PSYCHIATRY IN THE DIGITAL AGE (JS LUO, SECTION EDITOR)

Privacy in the Digital World: Medical and Health Data Outside of HIPAA Protections

Tasha Glenn · Scott Monteith

Published online: 14 September 2014

© Springer Science+Business Media New York 2014

Abstract Increasing quantities of medical and health data are being created outside of HIPAA protection, primarily by patients. Data sources are varied, including the use of credit cards for physician visit and medication co-pays, Internet searches, email content, social media, support groups, and mobile health apps. Most medical and health data not covered by HIPAA are controlled by third party data brokers and Internet companies. These companies combine this data with a wide range of personal information about consumer daily activities, transactions, movements, and demographics. The combined data are used for predictive profiling of individual health status, and often sold for advertising and other purposes. The rapid expansion of medical and health data outside of HIPAA protection is encroaching on privacy and the doctor-patient relationship, and is of particular concern for psychiatry. Detailed discussion of the appropriate handling of this medical and health data is needed by individuals with a wide variety of expertise.

Keywords Privacy · HIPAA · Data broker · Privacy policy · Predictive analytics · Trust · Doctor-patient relationship · Mobile apps

This article is part of the Topical Collection on Psychiatry in the Digital Age

T. Glenn ChronoRecord Association, Inc., Fullerton, CA 92834, USA

S. Monteith () Michigan State University College of Human Medicine, Traverse City Campus, 1400 Medical Campus Drive, Traverse City, MI 49684, USA

e-mail: monteit2@msu.edu

Introduction

Trust between doctor and patient is fundamental to the practice of medicine. A patient must trust the physician sufficiently to share personal details that may be stressful, embarrassing, or potentially damaging. A physician must trust that a patient is sharing enough information to make an accurate diagnosis, and that a patient is able to give informed consent about treatments that may pose significant risks. Trust in psychiatrists may be more important to patients with mental disorders than to patients with other serious illnesses [1]. An essential component of the trust between doctor and patient is privacy. Over two thousand years ago, Hippocrates emphasized the importance of privacy, and the practice of medicine has recognized and valued the importance of privacy ever since.

Privacy of medical data is regulated by federal and state laws but primarily HIPAA. HIPAA regulates patient data that is collected by providers and their business associates in relation to treatment, payment or healthcare operations. Most privacy discussions relate to concerns about HIPAA, such as the relative ease of re-identification of deidentified data [2, 3]. This review will focus on the medical and health data that are increasingly being collected outside of HIPAA protections. Medical and health data outside of HIPAA can be volunteered by consumers directly, observed by corporations recording consumer actions, and inferred by calculated models [4]. The rapidly expanding stores of data collected outside of HIPAA are encroaching on the traditional doctor patient relationship and eroding medical privacy.

Digital World

To understand the implications of medical and health data collected outside of HIPAA, it is necessary to review the scope and complexity of the rapidly expanding digital world. The



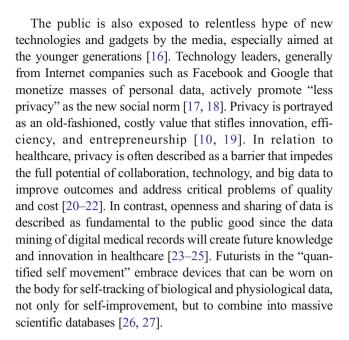
percent of the world's stored information that is in digital format has dramatically increased from 25 % in the year 2000 to more than 98 % in 2013 [5]. In the US, the amount of digital data is doubling every three years, driven by increased consumer use of smartphones, Internet, social networks and picture-taking, metadata (information about information), conversion from analog to digital (film, TV, voice), and the growth of machine generated data including RFID tags, sensors, and surveillance cameras [6]. Metadata for online transactions contains information such as account numbers, login IDs, passwords, phone numbers, browser types, IP addresses, date, time, email sender and recipient, search terms and results, cookies and device fingerprints [7••].

Eighty % of the digital data stored in the US is consumer related and the majority of this is data about consumers' lives, such as metadata, medical records and imaging, rather than data explicitly created by consumers such as emails sent or pictures taken [6, 8]. This personal detail is valuable because it can be combined, indexed and searched in databases, used to create individual digital dossiers and used for predictive modeling or profiling. Indeed, the digital databases about consumer daily activities, transactions and movements are considered to be a new asset class and the primary source of competitive advantage in the twenty first century [4].

Digital data does not reside where it was generated. Data moves and is serviced by many corporations and devices, including Internet service providers, communications companies, mail servers, database servers, web site owners, Internet retailers, data brokers, analytics firms and advertising networks. Every organization along this journey has the ability to copy and store data, including in countries with different regulations. About one-fourth of all digital data are original information, while the remaining three-fourths are duplications such as email attachments and backup copies [9].

Changing Public Perceptions of Privacy

Along with the expansion of the digital world, the public attitude toward privacy is evolving [10]. Although surveys completed outside of healthcare find that consumers still value privacy, there is a well documented "privacy paradox" showing inconsistencies between peoples intentions and behaviors relating to disclosing personal information [11]. Most consumers are willing to pay for online services with personal information rather than money [12, 13], or to disclose personal information for monetary rewards of less than \$50 [14]. Personal information is willingly and routinely disclosed in daily life to save time and money with the use of credit cards, cell phones, social media, search engines, and loyalty cards, and because the use of many digital technologies is no longer optional [15••].



Sources of Medical and Health Data Outside of HIPAA

Daily Sources

There are numerous daily sources of medical and health data outside of HIPAA protection. These include credit card payments for physician visit co-pays, purchase of over the counter (OTC) medications, home testing products, tobacco products, health foods, items related to disabilities, and visits to alternative practitioners [28, 29•, 30]. People also volunteer medical information online by searching for disease information, discussing their medical experiences in emails, blogs, chat groups, or social media sites including those dedicated to specific illnesses, or by calls to toll-free numbers. Other online activities that reveal medical information include registering for coupons on pharmaceutical direct-toconsumer advertising sites, registering for free trials of OTC products or online health services, registering for disease advocacy sites or to view patient support forums, "liking" web pages about diseases, completing online health and symptom checkers, and donating to health causes [30-33]. About three-fourths of consumers who use the Internet search for health information [34], and about three-fourths of health web sites contain third party tracking elements [35, 36]. Furthermore, one-third of U.S. consumers use YouTube, Facebook and Twitter for medical related discussions such as to check consumer reviews [37]. See Table 1 for an example of how a patient with depression may potentially disclose personal medical and health data outside of HIPAA protections.



