

Big Data and the Law

PLSC 597D

Professor Christopher Zorn
Pennsylvania State University
Fall 2013

Contact Information

Dr. Christopher Zorn
Department of Political Science
Pond Lab 226
Pennsylvania State University
University Park, PA 16803
E-mail: zorn@psu.edu
Phone: 814-863-9446
Phone (alternate): 803-553-4077

Course Time: Wednesday, 2:00-5:00 p.m. ET
Course Location: Ford Hall 202
Office Hours: By appointment

Course Description

This graduate-level course examines the social, political, and legal implications of 'big data.' Topics will include questions of agency and consent regarding personal information, privacy and confidentiality, and ethical issues raised by ubiquitous, real-time digital information. We will also discuss legal and judicial responses to big data, including its implications for standards of proof and liability, rules of evidence, criminal procedure, and intellectual property. Finally, we'll consider the impact of advances in data and analytics on the operation of law enforcement, the judicial system, and the legal services industry.

Items listed under "Materials" are effectively required reading for the course; those under "Of Interest" can be thought of as optional.

Grading, Etc.

Grading will be based on a collaborative class project. More details will be provided during class.

Academic Dishonesty

The Department of Political Science, along with the College of the Liberal Arts and the University, takes violations of academic dishonesty seriously. Observing basic honesty in one's work, words, ideas, and actions is a principle to which all members of the community are required to subscribe.

All course work by students is to be done on an individual basis unless an instructor clearly states that an alternative is acceptable. Any reference materials used in the preparation of any assignment must be explicitly cited. Students uncertain about proper citation are responsible for checking with their instructor.

In an examination setting, unless the instructor gives explicit prior instructions to the contrary, whether the examination is in class or take home, violations of academic integrity shall consist but are not limited to any attempt to receive assistance from written or printed aids, or from any person or papers or electronic devices, or of any attempt to give assistance, whether the one so doing has completed his or her own work or not.

Lying to the instructor or purposely misleading any Penn State administrator shall also constitute a violation of academic integrity.

In cases of any violation of academic integrity it is the policy of the Department of Political Science to follow procedures established by the College of the Liberal Arts. More information on academic integrity and procedures followed for violation can be found [here](#). More generally, if you are unfamiliar with general standards regarding plagiarism, **learn them**; a good place to start is [here](#). The professor is not responsible for students' lack of understanding of standards regarding academic misconduct.

Statement on Disabilities

Penn State welcomes students with disabilities into the University's educational programs. If you have a disability-related need for reasonable academic adjustments in this course, contact the Office for Disability Services (ODS) at 814-863-1807 (V/TTY). For further information regarding ODS, please visit their web site [here](#). Instructors should be notified as early in the semester as possible regarding the need for reasonable accommodations.

Course Schedule

August 28 – Introduction: What Are We Talking About?

Materials (read for background):

- Mike Dewar (Data Scientist, NYT Data Lab), “Big Data: Understand and Visualize.” Presentation at 2013 Media Evolution conference.
- Doctor Decade, “Big Data” (warning: seizure risk).

September 4 – Foundations: “Big Data”

Materials:

- Bush, Vannevar. 1967. “Memex Revisited.” In William Montagna, *Science is Not Enough*.
- McKinsey & Co. 2011. *Big Data: The Next Frontier for Innovation, Competition, and Productivity – Executive Summary* (read this summary; the full 2011 McKinsey report is [here](#)).
- Anderson, Christopher. 2008. “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete” (the “end of theory” paper). *Wired*.
- Smith, Mark. 2013. “Don’t Worry: Big Data Hasn’t Jumped the Shark Yet.” *Wired*.
- boyd, dana, and Kate Crawford. 2012. “Critical Questions for Big Data. *Information, Communication & Society* 15:662-79.

Of Interest:

- Schradie, Jen. 2012. “The Trend of Class, Race and Ethnicity in Social Media Inequality: Who Still Can’t Afford to Blog?” *Information, Communication and Society* 15(April):555-71.

