

Privacy on Planet Google: Using the Theory of “Contextual Integrity” to Clarify the Privacy Threats of Google’s Quest for the Perfect Search Engine

I. PLANET GOOGLE

THE WEB SEARCH ENGINE, GOOGLE, HAS ESTABLISHED ITSELF as the prevailing interface for searching and accessing virtually all information on the Web. Google originated in 1996 as a Ph.D. research project by Larry Page and Sergey Brin at Stanford University.¹ Although it was a relative latecomer to the search engine industry,² Google’s Web search engine quickly rose to dominance, processing over 4.1 billion search queries in July 2007, over half of all Web searches performed.³ Google’s mission, stated quite simply and innocuously, is to “organize the world’s information and make it universally accessible and useful.”⁴ In pursuit of this goal, Google’s searchable index has expanded beyond websites to include other online documents as well, such as images, news feeds, Usenet archives, and video files.⁵

* Ph.D., Media, Culture and Communication, M.A., Media Ecology, New York University; Microsoft Resident Fellow, Information Society Project, Yale Law School; michael.zimmer@yale.edu.

1. See Sergey Brin & Lawrence Page, *The Anatomy of A Large-Scale Hypertextual Web Search Engine* (1998) (unpublished Ph.D. dissertation, Stanford University), available at <http://infolab.stanford.edu/~backrub/google.html>.

2. See Posting of Danny Sullivan to SearchEngineWatch.com, <http://searchenginewatch.com/showPage.html?page=2156481> (Jan. 28, 2005) (tracking the changes in search engine size from 1995 to 2003). As recent as December 2000, Google processed only 5 percent of Web searches. *Id.*

3. Press Release, Nielsen//NetRatings, Nielsen//NetRatings Announces July U.S. Search Share Rankings (Aug. 20, 2007), available at http://www.netratings.com/pr/pr_070820.pdf. At its peak in early 2004, Google handled upwards of 80 percent of all search requests on the Web through its own website and clients such as Yahoo!, AOL, and CNN relied on Google for their customer’s search engine results. While maintaining its dominant market position, Google’s share of searches performed fell significantly in 2004 when Yahoo! dropped Google’s search technology for its own. Evan Hansen & Jim Hu, *Yahoo, Google Primed for Search War*, CNET News.com, Jan. 14, 2004, <http://www.news.com/2100-1024-5141328.html>.

4. Google, Corporate Information, <http://www.google.com/corporate/index.html> (last visited Sept. 27, 2007) [hereinafter Google, Corporate Information].

5. It is estimated that Google has indexed nearly 70 percent of the total World Wide Web. Posting of Danny Sullivan to SearchEngineWatch.com, <http://blog.searchenginewatch.com/blog/050517-075657> (May 17, 2005 7:56 EST). They claim to have an index more than three times larger than that of any other search engine. Google Help Center, Sizing Up Search Engines, <http://www.google.com/help/indexsize.html> (last visited Sept. 13, 2007).

PRIVACY ON PLANET GOOGLE

Additionally, Google has begun digitizing the material world, adding the contents of popular books, university libraries, maps, and satellite images to their growing index.⁶ Further, users now can search the files on their hard drives, send e-mail and instant messages, shop online, engage in social networking, organize photos, share videos, collaborate on projects, and publish blogs through various Google offerings.⁷ Consequently, a growing contingent of Internet users search, find, and relate to information, as well as communicate, collaborate, navigate, and organize their lives through Google's growing infrastructure of search-related services and tools, affectionately known as "Planet Google."⁸

The gravitational pull of Planet Google is both alluring and powerful. Lured by a tantalizing collection of innovative, user-friendly, and useful tools, consumers increasingly relish becoming citizens of "Planet Google," as explained in a recent *New York Times* article profiling a citizen of this brave new world:

As Dan Firger, a law student at New York University, strolls from class to class during the course of his day or pauses for a breather in Washington Square Park, his cellphone is routinely buzzing inside his messenger bag. He can often guess who it is: Google.

Six to eight times a day text messages pop up, courtesy of Google Calendar, a free daily organizer introduced this year. The program can scan appointments and send reminders of coming events.

Google is everywhere in Mr. Firger's life. He scours the Web with its search engine; he chats with friends in Bolivia using Google Talk; and he receives e-mail messages on a Google Gmail account.

"I find myself getting sucked down the Google wormhole," Mr. Firger said with equal parts resentment and admiration. "It's all part of Google's benign dictatorship of your life."

Planet Google's benign flavor stems from its carefully crafted ethos, distilled in its pithy, informal corporate motto, "Don't be evil."¹⁰ In their "Letter from the Founders" submitted in anticipation of Google's initial public offering, Brin and Page state that Google is "not a conventional company" and that they aspire to make "the world a better place" by "improv[ing] the lives of as many people as possible."¹¹ Elsewhere, Brin and Page have noted their desire to "have positive social effects"

6. See Edward Wyatt, *Google Adds Library Texts to Search Database*, N.Y. TIMES, Nov. 3, 2005, at C11; see also John Markoff, *Technology Briefs Services: Google to Enhance Map Service*, N.Y. TIMES, Apr. 5, 2005, at C13.

7. For a brief overview of Google's products and services, see *infra* tbl.1.

8. Alex Williams, *Planet Google Wants You*, N.Y. TIMES, Oct. 15, 2006, § 9, at 1.

9. *Id.*

10. Google Investor Relations, *Google Code of Conduct*, <http://investor.google.com/conduct.html> (last visited Sept. 27, 2007).

