THE GOOGLING OF ONLINE PRIVACY: GMAIL, SEARCH-ENGINE HISTORIES AND THE NEW FRONTIER OF PROTECTING PRIVATE INFORMATION ON THE WEB

by
Matthew A. Goldberg[*]

When Congress passed the Stored Communications Act ("SCA") in 1986, the Internet as it presently exists was barely imaginable. In the nearly twenty years since its passage, courts, scholars and privacy advocates have struggled mightily to apply the SCA to continuously evolving online technology. Two of the more problematic technologies, from the standpoint of applying the SCA, are Web-based e-mail and search engines. In 2004, now publicly-traded Internet giant Google underscored the importance of understanding the privacy implications of these two technologies when it launched its revolutionary Gmail Web-mail service, a technology that is capable of keeping an ongoing record of the contents of users' e-mail. Google's technologies give it the potential to maintain unprecedented electronic dossiers of personal information about users, which if not protected would be sought after by all manner of third-party marketers. This Comment explores several key privacy issues surrounding Google's Web-mail and search services, including the extent to which the SCA protects users' information from disclosure by Google to third-party marketers. Also discussed are nascent state law attempts to regulate these Web services and the role that the common law of contracts might play in this otherwise statutorily dominated realm.

[*] Student, Lewis & Clark Law School, J.D. expected 2005; B.A. State University of New York, Binghamton; M.A. University of Massachusetts, Dartmouth. The author wishes to thank Professor Lydia Pallas Loren for her comments and advice.

I. INTRODUCTION

Google is a ubiquitous presence in the online world. In May 2004, 36 percent of all Web users in the United States (51 million users) visited Google's Web site.[1] This figure is in a way a conservative estimate of Google's reach because Google's search technology powers the search engines of other popular sites such as AOL.[2] The Time Warner Network, of which AOL is a major part, reached 58 percent of all U.S. Web users in May 2004 (almost 83 million users).[3] Back in February 2003, Google already handled a staggering 250 million Web searches per day, the total of all searches submitted through its own site as well as the sites of its partners.[4] Google's status as a giant among Internet companies was indelibly fixed when the company completed a successful initial public offering in August 2004. At the start of its first day of trading, Google was worth upwards of $23 billion, roughly equivalent to the value of General Motors.[5]

Beginning in the spring of 2004, Google began to receive even more attention than usual when it launched a beta test of a revolutionary Web-based e-mail service called Gmail.[6] The so-called soft launch of Gmail turned out to be one of the most controversial product launches in the roughly ten-year history of the commercial Web and placed Google at the center of an, at times, fierce debate over online privacy. A global coalition of privacy advocacy groups has rallied against Gmail service and Google itself.[7] The press has overflowed with divergent analyses of Google and online privacy. The debate has seen the passage of the nation's first state law governing online privacy

---

[1] *U.S. Web Usage and Traffic, May 2004,* CLICKZ NETWORK (June 16, 2004), *at* http://www.clickz.com/stats/big_picture/traffic_patterns/article.php/3369221.

[2] Danny Sullivan, *Who Powers Whom? Search Providers Chart*, SEARCH ENGINE WATCH (July 23, 2004), *at* http://searchenginewatch.com/reports/article.php/2156401.

[3] *U.S. Web Usage and Traffic, May 2004, supra* note 1, *at* http://www.clickz.com/stats/big_picture/traffic_patterns/article.php/3369221.

[4] Danny Sullivan, *Searches Per Day,* SEARCH ENGINE WATCH (Feb. 25, 2003), *at* http://searchenginewatch.com/reports/article.php/2156461.

[5] *Google Closes Up 18%*, CBSNEWS.COM, Aug. 19, 2004, *at* http://www.cbsnews.com/stories/2004/08/17/tech/main636561.shtml.

[6] *See* http://gmail.google.com (homepage of Gmail service).

[7] *See* Press Release, Privacy Rights Clearinghouse, Thirty-One Privacy and Civil Liberties Organizations Urge Google to Suspend Gmail (Apr. 19, 2004), *at* http://www.privacyrights.org/ar/gmailletter.htm (open letter to Google regarding its proposed Gmail service).

policies[8] and the partial passage of a state law specifically inspired by one of Gmail's most unique and divisive features.[9]

In addition to raising questions about the efficacy of state regulation of the Internet, the debate over the privacy implications of Google's current and future suite of services also has underscored the importance of ascertaining the extent to which such services are governed by existing federal statutes protecting the privacy of electronic communication, particularly the Stored Communications Act of 1986. Another critical component of solving the Google privacy puzzle is developing an understanding of the enforceability of certain aspects of Web site privacy policies and terms of use agreements. In the absence of federal or state statutory protections, arguments based on the common law of contracts may provide an avenue of last resort for users seeking to protect their privacy in the online world as it evolves at the hands of Google and the companies following its lead.

This Comment proceeds as follows: Part II introduces the functionality of the Gmail service and the resulting privacy concerns raised by the service's unique features. Part II also provides an overview of the somewhat equivocal stance Google has taken with respect to these privacy issues, which has underscored the need to determine the extent to which the Stored Communications Act as well as state contract law might offer legal protections to users. Part III begins with an analysis of the so-called Gmail Bill, which the California Senate passed in the spring of 2004. After considering the potential weaknesses of the Bill from the standpoint of regulating Google's broad product offerings, Part III addresses at length the applicability of the Stored Communications Act to the type of Web services provided by Google.

II. GMAIL, SEARCH HISTORIES AND PRIVACY ISSUES

The three main attributes of Gmail that distinguish it from other popular Web-mail services like Hotmail are (1) a method of delivering contextually relevant advertising to users based on an ability to scan e-mail to determine its meaning; (2) an unprecedented amount of online storage space; and (3) the ability to index and make searchable a user's e-mail in a way commensurate with Google's expertise in indexing and searching the Web.[10] The first of these features, the delivery of contextually relevant advertising alongside a user's e-mail based on scanning for meaning, ignited the privacy backlash (with the second, the practically unlimited storage, adding fuel to the fire).

---

[8] *See, e.g.*, Stefanie Olsen, *California Privacy Law Kicks In*, CNETNEWS.COM, July 6, 2004, *at* http://www.news.com.com/2100-1028-5258824.html (describing passage of California Online Privacy Protection Act). For discussion of this law, see *infra* text accompanying notes 26-28 .

[9] *See, e.g.*, Michael Bazeley, *Senate OKs Limits on E-mail Data*, SILICONVALLEY.COM (May 28, 2004), *at* http://www.siliconvalley.com/mld/siliconvalley/8781969.htm (describing California Senate's passage of the so-called Gmail Bill). For an at length discussion of this Bill, see *infra* Part III.

[10] *See, e.g.,* BRAD TEMPLETON, PRIVACY SUBTLETIES OF GMAIL, *at* http://www.templetons.com/brad/gmail.html (Aug. 23, 2004) (describing noteworthy aspects of Gmail service).

Google's Gmail service employs a computer program to scan a user's incoming and outgoing e-mail and extract the semantic thrust of the communications. Gmail "examines the entire content of the e-mail message including the header and addressing information in order to derive the 'concepts' contained in the e-mail."[11] Once these concepts are derived, the service can display a contextually relevant advertisement on the same Web page as the e-mail being viewed by the user.

Some critics of the content-scanning technology have found it unpalatable to imagine their private e-mail exchanges being examined and analyzed in this way. One writer described this uncomfortable feeling as the "ickiness factor," an idea that incorporates people's unease with the perceived transgression of normative boundaries and the fear that their private information might be made available to an "undesired audience" or "authorities with power over the individual."[12]

More specifically, critics of Gmail worry that the scanning technology puts Google in a position where it could maintain an ongoing record of the subject matter contained in all of a user's incoming and outgoing e-mail. Such a record could contain "a log of which ads went to which users" as well as "a record of keywords that appear often in an individual's e-mail."[13]

There is little doubt that Google's system technically allows for the creation of such profiles. An analysis of Google's Gmail patent by the Electronic Information Privacy Center revealed that the "content extraction" system was designed to incorporate information gleaned from sources including past e-mails sent and received by users of the service.[14] The scanning of e-mails does not "take place in isolation."[15] Rather, as suggested by the description in the patent, the technology "requires a substantial supply chain of directory structures, databases, logs, and a long memory," which among other things keep "auditing trails of the ad text" delivered.[16]

As the privacy implications of Gmail continued to be considered following the launch of the beta test, privacy advocates realized that the existence of a content profile derived from a user's e-mail activity was just the tip of a much larger privacy iceberg. Not only could Google keep records on the subject matter of a user's e-mail, but it could also correlate this information with a user's search history (the list of terms a user has typed into Google's search engine over time), creating "a giant dossier of all your personal information in a central place."[17]

---

[11] ELECTRONIC PRIVACY INFORMATION CENTER, GMAIL PRIVACY PAGE, *at* http://www.epic.org/privacy/gmail/faq.html (last modified Aug. 18, 2004).

