F.3d 455, 471 (6th Cir. 2007). It would also be contrary to Congress's definition of "contents" in the ECPA, which emphasizes the "substance, purport, or meaning" of a communication. 18 U.S.C. § 2510(8) (2006). This body of law reflects an "analogy" approach to Internet communications, in which courts seek to determine the content status of information based on whether it concerns or reveals the meaning of a communication; in other words, courts seek to determine if the information is more like letter text or envelope information. *See Forrester*, 495 F.3d at 1049-50. This Article follows this thus-far dominant means of content analysis. The "hyper-literal" approach would also be entirely technology dependent and confined to current Internet technologies. For all of these reasons, no scholar has, to my knowledge, proposed a packet-based approach.

On the wisdom of applying existing legal frameworks rather than creating new legal regimes for each new technology, see Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207, 208-10. For additional discussion on the issues of analogy and perspective in Internet law, see A. Michael Froomkin, *The Metaphor Is the Key: Cryptography, the Clipper Chip, and the Constitution*, 143 U. PA. L. REV. 709 (1995); Orin S. Kerr, *The Problem of Perspective in Internet Law*, 91 GEO L.J. 357 (2003).

100. *See* Kerr, *supra* note 49, at 643. The Mosaic theory of data gathering suggests that, given a sufficient number of disparate bits of data, the government may be able to glean a great deal of information about an interaction between two parties under investigation. *See* Berman v. CIA, 378 F. Supp. 2d 1209, 1215 (E.D. Cal. 2005). Nonetheless, such information would not include the substantive content of the interaction, and there is good reason to be skeptical of any entity's ability to form concrete or reliable conclusions by gathering and combining numerous bits of innocuous data. *See* David E. Pozen, Note, *The Mosaic Theory, National Security, and the Freedom of Information Act*, 115 YALE L.J. 628, 654, 664-66 (2005).

101. *See, e.g.*, Freiwald, *supra* note 40, at 60.

102. *See, e.g.*, United States v. Hambrick, 55 F. Supp. 2d 504, 508 (W.D. Va. 1999). Note

information, which may include credit card or social security numbers, certainly implicates whatever privacy interests consumers have in their personal financial information.[103] But information about a subscriber has no bearing on the actual contents of messages, except perhaps as subscriber identity relates to government investigators making guesses about the subject matter of content based on email to/from information. Those concerns, which ultimately have no bearing on the classification of such information as noncontent, are discussed above.[104]

D. Email Size and Text Length

The government's pen register information collection software—formerly called "Carnivore" and now called "DCS-1000,"[105] does not collect an email's subject line, but it does replace the text in the subject line with Xs, allowing a viewer to know how many letters were used in the subject.[106] The government's pen register surveillance programs may also collect information about the total size of email transmissions.[107] To be sure, the length of an email or email subject line (or a phone call, to which the same argument would apply) can tell an observer something about the email itself.[108] The size of an email might also indicate that a file of that size has been attached to the email. Conceivably, authorities investigating a child pornographer might guess that he was sending child pornography images in captured emails of a certain size.[109] But the size of the file does not actually indicate or rule out any specific content; it does not reveal whether a file exists or whether the file

---

that free ISPs often do not require any personal information about their subscribers. *See, e.g.*, Dana L. Bazelon, Yun Jung Choi & Jason F. Conaty, *Computer Crimes*, 43 AM. CRIM. L. REV. 259, 290 n.241 (2006).

103. *See, e.g.*, Reserve Solutions, Inc. v. Vernaglia, No. 05 Civ. 8622 VM RLE, 2006 WL 1788299, at *2 (S.D.N.Y. June 26, 2006); WRIGHT ET AL., *supra* note 69, § 2459 n.7.

104. *See supra* text accompanying notes 96-101.

105. *See* Manton M. Grier, Jr., Comment, *The Software Formerly Known as "Carnivore": When Does E-Mail Surveillance Encroach Upon a Reasonable Expectation of Privacy?*, 52 S.C. L. REV. 875, 875 & n.2 (2001).

106. IITRI REPORT, *supra* note 57, at C-3.

107. *See id.*; Grier, Jr., *supra* note 105, at 886.

108. McPhie, *supra* note 82, at ¶ 33.

109. Grier, Jr., *supra* note 105, at 886-87.

is an image or a similarly sized file of another type, and it certainly does not reveal anything about the content of the hypothetical image.[110] Neither does the length of a line of text reveal the content of the text, nor its subject matter or purport. Guessing the content of even a three-letter subject heading would be impossible to do with any accuracy; not counting acronyms or punctuation (for example, "hi!"), there are still hundreds of distinct three-letter words.[111] Though it may be unnerving to know that the government can determine the length of one's emails with a pen register, mere information about the length of emails should not be considered content.

E. Email Subject Headers

The subject line of an email, unlike the Internet communication categories discussed above, cannot be directly compared to any feature of regular mail. It contains communicative writing and does not contain any routing information, but it is transmitted in the header portion of email packets.[112] As such, it would probably not qualify as content under the very narrowest definition of the "contents" of a message. It appears on the external header portion of the email; only the body of the email is "contained" in the packaging that is the email header. Nonetheless, both the Department of Justice and the one district court to have commented on the matter have concluded that the subject header, despite its location in an email transmission, should be treated as content.[113]

---

110. In this way it can be roughly analogized to the weight of a package sent through the mail. This information is not considered content, and it has been definitively excluded from Fourth Amendment protection since 1877. *See Ex parte* Jackson, 96 U.S. 727, 733 (1877); *see also* United States v. Forrester, 495 F.3d 1041, 1049-50 (9th Cir. 2007).

111. There are, for instance, 972 words on the Scrabble three-letter word list at San Jose Scrabble Club No. 21, Three-Letter Word List, http://www.yak.net/kablooey/scrabble/3letterwords.html (last visited Mar. 11, 2009).

112. Kerr, *supra* note 49, at 613.

113. CCIPS MANUAL, *supra* note 71, at 91; *In re* Application of the United States for an Order Authorizing the Use of a Pen Register, 396 F. Supp. 2d 45, 48 (D. Mass. 2005). Both the DOJ and the district court opinion refer to the ECPA, but their reasoning strongly suggests that email subject lines should also be considered content for the purposes of constitutional law. Again, this issue is wholly separate from the issue of whether subject line content is protected under the Fourth Amendment.

The House Report on the USA PATRIOT Act and the PATRIOT Act's amendments to the ECPA, state that subject lines are "clearly content."[114] Likely as a result, the Department of Justice's policy is to treat email subject lines as content.[115] Neither the Department's policy nor the House's legislative history give a reason for their interpretation; they simply consider email subject lines to be "letter" information despite their location. Under this rationale, the subject line would be just another form of email body text. This rationale is sound, and it provides an attractive bright line rule. Any information that is not used in routing a communication, and which itself may contain communicative text (for example, "The White Sox won"), is "letter" content, regardless of its location in the packet of data actually transferred between ISPs.

Email subject lines might also be considered contents to the extent that they obviously disclose the subject and purport of the body of the communication. In a recent district court case involving a pen register application, the magistrate judge found that, while the subject line was contained in the email header, the "information contained in the 'subject' would reveal the contents of the communication and would not be properly disclosed pursuant to a pen register or trap and trace device" under the ECPA.[116] This alternative ground is not necessary for the framework advocated here (which proposes that email subject lines are essentially the equivalent of body text), but the court's argument is sound as well, especially given the broad definition of "contents" in the ECPA.[117]

III. SEPARATING CONTENT FROM NONCONTENT IN WEB BROWSING COMMUNICATIONS

Like emails, the transmission of website data on the Internet occurs via packets of digitized information.[118] When an Internet user

114. H.R. REP. NO. 107-236, pt.1, at 53 (2001).

115. CCIPS MANUAL, *supra* note 71, at 112 (stating that the subject line "can contain content"); *see also* Kerr, *supra* note 49, at 613.

116. *In re Application*, 396 F. Supp. 2d at 48.

117. *See* 18 U.S.C. § 2510(8) (2006) (defining "contents" as "any information concerning the substance, purport, or meaning of [a] communication").

118. *See* Kerr, *supra* note 49, at 614-15. The following brief description of how website communication works is, of course, a simplified account. For a detailed description of how

types in or navigates to a URL, the URL's domain name (for example "www.nytimes.com") is translated during transmission into the IP address of the target website.[119] Specific website routing information, like the individual page URL, is generally placed inside the packet transmitted to the target's IP address,[120] just as email to/from address information is generally placed inside packets, along with body text.[121] The packet is sent to a website host computer's IP address.[122] The host computer processes the URL information in this packet and uses it to select the web page requested.[123] It then sends the website data to the user's computer.[124]

This relatively simple process gives rise to many complex legal and semantic questions as to whether courts should consider any or all Internet communication information content, and further, whether courts should protect this information under the Fourth Amendment. The first of these questions is whether communications between a person and a website host computer should be considered communications at all.

The question is essentially confined to constitutional law because the extremely broad definition of "electronic communications," which the ECPA defines as "any transfer of signs, signals, writing, ... data, or intelligence of any nature," appears to encompass web surfing activities.[125] As for the constitutional issue, it appears that web surfing transmissions will similarly be treated as a form of communication, analogous in constitutional law to telephone

---

Internet transmissions work, see for example, PETE LOSHIN, TCP/IP CLEARLY EXPLAINED (3d ed. 1999).

119. *See* Brain, *supra* note 38.

120. *See* Ditzion, *supra* note 96, at 1330-31.

121. *See supra* note 91 and accompanying text.

122. *See* Ditzion, *supra* note 96, at 1330-32.

123. *See* Brain, *supra* note 38; *see also* Ditzion, *supra* note 96, at 1331-32 (describing how Internet pen register software must filter through packets in order to obtain email and URL information).

124. *See* Brain, *supra* note 38; Ditzion, *supra* note 96, at 1331-32.

125. 18 U.S.C. § 2510(12) (2006) (defining "electronic communication" as "any transfer of signs, signals, writing, images, sounds, data, or intelligence of any nature transmitted in whole or in part by a wire, radio, electromagnetic, photoelectronic, or photooptical system" that is not a wire or oral communication); *see, e.g.*, *In re* Pharmatrak, Inc. Privacy Litig., 329 F.3d 9, 18 (1st Cir. 2003) (holding that a website transmission is an electronic communication under the ECPA); Konop v. Hawaiian Airlines, 302 F.3d 868, 876 (9th Cir. 2002) (same).

conversations. Communicative data (such as website content, personal information, and blog comments) is transferred back and forth between Internet users and automated website hosts.[126] Current law provides no reason to believe that the constitutional protection for the content of communications is diminished because one part of the communication is automated or available to the public.[127] Courts have already implicitly treated telephone calls to an automated, voice recognition-based service as communications in which callers may have a reasonable expectation of privacy as to the content of their calls.[128] In other words, there is no indication that websites sent to users are not considered communications just because they are transmitted by automated processes rather than other human beings—after all, a living person wrote them and posted them at some point.[129] Accordingly, the one court to have addressed the issue has applied the Fourth Amendment "communications" framework of *Smith* to web surfing activities.[130] A number of scholars have also considered web surfing activity to be a communication for the purposes of constitutional law.[131] Still, a far more difficult question remains: which parts of this communication are content and which are not?

A. Website Text and Other Data Sent to Users; Data Input by Users and Sent to the Web Host

The question of whether website text transmitted to an Internet user or an Internet user's input data transmitted to a web host computer represents "content" is essentially answered once the transmissions are recognized as communications. As with email

126. *See Konop*, 302 F.3d at 876.

127. *See, e.g.*, *In re* Application of the U.S. Authorizing (1) Installation of Use of a Pen Register, 441 F. Supp. 2d 816, 823 (S.D. Tex. 2006) (holding that digits dialed to communicate with an automated telephone system constitute call content). Presumably the government could not intercept the contents of phone calls to automated lines with only a pen register order just because the line is automated.

128. *See, e.g.*, *id.*

129. *See Konop*, 302 F.3d at 876.

