4th International Workshop on Conducting Empirical Studies in Industry (CESI 2016) – Post-workshop Report

Carlos Henrique C. Duarte
Brazilian Development Bank (BNDES)
Rio de Janeiro, Brazil
carlos.duarte@computer.org

Andreas Jedlitschka
Fraunhofer Institute for
Experimental Software Engineering (IESE)
Kaiserslautem, Germany
andreas.jedlitschka@iese.fraunhofer.de

Ayse Bener
Ryerson University
Toronto, Canada
ayse.bener@ryerson.ca

ABSTRACT

Few would deny today the importance of empirical studies in the field of Software Engineering. An increasing number of studies are being conducted involving the software industry, but, while literature abounds on idealistic empirical procedures, relatively little is known about the dynamics and complexity of conducting empirical studies in the software industry. How research results are put into action in industrial settings and how much cross company learning takes place through replication of empirical studies in different contexts? What are the impediments when attempting to follow prescriptive procedures in the organizational setting and how to best handle them? These drivers underly the organization of the fourth in a series of workshops, CESI 2016, held on 17th May, 2016 at ICSE 2016. This report summarizes the workshop details and the proceedings of the day.

Keywords

Empirical Studies, Software Industry.

1. INTRODUCTION

An *empirical study* is an investigation, using established procedures (also called "empirical methods"), for the purpose of gaining knowledge through observation. Empirical methods fall under the broad categories of case studies, scientific experiments and surveys. Investigative questions are determined and related data is gathered and analyzed to answer these questions. Briefly, with experiments [1], we are in search of quantitative, cause-and-effect relationships, involving control of treatment. Typically, experiments are carried out in laboratory settings where the few variables involved in the investigation can be manipulated as desired. With case studies [2], we are in search of qualitative or quantitative findings (or phenomena) among the identified variables in the case under study in a real-world setting. Because we are not looking for causal relationships in the case study, there is no "control" of treatment that forms a basis for such a causal relationship. With surveys [3], the aim is to seek qualitative or quantitative responses from a sample representative of the population under study. There are various "research designs" to cater for different investigative situations. Examples include: independent measures, repeated measures, matched pairs etc; exploratory case studies, longitudinal case studies, ethnographic studies, action research etc; and online surveys, focus groups, interviews etc. With empirical studies being widely entrenched in fields such as social sciences, psychology, management sciences, and medicine, there is obviously much more in the general literature on empirical studies than what is hinted above; still, this brief introduction suffices for our purposes here.

In so far as Software Engineering (SE) is concerned, empirical studies lie at the heart of this burgeoning field. The quality of these studies is a determinant of the validity of the research findings, including that of the comparative analysis of competitive methods, tools and techniques. With increased awareness, more and more researchers are conducting empirical research in SE and, increasingly so, involving the software industry.

While there are established empirical procedures in the general literature, relatively little is known about conducting empirical studies involving the software industry. What pitfalls should be avoided when investigating phenomena in an organization; what challenges should be anticipated when evaluating the efficacy of methods and tools in actual projects; what are the Dos and Don'ts when conducting practitioner surveys? Such questions abound and formed the primary trigger for organizing this series of workshops. The chosen theme was thus "conducting empirical studies in industry", yielding to the CESI acronym.

Experience suggests that empirical studies conducted in industrial settings are particularly challenging because the actual environments are complex and what is first observable by researchers (typically from academia) may only be a tip of an iceberg. Yet, relevant investigative questions must be formulated, valid constructs need to be defined, trust needs to be in place, relevant data must be gathered within the small time-frames available, inaccuracies in data gathered (including missing data) needs to be managed, appropriate interpretations of the findings need to be made fitting the industry contexts, results need to be delivered in real-time etc. In essence, researchers often need to be able to run while they are still learning how to walk.

2. WORKSHOP GOALS AND PROCEEDINGS

The goals of the CESI workshop series are:

- to deliberate on challenges and experiences in conducting empirical studies in industrial settings;
- to discuss strategies for overcoming impediments;
- to debate on the limitations of contemporary research methods;
- to project towards their resolutions; and
- to analyze results in the context of empirical studies conducted in the organizational setting;

Several mechanisms were used to realize these goals: the invitation of a keynote speaker and invited talks, paper presentations and the not-so-common "wall of ideas" session.

3. THE CHANGES IN THE CESI WORKSHOP SERIES IN 2016

Since the first CESI workshop, which was held in San Francisco, USA, as part of ICSE 2013 [4], the event had Xavier Franch and Nazim Madhavji as the main organizers. Each year, a different set of additional organizers was invited in order to involve new participants in the CESI community and bring new ideas to the workshop organization. Both the second workshop edition, held in Hyderabad, India, as part of ICSE 2014 [5], and the third edition, which was held in Florence, Italy, as part of ICSE 2015 [6], kept the same governance structure.

