DIFFERENTIAL PRIVACY AS A RESPONSE TO THE REIDENTIFICATION THREAT: THE FACEBOOK ADVERTISER CASE STUDY*

ANDREW CHIN** & ANNE KLINEFELTER***

Recent computer science research on the reidentification of individuals from anonymized data has given some observers in the legal community the impression that the utilization of data is incompatible with strong privacy guarantees, leaving few options for balancing privacy and utility in various data-intensive settings. This bleak assessment is incomplete and somewhat misleading, however, because it fails to recognize the promise of technologies that support anonymity under a standard that computer scientists call differential privacy. This standard is met by a database system that behaves similarly whether or not any particular individual is represented in the database, effectively producing anonymity. Although a number of computer scientists agree that these technologies can offer privacy-protecting advantages over traditional approaches such as redaction of personally identifiable information from shared data, the legal community's critique has focused on the burden that these technologies place on the utility of the data. Empirical evidence, however, suggests that at least one highly successful business, Facebook, has implemented such privacy-preserving technologies in support of anonymity promises while also meeting commercial demands for utility of certain shared data.

This Article uses a reverse-engineering approach to infer that Facebook appears to be using differential privacy-supporting technologies in its interactive query system to report audience reach data to prospective users of its targeted advertising system, without apparent loss of utility. This case study provides an opportunity to consider criteria for identifying contexts where privacy laws might draw benefits from the adoption of a

** Associate Professor, University of North Carolina School of Law.

*** Associate Professor and Director of the Law Library, University of North Carolina School of Law. The authors gratefully acknowledge the diligent and insightful contributions of their research assistants John Dougherty, Andrew Gregory, Susan Han, Tabitha Messick, and Dean Smith to the findings in this Article.

differential privacy standard similar to that apparently met by Facebook's advertising audience reach database. United States privacy law is a collection of many different sectoral statutes and regulations, torts, and constitutional law, and some areas are more amenable to incorporation of the differential privacy standard than others. This Article highlights some opportunities for recognition of the differential privacy standard as a best practice or a presumption of compliance for privacy, while acknowledging certain limitations on the transferability of the Facebook example.

INTRODUCTION

United States law relies heavily on anonymization techniques, such as redaction of information like names and social security numbers from shared data sets, in order to balance the privacy interests of individuals and utility of data. Regulations under the Health Insurance Portability and Accountability Act of 1996

("HIPAA"), for example, permit health care providers and their business associates to satisfy requirements to deidentify individuals by removing eighteen specific data elements.[1] Other privacy laws, like the Video Privacy Protection Act of 1988[2] and the California Reader Privacy Act,[3] reveal reliance on this idea of protecting anonymity by preventing the disclosure of personally identifying information,[4] even if the law avoids listing which data elements present the most risk to individual's privacy. In addition, many businesses' privacy promises to their customers, enforceable through consumer protection statutes, tie privacy to anonymization.[5]

In a recent article, however, Paul Ohm writes that anonymization[6] techniques, such as the redaction of personally-identifying information, have become ineffective as an approach to reconciling the utilization of data with privacy concerns.[7] This is bad news for the privacy law community. Summarizing fifteen years of

1. *See, e.g.*, 45 C.F.R § 164.514(b)(2)(i) (2010). The Privacy Rule Safe Harbor option for deidentifying individuals in health data requires removal of eighteen types of identifiers and no actual knowledge that the remaining data could be used to identify individuals. An alternative provided in the rule is a statistical methodology shown by experts to provide a low risk of identification of individuals. § 164.514(b)(2)(ii).

2. Video Privacy Protection Act of 1988, Pub. L. No. 100-618, 102 Stat. 3195 (1988) (codified at 18 U.S.C. § 2710 (2006)).

3. Reader Privacy Act of 2012, 2011 Cal. Legis. Serv. 424 (West) (codified at CAL. CIV. CODE § 1798.90.05 (West 2012)).

4. 18 U.S.C. § 2710(a)(3), (b)(2) (establishing liability for video tape service providers who knowingly disclose information which identifies a person as having requested or obtained specific video materials or services); CAL. CIV. CODE § 1798.90(b)(5), (c) (prohibiting book service providers from disclosing, except in certain circumstances, any "information that relates to, or is capable of being associated with, a particular user's access to or use of a book service or a book, in whole or in partial form"). North Carolina, like many other states, protects privacy of library use by prohibiting the disclosure of "any library record that identifies a person," N.C. GEN. STAT. § 125-19(a) (2011), yet allows the sharing of "nonidentifying material that may be retained for the purpose of studying or evaluating the circulation of library materials in general." § 125-18(2) (2011).

5. *See, e.g.*, *Data Storage and Anonymization*, YAHOO!, http://info.yahoo.com /privacy/us/yahoo/datastorage/ (last visited May 4, 2012) (promising to "de-identify search user log data within 18 months of collection, with limited exceptions," and defining anonymization/deidentification as "a process of removing or replacing personal identifiers in data records so that the resulting data is no longer personally identifiable").

6. Ohm defines anonymization as "a process by which information in a database is manipulated to make it difficult to identify data subjects." Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1707 (2010). Ultimately, Ohm questions the usefulness of the term because of the threat of reidentification. *Id.* at 1742. The term anonymity is used in this Article to describe the broader condition of secrecy of the identity of an individual or data subject.

7. *Id.* at 1703–44 (discussing the history of reidentification techniques and how their advancement has overcome the aims of most privacy regulation).

computer science research, Ohm concludes that "researchers have learned more than enough already for us to reject anonymization as a privacy-providing panacea."[8] In other words, even the most thorough redaction of personally identifiable information has generally been found insufficient to protect the privacy of individuals represented in data sets.[9] Ohm describes how standard relational database tools facilitate the linking of anonymized data with outside information through common data elements to reconstruct personally identifying profiles.[10] Due to increased access to public, commercial, and other information, reidentification is no longer difficult or expensive and is capable of undermining traditional anonymization approaches in startling ways.[11]

Ohm employs the term "database of ruin" to describe the collection of private facts about a person maintained in one computer database or another that could cause that person legally cognizable harm if more widely known.[12] Reidentification threatens everyone in the modern world with the possible construction of his own personal database of ruin.

As a hypothetical example of a problematic reidentification, suppose that Jane Public from zip code 27514 notices her neighbor John Doe's name and age (36) on the finisher's list for the 2011 Asheville AIDS Walk and 5K Run. She is curious about whether John is HIV positive. Jane visits the targeted advertising area of Facebook's Web site where she can obtain "audience reach" statistics regarding the number of Facebook users whose profiles match a specified combination of characteristics.[13] Jane finds that there is

8. *Id*. at 1716.

9. *Id*. at 1716–31. Much of the debate about anonymization concerns which data elements are either immediately identifying and which are most likely to facilitate reidentification. *Compare, e.g.*, Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814 *passim* (2011) (arguing for refinement, not rejection of anonymization approaches based on removal of personally identifiable information), *with* Ohm, *supra* note 6, at 1742 (rejecting reliance on removal of personally identifiable information as an ever-expanding "carnival whack-a-mole game"). The terminology of "personally identifiable information" and "personal information" is sometimes used to describe both or either category respectively. *See* Schwartz & Solove, *supra*, at 1826–28 (reviewing and explaining the history of "personally identifiable information" as a model for privacy compliance in United States law).

10. *See* Ohm, *supra* note 6, at 1717–31.

11. *Id.* at 1730–31.

12. *See id.* at 1748.

13. Facebook allows advertisers to estimate the size of the target populations for their campaigns through a public interface that answers count queries regarding combinations of various elements of user profiles, such as age, gender, geographic location, activities,

exactly one male Facebook user aged 36 from zip code 27514 who lists the "2011 Asheville AIDS Walk and 5K Run" as an interest. Even though John has not made public his list of interests on his Facebook profile, he is included in this count.[14] Jane then places an ad targeted to Facebook users having this combination of characteristics offering free information to HIV-positive patients about a new antiretroviral treatment. If Jane is charged by Facebook for having her ad clicked, she may infer with some confidence (though not certainty) that John is HIV positive.

Such a scenario would be representative of what Ohm referred to in his title as a "broken promise[] of privacy." Facebook's data use policy assures users:

> We do not share any of your information with advertisers (unless, of course, you give us permission).
>
> When an advertiser creates an ad on Facebook, they are given the opportunity to choose their audience by location, demographics, likes, keywords, and any other information we receive or can tell about you and other users. . . . [W]e serve the ad to people who meet the criteria the advertiser selected, but we do not tell the advertiser who any of those people are. So, for example, if a person clicks on the ad, the advertiser might infer that the person is an 18-to-35-year-old woman who lives in the US and likes basketball. But we would not tell the advertiser who that person is.[15]

As this hypothetical illustrates, Facebook's assurance that "we do not tell the advertiser who any of those people [within the targeted population] are"[16] does not necessarily preclude a party from combining Facebook's statistical data with outside information to infer their identities. The harm to John might be limited to the disclosure of private health information to a nosy neighbor, but John must hope Jane's intent is not malicious. One can easily imagine other examples of reidentification that would expose information triggering harassment, discrimination, or identify theft.

---

interests, education level, and workplace. *Creating an Ad or Sponsored Story*, FACEBOOK, http://www.facebook.com/help/?page=175624025825871 (last visited May 4, 2012) (click the arrow next to the question, "What are my targeting options for Facebook Ads or Sponsored Stories?").

14. *See infra* text accompanying note 15.

15. *Data Use Policy*, FACEBOOK, http://www.facebook.com/full_data_use_policy (last updated Sept. 23, 2011).

16. *Id.*

It bears noting, as Cynthia Dwork, a principal researcher at Microsoft Corporation,[17] has pointed out, that Facebook does not reveal to users the criteria specified for each ad that serves as the basis for targeting them.[18] As our hypothetical illustrates, if those criteria happen to be much more specific than a taste for basketball and the advertiser happens to possess sufficient outside information about the people who meet these criteria, Facebook's ad targeting statistics could provide a basis for inferring the identity of a user who clicks on the ad.[19] Facebook's database could even reveal sensitive personally identifiable information to a party with outside information who had no intention of advertising.[20]

Given these privacy concerns, it is fortunate that Facebook's ad targeting database is programmed never to reveal that exactly one Facebook user possesses a given combination of characteristics. In fact, Facebook's database appears to restrict the disclosure of statistical information even more carefully than is suggested by its data use policy. Based on the empirical observations of Facebook's ad targeting database described in Part III.A, Facebook has apparently

17. Cynthia Dwork has been described as "the world's foremost expert on placing privacy-preserving data analysis on a mathematically rigorous foundation" with the cornerstone of that work being differential privacy. *ICDM 2011 Invited Speakers*, INT'L CONFERENCE ON DATA MINING 2011, http://icdm2011.cs.ualberta.ca/invited-speakers.php (last visited May 4, 2012). Dwork is a Distinguished Scientist at Microsoft, winner of the Edsger W. Dijkstra Prize in Distributed Computing, a member of the U.S. National Academy of Engineering, and a Fellow of the American Academy of Arts and Sciences. Her work includes private data analysis, cryptography, combating of spam, complexity theory, web search, voting theory, distributed computing, interconnection networks, and algorithm design and analysis. *See* Cynthia Dwork, Curriculum Vitae 1, *available at* http://research.microsoft.com/en-us/people/dwork/cv.pdf.

