Roadmapping
Working Group 4 Results

Laurie Williams, Hakan Erdogmus, and Rick Selby

Our group[1] carefully considered the factors involved with the maturity of the field of Empirical Software Engineering (ESE). A graphical representation of the consensus of the group can be found in the polar chart of Figure 1. Each of the six axes represents a factor. Milestones points in the progress of the factors are delineated on each axis whereby a milestone indicates that the majority of work in ESE is at that level. The distances between the milestones are not drawn to scale. By this, we mean that the progress to move linearly along an axis is not proportional to the amount of work that must take place.

A point on each of the axes is joined via a line. The resulting shape of these lines is our indication of the current state-of-the practice of ESE. In general, the point the furthest away from the center of the chart indicates the ultimate goal in a mature field of ESE. However, due to their nature, the ultimate goal may be an interim point for certain families of study/domains. Each of the factors and the milestones will be discussed below, beginning with Domain/Coverage and progressing clockwise around the polar chart.

Domain/Coverage. Research results are only valid in the context in which the research was conducted. For this reason, we cannot assume *a priori* that the results of a study generalize beyond the specific environment in which it was conducted [1]. Researchers become more confident in a theory when similar findings emerge in different contexts [1]. Initial research results generally involve an isolated study of one artifact. Progression occurs with the examination of multiple artifacts in a system and then to multiple artifacts on a variety of systems in one domain. This research is then replicated in multiple domains. Finally, the family of empirical studies on a topic has been replicated in and/or ported to a comprehensive and representative set of domains.

Evidence Maturity. Researchers begin a line of research with a conjecture (or initial theory) which predicts an outcome. Empirical analyses are conducted to determine if the conjecture can be supported or refuted via observations. A set of analyses are then examined to extract common patterns from the results. The ultimate goal is to obtain a validated theory that is predictive or causal.

Research Methodology Maturity. We consider four general classes of techniques: *in vitro* (research conducted in a laboratory), *in vivo* (research conducted in a live setting), *in virtuo* (subjects interact with a simulated environment) and *in silico* (both subjects and objects are simulated). Initial research results will likely begin with one of these techniques. Ultimately, a respected empirical evaluation will contain results of an integration of all four of these techniques. We also consider a progression from mono-method qualitative or quantitative techniques to multi-method, combinations of qualitative and quantitative techniques.

---

[1] The group consisted of Hakan Erdogmus, Natalia Juristo, Marek Leszak, Sandro Morasca, Rick Selby, Elaine Weyuker, Laurie Williams, Claes Wohlin, Sira Vegas, Marv Zelkowitz.

V. Basili et al. (Eds.): Empirical Software Engineering Issues, LNCS 4336, pp. 181 – 183, 2007.

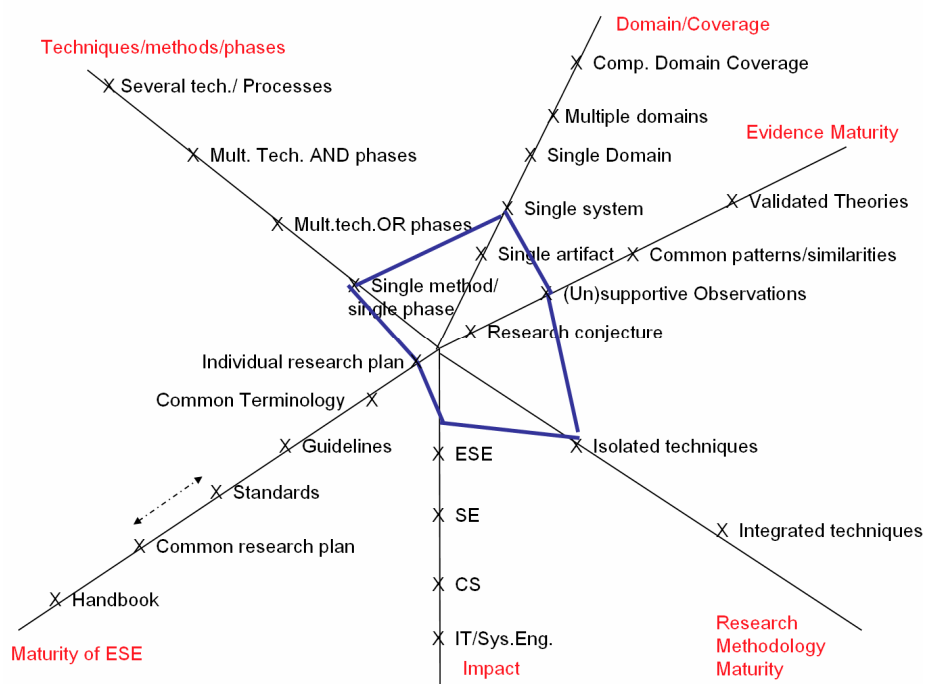

Fig. 1. Empirical Software Engineering Roadmap by Group #4

Impact. The ultimate goal of software engineering research is to impact the practice of software engineering. As the field of ESE matures, its practices will be progressively integrated and accepted by researchers in the field of ESE, then by researchers in software engineering in general, then by researchers in computer scientists, and finally by researchers in the industrial fields of information technology and systems engineering.

Maturity of ESE. Several artifacts must be developed and understandings must evolve for the field of ESE to be mature. As far as artifacts, first an agreed-upon glossary of terminology must be developed and approved by the community. This can be spearheaded by a small working group, whose output can then be reviewed and commented on by the larger community. Using this common terminology, researchers in the community should produce and publish guidelines on ESE practices which can be commented on and evolved by others in the community. A potential venue for developing and publishing these guidelines is a dedicated track of the Empirical Software Engineering journal. Over time, these guidelines can evolve to a set of accepted and published standards. The set of these standards can then be published in a handbook also containing the glossary of terminology. Simultaneously, the field will evolve from the point that every researcher/research group has his/her/their own research agenda to a point where the field has a set of known research agendas and has strategies for jointly tackling relevant topics to produce a validated set of theories in a comprehensive set of domains.

Techniques/methods/phases. In the simplest case, research will focus on a particular technique in one particular development phase (e.g. inspections during the coding phase). Next, this technique will be studied in a variety of phases (e.g. requirements inspections, design inspections, code inspections, test plan inspections). The research could at the next step extend to a comprehensive set of techniques across multiple phases of development. Ultimately, the research will involve several techniques and/or entire processes.

Reference

[1] V. R. Basili, F. Shull, and F. Lanubile, "Building Knowledge Through Families of Experiments," *IEEE Transactions on Software Engineering*, vol. 25, no. 4, 1999, pp. 456 - 473.