Table 1 Examples of data that may potentially be collected outside of HIPAA protection for a patient with depression

Patient activity	Data	Source of data
Schedule appointment with psychiatrist using cell phone	Cell phone call to a psychiatrist or mental health facility	Telephone metadata [38]
Look up driving directions to psychiatrist/mental health facility	Driving directions from home to psychiatrist	Map web site content provider
Co-pay for visit to psychiatrist using credit card	Payment for visit to psychiatrist or mental health facility	Credit card records
Purchase prescribed medication using pharmacy loyalty card.	Purchase of psychotropic medications	Pharmacy loyalty programs that waive HIPAA rights [39]; credit card records.
Online search about depression and psychotropic drugs	Search terms	General search engine
Enrolls on pharmaceutical web site for drug discount coupon	Specific medication use	Pharmaceutical company
Reads web pages on depression	Web page activity	Medical web page content provider
Purchases book on depression	Book purchase	Online retailer
E-mail family with symptoms	Patient entered content	Email provider
Reads depression chat room	Web page activity	Medical chat room provider
Purchase OTC drugs such as St John's Wort	Drug purchase	Credit card records
Selects Facebook Like button on web page about depression	Web site visited	Social media and third party sites [40]
Uses medication reminder mobile app	Daily medications taken	Mobile app vendor

Other medical information outside the HIPAA framework is held by gyms, fitness clubs, wellness providers, banks, medical researchers, health fairs, and transit companies [29•]. Employers who do not fall under HIPAA, including those with fewer than 50 employees, may obtain medical information such as to determine ability to perform duties required for employment [41]. Additionally, state and federal governments are excluded from HIPAA requirements, allowing the storage of Medicaid records offshore [42] and allowing 33 state governments to sell or share personal health data [43].

Mobile Medical Apps

A myriad of technologies are now available to monitor every aspect of daily life including physiological measurements, physical activity and behavior [44•]. There has been an explosion of applications for mobile devices to promote health and disease self-management. As of 2012, there were about 13,000 health apps for consumers on the Apple AppStore, of which 5.8 % were related to mental health, 4.13 % to sleep, and 11.44 % to stress and relaxation [45]. A 2013 study reported 14,000 health apps, of which, 558 were for mental health and behavioral disorders, with two-third being for autism, anxiety, depression, and attention deficit hyperactivity disorder [46].

The vast majority of these applications are not medical devices and do not require FDA approval. The data from most apps are managed by the software vendor, not accessible by healthcare providers, and are outside of HIPAA regulations. Patients may mistakenly assume

that mobile apps are under the scope of HIPAA since the same data, such as heart rate, may be collected by an application that is accessible to their physician and covered by HIPAA, or on a mobile app that is not accessible to the physician and not covered by HIPAA [47]. Even data from a prescribed medical device may fall outside of the scope of HIPAA if it is sent directly to the device manufacturer, who in turn provides a summary report to the physician [48]. Many consumers are not aware that data from medical apps are frequently sent to the software vendor, and to third party sites for analytics and advertising services [49].

Patient Control of Digital Medical Records

Many patients are obtaining digital copies of their medical records, such as with Blue Button from the Veterans Administration. Once downloaded from a provider's EHR system, the medical record data are outside of HIPAA protection, and the patient becomes responsible for stewardship of the data. Patients without a background in technology management may inadvertently become a large source of leaking medical records. Moreover, data posted to the Internet are effectively permanent, since data cannot be deleted with assurance due to the distributed and redundant storage of Internet data [9, 50]. For example, comments from patients with multiple sclerosis containing private health information were found on YouTube health videos after their accounts were deleted [51]. Another concern is that patients will combine data downloaded from their EHR



with unprotected data in a mobile app. There are also many online sites for maintaining personal health records (PHR) although these are rarely used today.

Data Brokers

Data brokers, also referred to as data aggregators or information resellers, are a multi-billion dollar industry that collect, analyze, and sell data on consumers [28, 52•]. As of 2012, about 4000 data brokers have data on about 300 million Americans [53]. Data brokers collect data from every aspect of our lives including public records such as property taxes and voter registrations, publicly available information such as phone numbers and Internet postings, and non-public information such as financial data, loyalty cards, and Internet transactions [28, 52•]. Additionally, consumers have accepted location aware mobile devices such as smartphones, which contain multiple sensors, are frequently always-carried and always-on, and provide tracking information [54•]. Data brokers link together data from all of an individual's online and offline accounts and devices [52, 55], and some store data indefinitely [30]. In general, consumers do not have the right to control what personal information is collected, maintained, used, and shared by data brokers or to correct errors [28, 30, 52•]. Furthermore, data brokers routinely purchase data from other data brokers, so a consumer could not realistically trace the source of incorrect data [30]. Most regulations that impact data brokers pertain to the financial sector such as under the Fair Credit Reporting Act. The primary products from data brokers are used to predict consumer behavior and are sold mainly to online marketers.

Medical and Health Products from Data Brokers

Data brokers sell a variety of products about health issues based on data collected outside of HIPAA. Consumer lists are available by diagnosis such as depression, ADHD, or anxiety [52•, 56, 57] and by medications taken such as antidepressants [58]. Data brokers also combine health data with data from consumer habits, assets, and demographics to use in consumer health scores, profiling, and predictive modeling [59]. Examples of scores that are used outside the HIPAA framework include the Brand Name Medicine Propensity Score from Acxiom [60] and the FICO Medication Adherence score [29•, 61]. Consumer health scores may be used as variables within predictive models by life insurers or actuaries as part of an evaluation process [62, 63]. Data collected by data brokers can also be purchased for re-identification. This is of great concern since the more information available about a person, the easier it is to re-identify the person in the future [3].