18. *See* Cynthia Dwork, *I'm in the Database, but Nobody Knows*, BERKMAN CENTER FOR INTERNET & SOC'Y LUNCHEON SERIES 32:40 (Sept. 28, 2010), http://cyber.law.harvard.edu/events/luncheon/2010/09/cdwork (discussing privacy attacks using ad-targeting criteria to exploit outside knowledge and indentify individuals).

19. To give another artificial but illustrative example, Cynthia Dwork states that she can be uniquely identified as (1) a Microsoft employee who is (2) a female (3) distinguished scientist with (4) very curly hair. *See* Dwork, *supra* note 18, at 6:00 (discussing two "large set" queries that "differ only in me"). Any advertiser with outside knowledge of this unique combination of characteristics of Dwork, if given the further information that exactly one Facebook user possessed this combination of characteristics, could target the combination and thereby identify Dwork as the person clicking on the ad.

20. As Dwork notes, anyone with outside knowledge that Dwork is the only person with characteristics (1)–(4), *see supra* note 19, could infer from the number of people having all of those characteristics and (5) "possesses the sickle cell trait" whether Dwork possesses the sickle cell trait. *See* Dwork, *supra* note 18, at 6:50. Facebook's database interface allows anyone to submit targeted advertising reach queries, and Facebook encourages users with privacy concerns to "[t]ry this tool yourself" even if they do not intend to place an ad. *See Data Use Policy*, *supra* note 15.

implemented a privacy mechanism that can be shown to achieve a relative notion of privacy, known as differential privacy, thereby ensuring that Facebook's database does not contribute significantly to the construction of any user's database of ruin. If this reverse-engineering analysis of Facebook's privacy technology is correct, the assurances in Facebook's data use policy need not represent "broken promises of privacy," but provable guarantees of differential privacy that also deliver sufficient utility of that shared data. As a consequence, Facebook should be able to claim compliance with consumer protection laws enforcing Facebook's policy promises relating to this sharing of data with advertisers.[21] In addition, this example can serve as a proof of the concept that other laws balancing privacy and utility of data could look to differential privacy as an achievable standard to replace or supplement fragile anonymization approaches.

The feasibility of achieving differential privacy while providing useful statistical database information would appear to be of particular interest to Facebook. Facebook's revenue model depends heavily on the ability of advertisers to target users with specified characteristics.[22] At the same time, Facebook's privacy policies and practices have faced continual scrutiny.[23] As one industry commentator wrote in 2010, "The company's future depends on finding just the right balance between the privacy expectations of its users and the quality of the social marketing data it can serve to its business partners."[24]

If Facebook has, in fact, successfully implemented mechanisms in its advertising reach database that achieve differential privacy for its users, then this Facebook case study may be instructive in other

21. Facebook's promises not to identify individuals in its advertiser audience reach system are enforceable by the Federal Trade Commission ("FTC"), which has the authority to investigate and issue a cease-and-desist order against a business for "unfair or deceptive acts or practices." 15 U.S.C. § 45 (2006). Notably, Facebook's audience-reach sharing is not the only type of sharing of member data that Facebook facilitates for its advertisers, and some of those other types of sharing have recently been the subject of a FTC investigation and settlement with Facebook. *See* Facebook, Inc.; Analysis of Proposed Consent Order to Aid Public Comment, 76 Fed. Reg. 75,883, 75,884 (Dec. 5, 2011) ("Facebook promised users that it would not share their personal information with advertisers; in fact, Facebook did share this information with advertisers when a user clicked on a Facebook ad.").

22. *See* Facebook, Inc., Registration Statement (Form S-1) 40, 55 (Feb. 1, 2012), *available at* http://sec.gov/Archives/edgar/data/1326801/000119312512034517/d287954ds1.htm (stating 2011 revenues of $3.711 billion, 83% of which came from advertising).

23. *See infra* text accompanying notes 125–26.

24. Mark Sullivan, *How Will Facebook Make Money?*, PCWORLD (June 15, 2010, 1:00 AM), http://web.archive.org/web/20100616061911/pcworld.com/article/198815.

privacy contexts where the reidentification threat may arise. Widely publicized events have demonstrated the risk of reidentification when statistical data has been released under the assumption that the redaction of personally identifiable information adequately protects individual privacy. Netflix attempted to deidentify individuals when it released subscriber data in a contest for developers to create new programs to help Netflix improve its movie recommendation service.[25] Although Netflix removed key elements considered personally identifying information, the surviving shared information was compared with publicly accessible data to reveal several individuals' identities, along with their movie rental habits.[26] This exposure led to a lawsuit against Netflix for violation of the federal Video Privacy Protection Act and several California laws[27] and caused the Federal Trade Commission ("FTC") to raise objections to a second such contest.[28] In reaching a settlement in the lawsuit and in resolving the FTC inquiries, Netflix announced it had agreed to

25. *See* Steve Lohr, *Netflix Cancels Contest After Concerns Are Raised About Privacy*, N.Y. TIMES, Mar. 13, 2010, at B3 (reporting that "supposedly anonymized data" from Netflix customer records was used by researchers to reidentify individuals).

26. Arvind Narayanan & Vitaly Shmatikov, *Robust De-anonymization of Large Sparse Datasets*, 29 INST. ELECTRICAL & ELECTRONICS ENGINEERS SYMP. ON SECURITY & PRIVACY 111, 111–25 (2008), *available at* http://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf (describing how researchers were able to identify some individuals in the released Netflix data by comparing names and dates attached to movie reviews posted to the publicly accessible Internet Movie Database, imdb.com).

27. The complaint pointed out that the shared information violated not only privacy promises made by Netflix to its customers, but also violated the federal Video Privacy Protection Act of 1988 and several California consumer laws, constituted unjust enrichment, and implicated the common law privacy tort of public disclosure of private facts. *See* Jury Demand, Class Action Complaint at 1, Valdez-Marquez v. Netflix, Inc., No. 5:09-cv-05903 (N.D. Cal. Dec. 17, 2009), *available at* http://www.wired.com/images_blogs/threatlevel/2009/12/doe-v-netflix.pdf (alleging violation of: "(1) Video Privacy Protection Act, 18 U.S.C. § 2710, 2) Video Privacy Protection Act, 18 U.S.C. § 2710, 3) California Consumers Legal Remedies Act, Civil Code § 1750, 4) California Customer Records Act, Civil Code § 1798.80, 5) California Unfair Competition Law, Business and Professions Code § 17200, 6) California False Advertising Law, Business and Professions Code § 17500, 7) Unjust Enrichment, 8) Public Disclosure of Private Facts"). The Video Privacy Protection Act of 1988 created liability for videotape service providers that knowingly disclose "personally identifiable information concerning any consumer" with certain exceptions that did not cover the Netflix context. *See* 18 U.S.C. § 2710(b) (2006).

28. The FTC contacted Netflix and warned that the "risk of re-identification and the extent to which Netflix's previous representations to its customers about disclosure of their information would raise concerns under Section 5 of the FTC Act." *See* Letter from Maneesha Mithal, Assoc. Dir., Div. of Privacy & Identity Prot., Fed. Trade Comm'n, to Reed Freeman, Morrison & Foerster LLP, Counsel for Netflix, Inc. 2 (Mar. 12, 2010), *available at* http://www.ftc.gov/os/closings/100312netflixletter.pdf. Section 5 of the FTC Act grants the FTC authority to investigate and bring enforcement actions for unfair and deceptive trade practices. FTC Act § 5, 15 U.S.C. § 45 (2006).

"certain parameters" for how the company would use Netflix data and would cancel the second contest.[29]

Another high-profile reidentification occurred when Latanya Sweeney, then a graduate student at MIT, merged presumably anonymized Massachusetts state worker hospital records with voter registration records and was able to identify rather quickly the health records of then-Governor William Weld.[30] Sweeney later published a broader study finding that 87% of the 1990 U.S. Census population could be indentified using only gender, zip code, and full date of birth,[31] and others reproduced this work in the 2000 Census with 63% success in identifying individuals.[32]

The vulnerability of anonymization could undermine established compromises between privacy and competing interests reflected in several areas of the law.[33] Although computer scientists warn that removal of personally identifiable information is now a privacy fallacy, key areas of the law, notably health privacy, incorporate suppression of identifying data elements as privacy compliance.[34] In addition, many websites and companies promise to protect the privacy of customer data by anonymizing it before it is shared.[35]

29. Neil Hunt, *Netflix Prize Update*, NETFLIX U.S. & CANADA BLOG (Mar. 12, 2010), http://blog.netflix.com/2010/03/this-is-neil-hunt-chief-product-officer.html (reporting settlement of the lawsuit and FTC investigation, but not revealing the amount of any financial payment or acknowledging violation of any law).

30. *Recommendations To Identify and Combat Privacy Problems in the Commonwealth: Hearing on H.R. 351 Before the H. Select Comm. on Info. Sec.*, 2005 Gen. Assemb., 189th Sess. (Pa. 2005) (statement of Latanya Sweeney, Associate Professor, Carnegie Mellon University), *available at* http://dataprivacylab.org/dataprivacy/talks/Flick-05-10.html.

31. *See* Latanya Sweeney, *Simple Demographics Often Identify People Uniquely* 2, 3 (Carnegie Mellon Univ., Working Paper No. 3, 2000), *available at* http://dataprivacylab.org/projects/identifiability/paper1.pdf (finding that 216 million of 248 million persons represented in the 1990 Census were identifiable with only these three characteristics).

32. Philippe Golle, *Revisiting the Uniqueness of Simple Demographics in the US Population*, 5 ASS'N FOR COMPUTING MACHINERY WORKSHOP ON PRIVACY ELECTRONIC SOC'Y 77, 78 (2006) (testing the findings of Latanya Sweeney for the 1990 Census and extending those tests generally to the 2000 Census with identification rates of 61% and 63% respectively).

33. *See* Schwartz & Solove, *supra* note 9, at 1826–28.

34. *See* HITECH Act of 2009, Pub. L. No. 11-5, § 13424, 123 Stat. 226, 276–79 (codified at 42 U.S.C.A § 17953 (Supp IV 2010) (recognizing privacy compliance when eighteen data points, considered personally identifying, are removed from health data that is to be shared as required by 45 C.F.R. § 164.514(b)(2)(i) (2010)). Similarly, the Video Privacy Protection Act of 1988 requires suppression of personally identifiable information except under limited circumstances. *See* 18 U.S.C. § 2710(b) (2006).

35. For example, Amazon.com's privacy policy says the company avoids "selling, renting, sharing, or otherwise disclosing personally identifiable information from customers for commercial purposes." *See Amazon.com Privacy Notice*, AMAZON.COM,

A number of legal observers have agreed with Ohm that reidentification threatens the viability of common practices that attempt to reconcile data utility with individual privacy.[36] The FTC has reported that the threat of reidentification is at the heart of its demand for new types of "privacy by design."[37] Debate continues, though, over how to address this threat, both as a functional matter and as a matter of law. Ohm recommends several approaches including releasing deidentified data only to trusted researchers, with contractual or regulatory restrictions on uses beyond those deemed beneficial and presumably privacy respecting.[38] Paul Schwartz and Daniel Solove concede the difficulty of perfecting and enforcing these approaches, but protest big shifts that would restructure both the law and the habits of those handling sensitive data.[39] Schwartz and Solove recommend retention but refinement of the concept of "personally identifying information," with some categorical standards and

http://www.amazon.com/gp/help/customer/display.html/ref=footer_privacy/177-4355798-3623704?ie=UTF8&nodeId=468496 (last updated Apr. 6, 2012); *see also Data Storage and Anonymization*, *supra* note 5 ("Yahoo! takes additional steps so that data collected and used to customize interest based advising . . . are not associated with personally identifiable information.").