September 11 – Foundations: “The Law” (Privacy)

Materials:

- Warren, Samuel D. and Louis D. Brandeis. 1890. “The Right to Privacy.” *Harvard Law Review* 4:193-220.
- *NAACP v. Alabama* 357 U.S. 449 (1958).
- *Katz v. U.S.* 389 U.S. 347 (1961) (majority opinion).
- *United States v. Miller* 425 U.S. 435 (1976) (majority opinion).
- Electronic Privacy Information Center. 2013. “The Privacy Act of 1974” (summary).
- Solove, Daniel A. 2006. “A Taxonomy of Privacy.” *University of Pennsylvania Law Review* 154:477-560.
- Facebook. 2013. “Statement of Rights and Responsibilities” (Terms of Service).

Of Interest:

- *Olmstead v. U.S.* 277 U.S. 438 (1921).
- Prosser, William L. 1960. “Privacy.” *University of California Law Review* 48:383-423.
- *Griswold v. Connecticut* 381 U.S. 479 (1965).
- Westin, Alan F. 1966. “Science, Privacy, and Freedom: Issues and Proposals for the 1970s.” *Columbia Law Review* 66:1003-1050
- Gormley, Kenneth. 1992. “100 Years of Privacy.” *University of Wisconsin Law Review* 1335-.
- Instagram. 2013. [Terms of Use](#).
- [Privacy International](#) (NGO).

September 18 – Foundations: “The Law” (Evidence, Intellectual Property, and the Legal Profession)

Materials:

Evidence

- Dicarlo, Vincent. 2008. “Summary of the Rules of Evidence.” *FindLaw* (ignore the parts about California).
- [Federal Rules of Evidence](#), Articles I, VII, IX, and X (skim Notes).
- *Daubert v. Merrell Dow Pharmaceuticals* 509 U.S. 579 (1993).

IP

- Gordon, Wendy J. 2003. “Intellectual Property.” In the *Oxford Handbook of Legal Studies*, pp. 617-646.
- Resnick, D. B. 2003. “A Pluralistic Account of Intellectual Property.” *Journal of Business Ethics* 46:319-335.

The Legal Profession

- Henderson, William D. 2011. “Three Generations of U.S. Lawyers: Generalists, Specialists, Project Managers.” *Maryland Law Review* 70(1).

Of Interest:

- *Kumho Tire Co. Ltd. v. Carmichael* 526 U.S. 137 (1999) (applicability of *Daubert* to “non-scientists.”)
- Schwartz, David. 2011. “A Foundation Theory of Evidence.” *Georgetown Law Journal* 100:95-171.
- “Intellectual Property Law and Policy” page at the USPTO.
- Menell, Peter S. and Suzanne Scotchmer. 2005. “Intellectual Property.” In the *Handbook of Law and Economics*.
- *Mimi and Eunice*, a web comic about intellectual property.
- Galanter, Marc, and William D. Henderson. 2008. “The Elastic Tournament: The Second Transformation of the Big Law Firm.” *Stanford Law Review* 60.

September 25 – First Amendment Matters

Materials:

- [Sorrell v. IMS Health, Inc.](#) 131 S.Ct. 2653 (2011) (read the dissent, too).
- Bambauer, Jane R. 2013. “Is Data Speech?” *Stanford Law Review*, forthcoming.
- Benjamin, Stuart Minor. 2013. “Algorithms and Speech.” *University of Pennsylvania Law Review* 161:1445-1493.

Of Interest:

- “This Case Matters For You.” (IMS Health’s PR site for the *Sorrell* litigation.)
- Pasquale, Frank. 2011. “Rethinking *IMS Health v. Sorrell*: Privacy as a First Amendment Value.” *The Health Care Blog*, April 29, 2011.
- Ghosh, Shubha. 2012. “Informing and Reforming the Marketplace of Ideas.” *Utah Law Review* 2012:653-706.
- Wu, Tim. 2013. “Machine Speech.” *University of Pennsylvania Law Review* 161:1495-1533.
- Cherry, Miriam A., and Robert L. Rogers. 2008. “Prediction Markets And The First Amendment.” *University of Illinois Law Review* 2008(3).
- Tutt, Andrew. 2012. “Software Speech.” *Stanford Law Review Online* 65:73.
- Volokh, Eugene. 2013. “‘Liking’ a Facebook Page Is Presumptively Protected Speech.” *The Volokh Conspiracy*, September 18, 2013.

October 2 – Privacy: Data Collection

Materials:

- Federal Trade Commission. 2013. “Fair Information Practice Principles.”
- Ohm, Paul. 2012. “Don’t Build A Database Of Ruin.” *Harvard Business Review Blog*, August 23, 2012.
- Richards, Neil M. 2013. “The Dangers of Surveillance.” *Harvard Law Review* 126:1934-1965.