[12] Posting of Danah Boyd to Zephoria Ipseity, *at* http://www.zephoria.org/thoughts/archives/2004/04/index.html (Apr. 14, 2004).

[13] Kim Zetter, *Free E-Mail With a Steep Price?*, WIRED NEWS (Apr. 1, 2004), *at* http://www.wired.com/news/business/0,1367,62917,00.html.

[14] ELECTRONIC PRIVACY INFORMATION CENTER, *supra* note 11, *at* http://www.epic.org/privacy/gmail/faq.html (last modified Aug. 18, 2004).

[15] Press Release, *supra* note 7, *at* http://www.privacyrights.org/ar/gmailletter.htm.

[16] *Id.*

[17] TEMPLETON, *supra* note 10, *at* http://www.templetons.com/brad/gmail.html (*Privacy Subtleties of GMail*).

Each Web browser accessing Google is associated with a unique identifying number, which resides in a "cookie" file located on the computer running that copy of the browser. This cookie lets Google know that the same browser is accessing the site even when the computer on which the browser is running is connected to the Internet via different IP addresses (as it would be if a laptop user connected to the Internet sometimes from home, sometimes from the office, and sometimes while traveling). But while Google can recognize that a user conducting a search is doing so via a browser that has previously visited the site, and as a result can amass a search history for that browser, Google cannot associate that browser with a particular human being.[18] However, because the cookie that identifies a particular browser to Google so it can keep a log of a search history is the same one that tells Google that a given user is logging on to her Gmail account, Google "retains a powerful ability to create incredibly detailed profiles on users" who use both Google's search engine *and* Gmail.[19] Now, in addition to knowing a user's name and the subject matter of all of her e-mail, Google can associate with this information every search term she has ever entered into Google's search engine.[20]

The Gmail Privacy Policy states that Google "will *never* rent, sell or share information that personally identifies you for marketing purposes without your express permission."[21] Google also has stated that it does not reveal or "share . . . email content[] with any third parties."[22] Moreover, at the Computers, Freedom and Privacy conference in April 2004, Google said that it had no plans to correlate e-mail and searches.[23]

---

[18] Danny Sullivan, *Search Privacy at Google & Other Search Engines*, SEARCH ENGINE WATCH (Apr. 2, 2003), *at* http://www.searchenginewatch.com/sereport/article.php/34721_2189531 (functioning of search engine cookies).

[19] ELECTRONIC PRIVACY INFORMATION CENTER, *supra* note 11, *at* http://www.epic.org/privacy/gmail/faq.html (last modified Aug. 18, 2004).

[20] Two caveats are in order at this juncture. First, all search engines use cookies, and all Web sites offering search engines in addition to services that require registration (like Web-mail) are technically in a position to correlate search terms with personally identifiable information. Google is the gold standard of Web search, however, and its practices are likely to be indicative of emerging trends in the Web-services industry. Moreover, because of the breakthrough content-scanning capabilities of Gmail, Google has upped the ante in terms of how much personal information a site can maintain in its profile of a particular user. With the launch of Gmail, Google is responsible for a wave of heightened scrutiny into the privacy issues raised when a Web-services provider combines e-mail and search offerings into an integrated whole so that personally identifiable information can be correlated with a search history. Second, it is not my intention to portray Google as an evil company bent on flouting users' privacy rights. I think Google is an admirable company, and I am a big supporter of its search engine and its approach to product development. Indeed, without Google's search engine, I likely would not have been able to complete the research that informs this Comment.

[21] GOOGLE, GMAIL PRIVACY POLICY, *at* http://gmail.google.com/gmail/help/privacy.html (last modified Apr. 8, 2004) [hereinafter GMAIL PRIVACY POLICY] (emphasis added).

[22] *Id.*

[23] Ryan Singel, *Gmail Still Sparking Debates*, WIRED NEWS (Apr. 24, 2004), *at* http://www.wired.com/news/infostructure/0,1377,63204,00.html.

While Google has said it does not plan to correlate search and e-mail, the company will not rule out the possibility that it will do so in the future.[24] The current version of Gmail's Privacy Policy explicitly reserves for Google to right to so correlate: "Google may share cookie information among its other services for the purpose of providing you a better experience."[25]

This clear reservation of the right to correlate search and e-mail is actually an addition to the original Gmail Privacy Policy. On July 1, 2004, the California Online Privacy Protection Act ("OPPA") went into effect, making California the first state in the United States to enact a law governing online privacy policies.[26] The law requires operators of commercial Web sites that collect personally identifiable information from users residing in California to conspicuously post privacy policies. Such policies must "[i]dentify the categories of personally identifiable information that the operator collects through the Web site . . . and the categories of third-party persons or entities with whom the operator may share that . . . information."[27]

Google added the language about sharing cookie information across its various services to its Gmail Privacy Policy just as the new law took effect.[28] That Google clarified its stance with respect to correlating search and e-mail immediately following the passage of the OPPA suggests some duplicity on the company's part. Google representatives had stated in the past that the company has no intention of correlating personal data among services. But if this is so, "why does it need to explicitly reserve the right to do so?"[29]

In addition to the equivocal protections offered by the Gmail Privacy Policy, there are further reasons to be wary of relying solely on Google's promises that it will guard your online privacy. The Terms of Use agreement governing the use of Gmail, which by its own terms is the controlling document in the event of any inconsistency with the Privacy Policy,[30] eviscerates any protection purportedly offered by the promise not to disclose personal information. The Terms of Use agreement states that "Google may, in its sole discretion, modify or revise these terms and conditions and policies at any time."[31] Similarly, the Terms of Service agreement for Google's search engine

---

[24] *Id.*

[25] GMAIL PRIVACY POLICY, *supra* note 21, *at* http://gmail.google.com/gmail/help/privacy.html.

[26] Olsen, *supra* note 8, *at* http://www.news.com.com/2100-1028-5258824.html.

[27] CAL. BUS. & PROF. CODE § 22575(a),(b)(1) (Deering 2004).

[28] Olsen, *supra* note 8, *at* http://www.news.com.com/2100-1028-5258824.html.

[29] ELECTRONIC PRIVACY INFORMATION CENTER, *supra* note 11, *at* http://www.epic.org/privacy/gmail/faq.html.

[30] GOOGLE, GMAIL TERMS OF USE, *at* http://gmail.google.com/gmail/help/terms_of_use.html (last modified Apr. 6, 2004) ("In the event of an inconsistency between the Gmail Terms of Use and . . . the Gmail Privacy Policy . . ., the Gmail Terms of Use shall control.").

[31] *Id.*

states that Google "reserve[s] the right to modify these Terms of Service from time to time without notice."[32]

Whether these Terms of Use agreements—or at least the unilateral modification terms—are enforceable is a separate legal question.[33] For present

---

[32] GOOGLE, GOOGLE PRIVACY CENTER: GOOGLE TERMS OF PRIVACY FOR YOUR PERSONAL USE, *at* http://www.google.com/terms_of_service.html (last visited Nov. 13, 2004).