Predictive modeling, referred to by the advertising community as behavioral targeting, is used to bring specific advertisements to online users based on their perceived interests. Behavioral targeting is about twice as effective as other forms of online advertising, and is viewed as critical for a business model that provides free online content and services [64]. The data used to create behavioral targeting algorithms includes detailed activity at websites from content providers (such as search terms, search histories and content selected), clickstreams (route navigated across the Web), and a wide range of data purchased from data brokers. Many analysts believe that the more data that can be combined, the more precise the profile that can be generated about our habits. Acxiom offers "over 3000 propensities for nearly every U.S. consumer" [65].

Most algorithms used for profiling and targeted marketing are not publicly available but medical and social science researchers have identified a wide variety of individual traits and behaviors based on Internet data. Researchers have investigated patterns of activity, linguistic style, and emotional expression in the content of social media [66]. For example, personality was predicted from data in Twitter [67], personal web sites [68], and Facebook [69]. Data from Facebook were used to identify depression in college students [70], ethnicity and sexual orientation [71••], and schizotypy personality [72]. Data from Twitter were used to predict postpartum emotional changes [66].

Predictive modeling is also used to estimate health status and may have the same consequences for an individual as if the information came from an electronic medical record (EMR). For example, when Target predicted that a customer was pregnant due to purchasing patterns [73], it caused as much distress as if this was based on actual data from a healthcare provider [74]. This incident also highlighted that personal health information can be created by combining seemingly innocuous data, and that a predictive model outside of HIPAA protection can cause harm whether or not it is accurate. Health predictions may seriously impact a person's life including getting and keeping a job, and the ability to get life insurance [74]. Although health predictions may be incorrect or disclose information people want kept private [50], the current legal framework does not address predictive models using data outside of HIPAA [74, 75]. Adverse consequences of health profiling may affect members of certain groups disproportionately [50], such as those with mental illness. Health profiling is accurate enough to use to recruit patients for clinical trials [76].

Since data outside of HIPAA are easy to obtain and subject to minimal regulation, the use of predictive models of health status as a substitute for actual individual medical data may increase [75]. Predictive health models can also be combined



with traditional medical data, such as that leaked by a patient controlling data downloaded from a provider's EHR system. This could lead to a future in which data brokers have more detailed information about a patient than that directly disclosed to their physician. It is important to remember that the results of predictive models are not based on physician judgment or on a directly measured value, but are calculated values often by disciplines outside of medicine. The accuracy of commercial predictive models is not published and replicated like the results of a scientific study. Additionally, the data brokers who sell predictive health models are not involved in patient care and have no training in medical ethics.

Selling Patient Experience

One area of particular concern involves the health web sites at which users create as well as read content, such as online patient support communities. This data often consists of self-reported diagnoses, medical history, symptoms, treatments, drug reactions, and patient opinions about providers. These web sites commonly have a business model based on aggregating, mining, and selling user generated content, often to pharmaceutical companies, device manufacturers or researchers [77, 78]. Patient generated data is particularly valued by marketing organizations because it reflects routine behavior rather than answers to solicited surveys [77]. Many companies behind these web sites actively encourage sharing of data in order to build larger databases [79].

Patients may not be aware of the commercial ownership of these web sites [79] or may not realize the extent of the third party involvement [78, 80]. For example, in a study of 69 patient support sites, pharmaceutical connections to the organizations were difficult to determine by end users [81]. People who are comfortable sharing data online for the betterment of the general good may not want to do so to enrich a company [79]. Additionally, there are a growing number of web scraper companies that automatically gather data from unstructured or semi-structured data pages of target websites to amass large databases. One healthcare example is Treato, which "automatically collects the massive amount of patient-written health experiences from blogs and forums", then processes the data and sells to pharmaceutical marketers [82]. Finally, there are technical privacy issues unique to social networking sites such that the data may be more difficult to anonymize than that in relational databases [2, 80].

Privacy Policies for Online Activities

Internet privacy policies are not succeeding at explaining the risks of data sharing to the public, and may serve more as liability disclaimers than as assurances of consumer privacy [83]. Most people do not even read online privacy policies, including at healthcare web sites, or understand that commercial organizations share, analyze and sell data [84–86]. Surprisingly, many people have unexpected reactions to privacy policies. Some consumers mistakenly believe that the mere presence of a "privacy policy" means that their information will be kept private, and that the web site will not share their information [85]. Additionally, the perception of control over the release of information from a privacy policy may increase consumers' willingness to disclose sensitive information, even if actual control is not increased [87]. In contrast, some people see the presence of a privacy policy as a warning of an unsafe environment, and will withhold more information than when there is no mention of privacy [88].

Multiple studies of healthcare websites have found that the privacy policies are difficult to understand. Most privacy policies are written at a reading level equivalent to two years of college [89–92] although half the US adult population has completed less than 1 year of college [93]. One study found that the privacy policies of 185 major health institutions were about as long as a research article in JAMA [94]. A comparison of privacy policies for nine healthcare websites before and after HIPAA legislation found that after the legislation, the policies were more descriptive but longer and more difficult to comprehend [90]. The readability issues may be more important for patients with mental illnesses since they may also have impaired reading abilities [95, 96].

On social media web sites, privacy policies apply only to the data that the social media companies collect from the users such as through registration forms or cookies, and not to the content that is posted directly by the users [86]. Although the Federal Trade Commission (FTC) states all mobile applications should have a privacy policy [97], a study of 43 popular mobile health and fitness apps for Apple and Android devices found that less than half posted a privacy policy, and less than half of these policies were accurate [49]. A review of privacy policies on 24 PHR systems reported that the descriptions of security and privacy measures were insufficient, and compliance with HIPAA regulations were low [98]. Many consumers lack the technical skills to control privacy online, such as to change the default privacy settings on social media sites or browsers, or to use advertiser opt-out sites [99–101]. Increased consumer training on technical skills is needed to maximize use of the existing online privacy options.

