

PJ Lamberson p-lamberson@kellogg.northwestern.edu
Office hours: TBA, Chambers Hall (600 Foster St.)

Description

Social Dynamics and Network Analytics (Social-DNA) covers cutting edge research on social media, network analytics, big data, and crowdsourcing, and provides you with the tools to practically apply this research in your own career. By the end of the course you will know how to: measure volume and location of Internet search data to understand and forecast trends; measure volume and sentiment of Twitter conversations; collect network data and create meaningful network visualizations; use the wisdom of crowds, including setting up a prediction market, to create better forecasts; and use Amazon Mechanical Turk for crowdsourcing. *The course counts towards a major in Entrepreneurship and Management and Organizations.*

Grading

Assignments (50%) A series of individual and group assignments will give you experience applying the tools from class. *All assignments are due by Midnight on the day indicated. Late assignments will be penalized 20%. Assignments more than 24 hours late will not be accepted. This policy will be strictly enforced, and absolutely no exceptions will be granted.*

Final exam (50%) A final exam will test your cumulative understanding of the course material. The final exam will include both a take-home and an in-class component. The in-class exam will be held using student affairs proctoring option. The take-home exam will be distributed on the last Friday of class (June 6) and is due Tuesday, June 10 at Midnight.

Readings

All readings are available on Blackboard. There is no course packet or textbook for the course. Readings marked as (Reference) are optional and repeat material that will be covered in class. They are provided for your reference if you need to go over something again. Readings marked as (Advanced) are more difficult and are provided for those interested in exploring topics in more depth. They are strictly optional.

Honor Code

As with all Kellogg courses, by enrolling in this course you agree to abide by the Kellogg Honor Code (http://www.kellogg.northwestern.edu/stu_aff/policies/honorcode.htm). In this course you may (and are encouraged to) discuss both the individual assignments and group assignment with your fellow students; however, the finished product that you submit should be entirely your own work.

Schedule

Warning: This schedule is subject to change

Week 1 — Social Dynamics

Session 1. Unpredictability and Inequality in a Connected World.

How social influence creates unpredictable successes, catastrophic failures, and radical transformations. How did Susan Boyle have the sixth bestselling album of 2010? The Music Lab experiment. Market inequality and unpredictability.

Samuel J. Palmisano, “A Note to Fellow CEOs,” *Capitalizing on Complexity: Insights from the Global Chief Executive Officer Study*, IBM, 2010.

Robert M. May, Simon A. Levin, and George Sugihara (2008) “Complex systems: Ecology for bankers,” *Nature* 451, 893-895.

Session 2. Superstars.

Power law distributions. Network effects and winner-take-all markets. Information cascades.

W. B. Arthur, “Increasing Returns and the New World of Business,” *Harvard Business Review*, 1996.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch, “Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades,” *Journal of Economic Perspectives*, 12(3): 151—170, 1998. (Advanced)

Week 2 — Predicting the Present

Session 1. Predicting the Present and the Measure and React Strategy.

The Billion Prices Project. The “Measure and React” strategy at Zara. Predicting box office success, the DJIA, and election outcomes with Twitter. Sentiment analysis with Amazon Mechanical Turk. Twitterbombs, Astrourfing, and Truthy.

Kurt Kleiner, “Bogus Grass-roots Politics on Twitter,” *Technology Review*, November 2, 2010.

Duncan J. Watts, “A Twenty-first Century Science,” *Nature*, 445: 489, February, 2007.

Session 2. Googling.

Google Flu Trends. Using Google Correlate and Google Insights to identify trends and target markets.

✳ *In Class Activity. Predicting the Present with Google.*

Hal Varian, “Predicting the Present,” *Google Think Quarterly: The People Issue*.

📌 Assignment 1. *Predicting the present with Google. Due Thursday, April 17 at Midnight.*

Week 3 — Going Viral

Session 1. Going Viral.

Why do some things take-off while others don't? Modeling contagion and the viral tipping point. Passive and active viral features. Big seed viral campaigns for subcritical contagions.