[33] Google's retention of the right to unilaterally modify any and all terms of its Terms of Use ("TOU") agreements, including its promise not to disclose private information to third parties without a user's consent, might well prove unenforceable as a matter of contract law. Web site TOU contracts fall into a category of online contracts known as "browsewrap" agreements. The actual text of the TOU is presented to the user only if the user clicks on a link usually located on the site's homepage. "These contracts generally provide that using the site . . . constitutes acceptance of the conditions contained therein." Robert A. Hillman & Jeffery J. Rachlinski, *Standard-Form Contracting in the Electronic Age*, 77 N.Y.U. L. REV. 429, 464 (2002). There are two main ways in which a term in a TOU can be deemed unenforceable by a court: (1) assertion that no contract was formed because of lack of assent; and (2) determination that the objectionable term is unconscionable. "Mutual manifestation of assent . . . is the touchstone of contract." Specht v. Netscape Communications Corp., 306 F.3d 17, 29 (2d Cir. 2002). In the context of a browsewrap contract like a TOU agreement, "[r]easonably conspicuous notice of the existence of contract terms and unambiguous manifestation of assent to those terms by consumers are essential if electronic bargaining is to have integrity and credibility." *Id.* at 35. In *Specht*, the Second Circuit held that a lack of such notice and the consequent absence of unambiguous assent with respect to a browsewrap license for downloadable software rendered a mandatory arbitration clause in the license unenforceable. *Id.* The *Specht* court distinguished the browsewrap contract before it with another type of online contract, the "clickwrap" agreement, whereby users are "required to review license terms . . . and to click 'I Agree' or 'I Don't Agree.'" *Id.* (quoting Barnett v. Network Solutions, Inc., 38 S.W.3d 200, 203-04 (Tex. App. 2001)). A user confronted with a clickwrap agreement therefore "affirmatively manifest[s] assent," which is not the case with a browsewrap agreement, where "if a manifestation of assent . . . exists at all[,] it is not the result of an affirmative act but can only be inferred from inaction." Sharon K. Sandeen, *The Sense and Nonsense of Web Site Terms of Use Agreements*, 26 HAMLINE L. REV. 499, 548-49 (2003). Clickwrap contracts generally are enforceable; "the pop-up presentation style of clickwrap terms constitutes reasonable notice of the terms contained therein." Hillman & Rachlinski, *supra*, at 488. By contrast, "although TOUs are generally accessible as a link from a [site's] homepage, there is no guarantee that they are noticed, let alone read, by . . . users." Sandeen, *supra*, at 549. One district court used this analysis to grant a motion to dismiss a breach of contract claim (with leave to amend), finding that no agreement existed between the Web site plaintiff and defendant user. Ticketmaster Corp. v. Tickets.com, Inc., No. CV 99-7654 HLH (BQRx), 2000 U.S. Dist. LEXIS 4553, at *7-8 (C.D. Cal. March 27, 2000). "Many customers . . . are likely to proceed to the . . . page of interest rather than reading the 'small print.' It cannot be said that merely putting terms and conditions in this fashion necessarily creates a contract with any one using the [W]eb site." *Id.* at *8.

Based on the assent-based approach of the courts in cases like *Specht* and *Ticketmaster*, Google likely would be able to enforce the terms of the Gmail TOU because it is a clickwrap agreement (as are the TOUs for other Web-mail providers). The terms of Google's TOU for its search engine, however, including the unilateral modification clause, likely would be deemed unenforceable as a browsewrap agreement. Notably, Google's search-engine TOU is not accessible directly from its homepage, unlike the TOUs of competitors like Yahoo, but is available only after a user clicks on an "About Google" link on the Google homepage. *See* http://www.google.com. While the assent-based approach eliminates the offending unilateral modification clause in the search engine browsewrap TOU, it could do away with the rest of the contract as well. Conversely, while the assent-based approach allows for enforceability of the Gmail clickwrap TOU, it allows for the enforceability of the unilateral modification

purposes, suffice it to say that Google is not really promising never to disclose a user's private information to third parties for marketing purposes because it is reserving the right to revoke that promise any time it wishes. If a user is to have some privacy protection for the contents of her e-mail and her search history, then, this protection must come not from the terms of the user's relationship with Google, but from the law.

Furthermore, an analysis of how current laws like the Stored Communications Act protect the sort of personal information collected by an integrated e-mail/search service does more than indicate whether a user is protected outside of the rather indefinite promises made by Google; such an analysis can also help establish the legal parameters governing data collection by current and future providers of similar services.

---

clause. Thus, the assent-based approach is at once both underinclusive and overinclusive in terms of its ability to regulate online contracting for the benefit of consumers without unduly "chilling" the development of Internet-based business. *See* Dan Streeter, Comment, *Into Contract's Undiscovered Country: A Defense of Browse-Wrap Licenses*, 39 SAN DIEGO L. REV. 1363, 1390-93 (2002) (arguing that not enforcing browsewrap contracts will hamper online commerce).

A better approach, one that could eliminate an offending term like the unilateral modification clause while leaving the rest of the contractual relationship intact, is to employ the well established doctrine of unconscionability. Under this doctrine, a court will refuse to enforce a specific term of a contract if the term is deemed unfair or oppressive. Sandeen, *supra*, at 551. Usually, courts inquire into "the manner in which the parties entered the contract to police the quality of assent (procedural unconscionability) and . . . the fairness of the resulting terms (substantive unconscionability)." Hillman & Rachlinski, *supra*, at 456. Because the "mutual assent to a browse-wrap TOU is marginal at best, and . . . in many cases a [W]eb site user will not have seen the TOU, the procedural unfairness of browse-wrap TOUs is clear." Sandeen, *supra*, at 552. Substantive unconscionability is present where there are "manifestly unjust terms, such as terms that are immoral, conflict with public policy, deny a party substantially what she bargained for, or have no reasonable purpose in the trade." Hillman & Rachlinski, *supra*, at 457. The notion of what constitutes substantive unconscionability is echoed in section 211 of the Restatement (Second) of Contracts: "Where the other party has reason to believe that the party manifesting such assent would not do so if he knew that the writing contained a particular term, the term is not part of the agreement." RESTATEMENT (SECOND) CONTRACTS § 211(3) (1981). Such an analysis will render unenforceable "terms a business should reasonably understand a consumer would resist, namely those terms that defeat the purpose of the deal, that are 'bizarre or oppressive,' and that conflict with bargained-for terms." Hillman & Rachlinski, *supra*, at 458 (quoting RESTATEMENT (SECOND) CONTRACTS § 211(3) cmt. f). The unilateral modification clause in Google's various TOUs—as applied to Google's promises never to share users' personally identifiable information without their consent—is just the sort of term that should properly be deemed unenforceable under unconscionability doctrine. The privacy terms themselves envision user consent to be a necessary prerequisite to the disclosure of a user's private information. This is compelling evidence that a contrary term allowing for the consent requirement to be unilaterally revoked by Google without notice is bizarre and oppressive.

III. GOOGLE AND PRIVACY LAW: THE "GMAIL BILL" AND THE SCOPE OF THE STORED COMMUNICATIONS ACT AS APPLIED TO MODERN WEB TECHNOLOGIES

A. The "Gmail Bill"

In the spring of 2004, the controversy surrounding Gmail's content-extraction technology spurred the California Senate to pass a bill that limits the ability of a provider of e-mail service to California residents to compile profiles based on the content of a user's e-mail and to divulge content or personally identifiable information to third-parties (or even to the provider itself). The bill passed by the California Senate provides:

> A provider shall not derive content from an electronic mail . . . being electronically stored by the provider for the provider's marketing purposes unless all of the following are true:
>
> (A) The derivation is automated.
>
> (B) The derivation does not associate the contents of an electronic mail . . . with personally identifiable information or user characteristics.
>
> (C) What is derived is not divulged to any person, including the provider.
>
> (D) The derivation is with the lawful consent of the customer.
>
> (E) What is derived is not retained by the provider or any other person.[34]

If the Gmail Bill is enacted into law as passed by the Senate,[35] Google would be prevented from building user profiles based on e-mail content. This would hamper Google's ability to profitably mine user data, in that the Gmail system is designed, as discussed above, to incorporate the subject matter of a user's past e-mail into the content-extraction process when applied to new e-mails.

Though criticized by many journalists and even the stalwart privacy advocacy group, the Electronic Frontier Foundation ("EFF"),[36] the Bill squarely addresses a hole in Google's Privacy Policy. While Google has said that it will not keep logs of the concepts extracted from user's e-mail,[37] the Gmail Privacy Policy does not offer a similar guarantee. In the event the Bill does not become law, Google will be able to amass content-based user profiles.

