

The Use of iStar in Situational Method Engineering: An Ongoing Study

Marcela Ruiz¹[0000-0002-0592-1779], Jolita Ralyté²[0000-0001-8561-3567] and Xavier Franch³ [000-0001-9733-8830]

¹ Zürich University of Applied Sciences, Winterthur, Switzerland
marcela.ruiz@zhaw.ch

² University of Geneva, Geneva, Switzerland
Jolita.Ralyte@unige.ch

³ Universitat Politècnica de Catalunya, Barcelona, Spain
franch@essi.upc.edu

Abstract. Context. Situational Method Engineering (SME) is the discipline that aims at the systematic definition of methods adapted to specific contexts of use (situations). The use of goal-oriented methods for supporting SME is an active research line where the iStar 2.0 language is applied. **Objective.** We plan to conduct an experiment to investigate some designated pragmatic qualities, namely the perceived usefulness, ease of use and accuracy of iStar 2.0 when used in the SME context. **Method.** This paper presents our current work on designing an empirical study for the use of iStar 2.0 in SME. **Next steps.** We plan to refine our current study and run pilots until our measurement tools and sample population are ready for experimental execution.

Keywords: Situational Method Engineering, iStar 2.0, Experimental Design.

1 Introduction

Situational Method Engineering (SME) [1] is the discipline that aims at the systematic definition of methods adapted to specific contexts of use (situations). The term was coined in the early 90s and has been subject of much research since then, which is summarised in Section 2. In that section, we also present one line of research in SME, namely the use of goal-oriented methods for supporting the construction of methods driven by the goals sought. Among the several alternatives, we are interested in the use of the i* framework and more precisely, the iStar2.0 language [2].

Although we have already reported some results in this direction [3] [4] [5], we have focused more on the methodological aspects of SME than on the value that iStar 2.0 brings to our SME approach. This paper is the first step towards the search of evidence of this value. Following the recommendations given in the iStar 2.0 empirical evaluation roadmap [6], we are planning to conduct an experiment to investigate some designated pragmatic qualities, namely the perceived usefulness, ease of use and accuracy of iStar 2.0 when used in the SME context. We present the current protocol in Section 3 and discuss the next steps in Section 4.

2 The Object of Study

The discipline of Situational Method Engineering (SME) advocates the construction of a situation-specific method for each software development project in order to better fit its context and requirements [1]. The building of a new method is based on the reuse of parts of existing software engineering methods and approaches, redefined as method chunks, and stored in a method repository [7]. Building the method chunk repository is called *method engineering for reuse*. Method chunks are selected and assembled to compose a project-specific method, which is called *method engineering by reuse*.

Naturally, the notion of situation is a key concept in SME. It can be described by a set of context criteria. These criteria are used to specify, on one hand, the reuse context of the method chunks, and on the other hand, the characteristics of the software project. The method, to be built for the project, is expected to support the realization of different software development activities, and so, to help the engineer in achieving project goals. Each goal can be achieved with the help of one or several method chunks. Therefore, the concept of goal is also very important in SME as it is used to specify the intention of method chunks, as well as the requirements for the project-specific method.

The selection of the best fitting method chunks for the project at hand is done by matching their goals and context criteria. To facilitate this matching, in our recent work [3] [4] [5], we have introduced contextual goal modelling as a means for specifying the reuse context of method chunks and for describing method requirements of software projects. We use an extension of iStar2.0 [2] where goal models are annotated with context criteria. The case studies documented in the aforementioned publications demonstrate the applicability of contextual goal modelling in our SME approach. However, a further evaluation is still necessary.

3 Designing an Empirical Evaluation

We plan to perform a qualitative experiment to measure perceived usefulness, ease of use and accuracy of iStar 2.0 when used in the SME context. This experiment has been designed according to Wohlin et al. [8].

3.1 Research Questions and Variables

The main research goal, according to the Goal/Question/Metric template is to **gain** empirical knowledge on the use of goal models in SME **for the purpose** of *understanding whether goal models help to build methods in new situations* **with respect to** perceived usefulness and ease of use, and built methods' accuracy **from the point of view of method engineers and the researchers** **in the context of** software engineering experts from industry and academia.

The main research questions are formulated below:

RQ1: When the subjects use iStar 2.0 to build methods for new situations, what is the *perceived usefulness*? To answer this research question, we evaluate the perceived

usefulness when subjects use iStar 2.0 models for scoping methods, specify context criteria, and finally select appropriate method chunks.

RQ2: When the subjects use iStar 2.0 to build methods for new situations, what is the *perceived ease of use*? To answer this research question, we evaluate the perceived ease of use when subjects use iStar 2.0 models for scoping methods, specify context criteria, and finally select appropriate method chunks.