130. United States v. Forrester, 495 F.3d 1041, 1048-49 (9th Cir. 2007).

131. *See supra* notes 80-81; *see also* Kerr, *supra* note 49, at 646-47 (discussing the novel category of human-computer interactions but assuming that such interactions are communications).

communications, the "body" of a web surfing communication is the nonrouting information contained in the digitized packets transferred over the Internet.[132] Data that is input by users, including search terms typed into a search "box" on a web page, has been recognized by courts as content, at least as that term is defined in the ECPA.[133] Because website transmissions are communications,[134] the text and other information sent by websites to users is very likely to be treated as the contents of a communication for the purposes of statutory and constitutional law. As even proponents of a narrow reading of "content" have acknowledged, the nonrouting information in Internet packets represents Internet communications content.[135]

B. URLs Containing Search Terms

A great deal of the confusion over the content/noncontent status of URLs in general[136] may be the result of a fixation among courts and commentators on the status of URLs that include search terms. This section examines the process of using search terms to navigate the Web and discusses the arguments concerning the content status of search term URLs.

When a user types search terms into a box on a search engine's website, the user is not actually searching the entire Internet.[137] Instead, the user is directed to a page displaying the results of a search of the website's database.[138] This occurs when the user's Internet browser adds the search terms into the URL sent to the website, which sends back the requested results page based on the

---

132. Colloquially, the text, video, or audio portion of a website is generally referred to as its "web content." *See, e.g.*, Web-Designz, What is Web Content and Why Is It So Important?, http://www.web-designz.com/tutorials/website_planning/web_content.shtml (last visited Mar. 11, 2009).

133. *See In re* Application of the U.S. for an order Authorizing the Use of a Pen Register, 396 F. Supp. 2d 45, 49 (D. Mass 2005).

134. *See supra* notes 118-29 and accompanying text.

135. *See* Kerr, *supra* note 49, at 615 tbl.1.

136. *See supra* notes 77-82 and accompanying text.

137. *See* SHARI THUROW, SEARCH ENGINE VISIBILITY 11 (2003).

138. *Id.* at 15.

URL.[139] Indeed, typing the URL containing the search terms directly into an Internet browser takes one to the same results page.

As in the case of email subject lines, the arguments that search terms embedded in a URL are content could take two forms, one reflecting a very narrow definition of content, and the other reflecting the broad definition used in the ECPA. First, one could argue that the terms in the URL are actually the body text of a communication contained by the rest of the URL. Yet the terms in the URL are placed on the "envelope" of the Internet communication, and are not technically typed by the user (they are placed in the URL by the browser, which copies the text typed into the web page form). They are used for website routing just like any other URL information.[140] It would be difficult to argue that the search terms in a URL are literally contained like the text of a letter.[141]

Instead, both courts and scholars have focused on the argument that URLs containing search terms reveal the content of the user's input to the website, thus exposing that input and the subject and purpose of the entire communication.[142] In one recent district court ruling, the magistrate held that URLs containing search terms had the potential to reveal the subject and purpose of communications, and therefore such URLs must be treated as contents under the broader definition used in the ECPA.[143] Otherwise, pen registers

---

139. *See* Robert Berkowitz, *Packet Sniffers and Privacy: Why the No-Suspicion-Required Standard in the USA Patriot Act Is Unconstitutional*, 7 COMP. L. REV. & TECH. J. 1, 16 (2002).

140. By contrast, an email subject line is directly typed by the user rather than translated by a browser, and is not used to route the email communication.

141. *See* Matthew A. Goldberg, Comment, *The Googling of Online Privacy: Gmail, Search-Engine Histories and the New Frontier of Protecting Private Information on the Web*, 9 LEWIS & CLARK L. REV. 249, 265 (2005) ("[T]he URL that results from the search, the one containing your search term, is unmistakably a Web site *address* that tells the computer where to go, or at least what to do. The fact that one can easily copy the URL resulting from a particular search and re-enter it at a later time to retrieve a substantially similar page of search results supports a view of the URL as routing or addressing information.").

142. *See, e.g.*, *In re* Application of the U.S. for an Order Authorizing the Use of a Pen Register, 396 F. Supp. 2d 45, 49 (D. Mass. 2005); Bellia, *supra* note 45, at 1429 ("Though the URL only represents the location on [a web] server of a file generated in response to the search and containing the results of the search, it gives significant clues as to what that file contains."); Solove, *supra* note 77, at 1287 (stating the URLs can reveal user searches); *see also In re* Pharmatrak, Inc. Privacy Litig., 329 F.3d 9, 18 (1st Cir. 2003) (suggesting that a company records "contents" under the ECPA when it records URLs that include personal information input in a form).

143. *In re Application*, 396 F. Supp. 2d at 49-50.

could capture all of a user's search term information simply by capturing the URLs sent from his IP address.[144]

Perhaps because it is so intuitive that search terms in a URL should be considered content, the treatment of content-revealing communications data is undertheorized in computer surveillance scholarship. Indeed, by focusing on URLs containing search terms, privacy scholars may be promoting a sort of "search term exception." Courts and commentators developing such an exception might conclude that URLs containing search terms are contents, but such URLs present a relatively rare and exceptional case.[145]

This Article argues that the creation of a search term exception[146] would be the product of an erroneous understanding of the legal status of URLs. As discussed below,[147] a focus on search terms obscures the functional and legal similarities between URLs containing search terms and those that do not.

C. URLs

Setting aside the question of Fourth Amendment protection, this Section argues that standard URLs qualify as content information for the purposes of constitutional law. They reveal every bit as much content as do URLs containing search terms. And peripheral, noncontent information that inevitably reveals underlying content is treated as content itself, under existing common and constitutional law.[148] The failure to introduce this "content revealing" principle to Internet law would result in a drastic underprotection of private communications. Further, the recognition that Internet communications (such as URLs) represent content has the potential to clarify other aspects of the constitutional and statutory law that governs Internet privacy. This section examines first the practical

144. *See id.* at 49.

145. *See, e.g.*, Orin S. Kerr, *A User's Guide to the Stored Communications Act, and a Legislator's Guide To Amending It*, 72 GEO. WASH. L. REV. 1208, 1228 n.142 (2004); *see also* Solove, *supra* note 77, at 1286-88 (criticizing Kerr's views on Internet privacy, but accepting the idea that only URLs with search terms need be protected as content); Goldberg, *supra* note 141, at 266-67 (arguing that users have a unique privacy interest in search terms).

146. *Cf. In re Application*, 396 F. Supp. 2d at 49-50 (specifically including search terms in a list of "contents" that are off limits to an approved pen register).

147. *See* discussion *infra* Part III.C.1.

148. *See* discussion *infra* Part III.C.2.

arguments, and then the legal arguments, that lead to the above conclusions.

1. What Standard URLs Reveal

By focusing on the more obvious fact that URLs containing search terms reveal user input communications, many scholars have ignored or given insufficient attention to the fact that all URLs reveal underlying web surfing communications,[149] exposing the website content requested by and sent to users.

A conventional URL inevitably discloses its corresponding website content.[150] Government investigators (or hackers, for that matter) using a pen register to capture URLs as they are transmitted, or even receiving the URL data a short amount of time after the user herself transmits it, have surely obtained as a practical matter the content transmitted and viewed. The fact that they have to copy the URL from their pen register program and paste it into their Internet browser to extrapolate the website text itself does not change the fact that, in practice, the secret is out once the URL is obtained.[151]

---

149. Those scholars that have mentioned the fairly well-known fact that URLs point anyone who obtains them to the website that a user is viewing have not analyzed the legal implications of this fact. Instead, they have offered this fact as evidence that either the current constitutional surveillance framework is flawed, *see, e.g.*, Solove, *supra* note 77, at 1287 (focusing on search terms and the need for a rethinking of *Smith*), or merely as evidence that the current constitutional and statutory law dealing with pen registers is not up to the task of determining the content status of Internet communications. *See, e.g.*, Bellia, *supra* note 45, at 1429-30; Ditzion, *supra* note 96, at 1338-39.

150. Note that some website content, such as the text a user has entered into a website form, is encrypted when transmitted as URL data and is not revealed by the encrypted URL. Also, an Internet user exploring a web page may be able to activate some forms of web content without navigating to a new page and generating a new URL—although the user's web history would include the original page's URL. This section focuses on the standard URL/website configuration, in which a unique URL corresponds to a specific web page (or, as with search pages or other pages with dynamic content, a specific result generated by the host's database). Unconventional website configurations do not significantly impact the practice of Internet navigation or alter the fact that a URL history reveals the content requested by and sent to an Internet user. As such, they do not affect this section's analysis. These unique website forms and their implications for Internet surveillance law are discussed in Appendix 1.

151. Given the ability of virtually any American to access the Internet in his or her local public library, disclosing the URL of a website can be said to reveal to anyone what that website contains in roughly the same way that revealing a phrase coded into Pig Latin (for

Indeed, even URL data obtained from ISP records days, months, or years after the user engaged in the recorded web surfing generally reveals the contents of the web pages sent and viewed. First, many article-specific URLs will not change for years, and can lead investigators directly to the article viewed.[152] Second, even for websites like blogs whose content is updated more frequently, private and governmental organizations (including the Library of Congress) have archived previously posted websites and made their content available to the public.[153] The most extensive and easily accessible means of determining the content of a website on a specific past date is the Internet Archive, a website that provides free public access to a database of archived websites.[154] Although the Archive cannot capture every single web page on the Internet every day, it frequently (from multiple times per day to every other day) archives the most popular sites and infrequently (roughly once per two months) archives smaller, less popular (and probably less likely to be updated) websites.[155] Thus, an observer who knew that a user entered a URL into her Internet browser on a certain date could enter the URL into the Internet Archive and search for that date in the results.[156] Using this method, an observer could often determine precisely the content an Internet user viewed at a given URL—for example, the URL of the Libertarian Party's blog on February 2, 2006.[157] Finally, in the rare case that the content of a web page

---

example, "ill-bay s-iay opping-shay t-ay ome-hay epot-day") reveals that phrase to anyone who can decode Pig Latin. Similarly, disclosing a word in a code substituting the numbers 1 to 26 for the 26 letters of the alphabet reveals the word to anyone who knows the code (e.g., "[8][5][12][12][15]"). The coded phrase itself is nonsense, but decoding it is obvious and takes only a few seconds. Cutting and pasting a URL from a spreadsheet into a web browser would be equally obvious to any Internet user, and would take roughly the same amount of time.

152. For example, URLs cited in a 2003 Orin Kerr law review article direct an Internet user to the same government, educational, and private web pages that they did in 2003. *See, e.g.*, Kerr, *supra* note 49, at 610 n.13 (http://www.usdoj.gov/criminal/cybercrime/cyberstalking.htm, last modified in 2003); *id.* at 612 n.20 (http://www.cs.indiana.edu/docproject/bdgtti/bdgtti_6.html); *id.* at 614 n.32 (http://www.ictp.trieste.it/~radionet/nuc1996/ ref/tcpip/, a paper written in 1996).

153. Beryl A. Howell, *Proving Web History: How To Use the Internet Archive*, J. INTERNET L., Feb. 2006, at 3-4.

154. *See id.*; Internet Archive, http://www.archive.org/index.php (last visited Mar. 11, 2009).

155. *See* Howell, *supra* note 153, at 4.

156. *Id.* at 5.

157. *See* Archive of LP Blog: The Official Blog of the Libertarian Party, Feb. 2, 2006,

could not be ascertained using the above methods, the government would likely be able simply to subpoena the web host for its archived content records, requesting that it disclose the information that it displayed on a certain date when a certain URL was entered.[158] In virtually all cases, there would be little practical obstacle to a government investigator obtaining the content information associated with even years-old URLs.

Further, the information revealed about user inputs is generally the same regardless of whether URLs containing search terms are recorded.[159] Standard URLs frequently reveal the same information.[160] As an example, imagine that Alice Sebold's recent bestseller *The Lovely Bones* was not a novel, but rather the manifesto of a subversive political group disfavored by the government. Suppose that a user under investigation by the government for her political activities[161] searched for the book and then clicked on the book's web page in order to see its publication details or to purchase it. Government investigators (or hackers) obtaining a URL with search terms sent to Amazon.com would know that the user was looking for "lovely bones" and would be able to visit the book's page by clicking on it in the Amazon results page that the following URL

---

http://web.archive.org/web/20060203121340/http://www.lp.org/yourturn/ (last visited Mar. 11, 2009).