---

[34] S.B. 1822, 2003−2004 Reg. Sess. (Cal. 2004), http://www.leginfo.ca.gov/pub/bill/sen/sb_1801−1850/sb_1822_bill_20040729_status.html.

[35] As of this writing, S.B. 1822 is awaiting action by the full California Assembly. *See* http://www.leginfo.ca.gov/pub/bill/sen/sb_1801-1850/sb_1822_bill_20040729_status.html (last visited Oct. 6, 2004).

[36] Electronic Frontier Foundation, Deep Links, *It's the Privacy Law, Stupid*, *at* http://www.eff.org/deeplinks/archives/001468.php (Apr. 26, 2004) (describing belief that the narrowly focused Gmail Bill is insufficient to address the broad privacy issues faced by Internet users).

[37] Electronic Frontier Foundation, Deep Links, *Google's Gmail and Your Privacy−The Scoop*, *at* http://www.eff.org/deeplinks/archives/001398.php (Apr. 9, 2004).

And if Google in the future decides to unilaterally change its policies to allow for the transfer of collected information to third parties for marketing purposes (assuming such a change to the Terms of Use would be enforceable),[38] these profiles would be available for those purposes.

Additionally, even if the Bill does become law, it addresses only those providers based in California and those not based in California but that have California residents as users,[39] leaving gaps in the law's geographic coverage. One further problem with the Bill is that it too narrowly focuses on e-mail in general, and on profiling users based on the extraction of content from their e-mail in particular. This focus stems from the Bill being a direct and immediate reaction to the release of Gmail. Because of this narrow focus, the Bill fails to address one of the major privacy concerns discussed previously—that once a Google user has registered to use Gmail, Google can correlate that user's personally identifiable information with her entire search history.

This type of correlation is not covered by the Gmail Bill's language. The Bill states that "[a] provider shall not derive personally identifiable information or user characteristics *from electronic mail* . . . being electronically stored by the provider for the provider's marketing purposes."[40] Electronic mail is defined as "an electronic message that is sent *to an e-mail address* and transmitted between two or more telecommunications devices, computers, or electronic devices."[41]

But the personally identifiable information that Google could correlate with a user's search history does not come from e-mail in the user's Gmail account; rather, it is available to Google via the cookie file that Google uses to recognize a user visiting the site (assuming the use of the same browser for repeat visits). Though it is possible to construe the cookie as an electronic message, it is never sent to an e-mail address, thereby exempting from the reach of the Gmail Bill any derivation of personally identifiable information from a cookie and subsequent correlation of that information with a search history.

For several reasons, then, it is necessary to look beyond the Gmail Bill to determine if other laws are in place that can protect the privacy of users of integrated e-mail/search services. First, the Gmail Bill might not become the law in California. Even if it does become the law in California, (1) it could be challenged by affected companies as a violation of the Dormant Commerce Clause of the United States Constitution,[42] (2) the dealings of online service

---

[38] *See supra* note 33 for discussion of the contract issues involved in assessing the protections afforded by the Gmail Privacy Policy.

[39] Cal. S.B. 1822 (legislative counsel's digest).

[40] Cal. S.B. 1822 (§ 1798.88.2(b)(1)) (emphasis added).

[41] *Id.* (§ 1798.88(f)) (emphasis added).

[42] It is well established that the Commerce Clause "contains a negative or 'dormant' aspect that 'denies the States the power unjustifiably to discriminate against or burden the interstate flow . . . of commerce.'" Brown & Williamson Tobacco Corp. v. Pataki, 320 F.3d 200, 208 (2d Cir. 2003) (quoting Oregon Waste Sys., Inc. v. Dep't of Envtl. Quality, 511 U.S. 93, 98 (1994)). Discrimination against interstate commerce occurs when a state statute "has the practical effect of requiring out-of-state commerce to be conducted at the regulating state's discretion." *Id.* (quoting Nat'l Elec. Mfrs.' Ass'n v. Sorrell, 272 F.3d 104, 110 (2d

Cir. 2001)). This principle served as the basis for the Supreme Court's invalidation of a Connecticut statute that required beer sellers in Connecticut to certify that they sold beer for the same price in Connecticut as in neighboring states. *See* Healy v. Beer Inst., 491 U.S. 324, 336-37 (1989). "[T]he 'Commerce Clause . . . precludes the application of a state statute to commerce that takes place wholly outside of the State's borders, whether or not the commerce has effects within the State.'" *Id.* at 336 (quoting Edgar v. MITE Corp., 457 U.S. 624, 642-643 (1982) (plurality opinion)).

When reviewing a state statute that indirectly regulates out-of-state commerce in this fashion, the Court will examine "whether the State's interest is legitimate and whether the burden on interstate commerce clearly exceeds the local benefits." *Id.* at 337 n.14 (quoting Brown-Forman Distillers Corp. v. N.Y. State Liquor Auth., 476 U.S. 573, 579 (1986)). State laws geared toward the protection of state citizens and passed pursuant to the "traditional police powers" of the states often are "given special deference in this balancing test." Dan L. Burk, *Federalism in Cyberspace*, 28 CONN. L. REV. 1095, 1124 (1996). The interest of the California legislature in protecting the online privacy of its citizens would almost certainly be deemed legitimate for the purposes of a Dormant Commerce Clause analysis. Whether there is a significant burden on interstate commerce largely turns on the extent to which California could be seen as "exporting [its] law . . . into the local markets of sister states." *Id.* at 1127. A big part of this analysis concerns technology; if Google is unable to efficiently discern which users are coming from which states, a likely solution would be to comply with California's proscription against building content-based profiles from a user's e-mail for all users, even if it would be legal for the company to derive such profiles for users located in every state but California. "By complying . . . with . . . the most demanding . . . regulatory regime, a business might satisfy the lesser requirements of all the other jurisdictions as well." *Id.* at 1132. In the absence of a viable way to distinguish users based on geography, an online company dealing with at least some California residents is forced to obey California law in its interactions with all users, including those who are not residents of California, thereby foreclosing the ability to make otherwise legal use of a potentially lucrative business strategy involving users from across the country and around the world.

Analyses based on this technological-hurdle argument are characteristic of early attempts to apply the Dormant Commerce Clause to state legislation concerning the Internet. *See, e.g.,* Am. Libraries Ass'n v. Pataki, 969 F. Supp. 160, 169 (S.D.N.Y. 1997) (stating that "geography . . . is a virtually meaningless construct on the Internet); Burk, *supra*, at 1131-32 ("Internet businesses . . . simply cannot tell with any degree of assurance the geographic location from which access to data has been requested, and there is no practical way to screen out contacts from particular jurisdictions."). More recently, some commentators have soundly criticized this view. *See, e.g.,* Jack L. Goldsmith & Alan O. Sykes, *The Internet and the Dormant Commerce Clause*, 110 YALE L.J. 785, 808-16 (2001) (describing increasing accuracy of geographical filtering technology and attendant decrease in burdens on interstate commerce for purposes of Dormant Commerce Clause).

In any event, Google would be in a particularly tough spot trying to make the "geography is meaningless" argument with respect to Gmail, because users register for the service with personally identifiable information, giving Google the option to request that users identify their state of residence. Just knowing who is from where, Google might counter, is not enough to sufficiently reduce the burden; perhaps the burden is maintaining separate "versions" of the service for users depending on the particular regulations on content-profiling imposed by their state of residence. A fact intensive inquiry into the capabilities of current geographical filtering technology and of the detailed workings of the Gmail profiling system's architecture would be required to conclusively decide how to properly balance these countervailing forces. These analytical problems likely would not be present to the same degree were the California Online Privacy Protection Act (OPPA) to be challenged on Dormant Commerce Clause grounds. *See supra* text accompanying note 8 for discussion of the OPPA. Because the OPPA requires only that a Web site post a conspicuous statement regarding the nature of the personally identifiable information it collects from users, an argument that requiring truthful disclosure burdens interstate commerce in a way that clearly outweighs the local benefit of protecting users' online privacy will be hard to

providers not based in California with users not residing in California will be unaffected by the proscriptions of the law, and (3) the narrow drafting of the Bill (serving as the model for similar legislation in other states) excludes from its reach the potentially serious privacy breach represented by the correlation of search histories with personally identifiable information.