11. Google Investor Relations, *Letter From the Founders*, http://investor.google.com/ipo_letter.html, (last visited Sept. 27, 2007).

MICHAEL ZIMMER

and to make Google a “social good.”¹² This apparent benevolence is embraced by many who seem ready to give themselves up to Planet Google’s gravitational pull, noting how it is “so much of an improvement on how life was before”¹³ Google, armed with millions of users and a 71 percent brand loyalty rating,¹⁴ “is poised to become the perfect, all-seeing, all-knowing, all-powerful force of the 21st century.”¹⁵

II. A FAUSTIAN BARGAIN

Planet Google reflects an omniscient and omnipresent ideal—a “perfect search engine”¹⁶ that promises to provide a new means of accessing information, communicating with others, and organizing our lives. Yet, while this new information ecosystem is being delivered by a company whose motto is “Don’t be evil,” we are reminded by cultural critic Neil Postman that the true relationship between a society and its technology is often not purely benevolent, but instead, may require a sacrifice for society to enjoy its benefits, what Postman recognizes as a Faustian bargain:

*[A]nyone who has studied the history of technology knows that technological change is always a Faustian bargain: Technology giveth and technology taketh away, and not always in equal measure. A new technology sometimes creates more than it destroys. Sometimes, it destroys more than it creates. But it is never one-sided.*¹⁷

History has revealed how such a Faustian bargain persists with many technologies designed to enhance our informational lives, especially regarding the privacy of personal information related to information-seeking activities. The introduction of computer systems in libraries, for example, has improved management of collec-

12. Google Inc., Amendment No. 7 to Registration Statement (Form S-1), at B-1 (Aug. 13, 2004).

13. Williams, *supra* note 8.

14. Posting of Jim Hedger to StepForth SEO News Blog, <http://news.stepforth.com/blog/2006/02/google-users-display-biggest-brand.php> (Feb. 22, 2006, 16:31 EST).

15. Chris Ayres, *Google: Could This Be the New God in the Machine?*, *TIMES* (London), Nov. 1, 2003, at 4; see also Thomas L. Friedman, *Is Google God?*, *N.Y. TIMES*, June 29, 2003, § 4, at 13; Michael Gorman, Commentary, *Google and God’s Mind*, *L.A. TIMES*, Dec. 17, 2004, at 15.

16. For more on the “perfect search engine,” see Searchblog, <http://battellemedia.com/archives/000878.php> (Sept. 8, 2004, 20:24 EST). Early on, Google recognized the importance of designing a perfect search engine: the company’s very first press release noted that “[a] perfect search engine will process and understand all the information in the world. . . . That is where Google is headed.” Press Release, Google, Google Receives \$25 Million in Equity Funding (June 7, 1999), available at <http://www.google.com/press/pressrel/pressrelease1.html>. Google co-founder Larry Page later reiterated the goal of achieving the perfect search: “The perfect search engine . . . would understand exactly what you mean and give back exactly what you want.” Google, Corporate Information, Our Philosophy, <http://www.google.com/intl/en/corporate/tenthings.html> (last visited Sept. 27, 2007).

17. Neil Postman, *Informing Ourselves to Death*, Address Before the German Informatics Society (Oct. 11, 1990), available at http://www.iff.org/Net_culture/Criticisms/informing_ourselves_to_death.paper.

tions and circulation and also facilitated the recording of each patron's borrowing activities into databases that might be subpoenaed by government authorities.¹⁸ Similarly, while originally developed to facilitate Web-based commerce and enhance online information services, the widespread use of Web cookies also has facilitated the tracking of users as they navigate the Web, often without their awareness or consent.¹⁹ Furthermore, early digital rights management systems were intended to help increase the flow of copyright-protected content online, but resulted instead in a system capable of monitoring and tracking users' consumption of intellectual materials, thereby diminishing the freedom to consume copyright-protected content, while also posing a threat to the privacy of an individual's reading, viewing, and listening habits.²⁰

Concerns of a similar Faustian bargain persist with the rise of Planet Google as the expansion of Google's reach into so many areas of people's lives has left some uneasy, such as one paradigmatic Google user who expressed "feeling a 'weird tension' about his love of Google's products and his fear about its omnipresence in his life."²¹ Such anxieties are the result of many different fears, but tend to coalesce on the depth of knowledge that Google can gain over individual user's online intellectual and social activities. The amount of information Google possesses about individual users became most apparent, ironically, when the search engine refused to comply with a U.S. Department of Justice request to gain access to their search history logs.²² Given this event, combined with AOL's release of over twenty million search queries from 658,000 of its users that were insufficiently anonymized and exposed user queries,²³ anxiety has been on the rise among many searchers about

18. See, e.g., HERBERT N. FOERSTEL, *SURVEILLANCE IN THE STACKS: THE FBI'S LIBRARY AWARENESS PROGRAM 5*, 10 (Greenwood Publishing Press 1991); Bruce M. Kennedy, *Confidentiality of Library Records: A Survey of Problems, Policies, and Laws*, 81 LAW LIBR. J. 733 (1989); Robert D. McFadden, *F.B.I. in New York Asks Librarians' Aid in Reporting on Spies*, N.Y. TIMES, Sept. 18, 1987, at A1.

