Empirical Methods for Socio-Technical Research

Instructors: James D. Herbsleb / Marcelo Cataldo
Class Meeting Time: Tuesdays 9:00a-11:50a
Location: NSH 3002
University Units: 12
Office Hours: by request

Course Description.
The past decades has witness the emergence, among others, of two important trends in software systems. First, an increasing number of software technologies have a significant social component (e.g. end-user programming, collaborative development tools, phishing detection tools, etc). Second, software systems have increased significantly in complexity and size challenging traditional development and testing approaches. Empirical research methods provide a basic set of tools for researchers interested in designing and evaluating such system. This course is designed to acquaint you with several basic types of empirical methods such as ethnography, interviews, surveys and experimental and quasi-experimental design. Ultimately, the goal of the course is to develop your skills to make appropriate methodological decisions and allow you to become a better and more critical consumer of research. You will be required to critique a number of examples of published research as well as apply one or more methods to your current research projects to address an important software engineering question.

Course Overview.
The course will have the following general sequence.

- **Introduction to Empirical Methods** – An overview of the nature of empirical research, the various methodological approaches, as well as examples of good and not so good empirical studies will be discussed.
- **Qualitative Methods** – The central characteristics of qualitative research as well as traditional methodological approaches such as ethnography and interviews will be discussed.
- **Experimental and Quasi-Experimental Design** – Various design variants for experiments and quasi-experiments will be discussed and evaluated.
- **Measurement** – Measures are quite often used as the basis for empirical claims. The issues of reliability, validity and generalizability will be defined and discussed.
- **Data Collection Methods** – A collection of approaches for data collection will be examined including survey methodologies, the use of archival data, and content analysis??.
- **Implementing a Research Design** – A critical element of success in the appropriate implementation of the study whether it is archival, in the laboratory, or in the field. This particular issue will be part of the discussions throughout the entire course.

Class textbook and additional readings.
There are additional books that are solid references in various topics in research methods. You are welcome to acquire them but it is not required. We will only read a few chapters or part of chapters from them which will be available in electronic form in the "Course Documents" folder. The books are:


Additional papers will be assigned.

**Class grading.**
The students will have to prepare a 1-page critique of the assigned research papers. There will be two small assignments where the students will have to apply a particular research method to address a research question and present the research design to class. The final project will consist on a research paper addressing a software engineering research question of interest to the student. One or more empirical methods presented in class have to be part of the research paper.

The grading for the class will be structured in the following way:
- Readings critiques 30%
- Two small assignments 30% (15% each)
- Final project 40%

**Paper Critiques.**
Students are expected to submit critiques of published papers we will discuss in class. There are 12 papers marked with (*) and students should select 10 of those papers. Each paper critique will have a limit of 500 words and will be due before class on the day the selected paper is discussed. We will examine how to prepare paper critiques in the first few classes of the course.

**Small Assignment.**
We will have two small assignments in the course. Each assignment will consist on applying a research method to address one of three research questions provided by the instructors. Alternatively, students can choose to apply the research method to their own research projects.
For each assignment, students will prepare two deliverables. A report describing the research question/s, the details of the research method used and how the design allows you to answer the research questions should be prepared. The reports are due at noon on the day prior the assignment discussion class. The second deliverable consists of a 10-minute presentation to be given in class. Each presentation will be followed by a 10 minute discussion. The reports will be made available to the class and students are expected to read them prior to class.

**Final Project.**

An important amount of effort will be dedicated to a semester-long project in which students go through the process of articulating a set of research questions and developing the research methods appropriate for answering those questions. The outcome of the project will be a research paper that will contain 3 sections: introduction, literature review + research questions, and research method. The content is expected to be of sufficient detail and quality to be reviewable as a research proposal by a panel of experts from a funding agency. Throughout the semester, we will have deadlines for intermediate deliverable in order to motivate a continuous progress of the project.

**Class Syllabus.**

**Class #1 - January 12th:**
Topics: Introduction to Course / Motivation for Empirical Methods / Testing theory versus testing an outcome / A few examples of bad and good use of empirical methods

Readings:
- R&R - Chapter 2 (*read pages 37 through 54*)
- KKV - Chapter 1 (*read pages 1 through 28*)
- Examples of paper critiques.

**Class #2 - January 19th:**
Topics: The nature of qualitative research / Introduction to Ethnography and Interviews

Readings:
- R&R - Chapter 6 (*read pages 160-177*)
- S&C - Chapters 1, 2, 3 and 4.
Class #3 - January 26th:
Topics: Introduction to Survey Research / Questionnaire Construction / Computer Mediated Data Collection / Longitudinal Surveys

Readings:

- DSC - Chapters 4 and 5.
- R&R - Chapter 6 (*read pages 177-186*).

Class #4 - February 2nd:
Topics: Introduction to Measurement / Reliability / Validity / Generalizability / Scaling

Readings:

- R&R - Chapter 4
- SCC - Chapters 2 and 3

February 9th: NO CLASS

Class #5 - February 16th:

*Assignment #1 DUE on February 15th at NOON*

Topics: Assignment #1 discussion

Readings: Assignment #1 submissions as basis for the class discussion

Class #6 - February 23rd:
Final Project Proposal DUE (max. 1000 words)

Topics: Introduction to Experimental Design

Readings:

- R&R - Chapters 7 and 8

Class #7 – March 2nd:
Topics: The Role of Statistics / Understanding Interactions

Readings:

- R&R - Chapter 12

March 9th: NO CLASS – SPRING BREAK

Class #8 - March 16th:
Topics: Quasi-Experimental Design

Readings:

- SCC - Chapter 4
- SCC - Chapter 5 (read pages 135-156)
- SCC - Chapter 6 (read pages 175-181)

Class #9 - March 23rd:
Final Project “Intro + Literature Review + brief outline of research method” DUE

Topics: Introduction to Doing Lab Experiments / Problems and challenges

Readings:

- R&R - Chapter 9
- SCC - Chapters 9 and 10

Class #10 - March 30th:

*Assignment #2 DUE on March 29th at NOON*

Topics: Assignment #2 discussion

Readings: Assignment #2 submissions as basis for the class discussion

Class #11 - April 6th:

Topics: Writing Scientific Papers

Readings:

- Acceptance and Rejection:
  - Case I: Cataldo et al from ICSE/CSCW
  - Case II: Frost & Taylor from Administrative Science Quarterly

Class #12 - April 13th:

Topics: Exploratory Data Analysis
Readings:

- R&R - Chapter 10

Class #13 - April 20th:
Topics: Analysis of Large-Scale Datasets / Analysis of Trace Data

Readings:


Class #14 - April 27th:
Topics: Multi-Method Research

Readings:


Class #15 - May 4th:
Topics: Final project presentations

No Readings

May 7th: *Final Project Report DUE at NOON*