B. The Scope of the Stored Communications Act as Applied to Modern Web Technologies

The reality of current Internet use is that "[o]ur most private information ends up being sent to private third parties and held far away on remote network servers."[43] Because of what one commentator calls the "disclosure principle," courts have generally held that such private information stored by various Web sites and Internet service providers does not receive Fourth Amendment protection.[44] It is a well-established tenet of Fourth Amendment law that "an individual has no reasonable expectation of privacy in information revealed to third parties."[45] Since everything a user stores on another entity's servers—from Amazon.com shopping carts to Gmail accounts—is in some sense revealed to a third party, the Fourth Amendment provides uncertain protection for a user concerned about the privacy of information stored online.[46]

As a result, federal statutes have "fill[ed] this possible gap,"[47] offering privacy protection to Internet users who may be unprotected by the Fourth Amendment. The federal statute that protects the privacy of stored Internet communications is the Stored Communications Act ("SCA"), passed as part of the Electronic Communications Privacy Act of 1986.[48] The Electronic Communications Privacy Act ("ECPA") "amended the Federal Wiretap Act to extend privacy protections to 'electronic' communications such as email."[49] The ECPA also "expressly regulated the use of pen registers,"[50] devices that as

---

make successfully. *See* Washington v. Heckel, 24 P.3d 404, 411-12 (Wash. 2001) (holding that a state law requiring truthful disclosures in commercial e-mail messages did not burden interstate commerce at all but facilitated it).

[43] 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, 1209-10 (2004).

[44] Orin S. Kerr, *Internet Surveillance Law After The USA Patriot Act: The Big Brother That Isn't*, 97 Nw. U. L. REV. 607, 627 (2003).

[45] *Id.*

[46] *Id.* at 629. In the case of a private company disclosing user information to another private company, Fourth Amendment protection is non-existent. Such a situation would not merit Fourth Amendment protection even in the absence of the third-party disclosure problem because of the "private search doctrine," which makes the Fourth Amendment inapplicable to the actions of private parties not acting on behalf of the government. *See, e.g.,* United States v. Jacobsen, 466 U.S. 109, 113 (1984).

[47] Kerr*, supra* note 43, at 1212.

[48] *Id.* at 1208.

[49] Robert A. Pikowsky, *The Need For Revisions to the Law of Wiretapping and Interception of E-mail*, 10 MICH. TELECOMM. & TECH. L. REV. 1, 39 (2003).

[50] *Id.* at 18.

originally conceived and used "capture only the 'numbers dialed or otherwise transmitted' on [a] telephone to which the device is attached."[51]

Because the e-mail from which Google extracts content and the search histories it maintains reside in electronic storage on Google's servers, the SCA determines whether statutory privacy protection exists for this electronic information. The SCA regulates two types of access to stored communications. First, the statute limits the government's ability to compel providers to disclose information they are storing.[52] Second, the statute limits the ability of providers to voluntarily disclose information to the government or to non-government entities.[53] The voluntary disclosure provisions are of particular relevance to Google's practices and policies regarding the protection its users receive against the disclosure of their private information to third parties for marketing purposes.

To decide if the contents of e-mail as extracted by Gmail and the correlated search histories are covered by the voluntary disclosure provisions (and if so, to what extent), it is necessary to examine the parameters of the SCA's coverage. Because the SCA is regarded as a "dense and confusing" statute,[54] the exercise is a challenging one, particularly when trying to apply the nearly twenty-year-old statutory framework to cutting-edge Web technologies likely not contemplated by the drafters of the law.

1. Gmail Concept Tags and Google Search Engine Histories as "Electronic Communications"

The SCA does not define the word "communications," though "electronic communication" is 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 that affects interstate or foreign commerce."[55] This definition certainly includes e-mail. Whether it includes the records of the concept tags that Gmail has extracted from a user's e-mail is a conceptually distinct question, as is whether electronic communication includes the file stored on Google's servers containing a user's search history.

Because the electronic communications rubric has been applied in a fairly broad way, the available precedent suggests that both the concept tags and the search history should be deemed electronic communications. A Web site, which is an amalgamation of computer files containing software code and various other types of information, has been found to be an "electronic

---

[51] CENTER FOR DEMOCRACY & TECHNOLOGY, CDT'S ANALYSIS OF S. 2092: AMENDING THE PEN REGISTER AND TRAP AND TRACE STATUTE IN RESPONSE TO RECENT INTERNET DENIAL OF SERVICE ATTACKS AND TO ESTABLISH MEANINGFUL PRIVACY PROTECTIONS (Apr. 4, 2000), *at* http://www.cdt.org/security/000404amending.shtml [hereinafter CDT'S ANALYSIS OF S. 2092] (quoting 18 U.S.C. § 3127(3) (2000) (amended 2001)).

[52] *See* 18 U.S.C. § 2703 (compelled disclosure provisions).

[53] *See* 18 U.S.C. § 2702 (voluntary disclosure provisions).

[54] Kerr, *supra* note 43, at 1208.

[55] 18 U.S.C. § 2510(12).

communication held in storage."[56] Without deciding the issue, another court assumed that the cookie files residing on a user's computer, which "capture certain parts of the communications that users send to . . . Web sites," could also be electronic communications.[57]

Similarly, the electronic records containing e-mail concept tags and search terms represent files that have captured elements of the communications users have sent to Web sites. The comparison is perhaps more apt with the search history than with the e-mail concept tags. A search term is literally a transfer of writing sent through the wires by a user to a Web site. But the e-mail concept tags themselves are not sent by users to Web sites; rather, they are extracted by a computer from e-mails that have been sent (or received) and are in storage on Google's servers. At the same time, though, there is still a transfer of data or intelligence from the information contained in the e-mail to the record of the concept tags. This transfer also should qualify as an electronic communication, along with the transfer of search terms, bringing both within the scope of the SCA.

Any cognitive dissonance caused by the fact that the transfer of information from the stored e-mail to the record of concept tags seems to take place solely between computers is resolved fairly easily. The SCA has been interpreted to envision computers as communicators—a construction of the statute that wisely accounts for our increasingly automated worldwide communications network. The SCA defines a "user" as "any person or entity" using an electronic communication service.[58] Several district courts have included Web servers among the entities that can be users.[59] Once computers are seen as users of communications networks, it becomes straightforward to classify the transfer of information and data between computers as an electronic communication.

2. The Distinction Between "Contents" of Communications and "Noncontent" Information

Having concluded that search histories and e-mail concept tags are properly classified as electronic communications, a thornier classification problem arises—determining how the voluntary disclosure provisions of the SCA apply to these communications. The SCA prohibits online service providers from knowingly divulging to "any person or entity the contents of a communication."[60] On the other hand, as long as the recipient is not a governmental entity, the SCA does not prohibit a provider from divulging to another person or entity "a record or other information pertaining to a

---

[56] Pikowsky, *supra* note 49, at 69 (citing Konop v. Hawaiian Airlines, Inc., 236 F.3d 1035, 1044-46 (9th Cir. 2001), *opinion withdrawn by* 262 F.3d 972 (2001), *superseded by* 302 F.3d 868 (9th Cir. 2002), *cert denied,* 537 U.S. 1193 (2003)).

[57] In re DoubleClick Inc. Privacy Litig., 154 F. Supp. 2d 497, 504, 513 (S.D.N.Y. 2001).

[58] 18 U.S.C. § 2510(13).

[59] *See*, *e.g.*, In re DoubleClick, Inc. Privacy Litig., 154 F. Supp. 2d at 508-09.

[60] 18 U.S.C. § 2702(a)(1).

subscriber or customer of such service (not including the contents of communications)."[61]

Thus, the SCA "draws an important line between 'contents' of communications and noncontent information."[62] While "[s]ection 2702(a) generally bans disclosure of contents," providers are "free to disclose noncontent information to nongovernment entities."[63] As a result, "a company can disclose records about how its customers used its services to a marketing company."[64] To determine whether Google can legally disclose to private parties for marketing purposes users' search histories or the concept tags derived from their e-mail, one must decide if this information constitutes "contents" of communications or noncontent information.