19. See, e.g., Colin J. Bennett, *Cookies, Web Bugs, Webcams and Cue Cats: Patterns of Surveillance on the World Wide Web*, 3 ETHICS & INFO. TECH. 197, 197 (2001); Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1226-27 (1998); David M. Kristol, *HTTP Cookies: Standards, Privacy, and Politics*, 1 ACM TRANSACTIONS ON INTERNET TECH. 151, 162 (2001); John Schwartz, *Giving Web a Memory Cost Its Users Privacy*, N.Y. TIMES, Sept. 4, 2001, at A1.

20. See Joan Feigenbaum et al., *Privacy Engineering for Digital Rights Management Systems*, in LECTURE NOTES IN COMPUTER SCIENCE 76 (2001); DEIRDRE MULLIGAN ET AL., *ACM WORKSHOP ON DIGITAL RIGHTS MGMT., HOW DRM-BASED CONTENT DELIVERY SYSTEMS DISRUPT EXPECTATIONS OF "PERSONAL USE"* (2003), available at <http://portal.acm.org/citation.cfm?id=947380.947391>; Julie E. Cohen, *DRM and Privacy*, 18 BERKELEY TECH. L.J., 575, 575-77 (2003); Electronic Privacy Information Center, *Digital Rights Management and Privacy*, <http://www.epic.org/privacy/drm/> (last visited Sept. 27, 2007).

21. Williams, *supra* note 8.

22. It should be noted that this revelation only became public after Google resisted the government efforts to obtain its own search records. Katie Hafner & Matt Richtel, *Google Resists U.S. Subpoena of Search Data*, N.Y. TIMES, Jan. 20, 2006, at A1; Howard Mintz, *Feds After Google Data*, SAN JOSE MERCURY NEWS, Jan. 19, 2006, at BU1.

23. See, e.g., Michael Barbaro & Tom Zeller Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, N.Y. TIMES, Aug. 9, 2006, at A1; Saul Hansell, *AOL Removes Search Data on Vast Group of Web Users*, N.Y. TIMES, Aug. 8, 2006, at C4; Declan McCullagh, *AOL's Disturbing Glimpse Into Users' Lives*, CNET NEWS.COM, Aug. 9, 2006, http://news.com.com/AOLs+disturbing+glimpse+into+users+lives/2100-1030_3-6103098.html.

the systematic monitoring of their online information-seeking activities on increasingly-indispensable search engine services.²⁴ These anxieties led news organizations to investigate and report on the information search engines routinely collect from their users,²⁵ which led to criticism by various advocacy groups,²⁶ which, in turn, has led to inquiries and investigations by European and U.S. government regulators,²⁷ all with Google as the primary target. Moreover, these concerns tend to focus solely on search query data retained in Google's server logs, ignoring the myriad of user information capturable by the robust infrastructure of products and services—linked together by cookies and user accounts—amassed within Planet Google.²⁸

A Faustian bargain has indeed emerged in order to “organize the world's information and make it universally accessible and useful.”²⁹ Planet Google gains access to vast amounts of personal information that is potentially threatening to the privacy of the very users it aims to serve.³⁰ Yet, despite this, users continue to be attracted to the gravitational pull of Planet Google, embracing and integrating it into their daily lives. Our previously-quoted faithful Google user interviewed by the *New York Times* puts it best: “I don't know if I want all my personal information saved on this massive server in Mountain View [Google's headquarters], but it is so much of an improvement on how life was before, I can't help it.”³¹ Fueling such acquiescence to the Faustian bargain implicit in the perfect search engine are the persistent claims by pundits, journalists, and search engine companies themselves that no personal information is ever collected,³² that “the privacy concerns are

24. See, e.g., Barbaro & Zeller, *supra* note 23; Katie Hafner et al., *After Subpoenas, Internet Searches Give Some Pause*, N.Y. TIMES, Jan. 25, 2006, at A1; Steven Levy, *Technology: Searching for Searches*, NEWSWEEK, Jan. 30, 2006, at 34; Kevin Maney, Column, *AOL's Data Sketch Sometimes Scary Picture of Personalities Searching Net*, USA TODAY, Aug. 9, 2006, at 4B.

25. See, e.g., Elise Ackerman, *What Do Google, Yahoo, AOL and Microsoft's MSN Know About You?*, SAN JOSE MERCURY NEWS, Aug. 20, 2006; Joanna Glasner, *Worker Privacy: You Have None*, WIRED.COM, Dec. 9, 2005, <http://wired.com/print/politics/security/news/2005/12/69732>.

26. See, e.g., PRIVACY INT'L, *A RACE TO THE BOTTOM: PRIVACY RANKING OF INTERNET SERVICE COMPANIES* (2007), <http://www.privacyinternational.org/article.shtml?cmd%5B347%5D=x-347-553961>; Electronic Frontier Foundation, *Privacy and Search Engines*, <http://www.eff.org/Privacy/search/> (last visited Sept. 28, 2007).

27. See, e.g., *EU Data Privacy Officers Launch Investigation into Google's Internet Search Engine*, INT'L HERALD TRIB., May 25, 2007, available at <http://www.iht.com/articles/ap/2007/05/25/business/EU-FIN-EU-Google-Privacy-Probe.php>; Steve Lohr, *Google Deal Said to Bring U.S. Scrutiny*, N.Y. TIMES, May 29, 2007, at C1.

28. See *infra* tbl.2 for a sample listing of personal information capturable across Google's suite of products and services.