- Tene, Omer, and Jules Polonetsky. 2012. "Privacy in the Age of Big Data: A Time for Big Decisions." *Stanford Law Review Online: Perspectives*, February 2, 2012. (Note that this is a summary version of Tene and Polonetsky's "Big Data for All: Privacy and User Control in the Age of Analytics," *Northwestern Journal of Technology & Intellectual Property* 11:239-273.)

Of Interest:

- *Clapper v. Amnesty International USA* 133 S. Ct. 1138 (2013).
- Hatcher, Jordan S. and Lilian Edwards. 2009. "Consumer Privacy Law 2: Data Collection, Profiling and Targeting." In L. Edwards & C. Waelde, eds., *Law and the Internet*. London: Hart Publishing.
- Citron, Danielle Keats, and David Gray. 2013. "Addressing the Harm of Total Surveillance: A Reply to Professor Neil Richards." *Harvard Law Review* 126: Online Forum.
- Hu, Margaret. 2013. "Biometric ID Cybersurveillance." *Indiana Law Journal* 88:1475-1558.
- Takahashi, Timothy T. 2013. "Drones and Privacy." *Columbia Science and Technology Law Review* 14: 73.
- Bamford, James. 2012. "The Black Box: Inside America's Massive New Surveillance Centre." *Wired*, March 30, 2012.
- <http://knowprivacy.org/> (a bit dated, but interesting).
- Electronic Frontier Foundation. 2013. *Do Not Track*.
- The Surveillance Self-Defense Project at EFF.

October 9 – No class.

No Readings Assigned

October 16 – Discussion of Class Project.

October 23 – Privacy: Data Processing/Analysis

Materials:

- Froomkin, A. Michael. 2009. “Anonymity and the Law in the United States.” In *Lessons From The Identity Trail: Anonymity, Privacy And Identity In A Networked Society*. New York: Oxford University Press.
- Yakowitz (Bambauer), Jane R. 2011. “Tragedy of the Data Commons.” *Harvard Journal of Law and Technology* 25:1-67.
- Spiekermann, Sarah, and Lorrie Faith Cranor. 2009. “Engineering Privacy.” *IEEE Transactions on Software Engineering* 35:67-82.

Of Interest:

- Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. “Private Traits and Attributes Are Predictable from Digital Records of Human Behavior.” *Proceedings of the National Academies of Science* 110(15):5802-5805.
- Kifer, Daniel, and Ashwin Machanavajjhala. 2011. “No Free Lunch in Data Privacy.” *Proceedings of ACM SIGMOD International Conference on Management of Data*.
- Narayanan, Arvind, and Vitaly Shmatikov. 2008. “Robust De-anonymization of Large Datasets (How to Break Anonymity of the Netix Prize Dataset).” *IEEE S&P*.
- McSherry, Frank, and Kunal Talwar. 2008. “Mechanism Design via Differential Privacy.” *FOCS 2008*.
- Tucker, Patrick. 2013. “Has Big Data Made Anonymity Impossible?” *MIT Technology Review* blog, May 7, 2013.
- Gardner, James A. 2011. “Anonymity and Democratic Citizenship.” *William & Mary Bill of Rights Law Journal* 19:927.

October 30 – Privacy: Data Dissemination

Materials:

- Blitz, Marc Jonathan. 2012. "The Right To Map (And Avoid Being Mapped): Reconciling First Amendment Protection For Information-Gathering In The Age Of Google Earth." *The Columbia Science & Technology Law Review* 14:116-202.
- Walker, Robert Kirk. 2012. "The Right To Be Forgotten." *Hastings Law Journal* 64:257-286.
- Lagone, Laura. 2012. "The Right to Be Forgotten: A Comparative Analysis." Typescript: Fordham University School of Law.