The SCA defines the "contents" of an electronic communication as "any information concerning the substance, purport, or meaning of that communication."[65] With e-mail, this definition "clearly covers the body of the e-mail" as well as the subject line, which "generally carries a substantive message."[66] But "logs of account usage, mail header information minus the subject line, lists of outgoing e-mail addresses sent from an account, and basic subscriber information all count as noncontent information."[67]

One hurdle to a clear classification of the concept tags derived by Gmail is that the extraction process as described in the patent can draw upon both types of information: contents (body and subject line of e-mail) and noncontents (e-mail addresses, time of transmission, geographic location of one or more recipients).[68] That the concept tags are derived in part by considering noncontent information, however, does not change the essential fact that the tags plainly concern the "substance, purport, or meaning" of the e-mails from which they are derived. Indeed, if the tags did not to some extent convey the meaning of a user's e-mail, the Gmail system could not function as intended—by placing contextually relevant advertising alongside a user's e-mail *based on* the concepts contained therein. Accordingly, the reasonable conclusion is that Gmail concept tags represent the contents of electronic communications under the SCA.

It is more problematic to settle on the proper classification of a user's search history. The SCA describes noncontent information as "a record or other information pertaining to a subscriber to or customer of such service (not including the contents of communications)."[69] There is no doubt that information such as the date, time, and duration of a user's visit to Google would fall into the category of noncontent records. Whether the statutory

---

[61] 18 U.S.C. § 2703(c)(1)(A).

[62] Kerr, *supra* note 43, at 1227.

[63] *Id.* at 1220.

[64] *Id.*

[65] 18 U.S.C. § 2510(8).

[66] Kerr, *supra* note 43, at 1228.

[67] *Id.*

[68] ELECTRONIC PRIVACY INFORMATION CENTER, *supra* note 11, *at* http://www.epic.org/privacy/gmail/faq.html.

[69] 18 U.S.C. § 2703 (c)(1)(A).

description of noncontents includes the terms a user types into Google's search engine is much less clear.

Additional guidance as to the meaning of noncontents can be found elsewhere in the ECPA, in the provisions regarding the use of pen registers. A pen register is a device that "records or decodes dialing, routing, addressing, or signaling information transmitted by an instrument or facility from which a[n] . . . electronic communication is transmitted, provided, however, that such information shall not include the contents of any communication."[70]

Pen register law is premised on the view that "every communications network features two types of information: the contents of communications, and the addressing and routing information that the networks use to deliver the contents of communications."[71] The latter type of information, the noncontent information, is also known as "envelope information," an analogy to the type of information one could glean from the outside of an envelope sent through the postal mail system, including "the mailing and return addresses, the stamp and postmark, and the size and weight of the envelope when sealed."[72] Similarly, the envelope information for a telephone call includes "the number the caller dials, the number from which a caller dials, the time of the call, and its duration."[73]

In the case of electronic communications like search terms, however, the "transactional or addressing data . . . can be much more revealing than telephone numbers dialed."[74] The extent to which this is so can best be understood by reference to a specific example of how a search for the word "cars" might appear in Google's logs:

> inktomi1-lng.server.ntl.com - 25/Mar/2003 10:15:32 - http://www.google.com/search?q=cars" - MSIE 6.0; Windows NT 5.1 – 740674ce2123e969[75]

Contained within this simulated (and simplified) log file are numerous pieces of information pertaining to the search being conducted. Included among these are the date and time of the search, the version of Microsoft's Internet Explorer browser being used (6.0), and the version of Microsoft's Windows NT operating system running on the user's computer (5.1). The last piece of information is the user's unique cookie identification number, which allows Google to associate this particular search with all the others submitted in the past by the browser being used for the current search. In this way, Google creates a search history for every unique cookie file.[76]

---

[70] 18 U.S.C. § 3127(3).

[71] Kerr, *supra* note 44, at 611.

[72] *Id.*

[73] *Id.*

[74] CDT'S ANALYSIS OF S. 2092, *supra* note 51, *at* http://www.cdt.org/security/000404amending.shtml.

[75] Sullivan, *supra* note 18, *at* http://www.searchenginewatch.com/sereport.php/34721_2189531.

[76] *Id.*

All of the information mentioned in the previous paragraph fits neatly into the noncontents category. None of it concerns the meaning of a communication. But what about the Uniform Resource Locator ("URL") contained in the log file? When a user types the word "cars" into Google's search engine, the site produces a page with that address—http://www.google.com/search?q=cars"—on which are displayed the search results Google has found that match the term "cars." The address is recorded in the log file as part of the record of that particular visit to the Web site.

Is the URL that incorporates a user's search term "dialing, routing, addressing, or signaling information," and hence noncontents? Or does the fact that the URL "can actually convey the substance or purport of a communication"[77]—in this case that the user is interested in cars—indicate that the search term is protected as "contents"? The answer is of critical importance to the users of Google and other search engines, because if search terms are no more than addressing or routing information, the search engine provider can disclose a user's search history to a marketing company without the user's consent. If, instead, the search terms are seen as the contents of communications, then search engine providers are prohibited by the SCA from voluntarily disclosing a users' search history to third parties. Part of the "conceptual difficulty"[78] involved in resolving this issue stems from wrestling with the notion of humans communicating with computers as compared to the more familiar construct of human-to-human communication that underlies traditional thinking about content and envelope information in a communications network. One commentator has suggested two perspectives one could take when looking at a communication like the entry of a search term into Google, with the choice one makes between the two perspectives dictating whether the communication is seen as contents or noncontents: "either the command is the 'content' of the communication between the user and [a] computer or it is merely 'addressing information' that the user entered into [the] computer to tell the computer where it should go and what it should do."[79]

The technical reality is that the search term fits into both of these categories. It is a substantive communication by a user to a computer that the user seeks information on a given subject. Functionally, this is the equivalent of calling the reference desk of the library on the telephone and asking the librarian to assist you in a search for articles on cars. It seems beyond argument that such a conversation is contents of a communication.

At the same time, the 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.

---

[77] CDT'S ANALYSIS OF S. 2092, *supra* note 51, *at* http://www.cdt.org/security/000404amending.shtml.

[78] Kerr, *supra* note 44, at 645.

[79] *Id.* at 646.

Perhaps the way to resolve the issue is to focus not on the technology but rather on "the nature of the privacy interest at risk."[80] Changing the focus of the inquiry in this way would protect the privacy of the terms a user enters into a search engine because of the user's interests in keeping that information private—even though those terms happen to get incorporated into Web site addresses because of the way the Web functions technically.

A decision by the Court of Appeals for the District of Columbia lends credence to this approach. In *United States Telecom Association v. Federal Communications Commission*, the court considered a challenge to FCC regulations requiring telecommunication providers to divulge, under certain circumstances, what are called "post-cut-through dialed digits," or "all digits dialed after calls are connected."[81] The government argued that since the pen register statute permits the disclosure of dialing information and since the post-cut-through dialed digits are just that—dialed numbers—the post-cut-through dialed digits must be noncontents. [82]

The court rejected this argument, finding that not all numbers dialed could be considered dialing or routing information. Granted, the court stated, some post-cut-through dialed digits are dialing information, "such as when a subject places a calling card, credit card or collect call by first dialing a long-distance carrier access number and then, after the initial call is 'cut through,' dialing the telephone number of the destination party."[83] But the court made clear that some post-cut-through dialed digits "can also represent call *content*."[84] For example, the court stated, "[people] calling automated banking services enter account numbers. When calling voicemail systems, they enter passwords. When calling pagers, they dial digits that convey actual messages. And when calling pharmacies to renew prescriptions, they enter prescription numbers."[85]

In essence, the court counseled against an overly formalistic application of the distinction between contents and envelope information in the context of human-to-computer communication. While no court has held that the *United States Telecom Association* approach applies to human-to-computer communication on the Internet, such an extension is entirely appropriate.