Data Breaches

Disclosures of HIPAA protected medical data are a major concern. The enforcement provisions in HIPAA were significantly strengthened by the 2009 HITECH Act, which included the first federal data breach notification, instigated security audits, significantly increased fines and authorized HIPAA



Table 2 National survey findings of adults in the US regarding privacy concerns about HIPAA protected medical records

Survey	N	Privacy findings
Harris Interactive 2007, [113]	2337	30 % are not satisfied with the way doctors and hospitals protect confidentiality of PHI
California Healthcare Foundation 2010, [114]	1898	68 % are concerned about privacy of personal medical records
California Healthcare Foundation 2005, [115]	2100	67 % somewhat or very concerned about privacy of medical records.
Ancker 2013, [116]	2013	48 % believed health IT would worsen privacy and security
National Partnership for Women & Families 2012, [117]	1961	About 60 % say widespread adoption of EHR will lead to more lost or stolen PHI About half say PHI not well protected by current laws
Westin/Institute of Medicine 2007, [118]	2392	58 % think privacy of medical records not protected well enough by current federal and state laws and organizational practices
Employee Benefit Research Institute 2008, [119]	1000	62 % do not think data in an EMR would remain confidential
Markle 2006, [120]	1003	80 % very concerned about identity theft or fraud, and 77 % very concerned about marketing firms getting their information
Deloitte 2012, [121]	4000	About 1/3 not comfortable with safeguards for personal health information
Agaku 2014, [122]	3959	About 2/3 concerned about data breaches with electronic records
NPR/Kaiser 2009, [123]	1238	76 % thought it likely that an unauthorized person would get access to EMR
FairWarning 2011, [124]	1265	Only 48 % believe their provider is committed to protecting their privacy.
Lowes 2012, [125]	2147	63 % fear a computer hacker will steal their personal data in EMR

enforcement by the states attorney generals [102]. Yet data breaches of HIPAA protected medical information are increasing in frequency [102]. In a 2014 survey of 91 healthcare organizations, 90 % reported at least one incident in the last two years while 38 % reported more than five incidents [103]. When including only breaches involving at least 500 individuals, over 29 million patient health records have been compromised since 2009. Medical data breaches are well publicized in the press [104] and at a web site from HHS for all breaches affecting more than 500 patients [105].

Many people require access to medical records, including doctors, nurses, technicians, administrators, clerical workers, and those working in business associates such as insurance companies, billing, coding and transcription companies, pharmacies, medical suppliers, care facilities, and government offices. This fragmented nature of the US healthcare system makes data breaches particularly difficult to control since the risk of a breach is the product of the risk at each of the organizations involved. About 20 % of the recent breaches involved a business associate [106], many of which lack technical expertise [107]. Most breaches involve portable

devices [106] and the most common cause is theft [102]. EMR are a prime target for theft, since they contain financial, credit, personal and insurance information, and medical identity theft is the fasting growing healthcare fraud [108].

It is harder to know how frequently breaches occur at data brokers as there is no current federal standard for breach notification by data brokers. However, large breaches have been reported, including at LexisNexis, Kroll Background America [109], Experian [110], and Acxiom [111]. PHR that are not associated with HIPAA-covered entities are regulated by FTC breach notification requirements [112].

Losing Trust

Although the public routinely gives away most personal information, medical privacy remains uniquely important to most, as underscored by the very existence of HIPAA and HITECH. The use of technology in medicine is widely supported but concern remains about the security of the medical information that is protected by HIPAA, such as in EMR, as

Table 3 National survey findings of adults in the US on withholding medical information due to privacy concerns related to technology

Survey	N	Findings on withholding information
California Healthcare Foundation 2010, [114]	1898	48 % would or may hide information from their doctor if it was shared through an EHR
Harris Interactive 2007, [113]	2337	17 % would withhold medical data because of worries about data disclosure
ONC 2014, [127]	2050	7 % withheld information from their doctors for privacy concerns, increasing five-fold among those thinking EHR inadequately protected
Agaku 2014, [122]	3959	12 % withhold information out of concern for a data breach
FairWarning 2011, [124]	1265	$25\ \%$ would withhold information or postpone seeking care if they had a sensitive medical condition



summarized in Table 2. In a study of psychiatric outpatients almost 90 % had concerns about confidentiality with the use of EMR, such as unauthorized access within a university healthcare system, inappropriate use of information, and stigmatization [126]. There are serious consequences when patients fear their privacy is at risk. Patients may become selective about the information they provide, offering an incomplete or misleading description of their condition. In recent surveys, a substantial number of people said they would withhold data from their physician due to privacy concerns related to technology, as shown in Table 3. Patients who are worried about privacy are also less likely to seek care or return for follow-up treatment, or may seek care outside of their provider network undermining the benefits of care coordination [126, 128].

Much of the general public is unaware of the large amount of medical and health data being amassed outside of HIPAA confidentiality protections. As the public becomes more informed about the secondary market for health data, concern about privacy and security of all medical data is likely to increase. This, in turn, may dissuade more people from seeking help or revealing the information to physicians. This is of particular concern to psychiatry, since patients with mental disorders are more likely to withhold information from their doctors than patients with other serious illnesses [1].

Protecting Data

The first step to protect medically related data from collection by data brokers and Internet companies outside of HIPAA protection is to recognize the scope of the problem. Actions needed to address this complex problem are outside the scope of this review. These include steps which are specific to individual activities, devices, and applications, and changes to federal and state laws.

Conclusions

Large quantities of health data are being created outside of HIPAA protection, primarily by consumers. Most of the data generated by consumers are controlled by data brokers and Internet companies that have no involvement in patient care and no training in medical ethics. Data brokers are combining health data with other consumer data to make health related profiles, which may increasingly be used to identify individual health status. The results of the predictive profiles may have adverse impact regardless of accuracy. As knowledge of data brokers becomes more widespread, more patients may avoid healthcare or withhold data from physicians due to privacy concerns, which may have especially serious consequences in psychiatry. The far reaching problems relating to the use and protection of medical and health data outside of HIPAA need

to be addressed by broad collaborations of medical, legal, consumer, and technical expertise. In the interim, measures to increase awareness of the growth of medical and health data outside of HIPAA protection are needed for both clinicians and patients.

Compliance with Ethics Guidelines

Conflict of Interest Scott Monteith declares no conflict of interest.

Tasha Glenn shares a patent for ChronoRecord software but does not receive any financial compensation from The ChronoRecord Association, a 501(c)(3) nonprofit organization.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any authors.