29. Google, Corporate Information, *supra* note 4.

30. See *supra* notes 19–23 and accompanying text.

31. Williams, *supra* note 8 and accompanying text.

32. See Danny Sullivan, *Search Privacy: An Issue?, Part 1*, CLICKZ, May 21, 2003, <http://www.clickz.com/showPage.html?page=2207951>; Danny Sullivan, *Search Privacy at Google & Other Search Engines*, SEARCHENGINEWATCH.COM, Apr. 2, 2003, <http://searchenginewatch.com/showPage.html?page=2189531> [hereinafter, Sullivan, *Google Search Privacy*]. This claim was disproven by the ease of identifying users from the “anonymized” AOL search records data release. See Barbaro & Zeller, *supra* note 23.

probably overblown,"³³ or that "[i]f you have nothing to hide when you use the internet, you have nothing to fear."³⁴

An equally common—and equally problematic—response to concerns about the information-gathering abilities of Planet Google is the argument that users knowingly share the information, and already share similar information with other people and institutions.³⁵ For example, addressing concerns that Google is able to track and collect all of a user's browsing activity through the Web Accelerator product, a company representative attempted to quell the privacy issue by asserting that "Web Accelerator receives much of the same kind of information [that] people already share with their Internet service providers . . . when surfing the Web."³⁶ Others have argued that Google logging book searches is no different than asking a librarian for help finding particular books,³⁷ and that Google's scanning of Gmail messages in order to place contextual advertisements is no different than spam filters.³⁸

Such appeals that the status quo simply has been maintained have clouded many discussions and concerns about user privacy. For those concerned with search engine privacy³⁹ and wishing to renegotiate the Faustian bargain that Planet Google represents, we must broaden the conceptual understanding of the privacy threats inherent in the new depth of knowledge gained by these powerful new information infrastructures. To bring conceptual clarity and a normative understanding of some of the ways in which Planet Google bears on user privacy, we can turn to Nissenbaum's theory of "privacy as contextual integrity"⁴⁰ to reveal the often-ignored (and sometimes openly refuted) privacy concerns implicit in Google's quest to build the perfect search engine.

33. See Elinor Mills, *Google Balances Privacy, Reach*, CNET NEWS.COM, Aug. 3, 2005, http://news.com.com/Google+balances+privacy,+reach/2100-1032_3-5787483.html.

34. Peter Griffin, *Big Brother Wants to Track Your Cybersteps*, N.Z. HERALD, Jan. 27, 2006.

35. Sullivan, *Google Search Privacy*, *supra* note 32.

36. Matt Hines, *Google Tool to Speed Web Surfing*, CNET NEWS.COM, May 6, 2005, http://news.com.com/Google+tool+to+speed+Web+surfing/2100-1032_3-5696496.html.

37. See *supra* note 18 and accompanying text.

38. See generally Grant Yang, Note, *Stop the Abuse of Gmail!*, 14 DUKE L. & TECH. REV. 1 (2005), available at <http://www.law.duke.edu/journals/dltr/articles/pdf/2005dltr0014.pdf> (arguing that Google's practice of scanning incoming messages for advertising purposes is no different than similar activities of its peers).

39. Concern about search privacy has prompted various symposia and workshops among lawyers, scholars, policymakers, and advocates, including the "Regulating Search: A Symposium on Search Engines, Law, and Public Policy" held at Yale Law School. See Symposium, Yale Information Society Project, *Regulating Search: A Symposium on Search Engines, Law, and Public Policy*, available at <http://islandia.law.yale.edu/isp/regulatingsearch.html>; Symposium, The Ethics and Politics of Search Engines, Santa Clara University Markkula Center for Applied Ethics, available at <http://www.scu.edu/sts/Search-Engine-Event.cfm>. New York University School of Law's Information Law Institute also hosted an informal panel discussion on this topic entitled "A Discussion About Privacy in Web-Search" on March 30, 2007.

40. See Helen Nissenbaum, *Protecting Privacy in an Information Age: The Problem of Privacy in Public*, 17 LAW & PHIL. 559, 582 (1998) [hereinafter Nissenbaum, *Protecting Privacy*]. See generally Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 WASH. L. REV. 119 (2004) [hereinafter Nissenbaum, *Privacy as Contextual Integrity*].

III. PRIVACY AS CONTEXTUAL INTEGRITY

Contextual integrity is a benchmark theory of privacy, a conceptual framework that links the protection of personal information to the norms of personal information flow within specific contexts. Through her development of this new theory of privacy, Nissenbaum argues that informational norms—specific to particular contexts—govern the flow of personal information from one entity to another.⁴¹ In doing so, the theory of contextual integrity provides a framework for evaluating the flow of personal information between agents to help identify and explain why certain patterns of information flow are acceptable in one context, but viewed as problematic in another.⁴²

Rejecting the traditional dichotomy of public versus private spaces—and its related clean division between public and private information—a key recognition within contextual integrity is that the multitude of information-sharing activities take place in a “plurality of distinct realms.”⁴³ Within each of these contexts, norms exist—either implicitly or explicitly—which both shape and limit our roles, behaviors, and expectations. For example, it might be acceptable for me to approach a stranger and offer her a hug at a moving religious service, but not in the grocery store. A judge might willingly accept birthday gifts from colleagues, but would hesitate to accept one from a lawyer currently arguing a case in her courtroom. It is deemed appropriate for a physician to ask me my age, but not for a bank teller. While it is necessary for an airline to know my destination city, it would be inappropriate for them to ask where I will be staying, with whom I will be meeting, or what we will be discussing.