158. *See* 18 U.S.C. § 2703 (2006).

159. Patricia L. Bellia, *Spyware and the Limits of Surveillance Law*, 20 BERKELEY TECH. L.J. 1283, 1311-12 (2005).

160. *Id.*

161. Recall that there is no judicial review of court order applications for pen registers. *See* 18 U.S.C. § 3123(a) (2006); Freiwald, *supra* note 40, at 48-49. And, if obtaining URLs is not a search under the Fourth Amendment, the government might obtain one's URLs for a variety of purposes. Government entities continue to show interest in monitoring the political and social activities of private citizens and interest groups. *See* Lisa Rein & Josh White, *More Groups Than Thought Monitored in Police Spying*, WASH. POST, Jan. 4, 2009, at A01. One particularly nefarious potential use of the ability to obtain URLs would be to leak to the press the URL history of a public figure whom the government wishes to discredit. Imagine that a government investigator obtained and leaked the URL history of a candidate from the opposition party for the purpose of discrediting him or her—if the candidate has ever visited any politically controversial or pornographic sites, or entered any suspicious sounding Google searches (even for wholly legitimate purposes), such a leak could cause massive embarassment or even a career-ending scandal. In other words, the government's ability to obtain our URL data for any purpose has implications far beyond the potential use of URLs in criminal prosecutions.

directs them to: http://www.amazon.com/s/ref=nb_ss_gw?url=search-alias%3Daps&field-keywords= lovely+bones.[162]

Imagine that the search term URL could not be obtained for technical or legal reasons. Investigators or hackers obtaining only the URL of the book's page at Amazon.com would also know from reading the URL that the user was looking for the novel "lovely bones," and would be able to visit the book's page once they obtained the URL: http://www.amazon.com/Lovely-Bones-Novel-Alice-Sebold/dp/0316666343.[163]

By contrast, Barnes and Noble's book page for *The Lovely Bones* does not provide the name of the book in the URL.[164] Nonetheless, investigators or hackers would easily be able to ascertain which book the user was looking for by visiting the book's web page at its URL: http://search.barnesandnoble.com/booksearch/isbninquiry.asp?ean=0316666343.[165]

The text in the three URLs differs; but, search terms or not, the three URLs all reveal to any observer with Internet access the same information: the user has requested and accessed the web page of *The Lovely Bones*. Only by blinding oneself to the realities of how URLs work could one argue that one of these URLs is meaningfully different from another in practice. Of course, this practical conclusion is not the end of the content inquiry under the law. The law may emphasize the form of communications over their function, and thereby require us to treat routing information as noncontent even when it reveals the entirety of a communication. The following sections argue that it does not.

2. Analogous Areas of Content/Noncontent Law

Internet surveillance law has not yet explicitly resolved the status of forms of information such as URLs, which do not themselves literally contain content (at least under the narrower definition of

162. Amazon.com: Books: Lovely Bones, http://www.amazon.com/s/ref=nb_ss_gw?url=search-alias%3Daps&field-keywords=lovely+bones (last visited Mar. 11, 2009).

163. Amazon.com: The Lovely Bones: Alice Sebold, http://www.amazon.com/Lovely-Bones-Novel-Alice-Sebold/dp/0316666343 (last visited Mar. 11, 2009).

164. The Lovely Bones, Alice Sebold - Barnes & Noble, http://search.barnesandnoble.com/booksearch/isbninquiry.asp?ean=0316666343 (last visited Mar. 11, 2009).

165. *Id.*

the term), but which completely reveal the content of the underlying communication.[166] It is entirely possible that courts may come to the practically absurd conclusion that URLs containing search terms that reveal content information should be treated as content, but URLs without search terms that reveal content information should not.[167] It should not come as a surprise, however, that in areas of law that are better developed than Internet surveillance, the law does not allow the government or private parties to obtain wholly noncontent information that inevitably reveals the contents of protected communications.[168]

Privilege law provides one such example. To be sure, privilege law is not directly applicable to Internet surveillance; Fourth Amendment protection and common law evidentiary privilege are, of course, different things. And while it is well-established that the contents of privileged communications are legally protected, it is still unclear whether the contents of web surfing communications will ultimately receive constitutional protection.[169] Yet privilege law faces the same problems of content/noncontent classification, and it can offer helpful guidance to a court attempting to determine whether URL information that reveals potentially protected communications content should itself be treated as content.

In privilege law, peripheral, noncontent facts about an attorney-client or doctor-patient communication (such as the identity of a client or attorney or the fact of a consultation) are generally not protected, even if the facts alone can be used as evidence against a defendant.[170] By contrast, the contents of communications are protected.[171] However, when the contents of a communication will obviously be exposed by the disclosure of the peripheral information, the peripheral information itself is treated as content, receiving the same protection as the underlying content that it reveals.[172] This is

166. *See* Foley, *supra* note 80, at 458.

167. *See* Kerr, *supra* note 49, at 646.

168. *See, e.g.*, Colton v. United States, 306 F.2d 633, 637 (2d Cir. 1962).

169. *See* discussion *infra* Part IV.A.

170. *See* Dole v. Milonas, 889 F.2d 885, 889 (9th Cir. 1989); *see also* Behrens v. Hironimus, 170 F.2d 627, 628 (4th Cir. 1948) (attorney-client privilege); City of Alhambra v. Superior Court, 168 Cal. Rptr. 49, 52 (Cal. Ct. App. 1980) (doctor-patient privilege).

171. *See, e.g.*, *Behrens*, 170 F.2d at 628.

172. *See In re* Grand Jury Subpoena (DeGuerin), 926 F.2d 1423, 1431 (5th Cir. 1991) ("If the disclosure of the client's identity will also reveal the confidential purpose for which he

true even where the "substance of a disclosure" (analogous to a web page's content) is already exposed but the peripheral information linking it to the defendant (analogous to the URL information sent from an Internet user's computer) is not yet known.[173]

For example, in *DeGuerin*, the Fifth Circuit found that disclosure of peripheral noncontent information—the identity of a third party who retained an attorney on behalf of an indigent drug dealer—would render obvious the content of the communication and the motivation of the third party in seeking the attorney.[174] In other words, it would reveal that the third party was a drug lord. As a result, the court held that the noncontent information was protected under the attorney client privilege, "because disclosure would allow the Government to obtain information given ... as part of a confidential communication."[175]

The protected content itself need not be the literal words exchanged; if the purport of the protected communication would be revealed by the peripheral information, then courts protect the peripheral information.[176] Much like URLs, peripheral information about attorney-client relationships that reveals underlying content is "connected inextricably with a privileged communication—the confidential purpose" behind the actual words exchanged.[177] The same general principles apply in the doctor-patient privilege context. The mere fact that a doctor-patient consultation occurred becomes privileged if that noncontent fact "discloses the nature of the condition for which the patient sought treatment."[178] The principle holds regardless of whether the mere fact of consultation itself discloses the protected communication, or if the communica-

---

consulted an attorney, we protect both the confidential communication and the client's identity as privileged."); *In re* Grand Jury Subpoenas (Anderson), 906 F.2d 1485, 1491 (10th Cir. 1990) (identifying a "confidential communication exception" to the general rule that noncontent information is unprotected, and citing supportive cases from the First, Second, Third, Fourth, Sixth, Seventh, Ninth, and Eleventh Circuits).

173. *Dole*, 889 F.2d at 889.

174. *DeGuerin*, 926 F.2d at 1431-32.

175. *Id.* at 1432.

176. *See, e.g.*, *Anderson*, 906 F.2d at 1492 (stating that the exception is generally applied when "the mere identification of the client would [disclose] the confidential communication from the client that he had committed the crime for which he sought advice").

177. *DeGuerin*, 926 F.2d at 1431.

178. City of Alhambra v. Superior Court, 168 Cal. Rptr. 49, 52 (Cal. Ct. App. 1980).

tion could obviously, and would inevitably, "be inferred" from the peripheral fact.[179]

The principles of the "confidential communications exception" for noncontent information in privilege law can be applied in the URL context. As discussed above, URLs obviously "allow the Government to obtain" underlying communicative information.[180] The well-established, analogous principles of privilege law strongly suggest that URLs should be treated as content and should receive the same protections that the actual text of Internet communications receive.[181] Further, like some peripheral information regarding attorney-client or doctor-patient communications, URLs reveal (and may even contain text indicating) the subject matter or purpose of an underlying communication.[182] Indeed, URLs reveal more, and more specific, information about the content of the underlying communications.[183] They expose not only the purport of communications but the very text of the communications themselves.[184]

Again, the authority of the privilege law cases dealing with content-revealing information is persuasive, not precedential, in the Internet communications context. Yet, if courts fail to apply these principles, Internet surveillance law will be the only area of the law in which peripheral information that entirely reveals the underlying content of a communication is not treated as content information itself. No other branch of law has thus far countenanced the practically absurd result that the government or a private party may discover potentially protected contents by obtaining noncontent information that reveals those contents.[185]

---

179. *Id.* (citing Marcus v. Superior Court, 95 Ca. Rptr. 545 (Cal. Ct. App. 1971)).

180. *DeGuerin*, 926 F.2d at 1432.

181. *See id.* at 1431; *Anderson*, 906 F.2d at 1491.

182. McPhie, *supra* note 82, at ¶ 29.

183. *Id.*

184. *Id.*

185. For an additional line of cases dealing with revealing noncontent information outside the communication context, see, e.g., Shelton v. Am. Motors Co., 805 F.2d 1323, 1329-30 (11th Cir. 1986) (holding that the mere fact of the selection of documents examined by an attorney would, if disclosed, reveal the content of the attorney's mental impressions about the case); United States v. Cook, No. CR05-0424-TSZ, 2006 WL 3474184, at *3 (W.D. Wash. Nov. 16, 2006) (holding that the mere fact of selection of documents in attorney's binder revealed the attorney's mental processes, and was therefore protected work product).

3. The Supreme Court's Fourth Amendment "Content" Jurisprudence

The Supreme Court, much like courts applying the law of evidentiary privilege, has had to decide how to treat peripheral information that is generally unprotected, but reveals protected information. The relevant cases are those that deal with the use of new technologies employed to surveil citizens in unique ways—*United States v. Karo*[186] and *Kyllo v. United States*.[187] It is important to clarify what is argued here. This section makes neither the debatable claim that the principles of *Kyllo* dictate that the Court must limit the holding of *Smith*;[188] nor does it make the wholly implausible claim that the *Kyllo* Court actually intended to abandon *Smith*.[189] Nor does it argue that *Kyllo* applies to URLs, which could be said to reveal the details of activities in a home.[190] *Kyllo* deals with "physical intrusion" (or its technological equivalents) and physical activity taking place in the uniquely protected area of the home.[191] It is not within the line of constitutional cases that specifically deal with communications directed outside of the home, and there is no indication that the Court intended it to be authoritative in the communications area.[192]

Nonetheless, it would be a serious error to think that the principles of *Kyllo* and *Karo* are wholly irrelevant to the constitutional status of Internet communications. Indeed, these cases stand for the proposition that peripheral information that reveals protected content must itself be treated as content information. In *Kyllo*,

186. 468 U.S. 705, 708 (1984).

187. 533 U.S. 27, 29-30 (2001).

188. *See* Tracey Maclin, Katz*,* Kyllo*, and Technology: Virtual Fourth Amendment Protection in the Twenty-First Century*, 72 MISS. L.J. 51, 97-101, 136-37 (2002) (arguing that obtaining telephone number or email address information should be held to violate *Kyllo*); Ric Simmons, *From* Katz *to* Kyllo*: A Blueprint for Adapting the Fourth Amendment to Twenty-First Century Technologies*, 53 HASTINGS L.J. 1303, 1341-43 (2002) (arguing that courts should disregard *Smith* and apply an "intimate nature" test derived from *Kyllo* to determine whether disclosed information should be protected by the Fourth Amendment). *But see* Solove, *supra* note 45, at 1152-53 (criticizing the "intimate nature" test as vague and unworkable).