References

Papers of particular interest, published recently, have been highlighted as:

- · Of importance
- Of major importance
 - Mechanic D, Meyer S. Concepts of trust among patients with serious illness. Soc Sci Med. 2000;51(5):657–68.
 - Narayanan A, Shmatikov V. Myths and fallacies of personally identifiable information. Commun ACM. 2010;53(6):24–6.
 - 3. Ohm P. Broken promises of privacy: responding to the surprising failure of anonymization. UCLA Law Rev. 2010;57(6).
 - World Economic Forum. Personal data: the emergence of a new asset class. 2011. http://www.weforum.org/reports/personal-dataemergence-new-asset-class. Accessed 31 May 2014.
 - Cukier KN, Mayer-Schoenberger V. The rise of big data: how it's changing the way we think about the world. Foreign Aff. 2013. http://www.foreignaffairs.com/articles/139104/kenneth-neilcukier-and-viktor-mayer-schoenberger/the-rise-of-big-data. Accessed 31 May 2014.
 - IDC. The digital universe in 2020: big data, bigger digital shadows, and the biggest growth in the far east - United States. 2013. http://www.emc.com/collateral/analyst-reports/idc-digitaluniverse-united-states.pdf. Accessed 31 May 2014.
 - 7.•• Guardian. A Guardian guide to your metadata. 2013. http://www.theguardian.com/technology/interactive/2013/jun/12/what-ismetadata-nsa-surveillance#meta=0000000. Accessed 31 May 2014. Clear tables on what is included in metadata for email, phone, Facebook, Twitter, search and web browser.
 - IDC. The diverse and exploding digital universe. 2008. http:// www.emc.com/collateral/analyst-reports/diverse-explodingdigital-universe.pdf. Accessed 31 May 2014.
 - IDC. The digital universe decade are you ready? 2010. http:// www.emc.com/collateral/analyst-reports/idc-digital-universe-areyou-ready.pdf. Accessed 31 May 2014.
 - Cohen JE. What privacy is for (November 5, 2012). Harv Law Rev. 2013;126.
 - Norberg PA, Horne DR, Horne DA. The privacy paradox: personal information disclosure intentions versus behaviors. J Consum Aff. 2007;41(1):100–26.
 - McDonald AM, Cranor LF. Americans' attitudes about internet behavioral advertising practices. In: Proceedings of the 9th annual



- ACM workshop on privacy in the electronic society. ACM; 2010. 63–72
- Bauer C, Korunovska J, Spiekermann, S. On the value of information—what facebook users are willing to pay. In: 20th European Conference on Information Systems proceedings (ECIS 2012). 2012.
- Hann IH, Hui KL, Lee SYT, et al. Overcoming online information privacy concerns: an information-processing theory approach. J Manag Inf Syst. 2007;24(2):13–42.
- 15. Abelson H, Leeden K, Lewis H. Blown to bits: your life, liberty, and happiness after the digital explosion. Addison-Wesley Professional; 2008. For those wanting background information, excellent introduction to the digital world.
- Black A, Gen Y. Gen Y: who they are and how they learn. Educ Horiz. 2010;88(2):92–101.
- Newman J. Google's Schmidt roasted for privacy comments. PC World. 2009. http://www.pcworld.com/article/184446/googles_schmidt_roasted_for_privacy_comments.html. Accessed 31 May 2014.
- Johnson B. Privacy no longer a social norm, says Facebook founder. The Guardian. 2010. http://www.theguardian.com/technology/2010/jan/11/facebook-privacy. Accessed 31 May 2014.
- Shapiro G. Op-Ed: don't let privacy concerns stifle innovation. Nextgov.com. 2013. http://www.nextgov.com/emerging-tech/ 2013/06/op-ed-dont-let-privacy-concerns-stifle-innovation/ 65195/. Accessed 31 May 2014.
- Kaye J. The tension between data sharing and the protection of privacy in genomics research. Annu Rev Genomics Hum Genet. 2012;13:415–31.
- Lane J, Schur C. Balancing access to health data and privacy: a review of the issues and approaches for the future. Health Serv Res. 2010;45(5 Pt 2):1456–67.
- Shachak A, Jadad AR. Electronic health records in the age of social networks and global telecommunications. JAMA. 2010;303(5):452–3.
- Groves P, Kayyali B, Knott D, et al. The 'big data' revolution in healthcare: accelerating value and innovation. McKinsey & Company; 2013. http://www.mckinsey.com/insights/health_ systems_and_services/the_big-data_revolution_in_us_health_ care. Accessed 31 May 2014.
- Institute of Medicine. Best care at lower cost. The path to continuously learning health care in America. 2012. http://www.iom.edu/~/media/Files/Report%20Files/2012/Best-Care/Best%20Care%20at%20Lower%20Cost Recs.pdf. Accessed 31 May 2014.
- Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA. 2013;309(13):1351–2.
- Swan M. The quantified self: fundamental disruption in big data science and biological discovery. Big Data. 2013;1:85–99.
- Ramirez E. How can we get more meaning out of our data? Quantified Self knowledge through numbers. 2013. http:// quantifiedself.com/2013/08/how-can-we-get-more-meaning-outof-our-data/ Accessed 31 May 2014.
- Government Accountability Office. Information resellers: consumer privacy framework needs to reflect changes in technology and the marketplace. 2013. http://www.gao.gov/assets/660/658151.pdf. Accessed 31 May 2014.
- 29.• Dixon P, Gellman R. The scoring of America: how secret consumer scores threaten your privacy and your future. World Privacy Forum. 2014. http://www.worldprivacyforum.org/2014/04/wpf-report-the-scoring-of-america-how-secret-consumer-scores-threaten-your-privacy-and-your-future/. Accessed 31 May 2014. A review of consumer scoring, describing scores and rankings created from consumer data such as for health, financial, identity and authentication.
- Federal Trade Commission. Data brokers: a call for transparency and accountability. 2014. http://www.ftc.gov/system/files/documents/

- reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf Accessed 31 May 2014.
- Monteith S, Glenn T, Bauer M. Searching the internet for health information about bipolar disorder: some cautionary issues. Int J Bipolar Disord. 2013;1:22.
- Sheehan KB. In poor health: an assessment of privacy policies at direct-to-consumer web sites. J Public Policy Mark. 2005;24(2): 273–83
- Mackey TK, Yagi N, Liang BA. Prescription drug coupons: evolution and need for regulation in direct-to-consumer advertising. Res Soc Adm Pharm. 2014;10(3):588–94.
- Fox S, Duggan M. Health online. Pew Res. 2013. http://www. pewinternet.org/Reports/2013/Health-online.aspx Accessed 31 May 2014.
- Krishnamurthy B, Naryshkin K, Wills C. Privacy leakage vs. protection measures: the growing disconnect. In: Web 2.0 Security and Privacy Workshop, 2011. http://www2.research.att.com/~bala/papers/. Accessed 31 May 2014.
- Huesch MD. Privacy threats when seeking online health information. JAMA Intern Med. 2013;173(19):1838–9.
- Pwc. Social media "likes" healthcare: from marketing to social business. 2013. www.pwc.com/us/en/health-industries/ publications/health-care-social-media.jhtml. Accessed 31 May 2014.
- Mayer J, Mutchler P. MetaPhone: the sensitivity of telephone metadata. http://webpolicy.org/2014/03/12/metaphone-thesensitivity-of-telephone-metadata/. Accessed 31 May 2014.
- Lazarus D. CVS thinks \$50 is enough reward for giving up healthcare privacy. Los Angeles Times. 2013. http://www. latimes.com/business/la-fi-lazarus-20130816,0,6519110,full. column. Accessed 31 May 2014.
- Valentino-DeVries J, Singer-Vine J. They know what you're shopping for. Wall Str J. 2012. http://online.wsj.com/news/articles/ SB10001424127887324784404578143144132736214#printMode.
- Tudor ML. Protecting privacy of medical records of employees and job applicants in the digital era under the Americans with Disabilities Act. North Ky Law Rev. 2013;40:635–65.
- Dickson V. Offshore health record storage may pose privacy risks. Mod Healthc. 2014. http://www.modernhealthcare.com/article/ 20140418/blog/304189995. Accessed 31 May 2014.
- Hooley S, Sweeney L. Survey of publicly available state health databases. Harvard University Data Privacy Lab. 1064-1. 2013. http://privacytools.seas.harvard.edu/files/privacytools/files/1075-1.pdf.
- 44.• Lowe SA, Ólaighin G. Monitoring human health behaviour in one's living environment: a technological review. Med Eng Phys. 2014;36(2):147–68. Review of technologies used for behavioral monitoring.
- Dolan B. Report:13K iPhone consumer health apps in 2012. MobileHealthNews. 2012. http://mobihealthnews.com/13368/report-13k-iphone-consumer-health-apps-in-2012/. Accessed 31 May 2014.
- 46. IMS. Patient apps for improved healthcare from novelty to main-stream. 2013. http://www.imshealth.com/portal/site/imshealth/menuitem.762a961826aad98f53c753c71ad8c22a/?vgnextoid=e0f913850c8b1410VgnVCM10000076192ca2RCRD. Accessed 31 May 2014.
- Landman Z. Debunking the most common myths about HIPAA. mHealthnews.com. 2013. http://www.mhealthnews.com/news/ debunking-most-common-myths-about-hipaa?single-page=true. Accessed 31 May 2014.
- 48. Marcus AD, Weaver C. Heart gadgets test privacy-law limits. Wall Str J. 2012. http://online.wsj.com/news/articles/ SB10001424052970203937004578078820874744076 Accessed 31 May 2014.



- Privacy Rights Clearinghouse. Technical analysis of data practices and privacy risks of 43 popular mobile health and fitness applications. 2013. https://www.privacyrights.org/mobile-medical-appsprivacy-technologist-research-report.pdf. Accessed 31 May 2014.
- President's Council of Advisors on Science and Technology. Big data and privacy: a technological Perspective. 2014. http://www. whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_ big data and privacy - may 2014.pdf.
- 51. Fernandez-Luque L, Elahi N, Grajales FJ. An analysis of personal medical information disclosed in youtube videos created by patients with multiple sclerosis. In: Adlassnig K-P, et al. (Eds.) Medical Informatics in a United and Healthy Europe: Proceedings of MIE 2009, the XXII International Congress of the European Federation for Medical Informatics. IOS Press; 2009. 150:292.
- 52.• US Senate Committee on Commerce, Science, and Transportation. A review of the data broker industry: collection, use, and sale of consumer data for marketing purposes. 2013. http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=0d2b3642-6221-4888-a631-08f2f255b577. A clearly written review of the data broker industry.
- Armour S. Data Brokers come under fresh scrutiny. Wall Str J. 2014. http://online.wsj.com/news/articles/ SB10001424052702303874504579377164099831516.
- 54. Michael K, Clarke R. Location and tracking of mobile devices: Überveillance stalks the streets. Comput Law Secur Rev. 2013;29(3):216–28. A review of how mobile devices are used for location tracking.
- Steel E. Acxiom to create 'master profiles' tying offline and online data. Financ Times. 2013. http://www.ft.com/cms/s/0/151d940e-2431-11e3-8905-00144feab7de.html. Accessed 31 May 2014.
- 56. Epsilon. Consumer data and data cards Ailments/health. http://lists.epsilon.com/market;jsessionid= E46C0F404A2FCB1EF6F0A24EE0DEC61A?page=research/ datacard&id=91407. Accessed 31 May 2014.
- 57. TargetSource. U.S. health and ailment database. http://lists.nextmark.com/market;jsessionid=1E89AC694197AB78C356A7B6672FD5BA?page=order/online/datacard&id=210939. Accessed 31 May 2014.
- DMDatabase.com. Ailments mailing list. http://dmdatabases.com/ databases/consumer-mailing-lists/ailments-lists. Accessed 31 May 2014.
- Garla S, Hopping A, Monaco R, Rittman R. What do your consumer habits say about your health? Using third-party data to predict individual health risk and costs. SAS Institute. 2013. http://support.sas.com/resources/papers/ proceedings13/170-2013.pdf.
- Acxiom Update Newsletter. Stay current with Acxiom product and industry alerts. 2009. http://www.mktgservices.com/ marketing/newsletter/myAcxiomUpdate/0509/v1/acxiom-alerts. html. Accessed 31 May 2014.
- FICO. Medication adherence score. http://www.fico.com/en/ products/fico-medication-adherence-score/. Accessed 31 May 2014.
- Scism L, Maremont M. Insurers test data profiles to identify risky clients. Wall Str J. 2010. http://online.wsj.com/news/articles/ SB10001424052748704648604575620750998072986. Accessed 31 May 2014.
- Hill T. Predictive modeling in life insurance underwriting. Society
 of Actuaries. The Future of Preferred Underwriting. 2013. http://
 www.soa.org/search.aspx?searchterm=tim%20hill%202013.
 Accessed 31 May 2014.
- 64. Network Advertising Initiative. Study finds behaviorally-targeted ads more than twice as valuable, twice as effective as non-targeted online ads. 2010. http://www.networkadvertising.org/pdfs/NAI_ Beales Release.pdf. Accessed 31 May 2014.