Sometimes a Web address is not just a Web address. That a search term gets stored by Google in the form of a URL should not turn it into the equivalent of a postal address or a telephone number when the search term is functionally analogous to a message communicated to a friend's pager using a telephone's alphanumeric keypad. Sure, the pager message could be represented formalistically as mere dialing information, as the government attempted to do in *United States Telecom Association*, but a "logical and consistent"[86] approach to the protection of private communications demands

---

[80] Pikowsky, *supra* note 49, at 22.

[81] United States Telecom Ass'n v. FCC, 227 F.3d 450, 456 (D.C. Cir. 2000).

[82] *Id.* at 458-59.

[83] *Id.* at 462.

[84] *Id.* (emphasis added).

[85] *Id.*

[86] Pikowsky, *supra* note 49, at 5.

treating the message, regardless of whether it is transmitted via numbers dialed on a telephone, as what the court found it to be—contents of communication.

Thus far, an analysis of how Google-like integrated e-mail/search services fit into the statutory framework of the SCA has suggested a favorable outcome for users concerned about the possible unauthorized disclosure of their private communications to third parties for marketing purposes. A sensible application of the statute to the technology and privacy interests at issue would result in both Gmail concept tags and Google search histories receiving protection under the provisions of the SCA. Not only should this information qualify as electronic communication, but it also should be considered contents of such communication (as opposed to noncontents). Consequently, Google would be prohibited from disclosing the information to third parties for marketing purposes without a user's consent.[87]

3. The contours of the statutory terms "electronic communications service" and "remote computing service"

To be certain that Google would be prohibited from voluntarily disclosing a user's Gmail concept tags and search history to third parties, one must undertake a final inquiry into the scope of the SCA as applied to modern Web technologies—whether a provider of integrated e-mail/search functionality is the type of provider covered by the provisions of the SCA. The SCA regulates two types of service providers: "providers of electronic communications service ("ECS") and providers of remote computing service ("RCS")."[88] ECS is defined as "any service which provides to users thereof the ability to send or receive . . . electronic communications."[89] RCS is defined as "the provision to the public of computer storage or processing services by means of an electronic communications system."[90] Lastly, an "electronic communications system" is defined as "any wire, radio, electromagnetic, photooptical or photoelectronic facilities for the transmission of . . . electronic communications, and any computer facilities or related electronic equipment for the electronic storage of such communications."[91]

The ECS and RCS distinction "freez[es] into the law the understandings of computer network use as of 1986."[92] A key question is how well the SCA's categories of service providers fit today's varied Web technologies. One commentator suggests that the ECS category has held up pretty well over the twenty years since the passage of the SCA. ECS's core function has not changed over the past two decades. Just as service providers provided "the ability to send or receive . . . electronic communications" in 1986, they do so today.[93]

---

[87] *See* 18 U.S.C. § 2702(a).
[88] Kerr, *supra* note 43, at 1214.
[89] 18 U.S.C. § 2510(15).
[90] 18 U.S.C. § 2711(2).
[91] 18 U.S.C. § 2510(14).
[92] Kerr, *supra* note 43, at 1214.
[93] *Id.*

Web-based e-mail systems like Gmail seem to fit nicely into the definition of ECS; clearly, they allow users to send and receive electronic communications. The 1986 Senate Report on the SCA described the functioning of an e-mail system in terms that can encompass a Web-based service: "[i]n its most common form, messages are typed into a computer terminal, and then transmitted over telephone lines to a recipient computer operated by an electronic mail company."[94]

Despite this apparent correspondence between the definition of ECS and the nature of the service provided by Gmail, there are reasons to be concerned about whether Gmail would qualify as an ECS provider. At least two district courts have grafted an additional requirement onto the statutory definition of ECS, one that would make it hard for any firm not in the business of actually providing Internet access to be an ECS provider.

In *Crowley v. CyberSource Corporation*, the plaintiff pled an ECPA claim based on the contention that Amazon.com was an ECS provider.[95] Because Amazon.com "receives electronic communications from customers," the plaintiff alleged, it must be an ECS provider.[96] The court disagreed, finding that Amazon.com did not necessarily provide ECS just because it receives e-mails from customers.[97]

On this specific point, the court reached the correct answer. Amazon.com is an online merchant, not the sort of service provider envisioned by the SCA drafters' definition of ECS. But the court did not stop there; instead, the court found that because Amazon.com itself "must purchase Internet access from an electronic communication service provider . . . it does not independently provide such services."[98]

In *In re Northwest Airlines Privacy Litigation*, the court used the same analysis to determine that Northwest was not an ECS provider with respect to its Web site.[99] The Northwest site communicates with users via e-mail, transacts business with users, and stores records of customer information. The court stated that "[d]efining electronic communications service to include online merchants or service providers like Northwest stretches the ECPA too far."[100] Because Northwest was not an "internet service provider," and had to "purchase[] its electronic communications service from a third party," the court found that Northwest was "simply not an electronic communications service provider."[101]

Both of these cases seem to require that any provider of ECS also be an Internet service provider ("ISP"). While it is true that many people have an e-

---

[94] S. REP. NO. 99-541, at 8 (1986), *reprinted in* 1986 U.S.C.C.A.N. 3555.

[95] Crowley v. CyberSource Corp., 166 F. Supp. 2d 1263, 1270 (N.D. Cal. 2001).

[96] *Id.*

[97] *Id.*

[98] *Id.* (quoting Andersen Consulting L.L.P. v. UOP, 991 F. Supp. 1041, 1043 (N.D. Ill. 1998)).

[99] In re Northwest Airlines Privacy Litig., No. 04-126 (PAM/JSM), 2004 U.S. Dist. LEXIS 10580, at *6-7 (D. Minn. June 6, 2004).

[100] *Id.* at *6.

[101] *Id.* at *6-7.

mail address associated with their ISP—user@comcast.net, for example—it is also true that many people have additional (or their only) e-mail addresses in connection with a Web-mail account provided by a service like Yahoo or Hotmail—or Gmail. Web-mail users send, receive, organize, and store e-mail with their Web-mail accounts. Functionally, the Web-mail account plays the same role in the user's communicative life as would an e-mail account maintained with an ISP that is accessed via a stand-alone e-mail application like Microsoft Outlook, which resides on the user's hard drive. In fact, Web-mail is arguably a more universal communications platform (perhaps more akin to the telephone system) in that it can be accessed using any computer, regardless of through which ISP that computer happens to be connecting to the Internet.

Concededly, Web-mail is a software application that requires an Internet connection to allow users to communicate, a connection not always provided by the operator of the Web-mail service. The same is true of instant messaging software. But to suggest that these services are not providing the ability to send and receive electronic communications because they do not provide Internet access is to focus, again, on the technological details instead of the nature of the privacy interests at stake.[102]

The district court judges who required a provider of ECS also to be an ISP missed the privacy forest for the trees of technological detail. While the above-cited cases involve e-commerce sites, not Web-mail services, their ISP requirement would preclude many Web-mail providers from receiving the ECS designation as well. This sets a dangerous precedent by casting doubt on whether electronic communications sent and received through a Web-mail provider are covered by the SCA.[103]

The definition of ECS in the statute should be revised to clarify that Web-mail and instant messaging services are included in the category of electronic communications services. The definition section of the California Senate's Gmail Bill provides a model for such a revision. A "provider of electronic mail or instant messaging service" is defined as "any person, including an Internet service provider and a provider of remote computing services, that is an *intermediary* in sending or receiving electronic mail or instant messages."[104] By

---

[102] The Internet is by design a network of layered technologies, where more complex applications run on top of less complex ones. Web-mail and instant messaging simply represent a new layer of electronic communications services running on top of another, older layer of electronic communications services, the provision of Internet access. Given the increasing variety of Internet applications and the growth of the online population, the development of such layered "stacks" of applications is entirely predictable and desirable. The Internet was designed to evolve in just this fashion, according to "end-to-end" principles whereby the most sophisticated components of network functionality are kept separate from the basic data-transport mechanisms of the network "plumbing." *See generally* LAWRENCE LESSIG, THE FUTURE OF IDEAS 23, 34-40 (2001) (discussing layered model of communications networks and the end-to-end design principle underlying the Internet).

[103] At least one district court opinion properly views Web-based e-mail as an example of ECS. *See* FTC v. Netscape Communications Corp., 196 F.R.D. 559, 560 (N.D. Cal. 2000) (stating than Netscape's provision of e-mail accounts through its Web site qualifies it as a provider of ECS).