36. *See, e.g.*, Kenneth A. Bamberger, *Technologies of Compliance: Risk and Regulation in a Digital Age*, 88 TEX. L. REV. 669, 716 (2010) (criticizing institutions' inertia in addressing new privacy threats to established systems of anonymization); Robert Gellman, *The Deidentification Dilemma: A Legislative and Contractual Proposal*, 21 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 33, 35 (2010) ("[T]he value of data for legitimate uses, such as research, may be significantly reduced when the data is processed without identifiers which were removed to protect privacy."). *But see* Jane Yakowitz, *Tragedy of the Data Commons*, 25 HARV. J.L. & TECH. 1, 8–10, 36–42, 48–50 (2011) (arguing that reidentification rarely occurs, that the social value of access to accurate data outweighs the threat of reidentification, and that criminalization of reidentification is the proper solution).

37. Edith Ramirez, Comm'r, Fed. Trade Comm'n, Keynote Address at the 28th Annual Institute on Telecommunications Policy & Regulation 2–3 (Dec. 9, 2010), *available at* http://www.ftc.gov/speeches/ramirez/101209fcbaspeech.pdf (explaining that reidentification threats are the main reason for the FTC's call for stronger privacy protections, including "privacy by design," though not specifically mentioning differential privacy); *see also* FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: A PROPOSED FRAMEWORK FOR BUSINESSES AND POLICYMAKERS 43 (Dec. 2010), *available at* http://www.ftc.gov/os/2010/12/101201privacyreport.pdf (noting that the FTC's proposals are "supported by a wide cross section of roundtable participants who stated that the traditional distinction between PII and non-PII continues to lose significance due to changes in technology and the ability to re-identify consumers from supposedly anonymous data"). "Privacy by design" is a concept promoted by Ann Cavoukian, Information and Privacy Commissioner, Ontario, Canada that calls for structural support for privacy protection. *See Privacy by Design*, PBD, http://privacybydesign.ca (last visited May 4, 2012).

38. *See* Ohm, *supra* note 6, at 1764–69.

39. *See* Schwartz & Solove, *supra* note 9, at 1883–86.

practices to match levels of risk for reidentification.[40] Robert Gellman proposes a federal statute that data disclosers and data recipients could invoke through contract to gain safe harbor protection by conformity with the statute's requirements for anonymization and prevention of reidentification.[41]

Few legal scholars or lawmakers have proposed differential privacy as a response to the threat of reidentification even in limited circumstances,[42] but the Facebook advertiser interactive reporting system suggests that differential privacy may have substantial promise for addressing a number of privacy threats from reidentification. A guarantee of differential privacy assures that presence or absence of any one individual in the database makes no significant difference in the likelihood of each possible response to a database query.[43] Differential privacy guarantees, therefore, provide meaningful protection from even the possibility of linkage to auxiliary data sets,

40. *See id.* at 1886–93.

41. *See* Gellman, *supra* note 36, at 47–61 (outlining a statutory proposal to prevent reidentification while allowing researchers useful access to data with "overt identifiers" that alone or in combination with other information could be used to identify a particular individual).

42. Paul Ohm has several criticisms of differential privacy mechanisms: they are less flexible than traditional anonymization, too expensive because of the need to have constant participation of a data administrator, and burdensome on utility of the data because of the introduction of randomization producing noise or inaccuracies in the responses to queries. Ohm, *supra* note 6, at 1756–57. Ohm also challenges the effectiveness of the privacy protection in certain circumstances, *id.*, and the applicability of these techniques in all situations. *Id.* at 1751. Thus, he argues instead for expanded prohibitions against and remedies for reidentification and facilitators of reidentification. *Id.* at 1759–61. In an overview of the problems of reidentification and potential solutions, Arvind Narayanan and Vitaly Shmatikov endorse differential privacy as a "major step in the right direction," but concede that it is not adaptable to all situations and must be "built and reasoned about on a case-by-case basis." Arvind Narayanan & Vitaly Shmatikov, *Privacy and Security Myths and Fallacies of "Personally Identifiable Information,"* COMM. ASS'N FOR COMPUTING MACHINERY, June 2010, at 24, 26, *available at* http://www.cs.utexas.edu/~shmat/shmat_cacm10.pdf. Jane Yakowitz critiques "noise-adding" techniques as overburdening utility. Yakowitz, *supra* note 36, at 46–47 (2011). Yakowitz, nonetheless, concludes her argument that reidentification risk has been overstated by alluding to the promise of systems that are probably differential privacy technologies, but she suggests implementation may take time. *Id.* at 66–67 (citing *American FactFinder*, U.S. CENSUS BUREAU, http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml) (describing "[s]tatistical software that allows the dataset to remain on a secure server while researchers submit statistical queries").

43. Cynthia Dwork, *A Firm Foundation for Private Data Analysis*, COMM. ASS'N FOR COMPUTING MACHINERY, Jan. 2011, at 86, 91 (defining differential privacy). Dwork is a major proponent and developer of differential privacy and tools in support of differential privacy. Dwork, with others, holds several patents for differential privacy related programs. *See Cynthia Dwork: Patents*, MICROSOFT RESEARCH, http://research.microsoft.com/en-us/people/dwork/patents.aspx (last visited May 4, 2012).

including ones that could be developed at some future time.[44] As Ohm acknowledges, a mechanism that achieves differential privacy "ensur[es] mathematically that even the most sophisticated reidentifier will not be able to use the answer to unearth information about the people in the database."[45] Prospects for practical implementation of differentially private database systems are continually improving, as an active community of computer science researchers has been refining mechanism designs[46] and releasing software development tools[47] in recent months.

With these prospects in mind, we believe our Facebook case study, to the extent that it may reveal the most commercially successful practical implementation of a differentially private database system, may helpfully inform data managers and policymakers in responding to the reidentification threat. The remainder of this Article is organized as follows. In Part I, we formalize the notion of differential privacy and descriptions of database mechanisms that achieve differential privacy. Part II presents our reverse-engineering analysis of Facebook's advertising reach database, concluding that the database's observed behavior is consistent with the hypothesis that Facebook has implemented differentially private mechanisms to protect individual user data. Part II also assesses the effectiveness and practicality of these mechanisms. Part III situates the Facebook case study among other contexts where there is a tension between privacy and utility and derives from the case study several criteria relevant to identifying those contexts where

44. Differential privacy avoids the problem of predicting which data elements are likely to be used with external data sets, especially future data sets, to achieve linkage attacks that would reidentify an individual. The debate over which data is "personally identifiable" is largely avoided.

45. *See* Ohm, *supra* note 6, at 1756.

46. *See, e.g.*, Xiaokui Xiao, Guozhand Wang & Johannes Gehrke, *Differential Privacy Via Wavelet Transforms*, 23 INST. ELECTRICAL & ELECTRONICS ENGINEERS TRANSACTIONS ON KNOWLEDGE & DATA ENGINEERING 1200, 1200–01 (2011) (describing a mechanism that achieves differential privacy while releasing accurate results for range-count queries).

47. *See* Frank McSherry, *Privacy Integrated Queries: An Extensible Platform for Privacy-Preserving Data Analysis*, COMM. ASS'N FOR COMPUTING MACHINERY, Sept. 2010, at 89, 89 (describing a new data-handling software development platform he has dubbed PINQ, for Privacy Integrated Queries, that he says makes deploying differential privacy easier for end users); Frank McSherry & Ratul Mahajan, *Differentially-Private Network Trace Analysis*, 2010 ASS'N FOR COMPUTING MACHINERY SPECIAL INT. GROUP ON DATA COMM. 123, 123 (testing systems using differential privacy and concluding that the error rates caused by the technique were low and that the technique held great promise).

the notion of differential privacy can usefully play a role in standards of privacy compliance.

I. ACHIEVING DIFFERENTIAL PRIVACY IN PRINCIPLE

At its core, the reidentification threat arises from the possibility that an attacker may have access to outside information that, when combined with information released by a database, allows the attacker to infer private information about an individual. Given the likelihood that a sophisticated attacker will have access to advanced computational tools and a vast supply of personal data, it seems prudent for both data managers and policymakers to operate from pessimistic assumptions about the performance of privacy technologies.

Such pessimism, however, forms only part of the landscape in the computer science research community. Confronted with the impossibility of providing *absolute* privacy against a powerful, well-informed adversary, researchers have turned their attention to developing database technologies that can at least guarantee a *relative* form of privacy. The idea is that no one can hide information that has already been made accessible to an attacker, but data managers can exercise care so that the release of information about a group does not further compromise any individual's private information.

Computer scientists, led by Dwork, have formalized this notion of relative privacy as a mathematical criterion known as differential privacy, which is defined as follows:

> A randomized function K gives ε-differential privacy if, for all data sets D_1 and D_2 differing on at most one element, and all $S \subseteq \text{Range}(K)$,
>
> $$\Pr[K(D_1) \in S] \leq \exp(\varepsilon) \times \Pr[K(D_2) \in S].$$[48]

In reading this definition, it is helpful to think of a database server containing private information about each individual in a database D. To protect this private information, the server is programmed not to respond to queries with the actual answer, but with a randomized response $K(D)$ (a function defined over the set of all possible databases) that is generally close enough to the truth to be useful. If K also satisfies the condition that it gives ε-differential privacy, then the server's response to any given database query (e.g.,

48. *See* Cynthia Dwork, *Differential Privacy: A Survey of Results*, *in* THEORY AND APPLICATIONS OF MODELS OF COMPUTATION 1, 3 (Manindra Agrawal et al. eds., 2008).

the query "Is $K(D) \in S$") is unlikely to be changed by the inclusion of any given individual in the database. The parameter $\varepsilon > 0$ serves to calibrate the stringency of the condition: if ε is very close to 0, then $\exp(\varepsilon)$ is very close to 1, giving a tight bound on multiplicative changes in the probability distribution of K resulting from any single-element change in D. As Dwork explains:

> Any mechanism [K] satisfying this definition addresses all concerns that any participant might have about the leakage of her personal information . . . . Even if the participant removed her data from the dataset, no outputs (and thus consequences of outputs) would become significantly more or less likely. For example, if the database were to be consulted by an insurance provider before deciding whether or not to insure a given individual, then the presence or absence of *any* individual's data in the database will not significantly affect her chance of receiving coverage.[49]

More tersely, differential privacy allows each user to be assured that "I'm in the database, but nobody knows," to borrow the title of a talk given by Dwork at Harvard Law School.[50]

Crucially, Dwork and her colleagues at Microsoft[51] have also demonstrated that it is possible to design a mechanism that not only guarantees ε-differential privacy for some $\varepsilon > 0$, but also provides

49. Dwork, *supra* note 43, at 91.

50. Dwork, *supra* note 18.

51. Claim 1 of U.S. Patent No. 7,562,071, assigned to Microsoft Corporation, appears to cover the use of Laplace noise addition to the output of a statistical database for the purpose of supporting a privacy guarantee. Claim 1 reads:

> 1. A method for producing a noisy output that reduces predictability of data inputs while increasing accuracy of said noisy output, the method comprising:
>
> formulating a query against data associated with a plurality of privacy principals and stored in a database;
>
> evaluating the query to determine a query diameter measurement;
>
> performing the query on one or more data inputs;
>
> calculating by a computer processor an output of the query;
>
> calculating by a computer processor a substantially exponential distribution of noise values as a function of the query diameter measurement and a privacy parameter, the substantially exponential distribution being calculated according to a Laplacian distribution;
>
> combining by a computer processor said output with a noise value selected from the calculated substantially exponential distribution of noise values to produce the noisy output; and
>
> disclosing the noisy output.

