



UNIVERSITÀ DI PISA
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
Dottorato di Ricerca in Ingegneria dell'Informazione

Doctoral Course

“Empirical Research Methods in Software Engineering”

Alessio Ferrari, Ph. D.

ISTI-CNR, Pisa, Italy

E-mail address: alessio.ferrari@isti.cnr.it

Short Abstract: Empirical studies in software engineering provide a systematic way of evaluating theories, languages, concepts, tools or methodologies, considering the industrial context in which they are applied [1]. The course will prepare students by examining how to plan, conduct and report on empirical studies in software engineering. The course will cover all of the principal methods applicable to software engineering (controlled experiments, case studies, surveys, systematic literature reviews, and ethnography) and will describe *quantitative* and *qualitative* methods of analysis, including hypothesis testing and grounded theory. To showcase the different methods, the course will critically review representative examples of published work. At the end of the course, the students will be able to approach real-world research problems in a scientifically sound way, and contribute to theory building in software engineering research.

Two versions of this course were delivered: 1) For the MSc and Ph.D students of the University of Florence, School of Mathematical, Physical and Natural Sciences in 2020; 2) for the Ph. D. students in Smart Computing at the University of Florence, in 2021. The course will be also delivered in Fall 2022, again at the University of Florence, and at the Ph. D. in “Ingegneria dell’Informazione” at the University of Pisa. A web version of the course has been made available on YouTube [2]. The reference book for the course is the handbook from Wohlin et al. [1].

Course Contents in brief:

- Overview of Empirical Research Methods in Software Engineering Research
- Formulating Research Questions
- Data Types, Measurements, Scale
- Data Collection Techniques
- Building Theories in Software Engineering
- Research Strategies: the ABC Framework for Software Engineering

- Controlled Experiments
- Hypothesis Testing and Statistical Tests
- Qualitative Research Methods: Ethnography, Interviews, Grounded Theory
- Survey Research in Software Engineering
- Case Studies
- Systematic Literature Reviews

Total # of hours of lecture: 20 hours

References:

[1] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in software engineering*. Springer Science & Business Media.

[2] A. Ferrari, *Empirical Software Engineering Lectures* (2020).
<https://www.youtube.com/playlist?list=PLSKM4VZcJv-P3fFJYMu2OhITjEr9BjI0>

CV of the Teacher

Alessio Ferrari is a researcher at CNR-ISTI since 2011. His research interests are primarily formal modelling and natural language processing applied to empirical software engineering. He received his Ph.D. in Computer Engineering in April 2011 from the University of Florence. The Ph.D. work was focussed on the introduction of the modelling and code generation technologies in a railway signalling company, namely General Electric Transportation Systems (GETS). In this context, he acquired experience with formal modelling and code generation by means of Matlab Simulink, and, during his Ph.D., he developed the core control algorithm of the ATP system for the metro of Rio de Janeiro. From September 2011, he is researcher at ISTI-CNR, where he studies the applications of natural language processing techniques to the analysis of natural language requirements, and other software engineering artifacts. As part of the FMT lab of ISTI-CNR, he participated in the TRACE-IT regional project, focussed on railway signalling systems, and took part in the design and development of a prototypical Automatic Train Supervision (ATS) system, as a railway domain expert. He also participated in the LearnPAD FP7 European project. He has been the work package leader for the Shift2Rail ASTRail project, actively conducting the work on comparing formal and semi-formal modelling approaches for the development safety-critical railway systems.

He is author of more than 60 papers in international journals and conferences and has been a program committee member of top conferences in software engineering in general and requirements engineering in particular, such as ACM/IEEE ICSE, IEEE RE, and REFSQ.

Publication record: https://dblp.uni-trier.de/pers/hd/f/Ferrari_0001:Alessio
Google Scholar: <https://shorturl.at/ezCD9>

Final Exam: the students will be asked to present selected papers in empirical software engineering in a seminar, and to discuss papers of their peers, in a workshop-like manner.

Room and Schedule

Room: *the course will be carried out online*

Schedule:

- Day 1 - 2 hours: Overview of Empirical Methods in Software Engineering Research and Formulating Research Questions
- Day 2 - 2 hours: Data Types, Measurements, Scale, Data Collection Techniques
- Day 3 - 2 hours: Building Theories in Software Engineering and ABC framework
- Day 4 - 2 hours: Controlled Experiments
- Day 5 - 2 hours: Hypothesis Testing and Statistical Tests
- Day 6 - 2 hours: Qualitative Research Methods: Ethnography, Interviews
- Day 7 - 2 hours: Coding and Grounded Theory
- Day 8 - 2 hours: Survey Research in Software Engineering
- Day 9 - 2 hours: Case Studies
- Day 10 - 2 hours: Systematic Literature Reviews