Jill Lepore, "It's Spreading," *The New Yorker*, June 1, 2009.

PJ Lamberson, "What it takes to 'Go Viral'"

<http://social-dynamics.org/what-it-takes-to-go-viral/> (Reference)

M. E. J. Newman. *Networks: An Introduction*. Sections 17.1-17.5. (Advanced)

Sinan Aral and Dylan Walker, "Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks," *Management Science*, 2011, 57(9): 1623—1639. (Advanced)

Session 2. Networks.

Modeling social interactions using networks. What is a network? How do we model networks? How can we visualize networks? Network vocabulary: Nodes, links, homophily, bipartite network, degree, degree distribution.

* *In Class Activity. The Degree Distribution of the Class.*

M. E. J. Newman. *Networks: An Introduction*. Chapter 1 and Sections 6.1 and 6.2. (Reference)

Week 4 — Mapping Network Data

Session 1. Mapping Networks.

An introduction to Gephi and mapping networks.

* *In Class Activity. Mapping Your Own Facebook Network.*

PJ Lamberson, "Visualizing Your Own Facebook Network with Gephi,"

<http://social-dynamics.org/gephi-facebook/> (Reference)

PJ Lamberson, "Gephi FAQs (and answers!),"

<http://social-dynamics.org/gephi-faq/> (Reference)

📌 Assignment 2. *Visualizing your Facebook network. Due April 24 at Midnight.*

Session 2. Network Data.

What is network data? What data should you collect? How do you collect it? What do real world networks look like?

* *In Class Activity. Mapping the Social Network of the Class.*

📌 Assignment 3. *Collecting and mapping network data. Part 1 (small data — individual) due May 1 at Midnight. Part 2 (big data — group) due May 8 at Midnight.*

Week 5 — Network Analytics

Session 1. Network Data Analytics.

How do you quantify networks and how do network metrics matter for contagion. Network metrics: path length, clustering, and degree distributions. Watts and Strogatz Small-world networks. The viral tipping point for a general degree distribution. Random networks, giant components, and percolation. Video: Six Degrees of Separation.

* *In Class Activity. Discovering the Fundamentals of Network Contagion.*

Session 2. Identifying Influentials.

Six degrees revisited. The influentials hypothesis. Identifying the most important people in a network. Degree centrality and eigenvector centrality.

* *In Class Activity. Seeding Wars.*

Malcolm Gladwell, “Six Degrees of Lois Weisberg,” *The New Yorker*, January 11, 1996.

David Easley and Jon Kleinberg. *Networks, Crowds, and Markets*, Chapter 14: “Link Analysis and Web Search.” (Advanced/Reference)

Week 6 — Influentials Cont'd and Presentations

Session 1. The Myth of the Influential.

Tracking cascades on Twitter. Influencer ROI. The Fundamental Attribution Error. Referral incentives versus seeding strategies.

Malcolm Gladwell, “Six Degrees of Lois Weisberg,” *The New Yorker*, January 11, 1999.

Clive Thompson, “Is the Tipping Point Toast?” *Fast Company*, January 28, 2008.

Sinan Aral, Lev Muchnik, Arun Sundararajan, “Engineering Social Contagions: Optimal Network Seeding and Incentive Strategies,” <http://ssrn.com/abstract=1770982>.

Session 2. No Class

Week 7 — Complex Contagions and Communities

Session 1. Network Presentations.

A selection of groups will present the networks that they collected in Assignment 3.

Session 2. Complex Contagions.

Threshold contagion, critical mass, and the cascade window.

Damon Centola and Michael Macy, “Complex contagions and the weakness of long ties,” *American Journal of Sociology*, 113(3): 702–34, 2007. (Advanced)

Vladimir Barash, Christopher Cameron, and Michael Macy, “Critical phenomena in complex contagions,” *Social Networks*, 34(4):451-461, 2012. (Advanced)

Week 8 — Network Analytics

Session 1. Network Communities and Social Segmentation.

Modularity and community detection. Identifying target markets and tracking cascades with Twitter and NodeXL.