189. *Cf.* Maclin, *supra* note 188, at 100 (stating that the *Kyllo* Justices did not appear to consider *Smith* while deciding *Kyllo* and would likely uphold *Smith*).

190. *See Kyllo*, 533 U.S. at 34-35.

191. *Id.* at 34.

192. *See* Maclin, *supra* note 188, at 100.

government agents used a thermal imaging device to scan heat emanating from a suspect's house, and determined based on the results that the suspect was "using halide lights to grow marijuana in his house."[193] The government argued that it had only observed heat radiating outside the house, which was not itself protected by the Fourth Amendment.[194] But the Court rejected this "mechanical interpretation," observing that the distinction between "off-the-wall" and "through-the-wall" thermal surveillance in such cases is without substance, because both types of surveillance reveal the same protected information.[195] The Court held that the use of the device allowed the government to explore details of the home that would generally only be discoverable upon physical intrusion, and had therefore effected a Fourth Amendment search of the suspect's home.[196]

Justice Stevens, in dissent, argued that the government's acquisition of external information—the heat emanating from a house—could not be protected by the Fourth Amendment, and therefore the majority had wrongly held that the "inference" made by the government agents was the unconstitutional search.[197] A similar argument could be made about URLs that do not contain search terms: the URLs themselves contain no body text, and it should not be a Fourth Amendment search to infer an Internet user's activity by visiting the public websites she visited. Yet the *Kyllo* majority, in a somewhat unclear rejoinder, explained that the search "is not the police's ... inferencing, but their ... thermal-imaging measurement of the emanations from a house."[198]

How can the observation of exposed, external information, entirely unprotected if observed with the naked eye,[199] be a search? The principle behind the privilege cases is at work once again—when noncontent peripheral information reveals content information, it must itself be treated as content. Thus, in *Karo*, *Kyllo*'s predecessor, the Court held that the use of an electronic beacon to

193. *Kyllo*, 533 U.S. at 30.
194. *Id.* at 35.
195. *Id.* at 35-36.
196. *Id.* at 40.
197. *Id.* at 44 (Stevens, J., dissenting).
198. *Id.* at 37 n.4 (majority opinion).
199. *Id.* at 43 (Stevens, J., dissenting).

monitor the location of a can of ether was itself a search because it revealed information about the inside of the suspect's home.[200] In *Karo*, whether noncontent information—the signal from a beeper—was protected under the Fourth Amendment depended entirely on whether it revealed protected content information.[201] When government agents did not know which locker the can was stored in, monitoring of the beeper did not constitute a search because it "revealed nothing about the contents of the locker."[202] But when monitoring of the beeper inevitably revealed the contents of a home (or locker for that matter[203]), the monitoring itself was a Fourth Amendment search.[204]

A similar, probably even more protective rule was applied in *Kyllo*. There, the government specifically measured external heat information that was consistent with the use of halide lamps, which led them directly to the conclusion that halide lamps were being used inside the house.[205] As the Court stated, the distinction between the external and internal information in such a situation would be a distinction without a difference.[206] In other words, the details of the house were exposed for all practical purposes as soon as the external information was obtained, and the Court would essentially be putting blinders on if it pretended otherwise.

In light of this principle, the meaning of the majority's statement about inference becomes clear.[207] The inference itself is not the search. The search is the acquisition of noncontent information that obviously, as in *Karo*, or for all practical purposes, as in *Kyllo*, will reveal protected content. The implications of this rule for the content status of URLs and other Internet communications are significant. Further, despite the differing contexts (home or locker content versus communicative content), the analogies become easy to draw. Like a tracking beeper that precisely discloses the contents

200. United States v. Karo, 468 U.S. 708, 715 (1984).

201. *See id.*

202. *Id.* at 720.

203. *See id.* at 720 n.6 (stating that "[h]ad the monitoring disclosed the presence of the container within a particular locker," use of the beeper would have constituted a search).

204. *See id.* at 715.

205. Kyllo v. United States, 533 U.S. 27, 29-30 (2001).

206. *Id.* at 36.

207. *See id.* at 37 n.4.

of a home or storage space, URL information reveals exactly the contents of underlying web surfing communications, and should therefore be treated the same as the contents themselves. Like information taken from the outside of a house that for all practical purposes reveals what is occurring inside the house, IP address information that reveals the subject matter or purpose of an underlying communication should receive the same protection as the communication itself.[208]

Karo and *Kyllo* stand for the principle that, in Fourth Amendment law, content-revealing information must be treated like content. While this principle has been disputed by a sizable minority of the Court (and has gone unrecognized by which Internet surveillance scholars),[209] it is implicit in the holdings of both cases, and the Court's acceptance of it should ultimately not be all that surprising. Like the law of privilege, the law of search and seizure is not fooled by technicalities. When the government obtains peripheral information that reveals protected content information, constitutional law will not put on blinders and pretend that no search has occurred. In accordance with this principle, URL information that reveals communications content should be treated as content itself for the purposes of constitutional law.

D. Website IP Addresses

In theory, website IP addresses, like URLs, might reveal the web page contents being viewed by an Internet user.[210] An IP address can often be used to determine the domain name of the website contacted, either by simply entering the IP address into a web browser like a URL or by looking it up in an IP address database.[211] As with URLs, observation of the IP addresses contacted by an Internet user might be used to determine the website viewed regardless of whether the IP address information is obtained

208. *See infra* Part III.D.

209. Justices Rehnquist, O'Connor, and Kennedy joined Justice Stevens in the *Kyllo* dissent. *See Kyllo*, 533 U.S. at 41 (Stevens, J., dissenting).

210. *See* Georgiton, *supra* note 80, at 1846; McPhie, *supra* note 82, at 38.

211. Georgiton, *supra* note 80, at 1846 n.73; Shawn C. Helms, *Translating Privacy Values with Technology*, 7 B.U. J. SCI. & TECH. L. 288, 296 n.45 (2001).

contemporaneously or in records of past web surfing communications.[212] On the other hand, IP addresses reveal significantly less about the content of web surfing communications than do URLs; they do not point to a specific page on a website, but instead may refer to any one of a number of pages. They do not even necessarily reveal that a user has viewed the main page of a website, as the user's computer would communicate with the same IP address even if she jumped immediately to a specific page from the web host's database[213] (perhaps directed there by a search engine). Further, multiple websites may share the same IP address.[214] It is generally difficult to determine whether a given website shares its IP address with another website, at least without directly contacting the web host.[215] Thus, in most cases, IP addresses will not reveal the underlying content or even subject matter of Internet communications with anything approaching the certainty of URLs.

Still, in cases in which a single website uses a single IP address, and when that website is either small enough or subject-specific enough, mere knowledge of the IP address contacted by the user could inevitably reveal the contents of the underlying communication. If so, under the above framework,[216] the IP address itself, though not technically content under a narrow definition of the term, should nonetheless be treated and protected as content.

For example, some websites, especially personal websites, may only contain one or two pages.[217] If such websites happened to have their own IP addresses, they would reveal as much or nearly as much about the content of the underlying web surfing communication as would URLs. Of course, smaller websites like these are also

---

212. The IP address can be used to determine the domain name of the website contacted, *see, e.g.*, Georgiton, *supra* note 80, at 1846 n.73; Helms, *supra* note 211, at 296 n.45, which can in turn be entered into Internet Archive to determine the content of the main page of the website on a given day.

213. *See* Ctr. for Democracy & Tech. v. Pappert, 337 F. Supp. 2d 606, 617-18 (E.D. Pa. 2004).

214. *Id.* at 618.

215. *Id.* at 619.

216. *See supra* Part II.

217. For example, the now-defunct website amyboyer.org, which memorialized a young girl killed by an Internet stalker, apparently contained only two pages. The archived site is available at Amy Lynn Boyer, http://web.archive.org/web/20021204212240/http://www.amyboyer.org/index.html (last visited Mar. 11, 2009).

likely to share an IP address with multiple other sites.[218] IP addresses assigned to larger websites might also reveal the subject matter and purpose of the underlying Internet communications. Again, the framework outlined above would dictate that such IP addresses should be treated as the content they inevitably reveal.[219] Still, such a situation would only arise if a website had a unique IP address, if this knowledge were obtained by government or private entities, and if any page of the website would obviously reveal the purpose and subject matter of the communication. For example, if the website of a prescription drug was large enough to have its own IP address, and its identity clearly revealed the purpose of the Internet user's communication, a court should treat the IP address itself as it would treat the actual content of the website.[220]

The few courts that have considered the content status of IP addresses have too readily dismissed the possibility that IP addresses will reveal content.[221] In *Forrester*, the Ninth Circuit expressed its concern about the ability of URLs to reveal content but concluded that IP addresses "reveal no more about the underlying contents of communications than do phone numbers" and are constitutionally indistinguishable from them.[222] In the ECPA context, a magistrate who expressed grave concerns about URLs and email subject lines revealing content nonetheless concluded that if "the government is seeking only IP addresses of the websites visited and nothing more, there is no problem."[223]

---

218. *See id.* (noting that the site is hosted by "Prospeed.net," a web server provider).

219. *See supra* Part III.D.

220. For example, if the Propecia website, which contains multiple pages, had its own unique IP address, the purpose of an Internet user's communications with the website could be revealed simply through the IP address. *See* Propecia, http://www.propecia.com/finasteride/propecia/consumer/index.jsp (last visited Mar. 11, 2009) (one can also access the website by entering the IP address 155.91.16.15 into a web browser). The cases involving noncontent information that reveals the medical condition of patients receiving a certain treatment are directly analogous. *See, e.g.*, City of Alhambra, 168 Cal. Rptr. 49, 52 (Cal. Ct. App. 1980).

221. *See, e.g.*, United States v. Forrester, 495 F.3d 1041, 1049 (9th Cir. 2007).

222. *Id.* The court provides the example of an IP address pointing to the New York Times's website, which would not disclose anything about the content of the underlying communications. *See id.* at 1049 n.6. This is surely correct, but the court overlooks the possibility of other websites' IP addresses revealing more about the subject matter of the underlying communications.

223. *In re* Application of the U.S. for an Order Authorizing the Use of a Pen Register, 396 F. Supp. 2d 45, 48 (D. Mass. 2005).

Although it is true that IP addresses probably do not reveal anything about the underlying content of web surfing communications, courts must recognize the possibility that the purpose or subject matter of certain communications might be exposed simply through the disclosure of an IP address. Of course, fact-intensive determinations of whether an IP address discloses content are probably not an efficient use of judicial resources. Courts can likely get around the difficulties inherent in such inquiries by prospectively barring the government from questioning web hosts about which domains use a given IP address or, if web surfing communications are protected by the Fourth Amendment,[224] retrospectively excluding evidence derived from IP addresses which reveal the contents of websites read by an Internet user.

E. Size of Information Accessed or Files Downloaded from Websites

The government's pen register software can and likely does collect information about the amount of data transferred between an Internet user and a host website.[225] Indeed, this may have occurred in *Forrester*, when the government's pen register captured the "total volume of information sent to or from [the defendant's] account."[226] This might refer to the total volume of the defendant's web surfing activity, but, given the function of the government's pen register software, more likely refers to the total volume of information sent to and from each website. If so, the court erred when it approved this capture as "constitutionally indistinguishable from the use of a [telephone] pen register."[227] Unlike the volume of data information that a pen register collects from email transmissions, volume information for individual websites may reveal the contents of web surfing communications, perhaps just as clearly as URLs do. Of course, for complex websites with many pages and/or many downloadable files, such total volume information is unlikely to yield much information beyond that already revealed by the IP

224. *See* discussion *infra* Part IV.A.
225. *See* IITRI REPORT, *supra* note 57, at C-5.
226. *Forrester*, 495 F.3d at 1044.
227. *Id.* at 1049.

address. However, for smaller, simpler websites, such information allows for a complete reconstruction of the content viewed and downloaded.