- Acxiom Annual Report. 2013. http://acxiom.com/wp-content/ uploads/2013/09/2013-Annual-Report.pdf. Accessed 31 May 2014
- De Choudhury M, Counts S, Horvitz E. Major life changes and behavioral markers in social media: case of childbirth. In: Proceedings of the 2013 conference on Computer supported cooperative work. ACM; 2013. 1431–42.
- Golbeck J, Robles C, Turner K. Predicting personality with social media. In: CHI'11 extended abstracts on human factors in computing systems. ACM; 2011. 253–62.
- Marcus B, Machilek F, Schütz A. Personality in cyberspace: personal Web sites as media for personality expressions and impressions. J Pers Soc Psychol. 2006;90(6):1014–31.
- Bachrach Y, Kosinski M, Graepel T, et al. Personality and patterns of Facebook usage. In: Proceedings of the 3rd Annual ACM Web Science Conference. ACM; 2012. 24–32.
- Moreno MA, Jelenchick LA, Egan KG, et al. Feeling bad on Facebook: depression disclosures by college students on a social networking site. Depress Anxiety. 2011;28(6):447–55.
- 71. •• Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci U S A. 2013;110(15):5802–5. Example of how a range of sensitive personal attributes can be predicted from Facebook Likes
- Martin EA, Bailey DH, Cicero DC, et al. Social networking profile correlates of schizotypy. Psychiatry Res. 2012;200(2–3):641–6.
- Duhigg C. How companies learn your secrets. New York Times 2, 16,2012. http://www.nytimes.com/2012/02/19/magazine/ shopping-habits.html?ref=general&src=me&pagewanted=all&_ r=0. Accessed 31 May 2014.
- Crawford K, Schultz J. Big data and due process: toward a framework to redress predictive privacy harms. Boston Coll Law Rev. 2014. http://bclawreview.org/files/2014/01/03_crawford_ schultz.pdf.
- Terry N. Protecting patient privacy in the age of big data. Univ Missouri-Kansas City Law Rev. 2012;81(2). http://ssrn.com/ abstract=2153269.
- Walker J. Data mining to recruit sick people. Wall Str J. 2013. http://online.wsj.com/news/articles/ SB10001424052702303722104579240140554518458. Accessed 31 May 2014.
- Lupton D. The commodification of patient opinion: the digital patient experience economy in the age of big data. Sociol Health Illn. 2014. doi:10.1111/1467-9566.12109.
- Li J. Privacy policies for health social networking sites. J Am Med Inform Assoc. 2013;20(4):704–7.
- Weigmann K. Health research 2.0: the use in research of personal fitness or health data shared on social network raises both scientific and ethical concerns. EMBO Rep. 2014;15(3):223–6.
- Williams J. Social networking applications in health care: threats to the privacy and security of health information. In: Proceedings of the 2010 ICSE Workshop on Software Engineering in Health Care. ACM; 2010. 39

 –49.
- Ball DE, Tisocki K, Herxheimer A. Advertising and disclosure of funding on patient organisation websites: a cross-sectional survey. BMC Public Health. 2006;6:201.
- Treato. Treato: patient intelligence based on real-life experiences. http://treato.com/about/. Accessed 31 May 2014.
- 83. Tene O, Polonetsky J. Privacy in the age of big data: a time for big decisions. Stanf Law Rev Online. 2012;64:63.
- Center for Democracy and Technology. Rethinking the role of consent in protecting health information privacy. 2009. https:// www.cdt.org/files/pdfs/20090126Consent.pdf. Accessed 31 May 2014
- Turow J, Hoofnagle CJ, Mulligan DK, et al. The Federal Trade Commission and consumer privacy in the coming decade. ISJLP.



- 2007;3:723. http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1934&context=facpubs. Accessed 31 May 2014.
- Determann L. Social media privacy: a dozen myths and facts. Stan Tech L Rev. 2012. 7–10. http://stlr.stanford.edu/2012/07/social-media-privacy/. Accessed 31 May 2014.
- Brandimarte L, Acquisti A, Loewenstein G. Misplaced confidences privacy and the control paradox. Soc Psychol Personal Sci. 2013;4(3):340–7.
- El Emam K, Moher E. Privacy and anonymity challenges when collecting data for public health purposes. J Law Med Ethics. 2013;41 Suppl 1:37–41.
- Savla P, Martino LD. Content analysis of privacy policies for health social networks." IEEE International Symposium on Policies for Distributed Systems and Networks. 2012;94–101.
- Anton A, Earp JB, Vail M, et al. HIPAA's effect on web site privacy policies. IEEE Secur Priv. 2007;45–52.
- Milne GR, Culnan MJ, Greene H. A longitudinal assessment of online privacy notice readability. J Public Policy Mark. 2006;25(2 (Fall)):238–49.
- Graber MA, D'Alessandro DM, Johnson-West J. Reading level of privacy policies on Internet health Web sites. J Fam Pract. 2002;51(7):642-5.
- Ryan C, Siebens J. Educational attainment in the United States: 2009. U.S. Census Bureau. 2012. http://www.census.gov/prod/ 2012pubs/p20-566.pdf.
- Breese P, Burman W. Readability of notice of privacy forms used by major health care institutions. JAMA. 2005;293(13):1593