In order to expand the community gathered around the theme, while trying to ensure the mid and long term governance and sustainability of the workshop series, a two tier organization structure was adopted from

2016 onwards: a new group of organizers was selected (respecting criteria of gender, location and affiliation type diversity), while the original workshop organizers established a workshop advisory board.

Building on the results and momentum of the previous CESI workshops, the fourth CESI workshop was conducted at ICSE 2016, in Austin, Texas, USA (http://sites.google.com/site/cesi2016). In addition to the methodological and research focus of the previous workshops, we sough contributions on how research results are put into action in industrial settings and how much cross company learning takes place through replication of empirical studies in different contexts. We were also interested in the impact of empirical studies conducted in industry, including successes and failures in the form of 'lessons learned'. The idea behind this move was to: (i) further precipitate empirical research in the SE community, and (ii) engage industry participants from the point of view of the utility of the results emanating from empirical studies. The following sections of this report describe the 2016 workshop.

4. THE SUBMISSIONS AND REVIEW PROCESS

There were 7 submissions to the 2016 workshop, from 28 different authors, from 8 countries, of 13 different affiliations, all of them pertaining to the categories: of vision papers, practitioner messages, technical papers and experience reports. One of the papers was desk-rejected for being out of scope. Each of the remaining papers was reviewed by at least three reviewers. Finally, we decided to accept 6 regular papers for presentation at the workshop.

Below, we analyze the accepted papers from various dimensions:

1. *Demographic data:*
   - *Region*. Each accepted paper was authored or co-authored by academics and practitioners from the same country. We had papers with authors from Canada [8], Croatia [10], Germany [11,12], Spain [7], Sweden [10] and Turkey [9].
   - *Industry or Academia*. Papers from universities and research centers were predominant: 6 papers had authors from academia; but 3 of these papers also had co-authors from industry [8,9,12].
2. *Type of study.* Half of the accepted papers [8,9,11] presented experience reports describing lessons learned, challenges, open issues etc, from a series of primary studies. The rest of the papers [7,10,12] were of technical nature, presenting visions or conclusions emerging from empirical studies.
3. *Discipline of the study*. Although in most cases the studied discipline may not have influenced the observations, all the works referred to specific software disciplines. We had papers concerning software tools [7], issue management [8,9] and software development in general [11,12]. One paper didn't specify any particular software engineering discipline [10].
4. *Type of studies analyzed.* Most of the papers were focused on one particular type of study: case studies [7,8,9,10,11]. One other paper addressed multiple types of studies: case studies and interviews [12].
5. *Own studies or studies from the community*. The majority of the papers reported on one's own work; one paper involved a vision of the studies performed by our community [10].
6. *Number of primary studies*. As one would expect, the vision paper [10] involved the greatest number of primary studies (4). Also, we had 4 papers reporting on one study conducted in one company (1-1) [8,9,11,12], one other reporting on a study repeated at many companies (1-n) [7].

The papers presented at the workshop are accessible through the ACM Digital Library (references appear below).

5. SUMMARY OF PRESENTED PAPERS

In [7], Valverde and Pastor describe the evaluation process of a capability modeling tool, which is currently in use by two industrial partners that develop software in-house. The empirical evaluation has been conducted using online issue tracking methods and web-based forms organized in a continuous feedback process. The lessons learned are that no-setup tools encourage participation, that unexpected contributions arise from open questions, and that not every issue should be traced, since their prioritization can be detected in form responses.

Karim and others [8] study factors that affect issue management at Plexia, a Canadian IT company that exploits the healthcare domain. They apply source code and run time analytics based on statistical methods and machine learning techniques based on domain tag identifiers to obtain suggestions on how to optimize software development processes based on agile methods. They conclude that efforts in logging estimates and categorizing issues pay off, since these enable the identification of factors affecting issue management.

A defect prediction model developed in partnership with Netas, a Turkish system integration company, is described by Koroglu and others in [9]. As a conclusion of applying the model on large scale legacy code, they suggest that development teams can focus with high accuracy on small portions of code, where the majority of faults are located, thus facilitating software maintenance.

Grbac and Runeson [10] argue that the continuous monitoring and coding of software engineering knowledge and practices will be critical for the evolution to empirical software engineering. In order to support this long term vision, they propose the development of a software framework to enable to formulation and dissemination of case study models to the community, through which modeling, data collection, recommendation and replication can be performed.