In short, norms of behavior vary based on the particular context. The latter examples above reveal the ways in which norms govern the flow of *personal information* in particular contexts. Whether in talking with a physician, purchasing items in a store, or simply walking through a public park, norms of information flow govern what type of and how much personal information is relevant and appropriate for sharing with others. The theory of contextual integrity is built around the notion that there are “no arenas of life *not* governed by *norms of information flow*”⁴⁴ These norms explain the boundaries of our underlying entitlements regarding personal information, and our privacy is invaded when these informational norms are contravened.

Within each context, the relevant *agents*, the *types of information*, and *transmission principles* combine to shape the governing informational norms.⁴⁵ Informational norms always include three relevant agents: the information subject, the one

41. Nissenbaum, *Protecting Privacy*, *supra* note 40, at 581–82.

42. *See id.* at 584–85; Nissenbaum, *Privacy as Contextual Integrity*, *supra* note 40, at 138–40.

43. Nissenbaum, *Privacy as Contextual Integrity*, *supra* note 40, at 137.

44. *Id.*

45. ADAM BARTH ET AL., *PRIVACY AND CONTEXTUAL INTEGRITY: FRAMEWORK AND APPLICATIONS* 184, 186 (2006 IEEE Symposium on Security and Privacy, 2006).

who has the information and is distributing it (who may or may not be the subject), and the one who receives the information.⁴⁶ The informational norms within a particular context dictate the roles of the agents, each “associated with a set of duties and privileges.”⁴⁷ For example, in the healthcare context, the personal information shared by the patient (the subject and sender) depends very much on who the recipient is—the physician, the receptionist, the claims processor, and so on. In turn, the rules governing the transmission of personal information by the physician depends on who the recipient is—the patient, a colleague, the insurance company, and so on.⁴⁸ The specification and roles of the various agents are key variables affecting the maintenance of contextual integrity within a particular context of informational norms.⁴⁹

The type of information in question is another defining aspect of informational norms.⁵⁰ Unlike most theories of privacy, contextual integrity rejects the notion that information types fit into a rigid dichotomy of public or private.⁵¹ “Instead, there is potentially an indefinite variety of types of information that could feature in the informational norms of a given context,” and whose categorization might shift from one context to another.⁵² Again, in a healthcare setting, different informational norms apply depending on whether the information is a patient’s medical condition, home address, or account balance. The notion of “appropriateness” is a useful “way to signal whether the type of information in question conforms to the relevant informational norms.”⁵³ Norms of appropriateness “circumscribe the type or nature of information about various individuals that, within a given context, is allowable, expected, or even demanded to be revealed.”⁵⁴ In some contexts, norms of appropriateness are very open, such as in a personal friendship where personal information tends to flow freely. In other contexts, such as the job interview or classroom, more explicit and restrictive norms of appropriateness prevail, and the flow of appropriate personal information is more highly regulated. Nevertheless, norms of appropriateness apply in all situations: among both strangers and loved ones, in personal and professional interactions, in private and public.

The notion of a transmission principle may be the most distinctive component of the informational norms that frame contextual integrity. Transmission principles place constraints on the flow or distribution of information from agent to agent within a context.⁵⁵ Confidentiality is an example of a transmission principle

46. *Id.*

47. *Id.*

48. *Id.*

49. *Id.*

50. *Id.*

51. *Id.*

52. *Id.*

53. *Id.*

54. Nissenbaum, *Privacy as Contextual Integrity*, *supra* note 40, at 138.

55. See BARTH ET AL., *supra* note 45, at 186.

where the agent receiving the information is prohibited from transmitting the information to other agents.⁵⁶ In some contexts, the information flow is bi-directional, representing the transmission principle of reciprocity.⁵⁷ In others, agents might be compelled by a legal authority to divulge information; in still others, the transfer of information might be voluntary, or made only when proper consent is provided.⁵⁸ For example, transmission principles outlined in professional codes of ethics dictate that my physician can share only some of my personal information with other doctors: she might share my symptoms or family history to aid in diagnosis, but not my name.⁵⁹ More restrictive principles have been codified in our legal systems, such as the burden necessary for law enforcement to obtain my detailed phone records.⁶⁰ Informational norms prescribe which transmission principles ought to govern the flow of information in particular contexts, and such norms are violated if the principles are not followed.

To summarize, within each context, informational norms are shaped by the identification of the relevant agents,⁶¹ the types of information,⁶² and the appropriate transmission principles.⁶³ With these components, contextual integrity generates a decision heuristic to help explain when privacy objections are likely to be aroused by the introduction of a new technology or practice.⁶⁴ Rather than aspiring to universal prescriptions for what is public information versus private information, contextual integrity works from within the normative bounds of a particular context. It is designed to consider whether, and in what ways, the introduction of a new practice or technology *into a given context* might breach the governing informational norms.⁶⁵ In order to determine if contextual integrity has been maintained, we must consider how the new technology or practice affects the agents involved, the appropriateness and type of information, and the transmission principles that constrain the flow of information from agent to agent. If the introduction of a new technology or practice in a given context is found to conflict with the standing informational norms, a red flag is raised, indicating that contextual integrity has been violated. Through such an analysis, potential privacy violations can be identified that might have otherwise gone overlooked. When applied to Google's

56. *See id.*

57. *See id.*

58. *Id.*

59. *See* AM. MED. ASS'N, CODE OF MEDICAL ETHICS § 5.05, 7.01 (1994).

60. *See generally* 18 U.S.C. § 2703 (2000) (codifying what a government agent must show to obtain a court order granting access to civilian communication records); *id.* § 3121 (stating that no tracking device may be installed without a court order); *id.* § 3123 (creating the procedure for issuance of a pen register or tracing device).

61. *See supra* notes 45–46 and accompanying text.