[104] Cal. S.B. 1822 (§ 1798.88(m)) (emphasis added).

including intermediaries in the online communications process within the scope of the Gmail Bill, the California Senate acknowledged the layered, end-to-end principles inherent in the Internet's architecture that the district court opinions in *Crowley* and *Northwest Airlines* erroneously overlooked.

Some privacy advocates question whether integrated e-mail/search services like those provided by Google would qualify as ECS even in the absence of a judicially imposed ISP requirement. Gmail stores and indexes its users' e-mail so it can be easily searched and incorporates contextually relevant advertising into the user experience of reading e-mail. According to the Chairman of the EFF, this arguably turns the application into a "database and shopping service," which "doesn't look as much like an e-mail provider service as it should according to the legal definitions in the ECPA."[105] The EFF worries that "different legal rules . . . may apply to mail that is indexed, searched, or keyword matched by a third party."[106]

To properly consider this possibility, one must understand the scope of remote computing services ("RCS"), the other category of SCA-regulated online service providers. If, as the EFF suggests, a Web-mail service cannot also be a "database and shopping service" without jeopardizing its status as an ECS provider, might it be swept into the RCS category and hence be brought into the ambit of the SCA?

As mentioned above, RCS consists of providing "computer storage or processing" services through a facility for the "transmission of . . . electronic communications."[107] The Senate Report on the SCA described RCS as follows:

> In the age of rapid computerization, a basic choice has faced the users of computer technology. That is, whether to process data in-house on the user's own computer or on someone else's equipment. Over the years, remote computer service companies have developed to provide sophisticated and convenient computing services to subscribers and customers from remote facilities . . . Data is most often transmitted between these services and their customers by means of electronic communication.[108]

What does it mean to provide "computer storage or processing"? Computer storage—in 1986 as well as today—is a pretty clear concept.[109] If a Web-mail provider, even one that indexes a user's mail and provides contextually relevant e-commerce services, stores e-mail for its users, the provider should be considered a provider of RCS and should be subject to the SCA.

But what if a user of Google's search engine does not use Gmail, but Google has the user's personally identifiable information because the user has registered for another Google service?[110] Could Google's provision of search

---

[105] TEMPLETON, *supra* note 10, *at* http://www.templetons.com/brad/gmail.html.

[106] Electronic Frontier Foundation, Deep Links, *Gmail: A Rough Guide to Protecting Your Privacy*, *at* http://www.eff.org/deeplinks/archives/001425.php (Apr. 15, 2004).

[107] 18 U.S.C. §§ 2711(2), 2510(14).

[108] S. REP. NO. 99-541, at 10-11, (1986), *reprinted in* 1986 U.S.C.C.A.N. 3555.

[109] Kerr, *supra* note 43, at 1229-30.

[110] Google's Orkut service, a social networking tool also in a beta test mode as of this

engine services to that user be classified as RCS under the SCA to protect the user's search history from unwanted disclosure?[111]

To the drafters of the SCA, computer "processing" meant "outsourcing," or the use of remote computers to perform complicated tasks that one's own personal computer could not perform.[112] In the 1980s, tasks that a modern computer user could accomplish with a spreadsheet program and an average PC needed to be outsourced to remote computing services with greater processing power.[113] RCS "raised privacy concerns because the service providers often retained . . . copies of their customers' files for long periods of time."[114]

RCS seems to have been forsaken as a viable category for online service providers under the SCA. Based on the reported decisions in *Crowley v. CyberSource Corporation* and *In re Northwest Airlines Privacy Litigation*, the litigants seeking to bring a Web site into the reach of the SCA appear not to have argued that the sites in question should be considered RCS providers as an alternative to their ECS-based theories. Moreover, at least one commentator has argued that the RCS label is an awkward description for a site like eBay because this "destination" site does not "process" information for its users in the way envisioned by the definition of RCS.[115]

The label might be less awkward, however, when applied to a search engine. Not only does the relationship between a search engine user and the search engine functionally resemble the relationship between a user of RCS and an RCS provider, but also the privacy implications raised by both relationships are quite similar. In a very real sense, users are outsourcing to remote computers with greater processing power tasks that the users' computers cannot easily perform by themselves.[116] A search engine is not a destination site where

---

writing, is an example of another Google service requiring registration. *See* http://www.orkut.com (last visited Nov. 13, 2004).

[111] While a user's search terms are stored on a search engine's servers, they are not really stored for the user (seeing that one cannot even access one's own search history). As a result, the possibility that the search engine is providing "processing" services to the user is the more fruitful avenue to explore by way of bringing a search engine into the realm of providing RCS under the SCA.

[112] Kerr, *supra* note 43, at 1214.

[113] *See, e.g., id.*

[114] *Id.*

[115] *Id. But see* Steve Jackson Games, Inc. v. United States Secret Serv., 816 F. Supp. 432, 434, 443 (W.D. Tex. 1993), *aff'd*, 36 F.3d 457 (5th Cir. 1994) (stating that an "electronic bulletin board system" where video game enthusiasts could "receiv[e] and pass[] on information" was a provider of RCS).

[116] One privacy lawyer I spoke to suggested that the applicability of the RCS label might turn on whether a user's computer was at least theoretically capable of performing the outsourced task itself, even at great inconvenience and expense, as compared to a situation, like with eBay or Amazon, where the desired result, e.g., shopping or participating in auctions, is flat out impossible for a user's computer to do on its own. Telephone Interview with Kevin S. Bankston, Attorney, Equal Justice Works & Bruce J. Ennis Fellow, Electronic Frontier Foundation (Sept. 23, 2004). This is a close reading of the legislative history, but I do not think it necessarily prevents a search engine from being an RCS provider. Though it is considerably less comprehensive and less efficient to do so, a user with an Internet connection can "search" the Web without a search engine. By browsing through a Web site directory, for example, such as the one maintained by Yahoo or the DMOZ Open Directory

users shop or read or engage in community activity. It is a service to which users send electronic communications (search terms); the service processes those electronic communications to produce results that are sent back to the users. Furthermore, in the course of providing this service, the search engine retains copies of the users' electronic communications, which is the aspect of the RCS relationship that originally raised privacy concerns for the drafters of the SCA.

Consequently, the best approach from a doctrinal perspective is to label a search engine provider a provider of RCS, thus bringing a user's stored search history (in addition to stored e-mail) into the realm of the SCA's privacy protections. Search histories then could not be disclosed to third parties for marketing purposes absent a user's consent.

IV. CONCLUSION

As a leader in the development of innovative Web services, Google finds itself in the perhaps unenviable role of lightning rod for a host of critically important online privacy issues, including the privacy status of search engine histories and the enforceability of Web site terms of use agreements. With the launch of Gmail, Google created a new privacy dilemma associated with computer software that can analyze a user's e-mail to ascertain its meaning and maintain records of the resulting data.

Though California's legislature has taken the lead in attempting to address some of these issues on a state level, its efforts ultimately are not far-reaching enough to resolve the privacy concerns of the United States Web-using population as a whole. Accordingly, the focal point of the online privacy debate is the nearly twenty-year-old Stored Communications Act. A methodical analysis of the SCA indicates that it likely protects the users of Google and other similar providers from the unauthorized disclosure of search histories and e-mail content profiles.

This analysis is by no means unassailable, however. The ideal solution is for Congress to update the SCA to properly account for today's pervasive Web technologies. To the extent that Congress does not soon take on the task of revising the SCA, the courts will have to supply the interpretations necessary for logical and consistent application of this aging statute. But for courts to be able to pass on such questions and hence evolve the law, litigants must be aware of the interpretive possibilities available when applying the SCA to contemporary Web technology. With any luck, analytical exercises like the one in this Comment will help create momentum for such progress.

Project (http://dmoz.org), a user can "manually" search through sites associated with a given topic, and can follow links on those sites to see where they lead, and so on. This is not the way most people explore the Web in the age of Google, but it is certainly possible that manual browsing could uncover at least some of the same sites that a nearly instantaneous Google search would produce. Given the huge time savings and broader reach associated with using Google instead of manually searching the Web, we outsource the searching function to a search engine.