For instance, imagine that a simple website consisting of only four web pages and minimal text (constituting roughly 100 kb of data in total) allows users to download three files: a pdf of the Gettysburg Address (500 kb), a pdf of the Constitution (1000 kb), and a pdf of the long, rambling manifesto of a subversive political group (10,000 kb). If the user downloads any of the files, capturing the total volume of website activity tells an observer where the user surfed and which documents he downloaded (and in all probability, read[228]): 600 kb of total activity means the user downloaded the Gettysburg Address; 1100 kb the Constitution; 1600 kb, both; 10,100 kb, the manifesto; and so on.[229] In such situations, the argument that total volume data that reveals the content of web surfing communications should be treated as content for the purposes of constitutional law tracks exactly the argument about URLs made above.

F. Summary

At this point, it is possible to map the discussion of Internet communications in Parts II and III onto the content/noncontent framework developed in *Karo* and *Kyllo*. URLs, IP addresses, and email to/from information differ from each other in terms of how much they reveal about their underlying content. Observers can form inferences about content from all of these types of information, but the level of content exposure differs in terms of (1) what kind of additional information is required in order to learn about the

228. *Cf.* United States v. Schaefer, 501 F.3d 1197, 1198 (10th Cir. 2007) (describing how police obtained a search warrant to search defendant's home based on a tip that he subscribed to websites containing images of child pornography); United States v. Gourde, 440 F.3d 1065, 1068, 1070 (9th Cir. 2006) (holding that membership in a website that allowed users to download photographs gave rise to a "fair probability" that the user had obtained and seen those photographs).

229. A similar hypothetical could be developed if a website had a limited number of different web pages of distinct data sizes. The total volume information would reveal which pages had been viewed by the user.

content,[230] and (2) how much content is ultimately exposed. Table 1 displays these differences between the types of Internet communications and matches the different types of information to analogous physical information in *Karo* and *Kyllo*.[231]

230. Recall that many of the arguments that email to/from information reveals content depend on the idea that the identities of email senders plus additional information about the senders (for example, that they are drug dealers) can allow investigators to make educated guesses about content. *See supra* notes 100-01 and accompanying text.

231. *See supra* Part III.C.3. A similar table could be developed for the analogous communications information in the privilege law context. *See supra* Part III.C.2.

Table 1

(Internet communications are described in plain text.
Physical analogies are described in italics.)

Type of Information	**What Can the Information Alone Reveal?**	**What Can the Information Reveal When Coupled with Additional Information?**	**To What Level of Certainty Does the Information Reveal Content?**
Body of email text. Content of website communications and user inputs. Email subject lines. *The interior of a locker.*	Content.	No additional information is necessary.	**Reveals content with certainty.** *Treated as content.*
URL information. *Beeper transmissions from a homing device.*	The content of the underlying web page; user search information. *Location of a can; the content of a locker.*	No additional information is necessary.	**Reveals content with certainty in virtually all cases.** *Treated as content (*Karo*).*
IP address of website contacted. Size of total transmission between an Internet user and a website. *Thermal imaging information from the exterior of a house.*	Numbers. May reveal the domain name of a website associated with the IP address, if the address is unique to the website. *The precise amount of heat radiating from the house.*	Coupled with information that the website has only a few pages, or files of distinct sizes, may reveal the content transmitted to the user**.** *Coupled with information about the heat output of halide lamps, may reveal that the interior of the house contains halide lamps.*	Coupled with the appropriate information, **reveals for all practical purposes what is going on.** *If so, treated as content (*Kyllo*).*
Email to/from information; ISP subscriber information; email length. *The outside of a locker.*	Name of an ISP subscriber or the sender of an email. Existence and size of a communication. *The appearance and location of the locker.*	Coupled with information about the sender or recipient (*or the owner of the locker*), may allow guesses as to content.	**Educated guess**, at best. *Not treated as content.*

The conceptual framework developed in Parts II and III above offers a means to resolve the difficult question[232] of whether certain Internet communications should be considered content or noncontent. Further, the "content revealing" principle applied to URLs above can very likely be used by courts to establish the content status of future communications technologies, no matter how complex or different from traditional postal mail. The discussion above has focused on the categories of communications that have been the most controversial and hardest to classify. According to the framework, URLs and email subject lines should be treated as content in all situations, whereas information about the total size of communications with websites and IP addresses should be treated as content only in certain situations. Though determining the content/noncontent status of these categories will not resolve the debates over the appropriate level of statutory and constitutional protection afforded to Internet communications, determining which Internet communications are content and which are not has the potential to clarify and even to change the terms of these debates.

IV. IMPLICATIONS

Knowing the content status of a certain type of Internet communication information does not automatically determine the level of constitutional or statutory protection that information will receive. As mentioned above, this paper has thus far separated out the question of protection from the question of content/noncontent status.[233] Yet, the above discussion has significant implications for the ultimate question of whether and to what degree Internet communications are protected by privacy law. This is especially true for the most controversial and hardest to classify categories of Internet communication information: web surfing data such as URLs and IP addresses.

232. *See* Kerr, *supra* note 49, at 645.
233. *See* discussion *supra* Part II.

A. Internet Communications in Constitutional Law

Though URLs and other content-revealing web surfing information should be treated as content for the purposes of the constitutional law of communications, it is far from clear that this content is protected by the Fourth Amendment. Ultimately, a defendant asserting that a government acquisition of her URL data without a warrant violates the Fourth Amendment must establish that she has a reasonable expectation of privacy in such data. But can Internet users have such an expectation in even the content of Internet communications? This section offers some preliminary conclusions and predictions about how *Smith* and other relevant cases might be applied to different types of Internet communications content.

Certainly, numerous courts and commentators have read *Smith* as holding that the content/noncontent distinction is crucial to determining whether an individual has a reasonable expectation of privacy in their communications information.[234] Perhaps the strongest argument that only noncontent information can be denied Fourth Amendment protection under *Smith* is that *Smith*'s predecessor *Katz* based its constitutional protection of telephone calls on the pervasiveness of telephone communications and a normative conclusion that individuals were entitled to privacy in the contents of their telephone conversations.[235] As the Court stated in *Katz*, one who places a telephone call "is surely entitled ... that the words he utters into the mouthpiece will not be broadcast to the world."[236] One could argue that the *Smith* opinion could only avoid

234. *See, e.g.*, Warshak v. United States, 490 F.3d 455, 471 (6th Cir. 2007); United States v. Hambrick, No. 99-4793, 2000 WL 1062039, **3-4 (4th Cir. Aug. 3, 2000); David J. Phillips, *Beyond Privacy: Confronting Locational Surveillance in Wireless Communication*, 8 COMM. L. & POL'Y 1, 6 (2003) ("In effect, the Court in *Smith* promulgated a distinction between the contents of telephone messages and other attributes of those messages, including their source, destination and duration. While preserving the strongest Fourth Amendment protection for content, it removed constitutional protection from noncontent attributes."); Georgiton, *supra* note 80, at 1841, 1846 n.73 (discussing *Smith* and stating that if IP addresses were held to be content, the Fourth Amendment would apply to them).

235. *See* Freiwald, *supra* note 7, at ¶ 29 ("To read the Constitution more narrowly is to ignore the vital role that the public telephone has come to play in private communication.'" (quoting Katz v. United States, 389 U.S. 347, 352 (1967))).

236. *Katz*, 389 U.S. at 352.

this normative conclusion on the basis of the fact that telephone numbers are not content.

Yet *Smith* itself provides several overlapping rationales for denying Fourth Amendment protection to dialed phone numbers. Any or all of these additional rationales could be the basis for denying Fourth Amendment protection even to email and web surfing content. The rationales are identified below in the order in which they appear in the *Smith* opinion.

After distinguishing *Katz* on the basis of the noncontent status of telephone numbers, the *Smith* Court first concluded that a telephone user likely has no subjective expectation of privacy in telephone numbers. Telephone users convey phone numbers to the telephone company.[237] The phone company has facilities for making permanent records of the numbers dialed, and telephone users are aware of this because they receive a bill each month listing the phone numbers contacted.[238] Further, phone companies regularly employ pen registers to identify fraud, violations of the law, defective dials, overbilling, or obscene callers.[239] In sum, users convey telephone numbers to the phone company, which can record them, and in fact does so for "a variety of legitimate business purposes."[240]

Second, even if the user has a subjective expectation of privacy, it is not one that society is prepared to recognize as objectively reasonable. The Court cited *United States v. Miller*,[241] which found no expectation of privacy in business records held by a third party, and concluded that a person has "no legitimate expectation of privacy in information he voluntary turns over to third parties" because she assumes the risk that the third party will turn over (or be forced to turn over) such information to the government.[242] The *Smith* court thus implicitly extended the "third party doctrine" of

237. Smith v. Maryland, 442 U.S. 735, 742 (1979).

238. *Id.*

239. *Id.* at 742-43.

240. *Id.* at 743.

241. 425 U.S. 435 (1976). The Court also cited several cases involving information disclosed to undercover government informants. *See, e.g.*, Hoffa v. United States, 385 U.S. 293 (1966).

242. *Smith,* 442 U.S. at 743-44.

Miller to carriers of communications intended for another private party.

Third, the Court pointed out that telephone call routing information was previously carried out by human operators, and that such information conveyed to a human operator would have clearly been disclosed.[243] The Court refused to find that an expectation of privacy in such numbers was preserved just because the phone companies had automated their routing systems. The rationale of *Miller*, that there could be no expectation of privacy in records "exposed to ... employees in the ordinary course of business,"[244] was apparently extended in *Smith* to cover information exposed to a telephone company's "equipment."[245]

It is, unfortunately, not clear in *Smith* which of its rationales was central to the Court's holding, and which represented a marshalling of justifications in support of an already-reached conclusion. As courts and commentators have pointed out, some of the rationales may apply equally well to the content of telephone calls that the Court explicitly deems protected.[246] Having determined the content status of various Internet communications, we can begin to examine how these rationales might apply when lower courts (or the Supreme Court) grapple with the question of how to apply *Smith* to determine Internet communications' Fourth Amendment status.

1. Subjective Expectations of Internet Users

The three reasons given in *Smith* to support the idea that telephone users have no subjective expectation of privacy in their phone numbers—(1) numbers are conveyed, (2) the phone company can record them, and (3) the phone company does record them for legitimate business purposes—should be considered together. Not only does the *Smith* opinion group them together,[247] but the first

243. *Id.* at 744.

244. *Miller,* 425 U.S. at 442.

245. *Smith,* 442 U.S. at 744.

246. *See, e.g.*, *id.* at 746 (Stewart, J., dissenting); Warshak v. United States, 490 F.3d 455, 470-71 (6th Cir. 2007); McPhie, *supra* note 82, at 62.

247. *Smith*, 442 U.S. at 743 ("Telephone users, in sum, typically know that they must convey numerical information to the phone company; that the phone company has facilities

two reasons alone are probably applicable to telephone calls themselves.[248] The distinction between phone calls and phone numbers appears only in the last of the three points: phone companies generally do not record the content of phone calls.[249]

How does this three-pronged inquiry apply to Internet communications content? In *Warshak*, the Sixth Circuit panel found that Internet users had a reasonable expectation of privacy in email content because, as with phone call content, ISPs did not access or record the content of emails as a matter of course.[250] On the other hand, *Warshak* has been vacated,[251] and future courts could hold that ISPs do in fact access or record emails in the ordinary course of business. First, ISPs generally make copies of emails during transmission.[252] Second, ISPs routinely scan users' emails for viruses, spam, and child pornography, and may even scan them for certain keywords in order to tailor advertisements to the user.[253] The en banc court, or any other court facing the issue of email content's Fourth Amendment status, could certainly find that email content is disclosed to ISPs in such a way as to erode any subjective expectation of privacy under *Smith*.

A court would be far more likely to find URL content to be unprotected under *Smith*'s subjective expectation analysis than it would email content. URL data can be recorded by ISPs, and the data may or may not be combined with personal information collected by the ISP or from other sources.[254] Although the actual practices of various ISPs are difficult to determine, news reports (and a review of ISP privacy policies) indicate that, most likely,

---

for recording this information; and that the phone company does in fact record this information for a variety of legitimate business purposes.")

248. *See supra* note 246 and accompanying text.

249. *See Warshak*, 490 F.3d at 471.

250. *Id.*

251. Warshak v. United States, 532 F.3d 521 (6th Cir. 2008) (en banc) (vacating 490 F.3d 455 (6th Cir. 2007) as unripe).

252. *See* GRALLA, *supra* note 26, at 87; Mulligan, *supra* note 26, at 1563 (noting also that unencrypted email messages could be read by ISP employees.).