 –4.
- Gralton E, Sher M, Lopez CD. Information and readability issues for psychiatric patients: e-learning for users. Psychiatr Bull. 2010;34:376–80.
- Goldston DB, Walsh A, Mayfield Amold E, et al. Reading problems, psychiatric disorders, and functional impairment from midto late adolescence. J Am Acad Child Adolesc Psychiatry. 2007;46(1):25–32.
- Federal Trade Commission. Mobile privacy disclosures. Building trust through transparency. 2013. http://www.ftc.gov/sites/default/ files/documents/reports/mobile-privacy-disclosures-buildingtrust-through-transparency-federal-trade-commission-staff-report/ 130201mobileprivacyreport.pdf1.
- Carrión Señor I, Fernández-Alemán JL, Toval A. Are personal health records safe? A review of free web-accessible personal health record privacy policies. J Med Internet Res. 2012;14(4): e114
- Hargittai E. Digital na(t)ives? variation in internet skills and uses among members of the "Net Generation". Sociol Inq. 2010;80:92– 113.
- Park YJ. Digital literacy and privacy behavior online. Commun Res. 2013;40(2):215–36.
- 101. Leon P, Ur B, Shay R, et al. Why Johnny can't opt out: a usability evaluation of tools to limit online behavioral advertising. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM; 2012. 589– 98
- 102. Solove DJ. HIPAA turns 10. J AHIMA. 2013;84(4):22-8.
- Ponemon. Fourth Annual Benchmark Study on Patient Privacy and Data Security. 2014. http://www.ponemon.org/blog/fourthannual-benchmark-study-on-patient-privacy-and-data-security. Accessed 31 May 2014.
- McCann E. HIPAA data breaches climb 138 percent. Healthcare IT News. 2014. http://www.healthcareitnews.com/news/hipaadata-breaches-climb-138-percent.
- 105. US Department of Health and Human Services. Breaches affecting 500 or more individuals. http://www.hhs.gov/ocr/privacy/hipaa/ administrative/breachnotificationrule/breachtool.html.
- Redspin. Breach report 2013: protected health information (PHI).
 2014. http://www.redspin.com/resources/whitepapers-datasheets/

- Request-2013-Breach-Report-Protected-Health-Information-PHI-Redspin.php.
- Johnson ME, Willey ND. Will HITECH heal patient data hemorrhages? In: System Sciences (HICSS), 2011 44th Hawaii International Conference on IEEE. 2011. 1–10.
- Figg WC, Kam HJ. Medical information security. Int J Secur (IJS). 2011;5(1):22.
- KrebsonSecurity. Data broker giants hacked by id theft service.
 2013. http://krebsonsecurity.com/2013/09/data-broker-giants-hacked-by-id-theft-service/. Accessed 31 May 2014.
- Schwartz MJ. Experian breach fallout: ID theft nightmares continue. Inf Week. 2013. http://www.darkreading.com/risk-management/experian-breach-fallout-id-theft-nightmares-continue/d/d-id/1112058? Accessed 31 May 2014.
- Rosencrance L. Acxiom database hacked. Computerworld. 2003. http://www.computerworld.com/s/article/83854/Acxiom_database hacked. Accessed 31 May 2014.
- Federal Trade Commission. Health privacy. http://www.business. ftc.gov/documents/bus56-complying-ftcs-health-breachnotification-rule. Accessed 31 May 2014.
- Harris Interactive. Many U.S. adults are satisfied with use of their personal health information. 2007. http://www.harrisinteractive. com/vault/Harris-Interactive-Poll-Research-Health-Privacy-2007-03.pdf. Accessed 31 May 2014.
- California HealthCare Foundation. Consumers and health information technology: a national survey. 2010. http://www.chcf.org/publications/2010/04/consumers-and-health-information-technology-a-national-survey. Accessed 31 May 2014.
- 115. California HealthCare Foundation. National consumer health privacy survey. 2005. http://www.chcf.org/~/media/ MEDIA%20LIBRARY%20Files/PDF/C/PDF% 20ConsumerPrivacy2005ExecSum.pdf Accessed 31 May 2014.
- Ancker JS, Silver M, Miller MC, et al. Consumer experience with and attitudes toward health information technology: a nationwide survey. J Am Med Inform Assoc. 2013;20(1):152–6.
- National Partnership for Women and Families. Making IT meaningful: how consumers value and trust health IT. 2012. http://go.nationalpartnership.org/site/DocServer/HIT_Making_IT_Meaningful National Partnership February 2.pdf.
- 118. Westin AF. Institute of Medicine project survey findings on health research and privacy. 2007. http://www.iom.edu/~/media/Files/Activity%20Files/Research/HIPAAandResearch/AlanWestinIOMsrvyRept.ashx. Accessed 31 May 2014.
- Employee Benefit Research Institute. Health confidence survey.
 2008. http://www.ebri.org/publications/notes/index.cfm?fa=notesDisp&content id=3987. Accessed 31 May 2014.
- 120. Markle. Survey finds Americans want electronic personal health information to improve own health care. 2006. http://www.markle. org/publications/1214-survey-finds-americans-want-electronicpersonal-health-information-improve-own-hea. Accessed 31 May 2014.
- 121. Deloitte. Survey of U.S. health care consumers: the performance of the health care system and health care reform. 2012. https://www.deloitte.com/view/en_US/us/Industries/US-federal-government/center-for-health-solutions/517f54995c0e7310VgnVCM2000001b56f00aRCRD.htm. Accessed 31 May 2014.
- 122. Agaku IT, Adisa AO, Ayo-Yusuf OA, et al. Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers. J Am Med Inform Assoc. 2014;21(2): 374–8.
- NPR/Kaiser Family Foundation. The public and the health care delivery system. 2009. http://www.npr.org/documents/2009/apr/ nprpoll_topline.pdf. Accessed 31 May 2014.



- Fair Warning. How privacy considerations drive patient decisions and impact patient care outcomes. 2011. http://www.fairwarning. com/whitepapers/2011-09-WP-US-PATIENT-SURVEY.pdf.
- Lowes R. Fear of data theft blunts public acceptance of EHRS. Medscape. 2012. http://www.medscape.com/viewarticle/769778.
- Flynn HA, Marcus SM, Kerber K, et al. Patients' concerns about and perceptions of electronic psychiatric records. Psychiatr Serv. 2003;54(11):1539–41.
- 127. Office of National Coordinator for HIT. Health care providers' role in protecting EHRs: implications for consumer support of EHRs, HIE and patient-provider communication. 2014. http://www.healthit.gov/sites/default/files/022414_hit_attitudesaboutprivacydatabrief.pdf.
- Sankar P, Moran S, Merz JF, et al. Patient perspectives of medical confidentiality: a review of the literature. J Gen Intern Med. 2003;18:659–69.