U.S. Patent No. 7,562,071 (filed Dec. 2, 2005).

usable information about the database's contents. One such mechanism K disguises the true value of a database query $f(D)$ by the addition of random noise taken from the Laplace distribution Lap(b) defined by the probability density function

$$p(x;b) = \frac{1}{2b} e^{-|x|/b}, -\infty < x < \infty \quad [52]$$

with mass at x, where the scale parameter b is defined by

$$b = \Delta f / \varepsilon = \max_{D_1, D_2} \| f(D_1) - f(D_2) \| / \varepsilon,$$

and where $D_1 D_2$ range over all possible databases differing on at most one element.[53] For any $r \in$ Range(K), the ratio

$$\begin{aligned} \frac{\Pr[K(D_1) = r]}{\Pr[K(D_2) = r]} &= \frac{\exp\left(-\left| f(D_1) - r \right| / \left(\Delta f / \varepsilon\right)\right)}{\exp\left(-\left| f(D_2) - r \right| / \left(\Delta f / \varepsilon\right)\right)} \\ &\geq \exp\left(- \frac{\left| f(D_1) - f(D_2) \right|}{\Delta f / \varepsilon} \right) \\ &\geq \exp(-\varepsilon), \end{aligned}$$

from which it follows that K provides ε-differential privacy.[54] This so-called "Laplace noise addition" mechanism K produces a reasonable proxy for the true value of a database query when it adds relatively little noise from Lap(b). The variance of Lap(b) is $2b^2$, which increases as $b = \Delta f / \varepsilon$ increases, so b needs to be kept relatively small. Thus, Laplace noise addition is especially suitable for *count databases*; i.e., those in which the queries $f(D)$ are all assumed to be "count" queries of the form "How many rows have property P?," since for these queries $\Delta f = 1$ and $b = 1 / \varepsilon$.[55]

Mechanisms that provide ε-differential privacy are so secure that there is no need for the database manager to obscure the mechanism's design. Thus, the fact that K's output incorporates random noise from Lap $(\Delta f / \varepsilon)$ can be made public without

52. *See* Dwork, *supra* note 48, at 4. *See generally* SAMUEL KOTZ ET AL., THE LAPLACE DISTRIBUTION AND GENERALIZATIONS (2001) (discussing the Laplace distribution in more depth).

53. *See* Dwork, *supra* note 48, at 4; *see also* Cynthia Dwork et al., *Calibrating Noise to Sensitivity in Private Data Analysis*, *in* THEORY OF CRYPTOGRAPHY 265, 270 (Shai Halevi & Tal Rabin eds., 2006) (presenting the original result).

54. *See* Dwork, *supra* note 48, at 4.

55. *See id.*

compromising the privacy of the database in any way.[56] As it would be immediately apparent that any fractional portion of a query result represented random noise, however, the manager of a count database may choose to restrict the range of query results to the set of whole numbers. For this purpose, K may be slightly modified by using a discrete version of the Laplace distribution, DL(p), in place of the continuous distribution Lap(b), where for $p \in (0,1)$, DL(p) has the probability distribution function

$$\Pr(Y = k) = \frac{1-p}{1+p} p^{|k|}, k \in Z$$ [57]

It is straightforward to verify that for $p = 1/b = \exp(-\varepsilon / \Delta f)$, a mechanism that adds noise from DL(p) provides ε-differential privacy. Again, for count databases, $\Delta f = 1$, and thus $p = \exp(-\varepsilon)$. The next Part will test the hypothesis, *inter alia*, that Facebook is using such a mechanism in connection with its ad targeting database.

II. ACHIEVING DIFFERENTIAL PRIVACY IN PRACTICE: FACEBOOK'S ADVERTISING REACH DATABASE

A. Reverse-Engineering Facebook's Privacy Technology

Facebook's ad targeting database is readily available for experimentation.[58] After designing an ad, a would-be advertiser (or someone posing as an advertiser) completes a form specifying the characteristics of Facebook's users to be targeted by the ad.[59] As the advertiser enters criteria into the form, Facebook's web interface automatically updates an "Estimated Reach" statistic corresponding to the number of users matching all of the selected criteria, as illustrated by the examples in *Figure*.[60]

56. *See id.* at 3 (noting the assumption that the parameter ε is public).

57. *See* Seidu Inusah & Tomasz J. Kozubowski, *A Discrete Analogue of the Laplace Distribution*, 136 J. STAT. PLAN. & INFERENCE 1090, 1092 (2006).

58. *Creating an Ad or Sponsored Story*, *supra* note 13.

59. *See Advertise on Facebook*, FACEBOOK, http://www.facebook.com/ads/create/ (last visited May 4, 2012).

60. *See id.*

Figure 1: Examples of Facebook's Reports of Advertising Reach Data

A casual observation of these "Estimated Reach" statistics reveals that the target audience is always reported as a multiple of twenty people, with estimates of less than forty reported as "fewer than 20 people."[61] The obvious conclusion is that Facebook has implemented a rounding mechanism that obscures the true value of the reach statistic in question.

Less obviously, Facebook's respective estimates for a targeted group and for the same group broken into disjoint subgroups often reveal small but substantial deviations from additivity.[62] For example, as of October 2011, among North Carolina Facebook users interested in Alzheimer's disease, Facebook reports 340 users aged fifty, 320 users aged fifty-one, and 620 users aged between fifty and fifty-one inclusive.[63] The discrepancy $(340 + 320) - 620 = 40$ is too large to be attributed to rounding error.[64] Moreover, these deviations appear to be persistent over the short term; repeating the same queries one hour later produces the same results.

These observations give rise to a two-part conjecture: first, that in addition to rounding, Facebook has implemented a mechanism that further obscures the true values of its reach statistics; and second, that if this mechanism is not deterministic, Facebook maintains a cache to ensure consistent responses to the same query over the short term.

To test this conjecture, we used a Perl script (developed by our research assistant Andrew Gregory) to submit queries automatically

61. *See id.*

62. Additivity is the mathematical characteristic describing any function f such that $f(x + y) = f(x) + f(y)$. ZALMAN USISKIN ET AL., MATHEMATICS FOR HIGH SCHOOL TEACHERS: AN ADVANCED PERSPECTIVE 14 (2002).

63. *See Advertise on Facebook*, *supra* note 59 (data obtained Oct. 24, 2011).

64. Assuming that Facebook's mechanism rounds values to the nearest multiple of twenty, the lowest possible true sizes of the fifty- and fifty-one-year-old groups are 330 and 310, respectively, which would give a population of at least 640, not 620, for the fifty- and fifty-one-year-old group. This observation also holds under the alternative assumptions that values are rounded down or up to the nearest multiple of twenty.

to Facebook's advertising reach database. Our queries focused on variations in two types of user characteristics—specified topics of interest and specified age ranges—across the U.S. population of Facebook users.

Given the heightened privacy concerns that may pertain to interests in medical topics,[65] we compiled our provisional list of topics from three sources: a list of diseases taken from the topics list on the Centers for Disease Control and Prevention website,[66] a list of psychiatric disorders from the Diagnostic and Statistical Manual of Mental Disorders,[67] and a list of branded and generic medications compiled by Medicinenet.com.[68] Since many of the items in these lists did not match a unique topic of interest in Facebook's advertising reach database, we used the interface's auto-suggest feature to identify, for each term in the provisional list, the related topic of interest associated with the highest number of Facebook users. Thus, for example, "Fibromyalgia" was converted to "Fibromyalgia Awareness," and "Lupus" was converted to "Lupus Foundation of America." These converted terms comprised a final list of 363 topics of interest.

Facebook's interface allows targeting of ages using intervals whose endpoints are between fourteen and sixty-four years of age.[69] We broke the fifty-year interval between fourteen and sixty-three years inclusive into one-, two-, five-, and ten-year subintervals (a total of ninety subintervals) to be specified in connection with each topic of interest. In all, therefore, we formulated $363 \times 90 = 32{,}670$ distinct queries.

To test the consistency of Facebook's responses, we repeated each query five times, several hours apart. In 633 of the 32,670 cases, the responses we received to the same query did not all agree. In these cases, we reran the query an additional twenty times. In each of these cases, the resulting distribution of responses allowed us to

65. *See generally* Health Insurance Portability and Accountability Act ("HIPAA"), 42 U.S.C. § 1320d(6)(B) (2006) (describing individually identifiable information as information collected from an individual relating to an individual's "past, present, or future physical or mental health or condition . . . , provision of health care . . . , or the past, present, or future payment for the provision of health care[,]" which either identifies or reasonably could identify the individual).

66. *CDC A-Z Index*, CTRS. FOR DISEASE CONTROL & PREVENTION, http://www.cdc.gov/az (last updated Mar. 14, 2011).

67. *Index of Psychiatric Disorders*, ALLPSYCH ONLINE, http://allpsych.com/disorders/disorders_dsmIVcodes.html (last visited May 4, 2012).

68. *Medications A-Z List - A*, MEDICINENET.COM, http://www.medicinenet.com/medications/alpha_a.htm (last visited May 4, 2012).

69. *See Advertise on Facebook*, *supra* note 59.

identify a consensus response for further statistical analysis, in that at least 21 of the 25 responses agreed. These results are summarized in *Table 1*. As we will see, this high level of agreement provides strong statistical evidence of short-term caching.

Table 1: Levels of Agreement Among the Facebook Database's Responses to Repeated Queries

Level of Agreement	Number of Queries
5 of 5	32,038
24 of 25	558
23 of 25	65
22 of 25	8
21 of 25	1

Our statistical analysis of the consensus responses focuses on the discrepancies between the query results for two-year intervals and the sums of the corresponding pairs of query results for one-year intervals covering the same age ranges, and the analogous discrepancies for ten-year versus paired five-year intervals (hereinafter referred to simply as "discrepancies").[70] Our conjecture, stated more formally and specifically as a null hypothesis, is that the observed distribution of the discrepancies reflects Facebook's use of both rounding (modulo 20) and discrete Laplace noise-addition mechanisms.[71] To test this hypothesis, we calculated the expected distribution of the discrepancies as a derived distribution based on the distributions of five discrete random variables: a, b, z_1, z_2, and z_3, where the z_1's are taken at random from DL(p) (where $p = \exp(-\varepsilon)$); and a and b, the remainders (modulo 20) of the true values, x and y, of the reach statistics in question (i.e., $x = 20m + a$ and $y = 20n + b$ for some integers m, n, a, b with $a, b \in [0,19]$) are each taken at random from DU(20), the discrete uniform distribution taking on values 0, 1, . . . , 19. The discrepancy f associated with the hypothesized mechanism's reporting of the reach statistics x, y, and $x+y$ can then be expressed as the function

70. For an example of this calculation, see *supra* text accompanying notes 63–64.

71. After rounding, the effects of Laplace noise addition and discrete Laplace noise addition are indistinguishable, so an equivalent conjecture is that Facebook is simply using Laplace noise addition. The discrete formulation is preferable here because it allows the use of generating functions to simplify the calculation of the derived distribution f.

$$f(x,y,z_1,z_2,z_3) = 20\left(\left\lfloor \frac{a+b+z_3+k}{20} \right\rfloor - \left\lfloor \frac{a+z_1+k}{20} \right\rfloor - \left\lfloor \frac{b+z_2+k}{20} \right\rfloor\right)$$

where $k \in \{1, 10, 11, 20\}$ is a parameter specifying the rounding discipline (i.e., the threshold remainder (modulo 20) at which the rounding mechanism switches from rounding down to rounding up to the next multiple of 20).