In [11], Fernandez and Wagner discuss practical challenges and lessons learnt in conducting several case studies in the software industry, hoping to increase the awareness of inexperienced researchers on the obstacles they might face in conducting empirical studies and paradigmatic ways to deal with them. They synthesize a list of success factors for conducting such studies.

Finally, in [12], Guzman and others propose the usage of multiple case-studies and interviews to evaluate the benefits and drawbacks of adopting project management techniques in large public sector software development projects. They report their plans to evaluate a project aiming to subsidize the development of a particular system that will equip the German armed forces, by asking specific stakeholders to try particular previously developed components of the system and by interviewing these persons afterwards, so as to obtain insights regarding the system acceptance.

6. SUMMARY OF KEYNOTE SPEECH AND INVITED TALKS

In CESI 2016, the keynote presentation was given by Natalia Juristo (Universidad Politécnica de Madrid, Spain, and University of Oulu, Finland) [13]. She addressed her research project on whether or not insights can be gained by performing experiments in the software industry in the same way that clinical trials are performed in medicine. Juristo argued that, if the answer is positive, not only companies could benefit from the elicited treatments, they could also use the resulting evidence in their decision processes. To that effect, she described a unique experiment that was carried out at seven sites of six different companies, together with the corresponding transferred results. Juristo concluded the presentation by reporting on the learned lessons and the perceived challenges ahead.

The workshop also had two invited talks, by Brendan Murphy (Microsoft Research Center, Cambridge, UK) [14] and by Guilherme Travassos (COPPE/UFRJ, Brazil).

Murphy reported his views on how software development processes in large scale organizations could be locally optimized, since it is virtually impossible to adopt the same uniform development process across the product development and service delivery businesses of the same organization. He mentioned that such optimizations can be performed based on the attributes that products and services possess. He then described the software development history of Microsoft, how it became critical to ensure correctness, the use of patches to insure this goal, the decisions to adopt agile software development processes and cloud infra-structures, together with the impact of such decisions in the whole software development process. Finally, he explained Microsoft's rationale for adopting feature models to predict software quality and performance based on their components size, coherence and coupling, allowing optimizations to be made on the fly.

Travassos presented his approach to the application of scientific methods in software engineering. He mentioned that customers need software technologies, but sometimes the research community pushes the adoption of immature ones. He also stressed that empirical and experimental methods are complementary and can be used together or in isolation to insure the required maturity levels and consequent scientific learning. Travassos argued that the evolving knowledge on software engineering can benefit from performing model building, experimentation and learning. In order to obtain scientific evidence, he suggested primary studies with *in vivo* or *in vitro* character (by analogy with natural sciences), as well as those with *in virtuo* or *in silico* nature (by recurring to simulations or observations of interactions between objects and their environment). Finally, he argued that this knowledge feedback loop would not be complete without secondary studies.

7. THE WALL OF TOPICS

During the workshop, we collected topics that came up during the talks and discussions at the "*Topic Wall*". All participants were invited to put their ideas, asynchronously, on the *Topic Wall* – an unstructured wall for capturing topics to be further discussed. The authors clustered the topics after the workshop as follows:

- **Expectations:** Relevance, Scale (generalization of the study results), **Impact**, Co-creation value utility;
- **Effort** (required to plan, conduct, participate in);
- **Recruitment** (of subjects);
- **Trust** (has to be created);
- Sustainability (of the results / impact);
- **Process** (empirical study process)**:** Pragmatism, Agile, Partial results (should be accepted and fed back early);
- Soft Skills: Optimism, **Motivation**, "Naivety";
- Methodology: methodological studies, Case study, Structure / Taxonomy, From Replication to Theory, Ethics;
- Results: Learn, Knowledge, Clarification, Negative Adaptation;

Due to time constraints, however, only some points (bold text) from the wall were discussed in more detail in a plenary discussion session.

8. SUMMARY

The organizers of the CESI (*Conducting Empirical Studies in Industry*) series of workshops started out with a premise that while an increasing number of empirical studies are carried out in the field of software engineering, relatively little was known about how the results of empirical studies influences practice. The workshop series was initiated to deliberate on pertinent matters on conducting studies in industry. CESI 2016 was the fourth in the series of workshops.

We had presentations from authors from many different parts of the world, reporting on their studies on distinct software engineering subjects, using a variety of empirical research methods (see Section 5 for more details). In her keynote, Natalia Juristo discussed experiences from conducting experiments in industry in course of the ESEIL FiDiPro project. Invited talks were given by Brendan Murphy on the industry perspective and by Guilherme Horta Travassos on the research perspective (see Section 6 for more details). The talks were complemented by an asynchronous session, called the "Wall of Topics", where participants posted their thoughts and ideas, in parallel, on a large sheet at the wall (see Section 7 for more details). Feedback from the attendees was that, for junior researchers, the CESI series of workshops was a valuable opportunity to expose their work and share ideas; whereas, for more experienced researchers the series depicted the evolution of the workshop's subject of "conducting empirical studies in industry".