RQ3: When the subjects use iStar 2.0 to build methods for new situations, to what extent all the elements that should appear in the resulting method are specified in the right way? To answer this research question, we measure the level of *accuracy* of the resulting method specification (iStar 2.0 models representing the scope, context criteria indicated in iStar 2.0 models, and selected method chunks).

3.2 Variables

We consider one *independent variable*:

- **iStar 2.0 in SME usage.** The use of iStar 2.0 in SME by reuse. This independent variable help us to understand the degree to which goal models help to build methods in new situations.

We consider the following *dependent variables*, which are expected to be influenced to some extent by the independent variable. We have adapted the Method Evaluation Model (MEM) [9] and the Technology Acceptance Model (TAM) [10] to structure the variables of this study.

- **Perceived usefulness (PU).** The degree to which the subject considers that using iStar 2.0 in SME is effective in achieving its intended objectives. This and the next variable are measured by means of the MEM [9] and TAM [10] questionnaires, which uses a 5-point Likert scale format to obtain subject perceptions.
- **Perceived ease of use (PEOU).** The degree to which a subject considers that using iStar 2.0 in SME is free of effort.
- **Accuracy (ACC).** The degree to which all the elements that should appear in specified methods (giving a certain situation) are actually contained in the method specification. To facilitate this calculation, the researchers take into account a reference method containing the minimum indispensable elements.

3.3 Subjects

Properly speaking, we plan to perform a quasi-experiment because the subjects were not sampled randomly across the population [11]. We use convenience sampling, a non-probability sampling technique for which nearest and most convenient persons are selected for our study. We select subjects with experience with the i* framework, and requirements engineering in general. For categorising the subjects, we plan to run a demographic questionnaire to investigate the following aspects:

- Experience in years with the i* framework
- Experience in years with requirements engineering as a discipline

- Experience in years with the use of method engineering, method engineering concepts, and SME.
- Profession
- Years of experience
- Current and previous roles
- Educational degree (PhD, postdoc, etc)

3.4 Procedure

As part of the experimental task, we provide the subjects with two different input cases for using iStar 2.0 in SME (A and B, see Fig. 2). Case A is intended to allow subjects to get familiarised with the use of iStar 2.0 for SME. For this, we present a full case to the subjects to explain how iStar 2.0 is used when applying method engineering for reuse and by reuse (see Section 2). The topic selected for case A is related to the design of a requirements elicitation method, since this case must be easy to follow by selected subjects. Once the training case has been presented to the subjects, we make use of the understandability form to evaluate to what extent subjects have familiarised themselves with the use of iStar 2.0 for SME. These results will allow us to better understand the results of specified methods when subjects actually use iStar 2.0 in SME to solve the experimental task B (see step 6 in Fig. 2). The topic selected for the experimental task B is the use of crowdsourcing for software engineering (Case B). Case A and Case B are different but are part of the software engineering domain. The design, according to [8] is a “multi-test within object study”, since it examines a single object (case B where iStar 2.0 is used for SME) across a set of subjects (see Table 1).

Table 1. Experimental design

	A set of subjects	Input object
Experimental task B	Use iStar 2.0 in SME	Case B

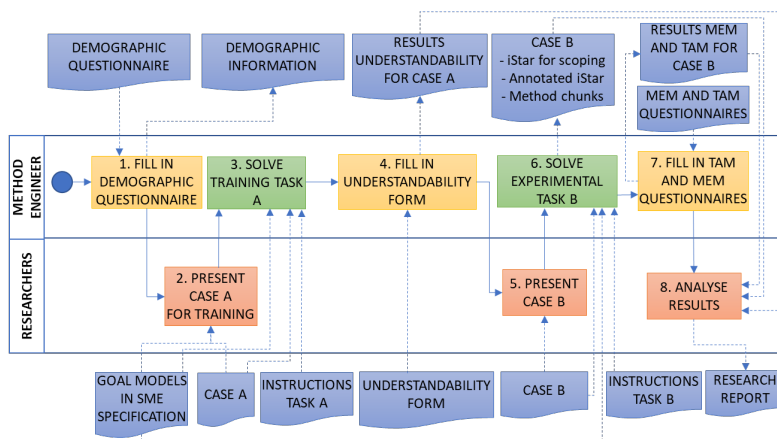


Fig. 1. Experimental procedure

3.5 Pre-analysis of the Threats to Validity of the Results

Conclusion validity. *Reliability of measures.* The understandability, MEM, and TAM questionnaires will be developed to avoid poor question wording. To mitigate possible threats to reliable measures, we will run a pilot to test the questionnaires and measurement tools. There is a risk that the variation in the experimental results related to *random heterogeneity of subjects* is larger than due to the treatment. We have decided to select a homogenous group of subjects to minimise the risk of having diverse results given individual differences. Nevertheless, we acknowledge that this fact reduces the external validity of the study.