The generating functions $A(x)$, $B(x)$, and $Z(x)$ respectively associated with the distributions for a, b, and each z_1 are given by

$$A(x) = B(x) = \frac{1}{20}\sum_{n\in[0,19]} x^n$$

and

$$Z(x) = \frac{1-p}{1+p}\sum_{n\in Z} p^{|n|}x^n.$$

From these functions, we can numerically calculate the generating function $F(x)$ associated with f with arbitrarily high precision for any choice of parameters ε and k. For $\varepsilon = 0.181$ and $k = 11$, we have

$$F(x) \approx 0.000314x^{-80} + 0.00552x^{-60} + 0.0587x^{-40} + 0.235x^{-20} + 0.379 + 0.254x^{20} + 0.0606x^{40} + 0.00587x^{60} + 0.00552x^{80}.$$

In the above analysis, our modeling of a and b as random variables uniformly distributed on [0,19] was based on the simplifying assumption that the empirical probability distributions of x and y are locally approximately uniform over every twenty-person interval. Since it is reasonably likely that these empirical probability distributions resemble a power law distribution,[72] we confined our calculation of discrepancies to those cases where each of the reported reach statistics was at least 1,000 (i.e., situated in the distribution's flat tail). Our dataset of consensus responses by Facebook's database to our 32,670 queries yielded 850 observed discrepancies of this kind. Table 2 compares the distribution of the 850 observed discrepancies with the hypothesized distribution f for the choice of parameters $\varepsilon = 0.181$ and $k = 11$.

72. *Cf.* Norman S. Matloff, *Another Look at the Use of Noise Addition for Database Security*, 1986 INST. ELECTRICAL & ELECTRONICS ENGINEERS SYMP. ON SECURITY & PRIVACY 173, 178 (showing that adding noise from a symmetric distribution to a numerically positive variable with a strictly decreasing density function tends to introduce negative bias). *See generally* CHRIS ANDERSON, THE LONG TAIL: WHY THE FUTURE OF BUSINESS IS SELLING LESS OF MORE (2006) (explaining the observed ubiquity of power law distributions in statistical measurements of cultural popularity).

Table 2: Comparison Between Expected and Observed Frequencies of Discrepancies for $\varepsilon = 0.181$ *and* $k = 11$.

Discrepancy	Observed	Expected
≤ -40	40	55 (6.5%)
–20	208	200 (23.5%)
0	319	322 (37.9%)
20	214	216 (25.4%)
≥ 40	69	57 (6.7%)

The null hypothesis that the observed discrepancies are a random sample taken from f is amenable to testing with the chi-square goodness-of-fit test.[73] Following standard procedures, in Table 2 we have combined categories of expected size less than five in each tail.[74] For these data, we find $\chi^2 = 6.984$ with four degrees of freedom, giving a two-tailed P value of 0.1368.[75] The hypothesized distribution f with the specified parameters is, therefore, a good enough fit for the observed data that we cannot find a statistical basis for rejecting the null hypothesis.[76] A fortiori, we cannot find a statistical basis for rejecting our less specific conjecture that Facebook's database employs both rounding (modulo 20) and discrete Laplace noise-addition mechanisms.[77]

Finally, we observe that independently generated (as opposed to cached) responses from a database employing rounding (modulo 20) with $k = 11$ and DL(–exp(0.181)) noise addition mechanisms tend to vary more widely than we found among Facebook's responses to

73. For an introduction and illustration of the chi-square goodness-of-fit test, see generally RONALD A. FISHER, STATISTICAL METHODS FOR RESEARCH WORKERS (1958).

74. *See, e.g.*, DAVID R. ANDERSON, DENNIS J. SWEENEY & THOMAS A. WILLIAMS, STATISTICS FOR BUSINESS AND ECONOMICS 489 (2011).

75. *See Compare Observed and Expected Frequencies*, GRAPHPAD SOFTWARE, http://www.graphpad.com/quickcalcs/chisquared1.cfm (last visited May 4, 2012) (providing an online tool for performing the chi-square goodness-of-fit test for a discrete distribution with up to twenty categories).

76. *Cf.* ANDERSON ET AL., *supra* note 74, at 487–90 (providing an analogous example of the chi-square goodness-of-fit test and concluding from a calculated P value of 0.1404 that the hypothesized distribution cannot be rejected).

77. *See supra* note 71 and accompanying text.

repeated queries. The probability of any five independently generated responses from such a database being equal is at most 0.409, achieved when the true value of the reach statistic is a multiple of 20. Performing similar calculations for the levels of agreement among Facebook's responses to twenty-five repeated queries (given at least one disagreement among the first five) yields the expected frequencies in Table 3.

Table 3: Comparison Between Observed and Expected Levels of Agreement Among the Facebook Database's Responses to Repeated Queries[78]

Level of Agreement	Observed Queries	Expected Queries
5 of 5	32,038	13,368 (40.9%)
24 of 25	558	367 (1.1%)
23 of 25	65	1,579 (4.8%)
22 of 25	8	3,257 (10.0%)
21 of 25	1	4,289 (13.1%)
≤ 20 of 25	0	9,810 (30.0%)

For these data, we have $\chi^2 = 44{,}963.972$ with five degrees of freedom, giving a two-tailed P value of less than 0.0001.[79] Assuming Facebook's database uses the hypothesized rounding and Laplace noise addition mechanisms, our findings in Table 1 provide a strong statistical basis for rejecting the null hypothesis that each of Facebook's responses to a repeated query is independently generated.

In summary, observations of Facebook's responses to an extensive range of potentially privacy-sensitive audience reach queries yield the following conclusions. First, the observed magnitude of discrepancies implies Facebook's database is not merely rounding the true audience reach statistics to the nearest twenty. Second, the observed distributions of discrepancies are consistent with our hypothesis that the database, in addition to rounding, is using a

78. The data for expected queries assumes that the responses are generated independently using the hypothesized rounding and Laplace noise addition mechanisms. The percentages do not add to 100% due to rounding.

79. *See Compare Observed and Expected Frequencies*, *supra* note 75.

(discrete[80]) Laplace noise addition mechanism. Finally, the observed frequencies of responses to repeated queries strongly support our hypothesis that any actual use of these hypothesized mechanisms is performed in combination with short-term caching.

B. Assessing Facebook's Apparent Solution

With advertising accounting for 83% of Facebook's worldwide revenues of $3.711 billion in 2011,[81] there can be little question that Facebook's advertising reach database provides a sufficient level of utility to advertisers to allow them to plan their campaigns with confidence. The more challenging question is the extent to which the level of differential privacy achieved by Facebook's ad targeting database can be extended to more general contexts.

As a possible framework for resolving privacy-utility tradeoffs in general, differential privacy has received mixed reviews from legal scholars and computer scientists. Some computer scientists have praised differential privacy as "a major step in the right direction."[82] Dwork asserts that differential privacy "has, for the first time, placed private data analysis on a strong mathematical foundation."[83] Critics, however, contend that differentially private mechanisms are impracticable in computationally intensive contexts[84] and place undue burdens on both the disclosers and recipients of data.[85] Such mechanisms are not as intuitive or simple as traditional procedures for removing specific data elements such as name, date of birth, and

80. *See supra* text accompanying note 57.

81. *See* Facebook, Inc., Registration Statement (Form S-1) 40, 55 (Feb. 1, 2012), *available at* http://sec.gov/Archives/edgar/data/1326801/000119312512034517/d287954ds1.htm.

82. Narayanan & Shmatikov, *supra* note 42, at 26.

83. *See* Dwork, *supra* note 43, at 95.

84. *See* Rathindra Sarathy & Krishnamurty Muralidhar, *Evaluating Laplace Noise Addition to Satisfy Differential Privacy for Numeric Data*, 4 TRANSACTIONS ON DATA PRIVACY 1, 15–16 (2011) [hereinafter *Evaluating Laplace Noise Addition*] (concluding that the use of Laplace noise additions as a differential privacy measure results either in lack of privacy or lack of utility, or both); Rathindra Sarathy & Krish Muralidhar, *Some Additional Insights on Applying Differential Privacy to Numeric Data*, PROCS. OF 2010 CONF. ON PRIVACY IN STAT. DATABASES 210, 212 (2010) [hereinafter *Some Additional Insights*] (finding that Laplace noise addition is suited only to numerical data where upper and lower bounds of query responses are known in advance).

85. *See* Xiao et al., *supra* note 46, at 1200–01(noting that Dwork's mechanism can potentially decrease the utility of data for researchers, especially with large data sets used in populations research); *see also* Ohm, *supra* note 6, at 1757 (contending that noise addition mechanisms require "complex calculations that can be costly to perform").

street address,[86] and can require custom programming prior to initial use of the data.[87] Dismissing differential privacy as a possible solution to the utility-privacy tradeoff,[88] Ohm concludes that "[u]tility and privacy are, at bottom, two goals at war with one another. In order to be useful, anonymized data must be imperfectly anonymous."[89]

This Facebook case study brings some needed concreteness to this discussion by providing an opportunity to assess the effectiveness and practicality of what we believe to be the most commercially successful implementation of a differentially private database system. In the remainder of this Section, we discuss four system-specific considerations pertaining to this assessment.

1. The Size of ε

By their nature, the above statistical tests say nothing about the probability that Facebook is actually using DL ($\exp(\varepsilon)$) noise addition, a mechanism that has been shown to achieve ε-differential privacy; we have shown only that the database's observed responses are not improbable if Facebook is indeed employing such a mechanism. The estimate $\varepsilon \approx 0.181$ represents the specific mechanism that provides the strongest support for this conclusion.[90]

Assuming that Facebook is indeed using our hypothesized mechanism or something close to it, we have no basis for an *a priori* view as to whether $\varepsilon \approx 0.181$ is "good enough" privacy for Facebook's users. What we can say is that by definition, 0.181-differential privacy implies a tolerance for up to a 20% change in the probability distribution of a mechanism resulting from the inclusion or exclusion of a user in the database. As Dwork notes:

86. Deidentification to comply, for example, with the HIPAA Privacy Rule is a straightforward matter of "suppressing" or "generalizing" personally identifiable information, such as names, addresses, and social security numbers. *See* Ohm, *supra* note 6, at 1711–16.

87. Ease of use is an important consideration. Studies have shown, for example, that uncertainty over computer technology has kept many doctors and healthcare practices from participating in national programs to speed the move to electronic health records. *See, e.g.*, Nir Menachemi et al., *Florida Doctors Seeing Medicaid Patients Show Broad Interest in Federal Incentive for Adopting Electronic Health Records*, 30 HEALTH AFF. 1461, 1464–67 (2011) (showing that, despite high interest in participating in federal incentive programs, doctors not participating cited uncertainty about technology, lack of specialized staff to implement technology, and the cost of new technology).

88. *See* Ohm, *supra* note 6, at 1756–58 (discussing limitations of differentially private mechanisms and other technological advances).

89. *See id.* at 1752.

90. *See supra* text accompanying notes 71–77.

