Science of Computer Programming ••• (••••) •••–•••

Contents lists available at ScienceDirect

Science of Computer Programming

www.elsevier.com/locate/scico

A case study of using grounded analysis as a requirement engineering method: Identifying personas that specify privacy and security tool users

Janna-Lynn (Weber) Dupree, Edward Lank, Daniel M. Berry *

David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada

ARTICLE INFO

Article history:
Received 25 July 2016
Received in revised form 27 July 2017
Accepted 14 August 2017
Available online xxxx

Keywords:
Interviews
Requirements analysis
Security
User categorization
User-space identification

ABSTRACT

This paper explains the importance (1) of full user-space identification with categorization in requirements engineering (RE) and of ensuring that the categorization is a partition of the user space, (2) of the creation and application of user-space-covering personas in RE, (3) of the use of grounded analysis to do RE to produce a specification as a grounded theory, and (4) of privacy and security features in computer-based systems. Then it gives the steps of a grounded analysis method for doing user-space identification with categorization and producing personas as a grounded theory that is describing the classes of users for a computer-based system. The paper summarizes a case study of an iterative application of this method to arrive at a set of user-space-covering personas for privacy and security features in computer-based systems, and it shows how these personas can be used to inform RE for these features. The full case study and the descriptions of the personas are found in the appendices.

1. Introduction

The primary goal of requirements engineering (RE)[1] is to elicit from a computer-based system's (CBS's) stakeholders, including its users, requirements for the CBS and to document these requirements. In order to elicit these requirements, a requirements analyst invests significant amounts of his[2] time identifying the CBS's users [98], identifying use cases for each of these users [34], and identifying requirements that support these use cases [70]. A large fraction of the RE literature is devoted to describing techniques for carrying out these activities more successfully.

* Corresponding author.
E-mail addresses: jannaweber@gmail.com (J.-L. (Weber) Dupree), lank@uwaterloo.ca (E. Lank), dberry@uwaterloo.ca (D.M. Berry).

[1] Glossary of acronyms:

ABB	attitude, belief, or behavior
ABBs	attitudes, beliefs, and behaviors
CBS	computer-based system
HCI	human–computer interaction
PAS	privacy and security
RE	requirements engineering
SE	software engineering

[2] To avoid both the use of plural pronouns to describe a single person and the heavy singular "he or she" and related constructions, this paper alternates the genders of arbitrary individuals by sections, starting with masculine in Section 1.

http://dx.doi.org/10.1016/j.scico.2017.08.010
0167-6423/

Weber, the first author of this paper, earned her master's degree by writing a thesis [108] under the supervision of Lank and Berry, the other authors of this paper. The thesis, titled "Privacy and Security Attitudes, Beliefs and Behaviours: Informing Future Tool Design", first defined a privacy and security (PAS) tool as any CBS or a part thereof, that has PAS concerns or that deals with enforcing PAS. It then categorized users of PAS tools by their attitudes, beliefs, and behaviors (ABBs) with respect to PAS into five categories,[3]

- that are believed to represent all users with any ABB with respect to PAS, including that of ignoring it, and
- each of which represents the subset of the PAS tool users that approach PAS with one particular configuration of ABBs, including that of ignoring it.

In other words, the five categories form a partition of the space of PAS tool users. The elements in this partition are the equivalence classes defined by the equivalence relation of the class members' sharing one particular configuration of ABBs with respect to PAS.[4] Finally, the thesis specified this categorization in terms of five personas, i.e., fictitious users, one per category.

When Weber did the research for her master's thesis, she used grounded analysis [24] to produce the five personas and the grounded theory [60] asserting that the five personas form a partition of the space of PAS tool users and cover the user space. When the thesis was filed, Berry realized, in retrospect, that the research that Weber had done amounted to an uncommonly thorough job of RE[5] for PAS tools. That is, the grounded analysis that she did to arrive at the personas ensured that all users of PAS tools had been accounted for and that their ABBs with respect to PAS had been considered. Moreover, the RE process she had effectively done seems to be usable for doing RE for any kind of CBS in which there are a variety of categories of users, each of which uses the CBS with a different, characterizing set of use cases. The categories that Weber generated illustrate how a thorough analysis of a CBS's users' widely differing ABBs with respect to the CBS can serve to divide the CBS's user space into equivalence classes according to those ABBs.

Weber's thesis shows how past categorizations [71] of PAS tool users fell short of describing in a predictable way users' ABBs with respect to PAS, and thus to PAS tools [47]. In general, an inability to predict a user's ABBs with respect to a CBS implies a less than complete understanding of the user space, which in turn, increases the likelihood that requirements identified for the CBS will be deficient in some way.

To get a better understanding of PAS tool users, Weber

1. identified a sufficiently representative group of prospective users of PAS features,
2. interviewed them,
3. collected data from the interviews, and
4. used grounded analysis to cluster the users around their self-described behaviors.

She determined that these five clusters, each taken as a category, represent the user space of the PAS features of CBSs better than do the past categorizations.

This paper describes the steps that Weber did as the steps of a CBS RE method that is focused on identifying the user space of the CBS and the categories of users in the space in order to understand the different ways to use the CBS. It first introduces the concepts that are necessary to understand in order to be able to carry out the method. Then it details the method itself. The main body of the paper describes the concepts and methods in general. The appendices describe the specifics of what Weber did in her research as a case study of the application of the method.

While the paper describes the method prescriptively and in general terms, the method was discovered by reflecting [65] on what Weber had done to solve one specific problem. Thus, we are certain that the method worked in that one case. That case, while not of a complete CBS, is of a substantial, focused collection of functionality present on many a real-world CBS, namely the PAS parts of the CBS. Thus, we believe that the method has received a substantial test of its effectiveness for at least one kind of CBS.

In the rest of this paper, Section 2 provides necessary background on user-space identification with categorization, personas, grounded analysis leading to a grounded theory, and PAS tools. Section 3 gives a detailed description of the grounded analysis method. Section 4 gives an overview of the case study in which the method was applied iteratively to produce personas that cover the user space of PAS tools. It discusses also the thinking that led to identifying the method, how to use personas to inform RE, and the benefits and drawbacks of doing so. Section 5 discusses past work that is related to the issues discussed in the subsections of Section 2, and Section 6 concludes the main body of the paper. Appendix A gives the

[3] As a matter of fact, the five categories are (1) the Marginally Aware, (2) the Fundamentalist, (3) the Struggling Amateur, (4) the Technician, and (5) the Lazy Expert.

[4] If each category consists of all users that share one particular configuration of attitudes, beliefs, and behaviors with respect to PAS, then the categories form equivalence classes according to the users' ABBs with respect to PAS. That is, each user is in one and only one equivalence class. A user, U, has to be in at least one equivalence class since U has *some* configuration of ABBs with respect to PAS, including that of ignoring PAS. Suppose that a user, U, is in two different classes. Then each class consists of all users that share U's ABBs with respect to PAS. Thus, the two classes are the same, and cannot be different as assumed.

[5] In general, RE for any new CBS, for which what the CBS does is a wicked problem [91], can be considered research in its fullest sense.

Please cite this article in press as: J.-L. (Weber) Dupree et al., A case study of using grounded analysis as a requirement engineering method: Identifying personas that specify privacy and security tool users, Sci. Comput. Program. (2017), http://dx.doi.org/10.1016/j.scico.2017.08.010