634: Empirical Methods

Spring 2017

Michelle Mazurek
Today’s class

• Introducing me
• Introducing you
• Course policies and syllabus
• Empirical methods / research design
• Some HCI background
Who am I?

- Michelle Mazurek (mmazurek@umd.edu)
- Assistant professor, CS and UMIACS
- Affiliated with MC2 and HCIL
- Office hours: Thurs 2:30-3:30 pm in AVW 3421
  – or by appointment
What do I measure, empirically?

• People
  – Interviews, surveys, task-based experiments
  – Needs, preferences, behaviors

• Artifacts
  – Real passwords from CMU
  – Code from github and android apps

• Systems/designs
  – Does this system / tool / design work? Why (not)?
Who are you?

- Preferred name
- Academic program, adviser if applicable
  - Area of research focus
- Background in HCI (a lot, a little, none)
- What do you (want to) measure?
- Why this course?
COURSE POLICIES
Course goals

• Understand, select, apply appropriate methods
  – For data collection
  – For data analysis

• Present your work accurately and effectively
  – Your methods
  – Your results

• Critically evaluate empirical results from others
Course topics

• Meta-topics
  – Research and experimental design
  – Ethics
• Data collection
  – People, systems, crowdsourcing, etc.
• Data analysis
  – Qualitative and Quantitative
Course website

- [https://sites.umiacs.umd.edu/mmazurek/634-s17/](https://sites.umiacs.umd.edu/mmazurek/634-s17/)
- Assignments posted and submitted via ELMS
  - Discussions on ELMS as well
Your grade

- Project: 35%
- Homework: 30%
- Final exam: 20%
- Critique (presentation): 5%
- Class participation: 10%
Reading and reports

• Usually 1-2 required readings per class
  – Sometimes additional optional readings
  – Complete BEFORE class!

• If you don’t do the reading, we can’t discuss
Class participation (10%)

• Contribute to in-class discussions, activities
  – Do the reading!
  – Go to relevant seminars and tell us about them

• Contribute to class discussion board
  – Share interesting empirical research news
  – Ask questions and spark discussion
  – Answer questions for other students
Critique (5%)

• Choose an empirical study to report on
• Describe and critique methods and approach
  – Good, bad, ugly
  – Not enough information to tell?
• 15-minute in-class presentation
  – Bidding for dates, soon
Homework (30%)

• Exercise skills we discuss in class
  – Collect data
  – Analyze data
  – Conduct a mini experiment
  – Propose possible studies
  – Etc.

• Six total, drop one

• Homework 1 out soon
Final exam (20%)

- Take-home, during the last week of class
- Much like a longer homework
Project (35%)

• Replicate a published empirical study
  – Pitch project ideas in class
  – Result in groups of (probably) 2-3

• With different data, different population, different analysis approach, etc.

• Deliverables: proposal, progress report, peer feedback, final paper and talk
  – Workshop-quality paper
  – Optional goal: Maybe a poster or a published paper?
Example projects (mostly CMU)

- An inconvenient trust: User attitudes toward key-directory encryption systems (submitting to SOUPS 2016)
- The post that wasn’t: Exploring self-censorship on Facebook (CSCW 2013)
- Exploring reactive access control (CHI 2011)
- How does your password meter measure up? The effect of strength meters on password creation (USENIX Sec 2012)
- Passwords gone mobile (CHI 2016)
- … and others!
Academic integrity

• Assignments **INDIVIDUAL** unless otherwise noted
  – Don’t look at other students’ assignments

• Zero-tolerance policy for plagiarism
  – Even one sentence is too much. Rewrite in your words.
  – Cite and reference
  – When in doubt, ASK BEFORE YOU SUBMIT

• Review university policies as needed
Other miscellaneous

• I expect you in class
  – Foreseeable family obligations, holidays, conferences, etc: send me email in the next week
  – Unforeseeable: let me know

• Consider joining:
  – hcil@cs
EMPIRICAL RESEARCH

What do we Want?
EVIDENCE BASED SCIENCE
When do we Want It?
AFTER PEER REVIEW

https://twitter.com/LDMay/status/82292212637330944
• What is science?
• Is computer science science?

http://dresdencodak.com/2008/05/21/i-will-do-science-to-it/
What is science?

• Both knowledge and a process
• Establish and organize facts: body of knowledge
• Facts should be falsifiable
  – Make predictions
  – Validate models
• Facts are often relationships
Example CS relationships

X
• Design of a cache
• Took a security class?
• Size and label of button
• Password composition

Y
• Cache performance
• Errors in crypto code
• Click speed, errors
• Rate of mistype errors
So, is computer science science?

• Sometimes!
  – Sometimes it’s math
  – Sometimes it’s engineering
  – Sometimes it’s art

• When appropriate, we would like it to be science!

• Historically, we aren’t great at this.
• “Computer science is not a science, and its ultimate significance has little to do with computers.” – Abelson and Sussman, 1996
For instance, functional programming, object-oriented programming, and formal methods are all thought to improve programmer productivity, program quality, or both. It is surprising that none of these obviously important claims have ever been tested systematically, even though they are all 30 years old and a lot of effort has gone into developing programming languages and formal techniques.

-- Walter Tichy, 1998
• Random sample of ACM papers in 1993: 40% with claims that needed empirical evidence had none
  – Vs. 15% in optical engineering, etc.
  – (Tichy et al., 1995)
What do we mean by empirical?

- “Relying on or derived from observation or experiment.” – American heritage dictionary
- Systematic
- Observation
- Controlled (?)
- Promotes validity (more next week)
Goals

- Representation: Learn something true
- Generalizability: That applies beyond this specific situation (study)
General approaches

• Inductive
• Deductive
Inductive research

Observe \rightarrow \text{Propose theory}

Warning: Can induce incorrect results from valid data!
Observation

• Ethnography
• Contextual inquiry
• Measure data or artifacts

• Self-reported data (interviews, surveys)
Correlation vs. causation

• Directionality
• Third variable

Deductive research

AKA: The scientific method
Variables and controls

• Treatments: Vary the thing you care about
  – Independent variable
• Control other things
• Measured outcomes
• If designed properly, can demonstrate causation!
• May not properly reflect complex phenomena

Much more next week