PJ Lamberson, “Identifying Network Communities.” (Reference)

🔗 Assignment 4: *Identifying influentials and influenceables with Twitter. Due May 26 at Midnight.*

Session 2. Causality in Network Data.

The spread of obesity, smoking, and happiness. Distinguishing selection, affiliation, and social influence.

David Easley and Jon Kleinberg. *Networks, Crowds, and Markets*. Chapter 4. (Reference)

Nicholas Christakis and James Fowler. *Connected*. p. 105—121.

Dave Johns, “Everything is Contagious,” *Slate*, April 7 and April 8, 2010.

Nicholas A. Christakis and James H. Fowler, “The Collective Dynamics of Smoking in a Large Social Network,” *The New England Journal of Medicine*, 358: 2248—2258. (Advanced)

Sinan Aral, Lev Muchnik, and Arun Sundararajan, “Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks.” *PNAS*, 2009, 106(51): 21544–21549. (Advanced)

Week 9 — The Wisdom of Crowds

Session 1. Big Data, Simple Models and The Wisdom of Crowds.

The relative performance of experts and models. Forecaster behavior and trend extrapolation. The averaging principle and the wisdom of crowds. The relative benefits of accuracy and diversity in forecasting (The Diversity Prediction Theorem).

* *In Class Activity. The Wisdom of the Class.*

Robyn M. Dawes, David Faust, and Paul E. Meehl, “Clinical Versus Actuarial Judgment,” *Science*, 1989, 243(4899): 1668—1674. (Advanced)

David Easley and Jon Kleinberg. *Networks, Crowds, and Markets*. Chapter 22. (Reference)

✍ Assignment 5: *Applying the tools to a new product idea. Due June 5 at Midnight.*

Session 2. Prediction Markets.

Prediction markets and how markets reveal and aggregate information. Prediction markets in theory and practice. The Iowa Elections Market, prediction markets at Google and HP. When to use experts, models, or crowds.

Sharad Goel, Daniel M. Reeves, Duncan J. Watts, and David M. Pennock, “Prediction Without Markets,” 2010, *Proceedings of the 11th ACM Conference on Electronic Commerce*, 357—366. (Advanced)

Week 10 — Crowdsourcing and Open Innovation

Session 1. Crowdsourcing.

How to tell if a model was photoshopped. Training computers with crowds. Running experiments with Amazon Mechanical Turk.

Jeff Howe, “The Rise of Crowdsourcing,” *Wired*, June, 2006.

Session 2. Open Innovation.

When and why diverse groups outperform high ability groups. The Netflix Prize. Fold-It.

Eliot Van Buskirk, “How the Netflix Prize Was Won,” *Wired*, September 22, 2009.

Eric Siegel “Casual Rocket Scientists: An Interview with a Layman Leading the Netflix Prize, Martin Chabbert,” *Predictive Analytics*, September, 2009.

Scott E. Page, “Making the Difference: Applying a Logic of Diversity,” *The Academy of Management Perspectives*, 21(4): 6-20, 2007.

K.R. Lakhani and J.A. Panetta, “The Principles of Distributed Innovation,” *Innovations: Technology, Governance, Globalization* 2, no. 3 (2007): 97-112.

Eric Bonabeau, “Decisions 2.0: The Power of Collective Intelligence,” *MIT Sloan Management Review*, Winter 2009, 50(2): 45—52.

Summary of Assignments

All assignments are due by Midnight on the day indicated.

		Title	Distributed	Due
1	<i>Individual</i>	<i>Predicting the present with Google</i>	April 11	April 17
2	<i>Individual</i>	<i>Visualizing your Facebook network</i>	April 22	April 24
3.1	<i>Individual</i>	<i>Collecting and mapping network data</i>	April 25	May 1
3.2	<i>Group</i>	<i>Collecting and mapping network data</i>	April 25	May 8
4	<i>Group</i>	<i>Identifying target markets and influentials with Twitter</i>	May 20	May 26
5	<i>Group</i>	<i>Applying the tools to a new product idea</i>	May 27	June 5