Johannes Kepler University Linz
Enterprise Business Intelligence: Bid Data and Privacy
Spring 2015

Instructor: Professor Dr. Noushin Ashrafi

I. Course Materials

Big Data and Privacy: Theoretical

- Privacy and Big Data: The players, Regulators, and Stakeholders'
  Terence Craig andMary E. Ludloff, O'Reilly, 2011, ISBN: 978-1-449-30500-0 (Required)
- Privacy, Big Data, and the Public Good: Frameworks for Engagement
  Edited by Julia Lne, Vitoria Stodden, Stefan Bender, and Helen Niessenbaum, Cambridge University Press, 2014, ISBN: 9781-107-06735-6. (Required)
- Privacy in Context: Technology, Policy, and the integrity of Social Life,
  Helen Nissenbaum, Stanford University, 2010, ISBN 978-0-8047-5236-7 (cloth : alk. paper)—ISBN 978-0-8047-5237-4 (pbk. :alk. paper). (Optional).
- Reading materials: key articles on Big Data and Privacy issues provided and are part of the course requirements

Big Data and Privacy: Practical/Technology (all application software will be provided free of charge to students (on the condition that appropriate lab infrastructure is available.)

- Hadoop for Dummies: Special addition' Robert D. Schnider, Jhon Wiley and Sons, Inc. 2012, ISBN: 978-1-118-25051-8 (Optional)
- IBM InfoSphere BigInsights : Hadoop, MapReduce
- Privacy Invasive Software (PIS): Cookies, Web Bug, Spyware
- Privacy Enhancing Technology (PET): Data Security, 256-bit AES Encryption, Data Reliability, File Shredding, Password Meter, Password Generator, Virtual Keyboard , History Cleaning, Stealth Mode, Manual Wipe.
- Tableau desktop 8.2 (Training videos and exercises).

II. Background

In a digital setting, such as the Internet, there are a wide variety of privacy threats, ranging from tracking user online activities, to mass marketing based on the retrieval of personal information, to the distribution of information on dangerous technologies used for, e.g., acts of terror. In the last few decades numerous privacy-enhancing technologies have been developed with mixed results- some have been successful while others have seen little adoption despite much hype and promise. We will study privacy technologies, their uses and limitations, the reasons for their success and failure, and think critically about their place in society. More broadly, we will try to understand the Privacy Paradox: Conflicting values in the age of Big Data. There is a call for restricting the flow of data or ensuring that the flow of the information on the Internet is managed appropriately. This is easier said than done because the subject of privacy in relation to information technology is deeply problematic; its definition, benefits, harms, and its conceptual morass have been debated and are notoriously controversial[i]. Free flow of data on the Internet diminishes control over personal information hence provoking anxiety and resistance to the concept of big data. But, big data also promotes intellectual development, health and well-being, and democracy- should it be feared and avoided or accepted and even celebrated? Let's find put.

III. Course Description and Learning Objectives

The course covers the concept of "big data" and its management through technology. It elucidates that ensuring protecting personal privacy becomes harder as information is multiplied and shared ever more widely around the world. Information regarding individuals' health, location, and online activity is exposed to scrutiny, raising concerns about profiling, discrimination, exclusion, and loss of control. The course is designed for graduate students and upon successful completion of this course students should be adept in the following areas:

- On the conceptual side:
  - Understand big data and its role as an instrument to guide business operations towards achieving strategic objectives,
  - Study the cost-benefit analysis of big data as it is utilized to disseminate knowledge and strengthen social cohesion while raising concerns regarding the extent to which ''user privacy' can be ensured'
  - Examine the state of privacy protection via technology and policies in Austria.
- On the technical side
  - Study Privacy Invasive Technologies (PIT) and Privacy Enhancing Technologies (PET)
  - Students will get a chance to get certified in Hadoop (Big Data Application) and fairly competent in using Tableau Software (BI application)
  - Time permitting, we will study MapReduced.

IV Course Format

The course will consist of lecture series, class discussions, Lab exercises, group assignments, and a final submission that could be a technology based project or a term paper. To simulate the real world work environment, students will work in groups and make class presentations on the topics assigned by the instructor. The course is highly interactive; students are required to study the course materials

provided by the instructor in advance and contribute to class discussions and presentations in the class. Lab exercises will be assigned and students will work on the existing Privacy Enhancing Technologies (PET) and Privacy Invasive Technologies (PIT) and their implementations. The extent of PET deployment as well as bylaws regarding data privacy protection in Austria will be examined. It is imperative that students come to class prepared: reading before each class and be prepared for discussing the revenant topics. Hand-outs will be given for lab exercises and students are expected to have studies them in advance in order to be ready to complete and submit their work at the end of each session- there would exceptions when students can complete lab exercises afterwards. Please note the syllabus materials are open to negotiation- since I knew nothing about the students, we will spend the first class to come up with work load, submission criteria, etc, that is doable, challenging, and hopefully enjoyable for us as a whole.