Of Interest:

- Fleischer, Peter. 2007. "The Need For Global Privacy Standards." *Peter Fleischer: Privacy...?* blog, September 14, 2007.
- European Commission. 2012. *Proposal for a Regulation Of The European Parliament And Of The Council On The Protection Of Individuals With Regard To The Processing Of Personal Data And On The Free Movement Of Such Data*. European Union. (long)
- Hendel, John. 2012. "Why Journalists Shouldn't Fear Europe's 'Right to be Forgotten.'" *The Atlantic*, January 25, 2012.
- Koops, Bert-Japs. 2011. "Forgetting Footprints, Shunning Shadows: A Critical Analysis of the 'Right to Be Forgotten' in Big Data Practice." *SCRIPTed* 8:229-256.
- Bennett, Stephen C. 2012. "The 'Right to Be Forgotten': Reconciling EU and U.S. Perspectives." *Berkeley Journal of International Law* 30:161-195.
- Beckles, Cherri-Ann. 2013. "Will the Right To Be Forgotten Lead to a Society That Was Forgotten?" *Privacy Perspectives*, May 14, 2013.
- Waldman, Katy. 2013. "California's Internet Eraser Law: Nice Idea, But It Won't Work." *Slate*, September 25, 2013.

November 6 – Challenges of Medical / Bio / Genetic Data

Materials:

- Hoffman, Sharona, and Andy Podgurski. 2013. "Big Bad Data: Law, Public Health, and Biomedical Databases." *The Journal of Law, Medicine & Ethics* 41(S1):56-60.
- Terry, Nicholas. 2012. "Protecting Patient Privacy in the Age of Big Data." *UMKC Law Review* 81(2).
- Kaye, Jane. 2012. "The Tension Between Data Sharing and the Protection of Privacy in Genomics Research." *Annual Review of Genomics and Human Genetics* 13:415-431.

Of Interest:

- National Institutes of Health. 2004. *Protecting Personal Health Information in Research: Understanding the HIPAA Privacy Rule*. U.S. Department of Health and Human Services, rev. 7/13/2004.
- National Institutes of Health. 2004. *Research Repositories, Databases, and the HIPAA Privacy Rule*. U.S. Department of Health and Human Services, rev. 7/2/2004.
- Office of Civil Rights, Department of Health and Human Services. 2012. "Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule." November 26, 2012.
- Anderlik, Mary R., and Mark A. Rothstein. 2001. "Privacy And Confidentiality Of Genetic Information: What Rules For The New Science?" *Annual Review of Genomics and Human Genetics* 2:401-433.
- Sankar, Pamela. 2003. "Genetic Privacy." *Annual Review of Medicine* 54: 393-407.
- Hoffman, Sharona, and Andy Podgurski. 2013. "The Use and Misuse of Biomedical Data: Is Bigger Really Better?" *American Journal of Law & Medicine* 39:forthcoming.
- Open Source Electronic Health Record Agent.
- 23AndMe.

November 13 – Big Data and the Legal Services Industry

Materials:

- Murphy, Tonia Hap. 2013. “Mandating Use of Predictive Coding in Electronic Discovery: An Ill-Advised Judicial Intrusion.” *American Business Law Journal* 50(Fall): forthcoming.
- Katz, Daniel Martin. 2013. “Quantitative Legal Prediction or How I Learned to Stop Worrying and Start Preparing for the Data Driven Future of the Legal Services Industry.” *Emory Law Journal* 62:909-966.
- Harbert, Tam. 2013, “Supercharging Patent Lawyers With AI.” *IEEE Spectrum* blog (article on [Lex Machina](#) – check out the website as well).

Of Interest:

- Oard, Douglas W., and William Webber. 2013. “Information Retrieval for E-Discovery.” *Foundations and Trends in Information Retrieval* 7:99-237. See also the “Development and Evaluation of Search Technology for Discovery of Evidence in Civil Litigation” project at the University of Maryland.
- Bassett, Debra Lyn. 2009. “E-Pitfalls: Ethics and E-Discovery.” *Northern Kentucky Law Review* 36:449-483.
- Ward, Burke T., Janice C. Sipior, Jamie Patrick Hopkins, Carolyn Purwin, and Linda Volonino. 2012. “Electronic Discovery: Rules for a Digital Age.” *Boston University Journal of Science and Technology Law* 18.
- “E-Discovery Rules.” *E-Discovery Resource Database*.
- Computational Legal Studies blog.
- *Legal Language Explorer*.
- The ReinventLaw Laboratory.
- TyMetrix and the RateDriver app.

November 20 – Big Data and Intellectual Property

November 27 – No Class – Fall Break.

No Readings Assigned

December 4 – Big Data, Courts, and the Judicial System

December 11 – Big Data: Social Implications