253. *See Warshak,* 490 F.3d at 474; Gmail, About Gmail, More on Gmail and Privacy, http://mail.google.com/mail/help/about_privacy.html (last visited Mar. 11, 2009).

254. *See* Jerry Berman & Deirdre Mulligan, *Privacy in the Digital Age: Work in Progress*, 23 NOVA L. REV. 551, 559-60 (1999); Skok, *supra* note 80, at 65-67. *See generally* Solove, *supra* note 45, at 1092-95 (discussing government acquisition of online data).

some ISPs record URL data (and likely transmit this data to third party advertisers[255]), and some do not.[256] In this environment of uncertainty, a court could find that Internet users do not retain a reasonable expectation of privacy in their URL data, which at least some ISPs likely record and even disclose to third-party advertisers.[257] A court might instead emphasize the objective/normative aspect of the reasonable exptectation test and decide that Internet users have a reasonable expectation of privacy in their URL data regardless of the current (often surreptitious) practice of their ISPs.[258] Alternately, a court could conduct a fact-specific inquiry and

---

255. *Cf. In re* DoubleClick Privacy Litig., 154 F. Supp. 2d 497, 503-04 (S.D.N.Y. 2001) (describing how a website disclosed URL data to a third-party advertiser's server, which in turn produced targeted advertising on the website).

Website-specific URL data is also frequently recorded by the websites themselves, generally via "cookies." *See, e.g.*, *In re* Pharmatrak, Inc. Privacy Litig., 329 F.3d 9, 13-14 (1st Cir. 2003). Aside from the fact that such data is far less revealing than the all-sites web surfing data that ISPs might record, data subpoenaed from an intended third-party recipient is likely not protected under Fourth Amendment law, unless that recipient herself has a Fourth Amendment interest in it. *See Warshak*, 490 F.3d at 471; *see also, e.g.*, *Ex parte* Jackson, 96 U.S. 727, 735 (1877) (stating that postal laws may be enforced by gathering evidence from the recipients of letters, but not by obtaining letters in transit). Note that the government would either have to obtain revealing IP address information that may legally be content from an ISP or have to guess that an Internet user had visited a certain website in order to subpoena such data.

256. Ryan Singel, *Which ISPs Are Spying on You?*, WIRED, May 30, 2007, *available at* http://www.wired.com/politics/onlinerights/news/2007/05/isp_privacy ("AOL, AT&T and Cox all say they don't store ... [users'] URLs at all, while Qwest dodged the question. Comcast, EarthLink, Verizon and Time Warner didn't respond."); *see, e.g.*, Comcast Customer Privacy Notice, *available at* http://www.comcast.com/MediaLibrary/1/2/CM/VanityURL/documents/customerprivacy/PrivacyPolicy_UniLegal_stnd_ENG_spk_comcastcom.pdf.

257. *But cf.* United States v. Forrester, 495 F.3d 1041, 1049 n.6 (9th Cir. 2007) (suggesting that URL data may be constitutionally protected because it reveals underlying communications content).

258. Such an approach may be necessary to avoid a potentially pernicious circularity in the reasonable expectation of privacy test. For instance, Internet users' subjective expectations may be shaped by the intrusive policies of ISPs taking advantage of the current ambiguity in privacy law to disclose their private information to third parties. Allowing ISPs' disclosure of their customers' information, often without their knowledge, to determine the level of constitutional protection such information receives could lead to an erosion in privacy protections for all types of electronic communications. Such an approach would enshrine the practices of early-adopting private entities into the constitutional law governing new technologies. Unless customers are well informed, have access to a variety of competing service providers, and are motivated to contract only with service providers that offer high levels of privacy, the privacy protections they secure from service providers in the early stages of a communications technology's development may be minimal. An overemphasis on the subjective aspect of the reasonable expectation of privacy test may consistently set privacy

base its finding on either the individual privacy policy of the ISP,[259] or whether the ISP combines URL data with personal information.[260] Indeed, some recent cases have suggested that a user's reasonable expectations of privacy should hinge upon her ISP's individual privacy policies, rather than the nature of the information captured.[261] Under this approach, an email system with a policy of never reading or scanning user's emails may create a reasonable expectation of privacy for its users, whereas a system such as Gmail, which informs users that the contents of their emails will be scanned and content-relevant ads placed in the margins of the web page,[262] may destroy any constitutional expectation of privacy. Fourth Amendment protection could depend on the "limited circumstances" in which the privacy policy allows the ISP to access email content,[263] or even on the level of spam protection chosen by the email user.[264]

2. Objective Reasonableness of the Expectations of Internet Users

At first blush, the third party doctrine of *Miller* and *Smith* appears to cover the disclosure to a third party of any information that can be recorded. Because this would also encompass telephone

---

protections for new communications technologies at this minimal level. The author plans to analyze this issue more fully in a subsequent paper.

259. *See* Smith v. Maryland, 425 U.S. 735, 742 (1979) (quoting a phone company policy statement); *see also id.* at 745 (stating that regardless of whether a telephone actually records such information, the fact that it was free to do so indicated the lack of an expectation of privacy).

260. *See* Comcast Customer Privacy Notice, *supra* note 256.

261. *See* Warshak v. United States, 532 F.3d 521, 526-27 (6th Cir. 2008) (en banc) (asserting that case-specific inquiries into individual privacy policies were required in order to assess email users' reasonable expectations of privacy); United States v. Maxwell, 45 M.J. 406, 417 (C.A.A.F. 1996) (describing how the uniquely protective email policy of America Online was a factor that strongly indicated a reasonable expectation of privacy in the instant case).

262. *See* More on Gmail and Privacy, *supra* note 253.

263. *Warshak*, 490 F.3d at 474.

264. Microsoft's Hotmail email service, for instance, allows users to select their level of spam protection. At the highest level, all mail is sent to the junkmail folder, and users select certain email addresses to be sent directly to the inbox. On this setting, no scanning of email content for spam would be necessary. *See* MSN Hotmail Help—Frequently Asked Questions, http://postmaster.msn.com/FAQ.aspx (last visited Mar. 8, 2009).

conversations, a more coherent reading of the Court's rationale is that any information disclosed to a service provider *and used* by that service provider's equipment in the ordinary course of business lacks any constitutional privacy protection.[265] If a court were to read *Smith* in this way, it would not matter whether the information was classified as content. URL data would obviously be disclosed to ISP routing equipment in the ordinary course of business. Email content would be exposed to ISP equipment in the normal course of business, though the content would generally not be used to route the email. On the other hand, spam filters that do not let emails with certain contents pass could easily be conceived of as using the content to route the emails. A court adopting this line of analysis from *Smith* may have to determine whether this "routing distinction" between URL content and email content is significant. Also, a court might distinguish email content from URL content on the basis that the disclosure of email content to spam filters is not intentional disclosure, because the Internet user only perceives herself as disclosing email addresses themselves to the ISP's equipment.[266] The rule resulting from this plausible distinction would be that any routing information, even if it contained content (including search terms), would be considered intentionally disclosed to an ISP for Fourth Amendment purposes. Alternately, the reasonableness of an email user's expectation of privacy might depend upon whether she uses a service like Gmail or whether she is aware that her email provider scans her emails with a spam filter.

3. Disclosure to Automated Systems

The disclosure identified in *Smith*'s objective reasonableness analysis hinges on the idea that phone companies' decisions to automate their routing services do not affect the reasonable expectation of privacy users have in their telephone numbers. Extending this concept to Internet communications would very

265. *See* Mulligan, *supra* note 26, at 1579 (arguing that *Smith* and *Miller* limit their holdings to records in which the third party has an independent interest; the phone company has an interest in phone numbers because it uses them to connect calls).

266. *Id.* at 1580.

likely be necessary if a court were to hold that Internet users intentionally disclose their web surfing information to third parties. Although ISPs might record users' URL data and even link that data with personal information when disclosing it to third party advertising services, the process is, at least in most situations, entirely automated.[267]

4. Four General Options for a Court Applying Smith *to Internet Communications Content*

Based on the above discussion, this Section evaluates several general approaches that a court deciding the constitutional status of Internet communications content might take.

First, a court could adopt the argument that the content of communications as widely used and relied upon as email and web surfing is protected by the Fourth Amendment under the implicitly normative analysis in *Katz*, and that the "third party doctrine" reasoning in *Smith* is limited to noncontent information. Under this approach, determining that a type of Internet communication information is content essentially decides the matter of a reasonable expectation of privacy. URL data, email content, and IP address and file download size information that reveal contents would all be protected under the Fourth Amendment. There is, in fact, some support for such a position in post-*Smith* Supreme Court cases. The Court has stated in several cases that a normative judgment under the "objectively reasonable" prong of *Katz* is generally the most important determinant of Fourth Amendment protection, and trumps any contrary conclusion based on subjective expectations of privacy.[268] Further, in one recent decision, the Court indicated that,

267. *See In re* DoubleClick Privacy Litig., 154 F. Supp. 2d 497, 503-04 (S.D.N.Y. 2001) (describing how a third-party advertising server automatically couples URL data with personal profile data and uses "a complex set of algorithms" to "determine which advertisements it will present to the user"); Orin S. Kerr, *Digital Evidence and the New Criminal Procedure*, 105 COLUM. L. REV. 279, 294 n.58 (2005) (describing how ISPs typically aggregate user information and how human ISP employees may find it extremely difficult to identify users personally).

268. *See* Hudson v. Palmer, 468 U.S. 517, 525 & n.7 (1984) (stating that the objective prong of *Katz* is of primary importance and that the objective inquiry contains a normative element); United States v. White, 401 U.S. 745, 751-52 (1971) (same); *see also* Smith v. Maryland, 442 U.S. 735, 740 n.5 (1979) (stating that when a subjective analysis would be insufficiently

at least in some special cases, individuals may retain a reasonable expectation of privacy in information voluntarily given to a third party if their reasonable expectation is that the party will not disseminate that information.[269] A court might determine that Internet users have a similar (albeit empirically incorrect) reasonable expectation as to their Internet communications data.

Second, a court could apply *Smith*'s third party doctrine analysis but distinguish some or all Internet communications content from telephone numbers. Subjectively, Internet communications content might easily be distinguished on the basis that Internet users, unlike telephone users, are likely unaware that their ISP may be keeping records of their web surfing data or making copies (in the course of transmission) of their emails.[270] A court might point to the intimate and private nature of emails and of some Internet activity, which can encompass everything from Internet voting[271] to viewing pornography,[272] as evidence that Internet users maintain at least a subjective expectation of privacy in their Internet use. Finally, a court might rely on the policy of the ISP in the instant case and determine that the policy is sufficiently protective as to create a subjective expectation of privacy in web surfing data.[273]

---

protective of Fourth Amendment values, courts should engage in a normative inquiry based on the objective prong).

269. *See* Ferguson v. City of Charleston, 532 U.S. 67, 78 (2001) (finding a constitutional privacy interest in diagnostic medical tests in which a patient's "reasonable expectation of privacy ... is that the results of those tests will not be shared ... without her consent").

270. *See, e.g.*, JOSEPH TUROW, AMERICANS & ONLINE PRIVACY: THE SYSTEM IS BROKEN (2003), *available at* http://www.asc.upenn.edu/usr/jturow/internet-privacy-report/36-page-turow-version-9.pdf (reporting the results of several polls indicating that Americans do not understand the extent to which their online activities are recorded).

271. Rebekah K. Browder, *Internet Voting with Initiatives and Referendums: Stumbling Toward Direct Democracy*, 29 SEATTLE U. L. REV. 485, 497 (2005) (describing five examples of Internet voting, in Alaska, Arizona, Washington, and Michigan, and by military personnel living abroad). Note that URL records could reveal for whom an Internet user voted, although this would only occur if the URL activated after the user clicked on Candidate A's name differed from the URL activated after clicking on Candidate B's name.

272. *See* David Crary, *Battle Brews as Porn Moves into Mainstream*, S.F. GATE, Apr. 1, 2006, *available at* http://www.sfgate.com/cgi-bin/article.cgi?f=/n/a/2006/04/01/national/a131130S42.DTL (reporting that 40 percent of Internet users in the United States visit pornographic sites each month).