Internal validity. *Single group threats.* Given the fact that we do not consider a control group for this study, there are other possible factors that could cause the possible observed effects. Since we plan to conduct this study in different points in time to maximise the number of subjects, there is a risk that *history* affects the experimental results. To mitigate this threat, we plan to control the point in time in which each experimental task was conducted. In this way we can determine if a possible threat related to history is present. Given the current experimental design, we acknowledge a *maturation* threat. Subjects are trained in the use of goals for SME by means of the requirements elicitation case (see Task A in Fig. 2). There is the threat that when solving the actual experimental task (see Task B in Fig. 2), the subjects are more experts and confident with the treatment. We plan to mitigate this threat by providing subjects with completely different cases that relate to two different contexts in SME.

Construct validity. *Mono-operation bias.* Our study includes a single independent variable evaluated through a single case. The cause of the effect might be underrepresented by the prepared experimental task. We plan to pilot the experimental task to ensure the different parts of our object of study are actually evaluated. From a social point of view, it is possible that our subjects could guess what we would like to find as a result of this study. This *hypothesis guessing* threat is mitigated by specifying various questions for each qualitative variable in the MEM and TAM questionnaires. For measuring accuracy, we plan to design an experimental task where subjects do not need to clearly identify what is the right solution.

External validity. Since we have selected a homogeneous group from a non-randomised sample, our ability to generalise the results are limited. We acknowledge this threat and we plan to conduct a demographic questionnaire to characterise our population in a proper manner.

4 Next Steps

The protocol presented in this paper needs to be refined and executed. For refinement, besides our own further work, we count on the feedback given by reviewers and also on the feedback given by experts from other communities on the quality of the instrument (e.g., using some members of the ISERN¹ network). As usual, we plan to pilot the

¹ International Software Engineering Research Network, <https://isern.iese.de/>

study before executing it. Once evaluation instruments are ready, we plan to make them available for further replication studies.

As for execution, the two most critical points to decide are: the population, and the setting. Concerning the population (that we want with knowledge on i*), one option to explore is to use the iStar workshop itself to execute a first round of the study and then allow for a second round after the workshop (as we did in the past when designing the iStarML interchange format [12]). Also, if we have the opportunity, we would like to replicate the study in a population without knowledge of i* in order to evaluate the impact of previous knowledge. Concerning the setting, given the current Covid-19 situation, it is difficult to predict how we will run the study, therefore we are keeping all doors open.

References

1. B. Henderson-Sellers, J. Ralyté, P. J. Ågerfalk, and M. Rossi, *Situational method engineering*. Springer Berlin Heidelberg, 2014.
2. F. Dalpiaz, X. Franch, and J. Horkoff, “iStar 2.0 Language Guide,” May 2016.
3. L. López, D. Costal, J. Ralyté, X. Franch, L. Méndez, and M. C. Annosi, “OSSAP - A situational method for defining open source software adoption processes,” in *Lecture Notes in Computer Science*, 2016, vol. 9694, pp. 524–539, doi: 10.1007/978-3-319-39696-5_32.
4. J. Ralyté and X. Franch, “Using contextual goal models for constructing situational methods,” in *Lecture Notes in Computer Science*, 2018, vol. 11157 LNCS, pp. 440–448, doi: 10.1007/978-3-030-00847-5_31.
5. X. Franch *et al.*, “A situational approach for the definition and tailoring of a data-driven software evolution method,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 10816 LNCS, pp. 603–618, doi: 10.1007/978-3-319-91563-0_37.
6. L. López, F. Başak Aydemir, F. Dalpiaz, and J. Horkoff, “An Empirical Evaluation Roadmap for iStar 2.0,” 2016.
7. I. Mirbel and J. Ralyté, “Situational method engineering: Combining assembly-based and roadmap-driven approaches,” *Requir. Eng.*, vol. 11, no. 1, pp. 58–78, Mar. 2006, doi: 10.1007/s00766-005-0019-0.
8. C. Wohlin and C. Wohlin, *Experimentation in software engineering*. Springer, 2012.
9. D. L. Moody, “The Method Evaluation Model: A Theoretical Model for Validating Information Systems Design Methods.”
10. F. D. Davis, “Perceived usefulness, perceived ease of use, and user acceptance of information technology,” *MIS Q. Manag. Inf. Syst.*, vol. 13, no. 3, pp. 319–339, 1989, doi: 10.2307/249008.
11. C. Robson, *Real world research : a resource for users of social research methods in applied settings*. Wiley, 2011.
12. C. Cares, X. Franch, A. Perini, and A. Susi, “Towards interoperability of i* models using iStarML,” in *Computer Standards and Interfaces*, 2011, vol. 33, no. 1, pp. 69–79, doi: 10.1016/j.csi.2010.03.005.