62. *See supra* notes 50–54 and accompanying text.

63. *See* BARTH ET AL., *supra* note 45, at 186, and text accompanying notes 55–58.

64. *See* BARTH ET AL., *supra* note 45; Nissenbaum, *Privacy as Contextual Integrity*, *supra* note 40, at 143–47 (stating that a presumption exists in favor of the status quo when new information technology becomes available).

65. *See* BARTH ET AL., *supra* note 45.

quest for the perfect search engine, contextual integrity provides the means to understand how Planet Google's vast infrastructure violates existing norms of information flows, countering the rhetorical claims to the contrary.

IV. CONTEXTUAL INTEGRITY AND THE DEPTH OF KNOWLEDGE OF INDIVIDUAL USERS

To determine the potential impact of Google's quest for the perfect search on the informational norms that dictate the flow of personal information when engaging in social and intellectual activities, we can create a thought experiment featuring two ideal typical information seekers, Elizabeth "Libby" Doe and Annette "Netty" Roe. Libby and Netty are nearly identical in their personal, social, political, cultural and economic characteristics. Both are thirty-year-old, single, gay south-Asian women. Both are Hindu, live in Brooklyn, New York and tend to vote for Democrats. Libby and Netty are graduate students at New York University, studying political science and feminist theory. They enjoy sports and cooking as hobbies; both are thinking of having a baby but have concerns due to being diabetic. They have similar investment portfolios, enjoy keeping in touch with friends, and like to share photos and stories from their travels.

The two differ, however, in how they navigate their "informational spheres." Libby prefers traditional, "old-fashioned" methods of information-seeking and communication: reading print newspapers, watching television news, relying on word-of-mouth, and writing correspondence. While not averse to using the Internet, when Libby needs to find information on a topic, she prefers visiting the library. Netty, on the other hand, relies heavily on the Internet to manage information and communicate with others. When Netty needs information about a topic, she "Googles" it. In fact, Netty relies on Google's broad array of products and services for virtually all of her online activities.

When navigating their respective "spheres of information," both Libby and Netty inevitably share bits of personal information with others. Appendix B describes these flows of personal information within each of the nine distinct contexts of information-seeking identified in the previous chapter, comparing the personal information shared by Libby in her traditional information-seeking methods, with the personal information shared by Netty, who relies almost exclusively on "Planet Google" to access and organize information. Building from this narrative of Libby and Netty's differing informational practices and flows, we can attempt to apply contextual integrity as a benchmark to determine if privacy violations exist. Assessing the information practices and flows from our thought experiment with any degree of certainty is not easy, but we can approximate the particular agents, information types, and transmission principles that govern information flows from our

thought experiment.⁶⁶ Within each context, there is evidence of shifts in each component of the governing informational norms.

For example, Libby's interactions are scattered among various agents, resulting in a fragmented dispersal of personal information. Rarely would any single receiver obtain information from multiple engagements with Libby, let alone across contexts. Indeed, given Libby's information practices in some contexts, there is no agent at all in receipt of her personal information. In contrast, *all* of Netty's information-seeking activities involve Google as an agent receiving personal information, allowing a level of consolidation not possible in Libby's scenario. The result of having one single entity act as a receiving agent across the various contexts represents a significant shift in informational norms.

The types of information shared by Libby tend to be incomplete, scattered verbal requests to librarians or booksellers, and the occasional transactional (but not window-shopping) data provided to retailers. Some agents have access to Libby's home address or financial data because she is a subscriber or repeat customer, but, in general, the information she divulges is only a fragment of the entire picture of her activities in each context. Netty, on the other hand, provides Google with much more complete sets of information through her interactions with Google products and services. The information is digital, allowing for simpler storage, processing, and sharing, and its accuracy is difficult to dispute.

Finally, the key difference in transmission principles for our two information seekers is that Libby voluntarily divulges information when she decides to interact directly with librarians, booksellers, and so on, while Netty is compelled to allow Google to track and collect her information browsing and usage habits as a condition of using its products and services. Further differences exist in terms of how these agents might share the information with other parties. In Libby's case, while some of the information divulged in commercial transactions might be used for marketing purposes, the librarians with whom she interacts are bound by a code of ethics. Also, the phone and financial companies who receive information must adhere to strict laws protecting consumer privacy. For Netty, in nearly all cases, use of the information by Google is dictated by its privacy policy, which states, in part: "We may combine the information you submit under your account with information from other Google services or third parties in order to provide you with a better experience and to improve the quality of our services."⁶⁷ Google further states that it will share personal information with third parties when "[w]e have a good faith belief that access, use, preservation or disclosure of such information is reasonably necessary to . . . satisfy any applicable law, regulation, legal process or enforceable governmental request"⁶⁸

66. *See infra* tbl.3.

67. Google Privacy Center, Google Privacy Policy, <http://www.google.com/privacypolicy.html> (last visited Sept. 27, 2007).

68. *Id.*

PRIVACY ON PLANET GOOGLE

To summarize, the shift from traditional information-seeking practices—represented by Libby—to the growing reliance on Google’s quest for the perfect search engine—typified by Netty’s information-seeking practices—represents a potentially significant shift of the existing informational norms. Libby’s divulgence of personal information is scattered, informal, and voluntary. While many of the details of her life are openly shared with various parties—a random checkout person at a bookstore might surmise she is gay and Democratic based on her purchases; she might have mentioned the fact she is Hindu to a librarian; her broker might know her age and address—no single person knows everything about her. Netty’s personal information, on the other hand, is concentrated with one agent—Google—who has amassed a comprehensive digital dossier of nearly every aspect of Netty’s life (her age, sexual orientation, political interests, address, and so on). Thus, a violation of the contextual integrity of the privacy of personal information across these various information-seeking contexts is revealed. It is no longer acceptable to hide behind the rhetoric that no private information is divulged when utilizing the tools that make up the perfect search engine, or that the information shared is simply the same as that provided in other information-seeking scenarios. Revealing this kind of transgression of informational norms helps to expose the Faustian bargain implicit within the quest for the perfect search engine.