273. *Cf.* United States v. Maxwell, 45 M.J. 406, 417 (C.A.A.F. 1996) (finding that America Online's protective email policy contributed to a reasonable expectation of privacy in the contents of a user's email).

As far as the objective prong of the analysis, a court ruling on the Fourth Amendment status of URLs or email content might distinguish *Smith* on the basis of its "automation" rationale.[274] The *Smith* Court's argument that disclosures to automated telephone equipment destroyed any reasonable expectation of privacy could be limited to the telephone context—the Court's argument is based largely upon the fact that human operators once obtained the numbers themselves[275] (and possibly that human employees in the billing department might view the information as well).[276] This is certainly not the case for email or URL data, which Internet users can reasonably expect will never be perceived by a human being at any point (unless perhaps that human being is a government investigator).[277] Indeed, a court might incorporate the same arguments about the importance of the normative aspect of the objectively reasonable prong as a court taking the first approach, holding that Internet users have a reasonable expectation of privacy in their web surfing data because it would be normatively unacceptable for their ISPs' human employees ever personally to observe users' content-revealing data.[278] Alternately, a court might find that only email content is protected by the Fourth Amendment, distinguishing email content from URL content on the basis of the distinctions traced above.[279]

Third, a court might decide, on the basis of the rationales discussed in Parts IV.A-B, that the principles of *Smith* suggest that all web surfing and email communications content is disclosed to a third party, and therefore users cannot have an objectively reasonable expectation of privacy in such data. Such an approach would

---

274. *See* Simmons, *supra* note 188, at 1338-39 (characterizing *Smith*'s automation rationale as an afterthought that is now highly problematic given modern communications technology).

275. Smith v. Maryland, 442 U.S. 735, 744-45 (1979).

276. *Id.* at 742.

277. *Cf.* Warshak v. United States, 490 F.3d 455, 473 (6th Cir. 2007) (holding that individuals maintain a reasonable expectation that ISPs do not generally access the contents of emails); *In re* DoubleClick Privacy Litig., 154 F. Supp. 2d 497, 504 (S.D.N.Y. 2001) ("DoubleClick's targeted advertising process is invisible to the user. His experience consists simply of requesting the Lycos.com homepage and, several moments later, receiving it complete with banner advertisements.").

278. *See supra* text accompanying notes 268-69.

279. *See* discussion *supra* Part IV.A.2.

likely rely on the language in *Smith* indicating that the individual recording policies of telephone companies are ultimately irrelevant to the Fourth Amendment analysis and establishing a bright line rule that telephone users have no privacy interest in their phone numbers regardless of whether they are monitored by the phone company.[280] Though predicting the course of the constitutional law of Internet communications is extremely difficult, it appears that most courts are likely to choose this third approach in upcoming Internet surveillance cases, given that the language of *Smith* suggests in many places a third party doctrine with an extremely broad scope.[281]

Finally, a court might take another approach, one that would further muddle an already unclear body of Fourth Amendment Internet surveillance law. That is, a court might fail to recognize that URLs and other content-revealing data should be treated as content in Fourth Amendment analysis. Under this approach, a court would simply determine that URLs are routing information, and thus not content, and would stop the Fourth Amendment analysis there, citing *Smith*'s content/noncontent distinction. Not only would this outcome be incorrect as a matter of law, but unlike the first three approaches, it would permit courts to dispose of Internet surveillance cases without shedding any light on the enduring meaning of *Smith*, a precedent which will surely only become more important as new forms of communication proliferate and become widely used.

Perhaps the most significant benefit of a clearer understanding of content and noncontent in Internet communications would be the way it would force courts to decide the difficult but important questions that persist in Internet surveillance law. Courts would be forced to confront the analysis of reasonable expectations of privacy head on, either clearly determining that some forms of Internet

---

280. *Smith*, 442 U.S. at 745 (expressing concern that case-by-case evaluation of disclosure would "make a crazy quilt of the Fourth Amendment").

281. For cases applying *Smith* and *Miller*'s third party doctrine broadly, see United States v. Jacobsen, 466 U.S. 109, 117-22 (1984) (finding no reasonable expectation of privacy in the contents of a box accidentally exposed to a third party carrier); United States v. Cormier, 220 F.3d 1103, 1108 (9th Cir. 2000) (finding no reasonable expectation of privacy in motel records); United States v. Daccarett, 6 F.3d 37, 50 (2d Cir. 1993) (holding that defendants had no privacy interest in personal records conveyed by their bank to another party).

communications content are protected by the Fourth Amendment, or clearly determining that there can be no reasonable expectation of privacy in any form of Internet communications content. Either outcome would be preferable to the current situation, in which the government could obtain web surfing and email communications content entirely unchecked by the Fourth Amendment *and* Internet users are largely unaware that their web surfing data and even email content receive dramatically less protection than the content of their phone calls.

Fortunately, the few courts to have considered the content status of URLs thus far have indicated without deciding that URL data might be treated as content, especially when the URL contains search terms.[282] Given this early precedent, it seems somewhat likely that future courts squarely confronting the issue will realize that URLs reveal the entirety of their underlying content, and thus must be treated like content themselves. The outlook is not so good, however, for web surfing communications that can be more easily analogized to envelope information: IP addresses and file size downloads. Courts commenting on the status of IP addresses have thus far erroneously assumed that they have no potential to reveal more about the content of communications than do phone numbers.[283] The volume of website transfers was only briefly mentioned in *Forrester*, in which the Ninth Circuit simply (and erroneously) assumed that such data could not reveal communications content.[284] Future courts might also be led astray by the same inaccurate analogy to the weight of mailed packages, which unlike file size downloads can reveal nothing specific about package contents.[285] Though the prospects are currently unfavorable, courts following the logic of privilege law and *Karo* and *Kyllo* (and the logic of their own arguments about URLs) should recognize that IP addresses and file download sizes have the potential to reveal just

282. *See* United States v. Forrester, 495 F.3d 1041, 1049 n.6 (9th Cir. 2007); *see also In re* Application of the U.S. for an Order Authorizing the Use of a Pen Register, 396 F. Supp. 2d 45, 49 (D. Mass. 2005) (discussing web surfing information that may reveal content in the context of the ECPA).

283. *See Forrester*, 495 F.3d at 1049; *In re Application*, 396 F. Supp. 2d at 48.

284. *Forrester,* 495 F.3d at 1049.

285. *See, e.g.*, *Ex parte* Jackson, 96 U.S. 727, 733 (1877) (stating that the outward form and weight of packages received no constitutional protection).

as much about web surfing communications content as URLs do in some situations.

B. Internet Communications in Statutory Law

In contrast with constitutional law, the ECPA clearly sets the level of protection afforded to URLs and other web surfing information deemed content.[286] And if, as seems fairly likely, courts determine that Internet communications content is not protected by the Fourth Amendment, the ECPA will be the only source of legal protection for such information. Yet, as mentioned above,[287] the fact that URLs are contents as a matter of constitutional law or general semantics does not necessarily dictate that they are contents within the meaning of the ECPA,[288] although there is legislative history strongly suggesting that Congress intended the statutory definition to track the constitutional definition.[289] Courts examining the plain language of the statutes, or the legislative intent behind them, might conclude that URL or other website routing information constitutes noncontent information under the statute. The question remains largely unresolved in the few cases dealing with the acquisition of web surfing data under the ECPA. Those cases have thus far explicitly determined only that URLs containing search terms constitute contents under the ECPA.[290] Courts dealing with the ECPA have yet to recognize that all URLs, and not just URLs containing search terms or identifiable file names, reveal content. This section examines whether all URLs do in fact meet the definition of "contents" in the ECPA.

The ECPA incorporates a broad definition of "contents," defining them as "any information concerning the substance, purport or meaning of [a] communication."[291] Contents are referred to again in

---

286. *See* discussion *supra* Part I.B.

287. *See supra* text accompanying notes 84-87.

288. *See, e.g.*, United States v. Bach, 310 F.3d 1063, 1066 (8th Cir. 2002) ("[I]nquiries under [the ECPA] and the Constitution are separate and distinct.").

289. *See infra* notes 298-99 and accompanying text.

290. *See In re* Pharmatrak, Inc. Privacy Litig., 329 F.3d 9, 16, 18 (1st Cir. 2003); *In re* Application of the U.S. for an Order Authorizing the Use of a Pen Register, 396 F. Supp. 2d 45, 49 (D. Mass. 2005).

291. 18 U.S.C. § 2510(8) (2006).

the Pen Register Act, which defines a pen register, which the government can operate with a mere court order,[292] as "a device or process which records or decodes dialing, routing, addressing, or signaling [(DRAS)] information transmitted by an instrument or facility from which a wire or electronic communication is transmitted, provided, however, that such information shall not include the contents of any communication."[293] The definition of pen registers very likely incorporates the definition of "contents" from § 2510, and, as a result, even routing information must be excluded from the permissible scope of pen registers if it also constitutes "contents."[294]

The plain text of § 2510(8)'s broad definition of contents[295] seems to incorporate, among other things, any information that discloses the underlying content of Internet communications. Thus far, courts applying the ECPA have interpreted "concerning" to encompass information that reveals or relates in any way to substance or subject matter.[296] Under this interpretation of "concerning," all URLs concern the substance of Internet electronic communications; indeed they reveal the substance entirely. Further, "contents" include any information concerning "the purport or meaning" of electronic communications. Information such as IP addresses that reveal which web pages were visited or website-specific total volume transmission data that reveals pages visited or files downloaded concern (and reveal) the purport of the underlying web surfing communications. Courts interpreting the ECPA should hold that such information, despite its status as routing information, is "content" within the meaning of the statute.

The history of the ECPA and the Pen Register Act amendments of 2001 also suggests that "contents" should be interpreted broadly

---

292. See *supra* notes 61-62 and accompanying text for a discussion of the ease with which the government can obtain such court orders.

293. 18 U.S.C. § 3127(3) (2006).

294. *See In re Application*, 396 F. Supp. 2d at 47 (stating that the government is not entitled to acquire DRAS information such as URLs containing search terms if such information reveals the contents of a communication); *see also* 18 U.S.C. § 3127(3)-(4).

295. 18 U.S.C. § 2510(8) (defining "contents" as "any information concerning the substance" of communications).

296. *In re Application*, 396 F. Supp. 2d at 48 (holding that subject headers are content because they reveal subject matter); *In re* Pharmatrak, Inc. Privacy Litig., 329 F.3d 9, 18 (1st Cir. 2003) (finding that search terms are information).

—or at least, interpreted as it would be for the purposes of constitutional law. The legislative history of the original Wiretap Act states, referring to "contents," that "[t]he privacy of the communication to be protected is intended to be comprehensive."[297] The current definition of "contents" in the ECPA was adopted in 1986 to clarify the legality of telephone pen registers and to mirror the content/envelope distinction of *Smith*.[298] Congress's intent to mirror the constitutional definition is made even clearer in the legislative history of the USA PATRIOT Act's amendments. The relevant House Report states that the amendments "reinforce the statutorily prescribed line between a communication's contents and noncontent information, a line identical to the constitutional distinction drawn by the U.S. Supreme Court in *Smith v. Maryland*."[299] As discussed above, URLs should be considered contents as a matter of constitutional law, regardless of whether they are protected under *Smith*'s third party doctrine.[300]

There is nothing in the legislative history of the USA PATRIOT Act involving the amendment of the Pen Register Act to indicate that Congress did not intend the definition of pen registers to incorporate fully the broad definition of contents from § 2510(8). Instead, Congress apparently intended that, "just as" under the 1986 version of the Pen Register Act, a pen register order under the amendments "could not be used to intercept the contents of communications protected by the wiretap statute."[301] To be sure, the House Report on the Act seems to contemplate DRAS information and content information as mutually exclusive categories.[302] But as an example of non-DRAS, content information, the Report mentions URL search terms and even the portion of a URL including "the name of a requested file or article" as an example of content information that could not lawfully be intercepted by a pen register;[303]

297. S. REP. NO. 90-1097, at 91 (1968).

298. *See id.*; H.R. REP. NO. 99-647, at 34 (1986).