> The choice of ε is essentially a social question . . . . That said, we tend to think of ε as, say, 0.01, 0.1, or in some cases, ln2 or ln3. If the probability that some bad event will occur is very small, it might be tolerable to increase it by such factors as 2 or 3, while if the probability is already felt to be close to unacceptable, then an increase by a factor of $e^{0.01} \approx 1.01$ might be tolerable, while an increase of e, or even only $e^{0.1}$, would be intolerable.[91]

In assessing whether 20% is a tolerable perturbation or risks revealing too much probabilistic information about the true value of a user reach statistic, it must also be kept in mind that the parameter ε is defined with respect to the behavior of the database in responding to a single query. As Dwork acknowledges and others have emphasized in critical responses to Dwork's work, the guarantee of differential privacy can rapidly deteriorate when confronted with a long series of queries.[92] If such degradation proves to be problematic, Facebook has available to it the technological option of restricting would-be advertisers to a number of queries deemed reasonably necessary to identify an appropriate target audience for their ad, and denying queries far in excess of that number as an abuse of the terms of service.[93]

2. Rounding and Caching

Facebook's apparent ancillary practices of caching responses to queries repeated over the short term and of rounding its database outputs to multiples of twenty have the salutary property that they preserve the guarantee of ε-differential privacy.[94] The rounding

91. Dwork, *supra* note 48, at 3.

92. *See id.* at 4; *see also Evaluating Laplace Noise Addition*, *supra* note 84, 9–15 (demonstrating vulnerability of Laplace noise addition to a "tracker attack" involving multiple queries).

93. *See generally Warning: Blocked from Using Feature*, FACEBOOK, https://www.facebook.com/help/?page=205619719473732 (click on the arrow next to "Why have I been blocked from using certain features?") (last visited May 4, 2012) (explaining that a user may be blocked from using certain features when Facebook determines that a user has been using a feature at rate that "is likely to be abusive," even though Facebook is unable to "provide any specifics on the rate limits that we enforce"). Fortunately, for purposes of the present study, Facebook has not taken such steps. *But cf. Statement of Rights and Responsibilities*, FACEBOOK, http://www.facebook.com/terms.php (last updated Apr. 26, 2011) ("You will not collect users' content or information, or otherwise access Facebook, using automated means (such as harvesting bots, robots, spiders, or scrapers) without our permission.").

94. To see this, note that for all D_1, D_2, the condition,

$$\Pr[r(K(D_1)) \in S] \leq \exp(\varepsilon) \times \Pr[r(K(D_2)) \in S]$$

is met whenever

discipline is also consistent with conventional understandings of precision and significant digits;[95] thus, by consistently outputting round numbers, the database tends to reinforce Facebook's notice to paying advertisers that its reported reach statistics are only estimates.

It is less clear whether these practices actually provide any privacy protections to Facebook's users beyond that already afforded by Laplace noise addition. In a 1989 survey paper on privacy mechanisms for statistical databases, computer scientists Nabil Adam and John Wortmann concluded: "Generally, rounding is not considered an effective security-control method. But combining rounding with other security-control methods seems to be a promising avenue."[96] Facebook's practice of reporting all user groups of less than forty as "fewer than 20 people" does seem to impede some simple and direct kinds of privacy attacks against individuals and small groups without significant loss of utility.

Symmetric noise addition mechanisms may be vulnerable to "averaging attacks," wherein an attacker simply repeats the same query and averages the responses; the Central Limit Theorem assures that the average will tend to converge to the true value.[97] Caching of responses to statistical queries may be deployed as an adjunct to noise addition mechanisms to defeat averaging attacks. As with other privacy techniques that require the logging of queries, there are significant time and storage overheads involved in storing and processing the accumulated logs.[98] Facebook could mitigate some of these overheads by maintaining its logs only over the short term, relying on rapid ongoing changes in its user population and their profile information to defend against averaging attacks over the longer term, and using probabilistic caching methods that improve efficiency at the cost of occasional cache misses. Such strategies may already be in use and reflected in the data in Table 1.

$$\Pr[K(D_1) \in r^{-1}(S)] \leq \exp(\varepsilon) \times \Pr[K(D_2) \in r^{-1}(S)]$$

where $r^{-1}(S)$ denotes the preimage of S under the rounding mapping r; the latter condition follows from the definition of differential privacy. *See supra* text accompanying note 48.

95. *See, e.g.*, NICHOLAS J. HIGHAM, ACCURACY AND STABILITY OF NUMERICAL ALGORITHMS 3–6 (2002) (explaining concepts of significant digits and precision in numerical analysis).

96. Nabil R. Adam & John C. Wortmann, *Security-Control Methods for Statistical Databases: A Comparative Study*, 21 ASS'N FOR COMPUTING MACHINERY COMPUTING SURVS. 515, 543 (1989).

97. *See id.*; Dwork, *supra* note 48, at 3.

98. *See* Adam & Wortmann, *supra* note 96, at 527.

Still, Dwork doubts that caching can provide a strong defense against averaging attacks:

> We do not recommend having the [database] curator record queries and their responses so that if a query is issued more than once the response can be replayed: If the query language is sufficiently rich, then semantic equivalence of two syntactically different queries is undecidable; even if the query language is not so rich, the devastating attacks demonstrated by Dinur and Nissim . . . pose completely random and unrelated queries.[99]

Despite Dwork's computability-theoretic reservations, Facebook may still be able to use short-term caching effectively in practice to defeat averaging attacks. Facebook should be able to detect semantically equivalent queries, because its advertising interface does not seem to support a very rich query language. For example, it seems to provide only one way to specify the set of 50- to 51-year-old North Carolina Facebook users interested in Alzheimer's disease. Also, Dinur and Nissim's "devastating attacks" rely on the attacker's ability to perform a very long series of queries,[100] which Facebook can defeat through technological restrictions.[101] While more study is needed to determine to what extent rounding and caching may contribute as adjuncts to Laplace noise addition generally in the practical implementation of differential privacy guarantees, Facebook's apparent use of these technologies in this specific context seems relatively easy to justify.

3. Extensibility to Social Network Data

The strength of our hypothesized privacy mechanism for Facebook's advertising reach database critically depends on the fact that it is a count database; this assures that $\Delta f = 1$, and thus $p = \exp(-\varepsilon)$. A database allowing queries about the characteristics of a group of c members could have $\Delta f = c$, so that the guarantee of ε-

99. Dwork, *supra* note 48, at 3 n.1 (citing Irit Dinur & Kobbi Nissim, *Revealing Information While Preserving Privacy*, 22 ASS'N FOR COMPUTING MACHINERY SPECIAL INT. GROUP ON MGMT. OF DATA-SPECIAL INT. GROUP ON ALGORITHMS & COMPUTATION THEORY-SPECIAL INT. GROUP ON ARTIFICIAL INTELLIGENCE SYMP. ON PRINCIPLES OF DATABASE SYS. 202, 202–10 (2003)).

100. *See* Irit Dinur & Kobbi Nissim, *Revealing Information While Preserving Privacy*, 22 ASS'N FOR COMPUTING MACHINERY SPECIAL INT. GROUP ON MGMT. OF DATA-SPECIAL INT. GROUP ON ALGORITHMS & COMPUTATION THEORY-SPECIAL INT. GROUP ON ARTIFICIAL INTELLIGENCE SYMP. ON PRINCIPLES OF DATABASE SYS. 202, 202 (2003) (using $n\log^2 n$ queries to infer the contents of a database, where n is the size of the database).

101. *See supra* note 93 and accompanying text.

differential privacy could assure only an $\exp(\varepsilon c)$ bound on dilation of the probability distribution of responses due to the inclusion or exclusion of a single participant's data.

To date, Facebook's advertising interface does not allow targeting to a user based on characteristics of the user's friends. Allowing queries to include friends' characteristics, however, would increase the value of Δf to the maximum size of a user's friends list, currently 5,000 friends.[102] While Dwork notes that an $\exp(\varepsilon c)$ bound on probability dilation "may be tolerable for small c,"[103] $c = 5{,}000$ is not "small," and $\exp(0.181 \times 5{,}000)$ is astronomical.

The observation that allowing queries involving friends' characteristics would effectively vitiate differential privacy provides further support for our hypothesis that Facebook is deliberately addressing its privacy-utility tradeoffs through the implementation of differential privacy mechanisms. Disallowing such queries represents a meaningful sacrifice of utility. For example, an airline presently cannot target an advertisement for discount fares from RDU to LAX to Raleigh-Durham area Facebook users who have 10 or more friends in Los Angeles. Privacy concerns provide a rational explanation for why Facebook has been willing to forego offering such a unique and potentially lucrative extension to its targeted ad platform.[104]

We have not examined Facebook's privacy practices with respect to the social network data it maintains; i.e., the graph-theoretic pattern of links formed between pairs of users who have identified themselves as friends on Facebook. We simply note here that Facebook routinely releases actual, personally identifiable social network data at the individual user level. Even though Facebook provides a setting that allows users to keep their friends list private,[105] friends lists are public by default,[106] and Facebook allows third-party software developers to crawl the public links of its social network

102. *See, e.g.*, Aimee Lee Ball, *Are 5,001 Facebook Friends One Too Many?*, N.Y. TIMES, May 28, 2010, at ST1 (discussing Facebook's 5,000-friend limit).

103. *See* Dwork, *supra* note 48, at 3.

104. *See infra* text accompanying note 123 (noting the alignment of Facebook's economic interests with its users' privacy interests).

105. *See Edit Profile*, FACEBOOK, http://www.facebook.com/editprofile.php?sk =relationships (last visited May 4, 2012) (providing a dropdown menu for visibility of friends list when logged in to Facebook).

106. *See* Jared Newman, *Facebook Beefs Up Security, Makes Captchas More Annoying*, PCWORLD (Jan. 26, 2011, 11:43 AM), http://www.pcworld.com/article/217844/ facebook _beefs_up_security_makes_captchas_more_annoying.html ("In Facebook's push to make users share more personal information, friends lists are now one of the things Facebook makes public by default.").

graph.[107] Thus, to the best of our knowledge, Facebook has made no effort to implement a differentially private mechanism to protect its social network data.[108]

In describing how the adoption and performance of Facebook's hypothesized privacy mechanisms may be predicated on various system-specific considerations, the preceding discussion might call into question the general applicability of differential privacy as a response to the reidentification threat. The following Part addresses the applicability of differential privacy technologies to other situations and ways that this standard for privacy could be incorporated into the law.

III. EXTENDING THE APPLICABILITY OF DIFFERENTIAL PRIVACY

The Facebook case study suggests a number of criteria for identifying contexts where the notion of differential privacy can play a useful role in standards of compliance with privacy laws. Although other solutions to the reidentification threat may also hold promise,[109] differential privacy can be an appropriate tool when these criteria are met. This approach can be flexible. When the criteria are present, a differential privacy standard may be appropriate; and when such a standard is appropriate, fulfillment of the standard may be recognized as evidence of compliance. Where privacy law already accommodates a fairly loose standard, such as a requirement of anonymization

107. *See* Eric Eldon, *Analysis: Some Facebook Privacy Issues Are Real*, *Some Are Not*, INSIDE NETWORK (May 11, 2010), http://www.insidefacebook.com/2010/05/11/analysis-some-facebook-privacy-issues-are-real-some-are-not/ (noting that public friends lists are "available to third parties through the Graph API, for services like search").