V. CONCLUSION

Proponents of the perfect search have succeeded in obscuring its privacy implications by presenting arguments that no real threat to privacy actually exists, as the information shared in the perfect search is not personally identifiable and often is already shared with other entities in other circumstances.⁶⁹ This Essay has utilized Nissenbaum’s theory of “contextual integrity” to provide clarity to the ways Google’s quest for the perfect search engine is altering personal information flows.⁷⁰ Viewing Planet Google through the lens of contextual integrity helps us to focus on the privacy threats represented by this burgeoning information infrastructure, and will arm us for future efforts to design laws, policies, and technologies to protect user privacy when searching the Web.

69. See *supra* notes 32–38 and accompanying text.

70. See *supra* Part III.

TABLE 1: GOOGLE SUITE OF PRODUCTS AND SERVICES (PARTIAL LIST)

Product	Description	Notes
<i>General Information Inquiries</i>		
Web search	- Query-based website searches	
iGoogle	- Customized Google start page with content-specific modules	- Use in conjunction with Google Account is encouraged
Alerts	- E-mail alerts of new Google results for specific search terms	
Image Search	- Query based search for website images	
Video	- Query based search for videos hosted by Google	- Google Video Player available for download
Book Search	- Full text searches of books scanned into Google's servers	- Google Account required in order to limit the number of pages a particular user can view
Blog Search	- Full text search of blog content	
Scholar	- Full text searches of scholarly books and journals	
News	- Full text search of recent news articles	- With a Google Account, users can create customized keyword-based news sections
<i>Communication and Social Networking</i>		
Gmail	- Free Web based e-mail service with contextual advertising	- Creation of Gmail account automatically results in activation of Google Account - Logging into Gmail also logs user into their Google Account
Groups	- Free Web based discussion forums	- Includes complete Usenet archives dating back to 1981 - Google Account required for creation of new Group
Talk	- Web-based instant messaging and voice calling service	- Google Account and Gmail e-mail address required
Blogger	- Web-based blog publishing platform	- Google Account required
Reader	- Web-based news feed reader	- Google Account required
Orkut	- Web-based social networking service	- Invitation-only - Google Account required
Dodgeball	- Location-based social networking service for cellphones	- Google Account required
YouTube	- Online video sharing service	- User account required to upload or comment
<i>Consumer Activities</i>		
Catalog Search	- Full text search of scanned product catalogs	
Product Search	- Full text search of online retailers	- Google Account required for shipping lists
Checkout	- Web-based payment system	- Google Account required

(Table continues)

PRIVACY ON PLANET GOOGLE

TABLE 1: GOOGLE SUITE OF PRODUCTS AND SERVICES (CONTINUED)

Product	Description	Notes
<i>Personal Data Management</i>		
Calendar	- Web-based time-management tool	
Finance	- Portal providing news and financial information about stocks, mutual funds; Ability to track one's financial portfolio	- Google Account required for posting to discussion board
Desktop Search	- Keyword based searching of computer files - Ability to search files on remote computer	- Downloaded computer application
Picasa	- Digital photo storage, editing, and sharing platform	- Downloaded computer application
<i>Productivity</i>		
Docs & Spreadsheets	- Online word processing and spreadsheet application - Capable of sharing and collaboration between users	- Google Account required
<i>Navigation</i>		
Maps	- Street maps and satellite images of global geographical locations - Depending on area, provides information for nearby businesses	
Street View	- In certain locations, can view and navigate within street-level imagery	
Earth	- View global satellite imagery, terrain, 3D buildings	- Downloaded computer application
<i>Internet Browsing</i>		
Bookmarks	- Online storage of website bookmarks	- Google Account required
Notebook	- Browser tool for saving notes while visiting websites	- Google Account required
Toolbar	- Browser tool providing access to various Google products without visiting Google websites	- Some features require Google Account
Web Accelerator	- Software to speed up page load times for faster Web browsing	- Downloaded computer application
Web History	- Archives history of all websites visited	- Google Account required

MICHAEL ZIMMER

**TABLE 2: SAMPLE OF PERSONAL INFORMATION CAPTURABLE BY
GOOGLE'S SUITE OF PRODUCTS**

Product	Information Collected	Notes
<i>General Information Inquiries</i>		
Web search	Web search queries Results clicked	Search for own name, address, social security number, etc is common
iGoogle	News preferences Special interests Zip code	Customizable for nearly any personal interest
Alerts	News preferences Special interests E-mail address	Alerts for a user's own name (vanity search) are common
Image Search	Search queries Results clicked	
Video	Search queries Videos watched/downloaded Credit card information for purchased videos E-mail details for shared videos	Google Video Player contains additional DRM technology to monitor off-site video usage
Book Search	Search queries Results clicked Pages read Bookseller pages viewed	
Blog Search	Search queries Results clicked	
Scholar	Search queries Results clicked Home library (Optional)	
News	News search queries Results clicked	
<i>Communication and Social Networking</i>		
Gmail	Text of email messages E-mail searches performed Email address or cellphone number (used for account creation)	
Groups	Search queries User interests Usage statistics Profile information	Users are encouraged to create detailed profiles, including name, location, industry, homepage, etc.
Talk	Contact list Chat messages Usage statistics	
Blogger	Weblog posts and comments Profile information Usage statistics	Users are encouraged to create detailed profiles, including name, location, gender, birthday, etc.
Reader	Feed subscriptions Usage statistics	
Orkut	Profile information Usage statistics E-mail address and content of invitations	Users are encouraged to create detailed profiles, including name, location, gender, birthday, etc.