299. H.R. REP. NO. 107-236, pt.1, at 53 (2001) (citation omitted).

300. *See supra* Parts III, IV.

301. H.R. REP. NO. 107-236 pt.1, at 53; *see* 147 Cong. Rec. S10372 (Oct. 9, 2001) (statement of Sen. Leahy) (describing how the Bush administration "agreed that the definition [of pen registers] should expressly exclude the use of pen/trap devices to intercept content").

302. *See* H.R. REP. NO. 107-236, pt.1, at 53.

303. *Id.*

of course, any such information appearing in a URL is used to route web surfing communications.[304] The Report also states that email subject lines are "clearly" content,[305] further suggesting an endorsement of the idea that envelope information that reveals subject matter should be interpreted as content under the Act.

Given the plain language of the definition of "contents" and the general legislative intent that the language be interpreted broadly, URLs and other content-revealing routing information should be treated as content under the ECPA. As far as other revealing information, Congress apparently did not contemplate the possibility that IP addresses or file size information could reveal the underlying content of web surfing communications—but then, no court has done so yet either. The fact remains that such information, if it reveals the contents of web surfing communications, should itself be considered "contents" under the broad definition of the ECPA.

CONCLUSION

The surveillance of Americans' Internet activities is likely to increase as both general Internet use and Internet-based crime become even more widespread. Meanwhile, confusion about the constitutional status of Internet communications, coupled with the lack of an exclusionary rule for illegally intercepted electronic communications, will continue to impede the development of clear Internet surveillance law.[306] The need for courts and scholars to clarify further the constitutional and statutory protections afforded to web surfing data is urgent. Courts should recognize that web surfing data that reveals the underlying content of communications should itself be treated as content. Not only would this be the correct legal outcome, it would also force courts to confront the difficult but necessary questions inherent in applying *Smith* and the ECPA to Internet communications. It would also provide courts with

304. Again, search terms are used as an example of protected content; the Report does not engage in a comprehensive review of web surfing communications. There is no mention anywhere of IP addresses or file download sizes. In other words, the Report cannot be read to indicate that Congress intended that only URL search terms should be protected. *See id.*

305. *Id.*

306. *See* Freiwald, *supra* note 7, ¶ 34; Kerr, *supra* note 8, at 807.

a guiding principle that could be applied in the future to establish the legal status of as-yet-undeveloped communications technologies.

Although there is reason to expect that courts will recognize that content-revealing web surfing communications are protected under the ECPA, and reason to doubt that courts will find they are protected under the Constitution, the actual course that Internet surveillance law will take remains extremely difficult to predict. What is certain is that misunderstanding the legal status of content-revealing communications data would lead courts to underprotect such data and to make a muddle of Internet surveillance law for years to come. Getting the law of Internet communications right will likely not be easy for courts—even the Supreme Court has historically had serious difficulties in applying old laws to new technologies. In its 1928 decision, *Olmstead v. United States*,[307] the Court failed to adapt Fourth Amendment law to the then-new technology of telephone wiretaps—a mistake that the court did not correct until nearly forty years later.[308] In the long interim, the FBI pervasively monitored the private communications of American citizens, secretly recording the conversations of countless politicians, activists, celebrities, writers, professors, and Supreme Court Justices.[309] We can only hope that it does not take another forty years for courts to determine that the contents of Internet communications are worthy of legal protection.

307. 277 U.S. 438 (1928).

308. *Olmstead*'s holding that the Fourth Amendment did not protect telephone conversations was overturned in the 1967 case *Katz v. United States*, 389 U.S. 347, 353 (1967).

309. *See* ALEXANDER CHARNS, CLOAK AND GAVEL: FBI WIRETAPS, BUGS, INFORMERS, AND THE SUPREME COURT (1992); Solove, *supra* note 77, at 1273-74.

APPENDIX

This appendix discusses several unconventional URL forms and website configurations that exist on the Internet. These configurations differ from conventional URL and website configurations in that they allow Internet users to access website content without generating a corresponding URL. The use of these unconventional forms is relatively limited. Further, website content that does not correspond to a specific URL is accessible only from web pages that do correspond to a specific URL. These unconventional forms do not significantly impact the standard URL-based practice of Internet navigation, nor does their existence alter the fact that a URL history reveals the content requested by and sent to an Internet user. These caveats aside, this appendix describes the extent to which these unconventional forms themselves may reveal or obscure Internet communications content.

A. AJAX Programming

AJAX (Asynchronous JavaScript and XML) applications are web applications that update portions of a web page without changing the underlying page (or generating a new URL).[310] For instance, a user visiting a web page could click on a streaming video file and request a partial page update via an AJAX application. The user would then receive the video from the web server without having to wait for an entirely new page to load. AJAX is typically used to create Rich Internet Applications (RIAs), including registration forms, web video, and interactive web applications (like web-based email and word processing) designed to act like traditional desktop applications.[311] Because the AJAX content loads without accessing a new web page, it is not associated with a unique URL. However, the original page, where the user finds the AJAX content, is associated with a URL. It is very likely that this original web page would reveal to an observer the subject matter of the user's

310. PAUL J. DEITEL & HARVEY M. DEITEL, AJAX, RICH INTERNET APPLICATIONS, AND WEB DEVELOPMENT FOR PROGRAMMERS 412-14 (2008).

311. *Id.* at 412-15.

communication with the website—that is, the URL of a website about computer programming manuals which contains AJAX content[312] would reveal that the user was reading about computer programming, although it would not be known for certain whether the user accessed the AJAX content or not. Further, in situations in which the AJAX application allows the user to receive one of several partial-page updates, the government may simply infer (perhaps based on other corroborating evidence) that the user likely viewed all of them—an inference that may very well hold up in court.[313]

B. Encrypted URLs

When a user completes and submits a web form (on, say, Amazon.com) the URL containing the input information is typically encrypted. Because Amazon has a key to decrypt the content, it can determine what was sent.[314] Third parties who intercept the URL, however, cannot determine the content of the webpage without the key. Similarly, websites may encrypt certain URLs in order to prevent hackers from gaining unauthorized access to secure web pages.[315] Encrypted URLs generally consist of an unencrypted domain name (which directs the request to the appropriate web server) and an encrypted "locator," "path," or "query string" (the final portion of a URL that refers to the particular directory and file sought).[316]

Evaluating the potential for encrypted URLs to reveal or obscure content is fairly straightforward. Obtaining an encrypted URL would generally allow an observer to obtain the domain name of the website visited by an Internet user—usually this would be sufficient to reveal the subject matter of the content the user requested and

---

312. See *id.* at 420 for an example of such a website.

313. *See* cases cited *supra* note 228.

314. HAROLD F. TIPTON & MICKI KRAUSE, INFORMATION SECURITY MANAGEMENT HANDBOOK 1131-32 (6th ed. 2007).

315. *Id.*

316. *See* ROLF OPPLIGER, SECURITY TECHNOLOGIES FOR THE WORLD WIDE WEB 337-38 (2d ed. 2003); Daniel Estermann, *URL Encryption, Phion Airlock Techzone* (June 25, 2008), https://techzone.visonys.com/url-encryption (last visited Mar. 11, 2009); Software Information Center, The Components of a URL, http://publib.boulder.ibm.com/infocenter/cicsts/v3r1/index.jsp?topic=/com.ibm.cics.ts31.doc/dfhtl/topics/dfhtl_uricomp.htm (last visited Mar. 11, 2009).

viewed. If the domain name somehow failed to reveal the subject matter communicated, then the encrypted portion of the URL would (if the encryption works properly) reveal no content whatsoever. Intercepting and entering the encrypted URL into a web browser would not produce the desired web page; the web host's encryption software would detect that the URL had been manipulated by a third party and would block the request.[317] Of course, if the government had reason to be interested in which specific encrypted files a user accessed at a given domain, it might subpoena the web host's encryption key, allowing it to translate the encrypted URL and to determine where it led.

C. Website Frames

Frames are an HTML layout feature that allows a web host to display more than one HTML document in the same web page.[318] A user visiting a website designed with frames might click on a menu item that directs them to text in a different frame, and then scroll through the text without altering the original page or changing the original URL.[319] Frames have become increasingly hard to find on the Internet—there are numerous technical and aesthetic disadvantages to using frames to design a website, and as a result they have fallen out of favor with many web designers.[320]

As with AJAX, the use of frames would allow the user to access some content without generating a unique URL. The original page containing the frame menu, however, is associated with a URL. As

317. *See* Estermann, *supra* note 316.

318. H.M. DEITEL, P.J. DEITEL & A.B. GOLDBERG, INTERNET & WORLD WIDE WEB: HOW TO PROGRAM 129-34 (3d ed. 2004); JAMES H. PENCE, HOW TO DO EVERYTHING WITH HTML AND XHTML 180-81 (2d ed. 2003).

319. PENCE, *supra* note 318, at 180-81.

320. *See, e.g.*, *id.*; Don't Use Frames To Design Your Website, http://www.hobo-web.co.uk/tips/41.htm (last visited Mar. 11, 2009); Shirley E. Kaiser, *To Frame or Not To Frame: That is the Question*, Website Tips.com, Nov. 2006, http://websitetips.com/articles/html/frames/ (last visited Mar. 11, 2009); Utah State University, HTML Frames, http://ocw.usu.edu/Instructional_Technology/producing-distance-education-resources/resource16.html (last visited Mar. 11, 2009); WWW FAQs: Should I Use Frames on My Website?, http://www. boutell.com/newfaq/creating/framesbad.html (last visited Mar. 11, 2009); *see also* PENCE, *supra* note 318, at 181-82 (describing the advantages and disadvantages of using frames).

with AJAX, it is very likely that this original web page would reveal to an observer the subject matter of the user's communication with the website. This is especially likely because websites that use menu frames often involve a narrow subject matter—for various practical and design reasons,[321] a general interest news website is very unlikely to employ a menu frame with associated content. For any web page containing a menu with associated frames, the government might also infer that a visitor viewed all of the frames, particularly if the menu is simple or the subject matter narrow.[322]

D. A Note on Dynamic URLs

"Dynamic URL" refers to a URL that directs a user to the results of a search or other query of a website's database.[323] For instance, the URL generated by entering search terms on Google is a dynamic URL, and the web page displayed reflects the results of the search of Google's database.[324] Dynamic URLs are widely used, and in practice they function much like "static" URLs, which direct users to a standard website with content encoded in the HTML coding language.[325] The primary difference between the two is that a web host would update a web page associated with a static URL by changing its HTML code, while the host would update a web page associated with a dynamic URL by changing the data in the database.[326] Both kinds of URLs correspond to a specific web page display. An observer who intercepts a dynamic URL can instantly obtain the content sent to the user by entering the URL into a web browser, just as with a static URL. Many web pages associated with dynamic URLs direct a user to a specific article or file on the database, which is never updated. Websites with dynamic URLs are also archived on the Internet Archive website, so, as with static

321. *See supra* note 320.

322. *See* cases cited *supra* note 228.

323. JAIMIE SIROVICH & CRISTIAN DARIE, PROFESSIONAL SEARCH ENGINE OPTIMIZATION WITH PHP: A DEVELOPER'S GUIDE TO SEO 39-40, 42-43 (2007); SHARI THUROW, SEARCH ENGINE VISIBILITY 151-52 (2003).

324. *See supra* text accompanying notes 137-39.

325. *Id.*

326. *See* SIROVICH & DARIE, *supra* note 323, at 42-43; THUROW, *supra* note 323, at 150-51; Dynamic URL (Jan 8, 2004), http://www.webopedia.com/TERM/D/dynamic_URL.html (last visited Mar. 11, 2009).

URLs, an observer could often view the content that an Internet user viewed at a specific site on a given date.[327] Further, as discussed above, if the content of a web page could not be ascertained using the above methods, the government could probably subpoena the web host for its archived database records—it could thereby obtain the specific information displayed on a given date.[328]

327. *See* Internet Archive Frequently Asked Questions, http://www.archive.org/about/faqs.php#11 (last visited Mar. 11, 2009).

328. *See supra* text accompanying note 158.