V. Evaluation

Item	Weight
Participation in class Discussion	15%
Applications Exercises	20%
Group assignments	30%
Group project or Term Paper	20%
Peer evaluation	15%
Total	100%

The weight and distribution structure may changes based on student performance and the class dynamics–So, expect changes to be announced around mid-semester.

Percentage	**performance**	**Grade**
91%-100%	**Very good**	1
81%-90%	Good	2
71%-80%	Satisfactory	3
61%-70%	Sufficient	4
<60%	fail	5

VI Weekly Schedule

Please note this schedule is just for the start-every semester I customize both the content and timing of assignments as the courses progresses and I get to know your strengths and weaknesses.

Week	**Readings**	**Topic**	**Deliverables/ assignments**
1		Syllabus, text books, rules of engagement. Efforts to evaluate students' command of English to determine adequate pedagogy and requirements of the course. Groups are formed.	Students and instructor introduce themselves and an agreement to be reached for course expectations, timing, and deliverables Group Assignment: Research German language literature to find an article related to 'privacy in digital age'. Write an abstract of no more than 150 words in English.
2	Introduction: Why Privacy Matters Even if You Have 'Nothing to Hide'	Why are we here? Who cares about privacy? What might the future look like? Lab: Learn to use BigData University	Groups make presentations on their assignment of week 1. Class Discussion on topic for week 2.
3	Text: Privacy and Big Data: Chapter 1&2. Watch video and/or read Transcript of Holistic BigData- parts 1,2&3	The perfect Storm & The right to privacy in the digital age. Discuss Holistic BigData	Class Discussion on topic for week 3 Groups will present examples of PIT
4	Text: Privacy and Big Data: Chap 3. Watch video and/or read Transcript of Holistic BigData- parts 4,5&6	The Regulators; US and EU. Discuss Holistic BigData	Class Discussion on topic for week 4 Groups will present examples of PET

5	Text: Privacy and Big Data: Chapter 4. Introduction to Hadoop	The Players. Lab exercise: Downloading IBM InfoSphere BigInsights Quick Start Edition, v2.1.2 VM Image (unit 1)	Class Discussion on topic for week 5 Group work: Research German language literature to find an article related to 'EU privacy regulation'. Write an abstract of no more than 150 words in English.
6	Text: Privacy and Big Data: Chapter 5. Video or transcript: Hadoop Architecture	Commodity Versus Right Lab: Hadoop Exercises (unit 2):	Groups make presentations on their assignment of week 5. Complete Lab exercise
7	Text: Privacy, Big Data And The Public Good- chap.6 Video or transcript: Hadoop Administration	A European Perspective on Research and Big Data Analysis. Lab: Hadoop Exercises. (unit 3)	Group work: Research German language literature to find an article related to 'Privacy, Big Data And The Public Good.' Write an abstract of no more than 150 words in English. Complete Lab exercise
8	Text: Privacy, Big Data And The Public Good Video or transcript: Hadoop Components IBM InfoSphere Data Privacy for Hadoop	Changing the Rules: General Principles for Data Use and Analysis Lab: Hadoop Exercises (unit 4)	Groups make presentations on their abstract for week 7 Complete Lab exercise
9	Text: Privacy, Big Data And The Public Good Data privacy protection for Hadoop	Extracting Information from Big Data: Issues of Measurement, Inference and Linkage Lab: Hadoop certificate	Group Project initiation Proposal for term paper Complete Lab exercise
10	Text: Privacy, Big Data And The Public Good Data Privacy Protection For Hadoop: Best Practices	Using Statistics to Protect Privacy Lab: Tableau Lesson 1	Work on Group Project Complete Lab exercise
11	Text: Privacy, Big Data And The Public Good Securing Hadoop	Differential Privacy: A Cryptographic Approach to Private Data Analysis Lab: Tableau Lesson 2	Class Discussion Complete Lab exercise
12	Article 1: Big Data: The End of Privacy or a New Beginning?	TBA Tableau Lesson 3	Class Discussion Complete Lab exercise
13	Article 2: it's not privacy, And it's not fair	TBA Tableau Lesson 4	Class Discussion Complete Lab Exercises
14	Article 3: prediction, preemption, Presumption: how big data Threatens big picture Privacy	TBA Lab: Work on group project	Class Discussion Complete Lab Exercises
15	Project Progress report		Presentation