108. Perhaps this is for the best because a design for such a mechanism has thus far eluded computer science researchers. *See* Vibhor Rastogi et al., *Relationship Privacy: Output Perturbation for Queries with Joins*, 28 ASS'N FOR COMPUTING MACHINERY SPECIAL INT. GROUP ON MGMT. DATA-SPECIAL INT. GROUP ON ALGORITHMS & COMPUTATION THEORY-SPECIAL INT. GROUP ON ARTIFICIAL INTELLIGENCE SYMP. ON PRINCIPLES OF DATABASE SYS. 107, 108–09 (2009) (noting that previous mechanisms "do not . . . provide quantitative guarantees of privacy and utility" and presenting a novel approach that does not guarantee "ε-indistinguishability" but only "a somewhat weaker adversarial privacy").

109. *See* Gellman, *supra* note 36, at 47 (suggesting that legislatures establish "a statutory framework that will allow the data disclosers and the data recipients to agree voluntarily on externally enforceable terms that provide privacy protections for the data subjects"); Ohm, *supra* note 6, at 1759 (urging regulators to focus on situations in which "harm is likely and . . . outweighs the benefits of unfettered information flow" and regulate only those situations); Schwartz & Solove, *supra* note 9, at 1879 (arguing for a privacy standard called "PII 2.0," which places private information on a "continuum of risk" so that privacy laws can be more specific in terms of legal protections for various types of information).

without specificity of a particular methodology, the notion of differential privacy can fit into the existing law. Where privacy compliance is a function of specific rule-oriented actions, a guarantee of differential privacy might be deemed to satisfy such rules (thereby elevating substance over form) or new regulation or legislation might be required.

Criteria for appropriateness of implementation of differential privacy are based on the Facebook advertiser audience reach case study and on limitations on the technology described by computer science experts. Some kinds of data sets and some uses of those data sets are better candidates for differentially private mechanisms than others. The best opportunities will have the following characteristics:

(1) The interest in privacy is strong, and the risk of reidentification is significant.

(2) The information to be released is in a database.

(3) The database is large.

(4) The uses of the database can tolerate some distortion in the information from the database.

(5) The uses of the database do not involve study of outliers, other individuals, or relationship networks between or among individuals.

(6) The upper and lower ranges of numerical information to be sought from the database can be anticipated.

A. Strong Privacy Interest/Significant Reidentification Risk

As a first criterion, a strong privacy interest is necessary because differential privacy mechanisms do trade some utility of data for privacy. An assessment of the relative weight of privacy and utility interests is appropriate. This balancing assessment involves both technical and normative evaluations, including an evaluation of the current requirements under the law. Differential privacy mechanisms introduce random "noise" that will result in rounding of numerical answers to queries to the database, so the interest in privacy must be strong enough to offset this reduction in accuracy of the information. In addition, current implementations require an investment in software to create an interactive query interface between the discloser

and the recipient of data. As standard software packages become available, this particular barrier may be lowered.[110]

The underlying assumption in this strong privacy interest criterion is that the risk of reidentification is significant. If the risk of reidentification is insignificant, the privacy mechanisms may not need to be robust. Schwartz and Solove make this argument,[111] as does Jane Yakowitz.[112] However, if one accepts Ohm's argument and that of many computer scientists,[113] determination of which data elements carry the most risk is challenging. Even if one might survey existing auxiliary data sets to determine which data elements would be most vulnerable to reidentifying linkage, this survey cannot anticipate future publicly accessible data sets.[114] Whether a data discloser is responsible for anticipating future risks of linkage with future data sets is an important question. Privacy advocates have warned about the dangers of unanticipated uses of information and futility of most remedies once sensitive information is revealed.[115] The potential for future reidentification should be a factor in determining the strength of the privacy interest or privacy risk, but it should not prevent consideration of other factors.

In the Facebook advertising reach database, the privacy interests are strong and the risk of reidentification is significant. There may be multiple "privacy" interests. First, Facebook has enticed users to post vast amounts of personal information,[116] and requires an accurate

110. *See, e.g.*, McSherry, *supra* note 47, at 89 (describing a program he has dubbed PINQ, for Privacy Integrated Queries, which he says does not require a high level of computer expertise by users).

111. *See* Schwartz & Solove, *supra* note 9, at 1877–79.

112. *See* Yakowitz, *supra* note 36, at 45–46.

113. *See* Ohm, *supra* note 6, at 1742–45 (drawing on computer science research to argue that the "list of potential PII [personally identifiable information] will never stop growing until it includes everything").

114. *See* Latanya Sweeney, *k-Anonymity: A Model for Protecting Privacy*, 10 INT'L J. UNCERTAINTY, FUZZINESS & KNOWLEDGE-BASED SYS. 557, 563–65 (2002) (describing the difficulty of measuring reidentification risk for particular data elements).

115. *See* DANIEL J. SOLOVE, THE FUTURE OF REPUTATION: GOSSIP, RUMOR, AND PRIVACY ON THE INTERNET 42–49 (2007) (discussing the permanence of access to information once released through the Internet).

116. *See, e.g.*, Emily Bazelon, *Why Facebook Is After Your Kids*, N.Y. TIMES MAG. (Oct. 12, 2011), http://www.nytimes.com/2011/10/16/magazine/why-facebook-is-after-your-kids.html (noting that Facebook's default settings allow more sharing of information than many users realize and that most users do not adjust privacy settings to make them more protective); Jessica Guynn, *Facebook Drawing Fire over Privacy*, L.A. TIMES, Sept. 27, 2011, at B1 (describing Facebook's "passive sharing" services, which allow third-party applications to share "every action users take" on the site).

name, gender, address, and date of birth of all subscribers,[117] while promising not to share that information in certain ways.[118] Although the privacy policies and access settings that Facebook offers its users are criticized as complex, confusing, and subject to change,[119] Facebook does make some promises. These promises indicate that Facebook seems to feel that the privacy concerns of its users are an important inducement for getting and keeping that customer base. These promises also mean that Facebook may have privacy obligations under contract,[120] state consumer protection laws,[121] and Section 5 of the Federal Trade Commission Act, which protects against "unfair or deceptive acts or practices."[122] Finally, Facebook has its own interest in protecting the identities of its users from its advertisers. If Facebook were to provide information to advertisers that could be used to create independent targeted marketing lists, these advertisers would no longer need Facebook. Facebook's own interests in protecting its proprietary information overlap with its customers' interests, forming a strong privacy interest.[123] Facebook's database has many data elements that could be linked to outside information, so it is likely that full access to its data, even with many identifying elements suppressed, would produce reidentification threats. Thus, Facebook's release of advertising reach data presents a context in which the first criterion for differential privacy is met.

117. *See Statement of Rights and Responsibilities*, *supra* note 93 (prohibiting users from "provid[ing] any false personal information on Facebook").

118. According to Facebook, "We only provide data to our advertising partners or customers after we have removed your name or any other personally identifying information from it, or have combined it with other people's data in a way that it is no longer associated with you." *Data Use Policy*, *supra* note 15.

119. *See, e.g.*, James Grimmelmann, *Saving Facebook*, 94 IOWA L. REV. 1137, 1168 (2009) (explaining that "one of the most disruptive things that a social network site can do is to change the ground rules of how personal information flows").

120. *See* Woodrow Hartzog, *Website Design as Contract*, 60 AM. U. L. REV. 1635, 1639–50 (2011).

121. For a chart of state consumer protection laws, see *Managed Care State Laws and Regulations, Including Consumer and Provider Protections*, NAT'L CONFERENCE OF STATE LEGISLATURES, http://www.ncsl.org/default.aspx?tabid=14320#comprehensive (last updated Sept. 2011) (listing comprehensive consumer rights statute citations in Table 6).

122. 15 U.S.C. § 45(a)(1) (2006).

123. Privacy scholars have struggled with definitions for privacy and confidentiality for quite some time. This analysis of strength of privacy interests does not give a nuanced treatment to the concepts. The key point here is that any type of privacy interest, and the overlap of multiple such interests, can suggest an opportunity to implement mechanisms to support differential privacy. For scholarship on defining privacy and privacy harms, see generally DANIEL J. SOLOVE, UNDERSTANDING PRIVACY (2008) and Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477 (2006).

B. Information Must Be from a Database

The second criterion for differential privacy is that the information to be released is in a database. The mechanisms now in development to support differential privacy were created to address the threats to privacy from the sharing of information contained in databases. These mechanisms have broad applicability to databases, but they do not address a number of other privacy concerns. Most fundamentally, differential privacy cannot "put the horse back in the barn": it cannot be used to solve the problem of reidentification based solely on data that is already publicly available.[124] Beyond this problem, Facebook has been embroiled in a number of privacy controversies that could not be covered by the interactive query programs that have been developed to prevent reidentification of individuals. In one instance, Facebook was criticized for failure to prevent advertisers from receiving detailed information about Facebook users who clicked on an advertisement.[125] Facebook corrected this problem, but the query mechanism $K(D)$ at the heart of the definition of differential privacy was not applicable.[126]

C. The Database Must Be Large

The third criterion for implementing differential privacy is that the database must be large enough that the introduction of randomization nonetheless yields useful information. If the database is small, the noise necessary for masking the presence of any one individual is likely to destroy the utility of the information shared. With more than 800 million active users as of December 2011,[127]

124. Alessandro Acquisti and Ralph Gross have reported that using only publicly accessible data and face recognition software, they could predict complete Social Security numbers for 8.5% of the people born in the United States between 1989 and 2003. Alessandro Acquisti & Ralph Gross, *Predicting Social Security Numbers from Public Data*, 106 PROC. NAT'L ACAD. SCI. 10975, 10975 (2009).

125. *See, e.g.*, Jim Puzzanghera & Jessica Guynn, *Study: User Data Sharing Is Pervasive*, L.A. TIMES, Oct. 12, 2011, at B1 (describing new research that showed how some websites transfer personal information to advertisers and other third parties every time users log onto the sites).

126. Matt Jones, *Protecting Privacy with Referrers*, FACEBOOK NOTES (May 24, 2010, 11:24 PM), http://www.facebook.com/notes/facebook-engineering/protecting-privacy-with-referrers/392382738919 (announcing a fix to "an unintentional oversight" in the data shared with advertisers when Facebook users clicked on ads from within Facebook). Because this practice was contrary to Facebook's promise, the FTC identified it as offensive to the Federal Trade Commission Act prohibition against "deceptive trade practices." *See* Complaint at 13, Facebook, Inc., No. 0923184 (F.T.C. Nov. 29, 2011), 2011 WL 7096348 at *8.

127. *Newsroom: -Statistics*, FACEBOOK, http://newsroom.fb.com/content/default.aspx?NewsAreaId=22 (last visited May 4, 2012).

Facebook's advertising reach database clearly satisfies this requirement.[128] National- and state-level census databases represent other clear examples.[129] Just what number of records is large enough is a function both of the information in the database and of the type of information sought from the database. But, in general, data bulk is an important prerequisite for the implementation of differentially private database mechanisms.

D. Use of the Data Must Be Able To Tolerate Some Distortion

The fourth criterion is that the uses of the data must be able to tolerate some distortion. Because current mechanisms for supporting differential privacy function by introducing noise, precision in the released data is not possible.[130] Facebook advertisers at least appear to be satisfied with reach statistics that are variable and imprecise. Given Facebook's dominance in the social networking market, one might question whether advertisers are truly satisfied with the utility of these counts or whether their inferior bargaining position leads them to accept inferior information. The reality, however, is that advertisers do accept this quality of information, so the utility must be adequate for commercial purposes.

