

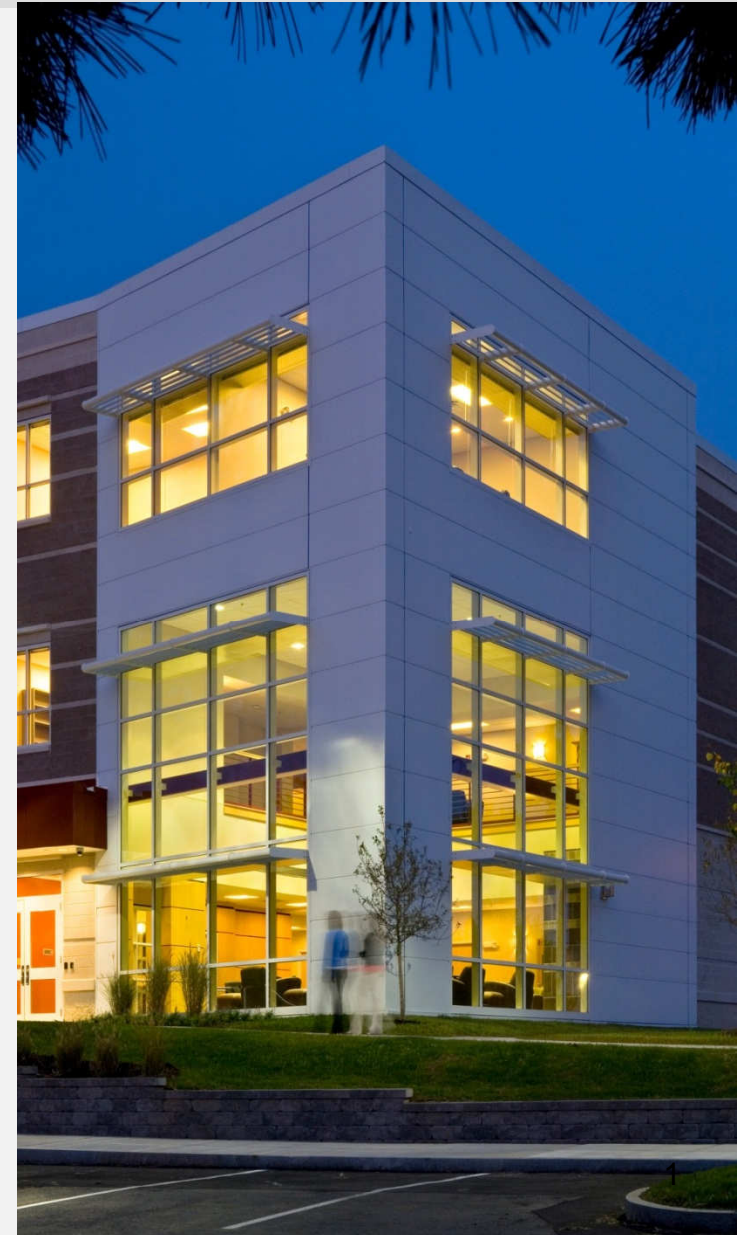
Teaching the Ethics of Big Data

Dean Poeth, Ph.D., P.E., C.Mfg.E.

poethd@uniongraduatecollege.edu

Adjunct Professor, Union Graduate College

eCOTS Conference, May 19-23, 2014



The problem

- Corporate ethics challenges
 - Cut-throat global competition
 - Rapid growth of employee ethics concerns
 - Ethics programs don't make companies ethical
 - *Enron had an established ethics program prior to its collapse*
 - Sarbanes Oxley Act
 - Department of Justice, FBI, and the Securities and Exchange Commission investigating high-frequency trading

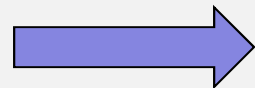
Prepare students for the ethical dilemmas they will face.

Is Teaching Big Data Just Statistics/Analytics?

- Social issues
- Legal issues
- Moral issues

- Privacy concerns
- Data security

- Informed consent
- Accountability
- Customer perception



Big data is neither ethical nor unethical

*Teaching big data requires more
than just teaching analysis*

Reality Check: Big Data Results Are Not Always Correct

- Google Flu Trends
 - Intended to predict flu outbreaks
 - Models based on key search term activity
 - Used to build time series predictive models
 - Started in 2008
 - Problem: performance has not met expectations
 - Overestimated cases of the flu for four consecutive years
- Data Brokers
 - Inability for customers to correct incorrect information
 - No feedback path to correct errors

*All models are wrong, but some are useful.
---George E.P. Box*

- Data brokers using big data to collect then sell personal identifying information
 - Name, income, purchase history, religious affiliation, sexual orientation, political affiliation, etc.
- Informed consent: getting customers permission first
 - Can they even find the privacy statement? And do they read it?
 - Often no provisions to view data or correct inaccuracies
 - Loss of customer's trust
- 39% of respondents said data mining is invasive (Infosys global survey)
- Website tracking now pervasive

Privacy fence?

Data Security Cases

- The hacking of 40 million credit card numbers from customers of Target in December, 2013
- 40% of South Koreans had personal data stolen and sold to marketing firms in 2013
 - Perpetrator arrested
 - Marketing firm executives who purchased the stolen data arrested
- The crowd-funding website Kickstarter was hacked in 2013, resulting in the theft of personal information

Cave-ins of the big data goldmine

Responsibility to protect data that has been collected, stored, transferred, analyzed.

Target Department Stores

- Target used big data and analytics to build a predictive model
- Objective: Identify customers who may be pregnant
- Intent:
 - 1) Anticipate customer needs
 - 2) Provide customers with coupons related to their future baby.
- **Did analysts consider this unanticipated consequence?**
 - An angry parent complaining to a store manager when his high-school aged daughter received a mailing containing advertisements for various new parent supplies

A big data backfire

Applying Analytical Results

- Code of ethics
 - For collecting, storing, transferring, and analyzing big data
 - For using analytical results
- At a minimum, apply analytical results consistent with organizational and professional values

Big data presents a new ethical landscape

Conclusions

- Big data use consistent with personal, professional, and organizational values
- How would actions or inactions look on the front page of a local or national newspaper?

The function of education... is to teach one to think intensively and to think critically. But education which stops with efficiency may prove the greatest menace to society. The most dangerous criminal may be the man gifted with reason, but with no morals.

--- Martin Luther King Jr.

- American Statistical Association, Ethical Guidelines for Statistical Practice, 2014
- Code of Ethics for Certified Analytics Professionals Prepared by the INFORMS Certification Task Force
- Declaration on Professional Ethics, Adopted by the International Statistical Institute (ISI) Council 22 & 23 July 2010, Reykjavik, Iceland
- Davis, Kord, Ethics of big data, O'Reilly Media, 2012
- Le Menestrel, M, Van Wassenhove LN. Ethics in Operations Research and Management Sciences: A never-ending Effort to Combine Rigor and Passion. Omega (2009), doi: 10:1016/j.omega.2008.12.009