A long-term goal of the series of CESI workshops is to create a vibrant research and practice community with a focus on conducting disciplined empirical studies in industry hoping that their results will lead to improved software engineering practices, techniques, methods, processes, technologies, products/systems and services. The empiricists in the SE community are invited to write to the authors of this report concerning future considerations for the CESI workshop.

9. ACKNOWLEDGMENTS

We greatly acknowledge the contributions from all the authors of papers submitted to CESI 2016; the valuable time and effort spent by the Program Committee members in reviewing these papers; the presentations made by the keynote speaker (Natalia Juristo), the invited speakers (Brendan Murphy and Guilherme Travassos); the presenters of the accepted papers and the workshop participants at large. We sincerely thank you all.

10. REFERENCES

[1] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B. and Wesslen, A. 2012. *Experimentation in Software Engineering*. Springer Verlag.

[2] Yin, R.K. 2009. *Case Study Research: Design and Methods* (4th ed.). SAGE Publications.

[3] Fowler, F.J. 2009. *Survey Research Methods* (4th ed.). SAGE Publications, 2009.

[4] Franch, X. Madhavji, N.H., Curtis, B., and Votta, L. 2013. 1st International Workshop on Conducting Empirical Studies in Industry (CESI 2013). In *Proc. 35th International Conference on Software Engineering* (San Francisco, USA, May 20, 2013). ICSE 2013. IEEE/ACM.

[5] Franch, X. Madhavji, N.H., Anitha, P.C. and Miranksyy, A.V. 2014. 2nd International Workshop on Conducting Empirical Studies in Industry (CESI 2014). In *Proc. 36th International Conference on Software Engineering* (Hyderabad, India, June 2, 2014). ICSE 2014. IEEE/ACM.

[6] Franch, X. Madhavji, N.H., Duarte, C.H.C. 2015. 3rd International Workshop on Conducting Empirical Studies in Industry (CESI 2015). In *Proc. 37th International Conference on Software Engineering* (Florence, Italy, May 16, 2015). ICSE 2015. IEEE/ACM.

[7] F. Valverde and O. Pastor. Continuous validation of a modeling tool in an industrial setting. Proc. 4th International Workshop on Conducting Empirical Studies in Industry (Austin, Texas, USA. May 17, 2016). CESI 2016. IEEE/ACM.

[8] M. R. Karim, S. M. D. A. Alam, S. J. Kabeer, G. Ruhe, B. Baluta, and S. Mahmud. Applying data analytics towards optimized issue management: An industrial case study. Proc. 4th International Workshop on Conducting Empirical Studies in Industry (Austin, Texas, USA. May 17, 2016). CESI 2016. IEEE/ACM.

[9] Y. Koroglu, A. Sen, D. Kutluay, A. Bayraktar, Y. Tosun, M. Cinar, and H. Kaya. Defect prediction on a legacy industrial software: A case study on software with few defects. Proc. 4th International Workshop on Conducting Empirical Studies in Industry (Austin, Texas, USA. May 17, 2016). CESI 2016. IEEE/ACM.

[10] T. G. Grbac and P. Runeson. Plug in software engineering case studies. Proc. 4th International Workshop on Conducting Empirical Studies in Industry (Austin, Texas, USA. May 17, 2016). CESI 2016. IEEE/ACM.

[11] D. M. Fernández and S. Wagner. Case studies in industry: What we have learnt. Proc. 4th International Workshop on Conducting Empirical Studies in Industry (Austin, Texas, USA. May 17, 2016). CESI 2016. IEEE/ACM.

[12] L. Guzman, S. Steinbach, P. Diebold, T. Zehler, K. Schneider, and M. Habbe. Evaluating the benefits of systematic project management in large public sector projects. Proc. 4th International Workshop on Conducting Empirical Studies in Industry (Austin, Texas, USA. May 17, 2016). CESI 2016. IEEE/ACM.

[13] N. Juristo. Experiences conducting empirical studies in industry: The ESEIL FiDiPro project. Proc. 4th International Workshop on Conducting Empirical Studies in Industry (Austin, Texas, USA. May 17, 2016). CESI 2016. IEEE/ACM.

[14] B. Murphy. Optimizing software development processes. Proc. 4th International Workshop on Conducting Empirical Studies in Industry (Austin, Texas, USA. May 17, 2016). CESI 2016. IEEE/ACM.