Other uses of large data sets might not accommodate even small inaccuracies in the data. For example, disclosures from census data that relied on introduction of noise have been criticized for containing too much noise for some types of analysis.[131] Certain studies of health care data from hospitals or from insurance companies might require perfect reporting of factors that differential privacy would obscure.[132] And, if Facebook had meaningful competition for advertisers, it might be pressured to provide more accurate reach data to secure advertisers' business.

128. *Id.*

129. *See, e.g.*, Johnnes Gehrke, *Technical Perspective: Programming with Differential Privacy*, COMM. ASS'N FOR COMPUTING MACHINERY, Sept. 2010, at 88, 88 (citation omitted).

130. *See Evaluating Laplace Noise Addition*, *supra* note 84, at 9–15.

131. Yakowitz, *supra* note 36, at 46–47 (noting that introduction of noise in samples of census data have resulted in demographic research errors).

132. *See* Douglas Peddicord et al., *A Proposal to Protect Privacy of Health Information While Accelerating Comparative Effectiveness Research*, 29 HEALTH AFF. 2082, 2083–84 (2010) (asserting that health data deidentified through customized and expensive statistical methods, presumably differential privacy methods, may meet the needs of the initial research query but would be "virtually impossible to combine with other deidentified data for new comparative effectiveness study").

E. Data Use Must Not Focus on Outliers

The fifth criterion for adopting a differential privacy standard precludes the use of the database to study outliers because that kind of information is inconsistent with differential privacy's goal of preventing identification of any one individual.[133] Differential privacy would obscure the outliers in any data set. In the Facebook advertiser example, some advertisers might prefer to have finer granularity in audience reach data, including knowing if they have selected audience reach characteristics that target a very small group or even one person. Advertisers seeking access to users of Facebook, though, have settled for less accurate feedback that specifically precludes the reporting of one or even a small number of individuals meeting advertiser-selected characteristics. In other contexts, data utility might require the study of outliers, so randomization technologies used to achieve differential privacy would be too severe a burden on utility of the data. For example, if a researcher of health data were interested in tracking individuals who were oddly immune to a particular pathogen, differential privacy might prevent identification of these small sets that could identify individuals.

At least for now, the obscuring of outliers through differential privacy technologies would also preclude the study of relationship networks between or among individuals. Noise-addition techniques that achieve differential privacy in statistical databases are generally unsuitable for reporting on relationships, essentially because each individual obscured produces ripple effects in the database and in the information reported.[134] Computer scientists have begun to develop privacy mechanisms specifically for social networks, but have not yet achieved the combination of reasonable utility and differential privacy.[135] As noted above, Facebook does not appear to have attempted to implement a differentially private mechanism to protect its social network data, but instead shares this information with third-party software developers.[136] Similarly, a database with genetic

133. Dwork, *supra* note 18, at 50:10–51:15 (noting the impossibility of studying outliers, social networking, or genetics through queries to a database using differential privacy technologies because by definition no individual may be identified).

134. *See* Dwork, *supra* note 43, at 91–95; *supra* notes 102–08 and accompanying text.

135. *See supra* note 108.

136. *See supra* text accompanying note 107; *see also Data Use Policy*, *supra* note 15 ("When you go to a game or application, or connect with a website using Facebook Platform, we give the game, application, or website (sometimes referred to as just 'Applications' or 'Apps') your User ID, as well your friends' User IDs (or your friend list).").

information that necessarily links individuals could not be well utilized if differential privacy were applied.[137]

F. Smallest and Largest Potential Numerical Answers Must Be Anticipated

The sixth criterion for the applicability of differentially private database mechanisms is specifically directed to Laplace noise addition and requires that all potential queries have a priori upper and lower bounds. This requirement is trivially met for count databases.[138] For example, in Facebook's advertising reach database, all queries take the form "How many users . . . ," so the lower bound is zero, and the upper bound is the number of Facebook users, a number that the database mechanism can monitor. Numerical databases where the results of queries can be of arbitrary magnitude are problematic, however, because there is no way to determine in advance the value of Δf, and, therefore, no way to identify the Laplace distribution Lap $(\Delta f / \varepsilon)$ from which the random noise is to be taken.[139]

G. Differential Privacy As Best Practice or Evidence of Privacy Compliance

Facebook's advertising reach database can serve as an exemplar for the adoption of differential privacy technologies in certain circumstances and recognition of differential privacy technologies as best practices for or evidence of compliance with privacy law. When the six criteria are met, as in the Facebook case study, differential privacy technologies can protect against reidentification without loss of utility. A number of kinds of data sets appear to exhibit our criteria in regulated areas such as health care, video rental, student records, and some government databases.

Explicit incorporation of differential privacy as a standard in privacy law would not be simple because U.S. privacy law is not simple. U.S. privacy law is a collection of piecemeal laws, including federal and state statutes that are largely sectoral, regulations that are standard- or rule-based, common law claims that vary by jurisdiction, and somewhat uncertain constitutional protections. Further,

137. *See* Rastogi et al., *supra* note 108, at 110 (demonstrating that joins lead to less accurate query results when using an ε-indistinguishable algorithm; Dwork, *supra* note 18 at 51:10–51:15 (explaining that study of outliers, social networks, and genetics is impossible using techniques designed to achieve differential privacy).

138. *Cf. supra* text accompanying note 55 (discussing the suitability of Laplace noise addition for count databases).

139. *See Some Additional Insights*, *supra* note 84, at 212.

incorporation of differential privacy as the standard may not be appropriate in some circumstances. A variety of approaches to privacy and prevention of reidentification are promoted by statisticians and computer scientists,[140] as well as by legal scholars.[141]

Differential privacy may, however, fit within existing privacy law requirements, even though these laws largely draw on assumptions that reidentification of individuals would be prevented through the suppression of a limited number of data elements. Recognition of differential privacy as a best practice for or evidence of compliance with privacy requirements can be most easily incorporated into areas of law that are grounded in standards, rather than a rule that might outline particular data elements for nondisclosure.[142]

The law governing privacy in social networking is largely standard-based, so this area is a strong candidate for incorporation of differential privacy. Contract, state, and federal consumer protection statutes, and negligence and privacy torts are generally amenable to acknowledging the evolution of best practices. When a social network makes a promise to users through its privacy policy, general promises to not share identifying information could be found to meet best practices when differential privacy techniques are implemented.[143] State consumer protection laws,[144] which often mirror the Federal Trade Commission Act's protection against unfair and deceptive trade practices, could also recognize differential privacy technologies as evidence of fulfillment of promises to protect individuals' privacy.[145] Similarly, a social network relying on differential privacy techniques should be able to defend itself against privacy torts and negligence claims for failure to take reasonable or effective steps to

140. Jerome Reiter et al., *Panel 2: Statistical Disclosure Control and HIPAA Privacy Rule Protections*, HEALTH & HUM. SERVICES WORKSHOP ON THE HIPAA PRIVACY RULE'S DE-IDENTIFICATION STANDARD 1:26 (Mar. 8, 2010), http://www.hhshipaaprivacy.com/conference_agenda.php?cid=1 (reviewing different approaches to balancing privacy and utility of data in health care and other industries and concluding that, at least with current technologies, "One-Size Will Not Fit All!").

141. *See* Schwartz & Solove, *supra* note 9, at 1870–72 (arguing for flexibility in the law to address rapidly changing social practices and technologies and to match differences in the types of behavior or information being regulated).

142. *See id.* at 1828–35 (categorizing privacy laws into two types of standard-oriented approaches and another category based on rules for protecting personally identifying information).

143. *Cf.* Hartzog, *supra* note 120, at 1635–39 (reviewing cases accepting website privacy policies as contracts and arguing that privacy settings on websites should also be enforced as contracts).

144. *See supra* note 121.

145. 15 U.S.C. § 45(a)(1) (2006) ("[U]nfair or deceptive acts or practices in or affecting commerce, are hereby declared unlawful.").

protect against reidentification of data that was promised to be kept nonidentifying. All of these areas of law are flexible enough to recognize differential privacy technologies as evidence of compliance with privacy promises.

Some other privacy laws, such as the federal Video Privacy Protection Act of 1988,[146] are based on standards that could also readily accept differential privacy applications.[147] This Act protects the privacy of video rentals and purchases by prohibiting videotape service providers from disclosing "personally identifiable information," which is defined as "information which identifies a person."[148] Student records and health records are protected under two statutes that may also incorporate differential privacy as evidence of compliance.[149] These statutes acknowledge the reidentification threat and prohibit disclosure that could be used to identify individuals.[150]

But, even if standard-oriented privacy laws adopted differential privacy as evidence of compliance, not all uses of those protected data sets would be served by differential privacy. The criteria for applicability exclude some uses that require more accuracy and access to outlier information. For example, computer programmers interested in the contest to develop the best improvement to the Netflix recommendation system might not find enough utility in an interactive query form of access to the customer database. Similarly, medical researchers might not find differentially private access to health care data sufficiently granular to advance knowledge of

146. Video Privacy Protection Act of 1988, Pub. L. No. 100-618, 102 Stat. 3195 (codified at 18 U.S.C. § 2710 (2006)).

147. *See, e.g.*, 18 U.S.C. § 2710 (forbidding video tape service providers from disclosing personally identifiable information, but not requiring a certain method to achieve this privacy protection); Family Education Rights and Privacy Act, 20 U.S.C. § 1232g(b) (2006) (stating that institutions possessing personally identifiable information of students must ensure that the information is not disclosed to third parties); Health Insurance Portability and Accountability Act, 42 U.S.C. § 1320d-2(d) (2006) (explaining that the Secretary shall set forth security standards for the protection of private health information, but not requiring that a specific means of privacy protection be utilized); 47 U.S.C. § 551(b)(1), (c)(1) (2006) ("[A] cable operator shall not disclose personally identifiable information concerning any subscriber without the prior written or electronic consent of the subscriber concerned and shall take such actions as are necessary to prevent unauthorized access to such information by a person other than the subscriber or cable operator.").

148. 18 U.S.C. § 2710(a)(3).

149. 20 U.S.C. § 1232g(b); 42 U.S.C. § 1320d-2(d).

150. *See* 20 U.S.C. § 1232g(b); 42 U.S.C. § 1320d-2(d)(6)(B).

prevention and treatment.[151] In the Netflix case, the normative answer may be that Netflix may not conduct another contest that would release customer data to the public, making reidentification likely.[152] Medical researchers, on the other hand, might be allowed more generous access to protected health care data than differential privacy would support, so long as they agreed to limitations on the use of that data.[153]

CONCLUSION

Facebook's apparent implementation of a commercially successful, differentially private database mechanism provides hope that in at least some contexts, the law can recognize best practices that go beyond traditional anonymization techniques to better protect privacy while maintaining utility of data. Differential privacy does not provide a solution to all problems of balancing privacy and utility, but privacy law should seek to enlist all the power it has to offer in the perpetual battle against the threat of reidentification.

151. *See* Douglas Peddicord et al., *A Proposal to Protect Privacy of Health Information While Accelerating Comparative Effectiveness Research*, 29 HEALTH AFF. 2082, 2083–84 (2010) (advocating that medical researchers be given full access to health care records and allowed to link data sets to conduct effective research).

152. As discussed earlier, Netflix did in fact cancel its second contest in the course of settling a lawsuit and an FTC investigation. *See supra* notes 27–29 and accompanying text.

153. *See* Gellman, *supra* note 36, at 58–59 (proposing new legislation that would support contracts allowing full access to government, non-profit, or research organizations with good records of data security if data is to be used "in research or in a public health activity").