(Table continues)

PRIVACY ON PLANET GOOGLE

TABLE 2: SAMPLE OF PERSONAL INFORMATION CAPTURABLE BY GOOGLE'S SUITE OF PRODUCTS (CONTINUED)

Product	Information Collected	Notes
Dodgeball	Profile information E-mail address Location Mobile phone information Text messages sent	User location when messages sent are tracked by Google
YouTube	Usage statistics User account information (age, e-mail address, etc)	
<i>Consumer Activities</i>		
Catalog Search	Product search queries Results clicked	
Froogle	Product search queries Results clicked Sites visited Shopping list	
Checkout	Payment information Shipping address	
<i>Personal Data Management</i>		
Calendar	Profile information Events Usage statistics	
Finance	Financial quotes Discussion group posts Discussion group views Portfolio (optional)	Names and e-mails are displayed with discussion posts
Desktop Search	Profile information Search queries Computer file index (Optional)	Search queries visible to Google under certain circumstances Desktop file index is stored on Google's services if using Search Across Computers
<i>Productivity</i>		
Docs & Spreadsheets	Content of documents Usage statistics	
<i>Navigation</i>		
Maps	Search queries within application Addresses searched Home location (optional)	
<i>Internet Browsing</i>		
Bookmarks	Favorite Websites When visited	
Notebook	Notes and clippings Sites annotated	
Toolbar	Search queries Websites visited	Use of some advanced features routes all browsing traffic through Google servers
Web Accelerator	Websites visited	All browsing traffic is routed through Google servers
Web History	Websites visited	All browsing activity logged by Google

MICHAEL ZIMMER

TABLE 3: DIFFERENCES IN INFORMATIONAL NORMS WITHIN VARIOUS INFORMATION-SEEKING CONTEXTS

Informational norm	Libby	Netty
<i>General Information Inquiries</i>		
Agent (receiver)	- Might interact with various librarians, booksellers, and other information sources	- Google
Information type	- Might verbally divulge personal interests due to interactions with agents - Booksellers might keep transaction logs	- All information queries logged in digital form
Transmission principle	- Information divulged to librarian voluntarily - Retailers might require certain information for purchase - Librarian bound by code of ethics to maintain patron privacy - Merchants might use/sell transaction data	- Information divulged to Google automatically and stored in server logs - Google privacy policy allows use to "provide a better user experience" - May share with third parties to "comply with legal processes"
<i>Communication and Social Networking</i>		
Agent (receiver)	- Recipients of messages - Various communication service providers	- Recipients of messages - Google
Information type	- Recipients see contents of messages - E-mail and phone providers track usage, might scan for spam, etc	- All message content and interactions logged in digital form by Google - Contacts and friends lists stored in databases at Google
Transmission principle	- Information voluntarily divulged to recipients - Recipients might share information; generally bound by norms of friendship	- Information divulged to Google automatically - Privacy policy
<i>Consumer Activities</i>		
Agent (receiver)	- Retailers receive some information for purchases	- Google
Information type	- Purchased items can be tracked - Browsing at select .com sites logged	- All browsing and purchase activity logged
Transmission principle	- Retailer might use/sell transaction data	- Information divulged to Google automatically - Privacy policy

(Table continues)

PRIVACY ON PLANET GOOGLE

TABLE 3: DIFFERENCES IN INFORMATIONAL NORMS WITHIN VARIOUS INFORMATION-SEEKING CONTEXTS (CONTINUED)

Informational norm	Libby	Netty
<i>Personal Data Management</i>		
Agent (receiver)	<ul style="list-style-type: none"> - Calendar information generally not shared with any third party - Financial data generally not shared with anyone outside of broker - No third party has access to computer file information 	<ul style="list-style-type: none"> - All data shared with Google
Information type	<ul style="list-style-type: none"> - Broker maintains financial portfolio information 	<ul style="list-style-type: none"> - Calendar information and queries logged in digital form within Google servers - Financial portfolio information - Some Desktop search terms could be logged via referrer field - All queries logged with "Search Across Computers" feature - Encrypted file index stored at Google with "Search Across Computers" feature
Transmission principle	<ul style="list-style-type: none"> - Broker privacy policy and federal laws restrict use of financial data 	<ul style="list-style-type: none"> - Information divulged to Google automatically - Privacy policy
<i>Internet Browsing</i>		
Agent (receiver)	<ul style="list-style-type: none"> - Individual websites visited keep server logs 	<ul style="list-style-type: none"> - Google
Information type	<ul style="list-style-type: none"> - Typical information collected by Web sites 	<ul style="list-style-type: none"> - Bookmarks, notes, etc - Web History tracks every Web site visited - Some Toolbar functions track every Web site visited - Web Accelerator tracks every Web site visited
Transmission principle	<ul style="list-style-type: none"> - Each site's privacy policy 	<ul style="list-style-type: none"> - Privacy policy