VII Suggested Reading

1. Armstrong, J (2004). Privacy in Europe: the New Agenda, *Journal of Internet Law* Nov. 1st.
2. Ashrafi, N. and Kuilboer, J-P. (2005). Online Privacy Policies: An Empirical Perspective on Self-Regulatory Practices. *Journal of E-Commerce in Organizations*, 3, 4, 61-74.
3. Ashrafi, N. and Kuilboer, J-P. (2007a). Implementation of Privacy Protection Policies: An Empirical Perspective. Utilizing and Managing Commerce and Services Online. CyberTech Publishing, 2007, Ch. IX: 187-204.
4. Ashrafi, N. and Kuilboer, J-P. (2007b). Is P3P an Answer to Protecting Information Privacy? E-Business Innovation and Process Management. CyberTech Publishing, 2007, Ch XV: 331-347.
5. Ashrafi, N. and Kuilboer J.P.(2004) "Platform for Privacy Preference: An Innovative Technology and Standard in E-Commerce." International Conference of the Information resources Management Association, IRMA 2004 , New Orleans 23-26, Louisiana, May 2004.
6. Ashrafi, N. and Kuilboer J.P. "Data Privacy, US Common Practices"13th International Conference on Database and Expert Systems Applications, The computer society (IEEE) DEXA 2002, Aix en Provence, pp. 488-492
7. Ashrafi, N. and Kuilboer J.P. (2000) "Managing Network Security" Proceedings of 12th International Conference of the Information resources Management Association, , May 20-24, 2000, Toronto, Canada. pp. 122-125.
8. Atkinson, R.D. and McKay, A.S. (2007) Digital Prosperity: understanding the economic benefits of the information technology revolution. The Information Technology & Innovation Foundation.
9. Atkinson, R.D. (2007) Boosting European Prosperity through the Widespread Use of ICT. The Information Technology & Innovation Foundation. Nov. 2007.
10. Atkinson R.D. and Castro D.D. (2008) Digital Quality of Life: understanding the personal & social benefits of the information technology revolution. The Information Technology & Innovation Foundation.
11. Clarke, R. (1999). Internet Privacy Concerns Confirm the Case for Intervention. *Communications of the ACM*, 42, 2, 60-67.

12. Culnan, M.J. and Armstrong, P.K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation, *Organization Science*, 10, 1, 104-115.

13. Culnan, M.J and Bies, R.J. (2003). Consumer Privacy: Balancing Economic and Justice Considerations. *Journal of Social Issues*, 59, 2, 323-342.

14. Committee on Consumer Policy (CCP). 2003. *Consumers in the Online Marketplace: The OECD Guidelines Three Years Later*. February. Retrieved December 12, 2003 from www.olis.oecd.org/olis/2002doc.nsf/LinkTo/dsti-cp(2002)4-final.

15. Desai, P., Ashrafi, N., Kuilboer, J.P., Koehler, W., With "Regulatory Privacy Practices in Europe," American Conference on Information Systems, AMCIS, 2009, August 10-13, San Francisco, CA

16. European Commission (2003). *Data Protection: Microsoft Agrees to Change its .NET Passport System after Discussions with EU Watchdog*. IP/03/151. Retrieved February 21st, 2009 from http://europa.eu/rapid/pressReleasesAction.do?reference=IP/03/151&format=HTML&aged=0&language=EN&guiLanguage=en.

17. European Parliament, (1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of individuals with Regard to the Processing of Personal Data and on the Free Movement of such Data, Retrieved January 12th, 2009 from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:NOT.

18. Fuchs, Christian (2013), Privacy and Security in Europe, The Privacy & Security Research Series, Issue #6, http://www.projectpact.eu.

19. Hearst, M.A. (1999), The Changing Relationship Between Information Technology and Society. IEEE Intelligent Systems, Jan-Feb 1999, pp. 5-17.

20. King, I. (2003). Online Privacy in Europe: New Regulation for Cookies. *Information and Communication Technology Law*, 12, 3: 225-236.

21. Kuner C. and Simpson, A. P. (2005). Managing Privacy Enforcement Risks in Europe. *Risk Management*, February, 1st: 42-44.

22. Laufer, R. S. and Wolfe, M. (1977). Privacy as a Concept and a Social Issue: A Multidimensional Development Theory. *Journal of Social Issues*, 33, 22-42.

23. Massa-Mias G., Ashrafi, N., Koehler W., and Kuilboer J-P. (2007). Privacy Policy Regulations- An Empirical Investigation. 38th Annual Meeting Decision Sciences Institute, November 17-20, Phoenix, AZ.

24. Milne, G. R. and Culnan, M. J. (2004). Strategies for Reducing Online Privacy Risks. Why Consumers Read (or Don't Read) Online Privacy Notices. *Journal of Interactive Marketing*, 18, 3, 15-29.

25. O'Brien, K. J. and Cramton, T. (2007). Berlin– European Union Warns Google on Possible Violations of Privacy Law. *New York Times*, May 26th: C3.

26. Organization for Electronic Co-operation and Development (1980). *Guidelines on the Protection of Privacy and Trans-border Flows of Personal Data*. Washington DC: OECD.

27. Warren, S. D. and Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review*, IV (5):193-220.

28. Westin, A. F. (1967). *Privacy and Freedom*. New York: Athenaeum.

[i] Nissenbaum (2009)