

# Notifying and Involving Users in Experimentation: Ethical Perceptions of Software Practitioners

Sezin Yaman, Fabian Fagerholm, Myriam Munezero, Hanna Mäenpää, Tomi Männistö  
Department of Computer Science  
University of Helsinki  
Helsinki, Finland  
firstname.surname@helsinki.fi

**Abstract—Background:** Experiment-driven development with the help of real usage data helps to build software products and services that are of high value to their users. As more software companies use experimentation in their development practises, ethical concerns are increasingly important. **Objective:** There is a need for understanding the ethical issues companies must take into account when practising experimentation as a development strategy. This paper examines how software development practitioners experience the need for notifying users when involving them in experimentation. **Method:** We conducted a survey within four software companies, inviting employees in different functional roles to indicate their attitudes and perceptions through a number of statements. **Results:** Employees working in different roles have different viewpoints on ethical issues. While managers are more conscious about company-customer relationships, UX designers appear more familiar with involving users. Developers think that details of experiments can be withheld from users if the results depend on it. **Conclusion:** Barriers to successfully conducting experiment-driven development are different for different roles. Clear and specific guidelines are needed for ethical aspects of experimentation.

**Index Terms—**User involvement, customer involvement, continuous experimentation, experiment-driven development, ethics, human factors, software development

## I. INTRODUCTION

Today’s software development environments are fast-changing, with competitive and unpredictable markets. Complex and uncertain customer requirements, and rapidly advancing technology, create pressure to deliver and evolve products fast. Hence, many software companies have adopted or are transitioning towards continuous experimentation, making product development decisions based on direct observation of user behaviour in their products [16], [4], [20], [23], [5].

Experimentation can be understood as an empirical activity which involves human subjects. Therefore it is subject to laws and ethical principles which should cover a wide spectrum of issues related, yet not limited to, fair treatment of individuals, personal privacy, and security of both their physical environment and data. Taylor [19] defines ethics as “inquiry into the nature and grounds of morality where the term morality is taken to mean moral judgements, standards, and rules of conduct”. Ethics define what is morally acceptable. As Ferrell et al. [8] point out, different practitioner roles and organizational cultures have different perspectives on ethical decision-making. This means that personal attitudes and perceptions are a factor in

the ethical decision-making process of the whole company [7], affecting how experiments are conducted.

Whereas ethical issues have been extensively considered in scientific experimentation, there is less research on the ethics of experimentation for commercial gain, and ethical guidelines of experiment-driven software engineering are only forming. This study examines the ethical attitudes and perceptions regarding notifying and involving users in experimentation. Our sample consists of people working in three different roles: software developers, managers and UX designers. It aims to contribute to opening the discussion on guidelines for practitioners who wish to involve users in experiments.

## II. BACKGROUND AND RELATED WORK

Experiment-driven development has gained popularity among software companies [17], [15], [13]. As experimentation, e.g., by A/B testing, has gained wide acceptance [14], it has also raised serious ethical questions (see e.g. [3], [11], [12] for illustrative examples). Neglecting ethical concerns can harm the customer’s trust for the company and its products, compromising the cornerstones of technology adoption and customer loyalty [6]. It can also invalidate findings of experiments, resulting in decisions based on unreliable and biased data [8]. This is especially risky, as yielding understanding from data and making managerial decisions based on this are two very different thought processes [17].

While using software products and services, users leave a digital trace of actions [9], [2]. From the viewpoint of customers, a company’s ability to inspect and analyse this data may be perceived to breach their basic human right of individual privacy [10]. This risk is often managed by requiring users to give permission to use their data; however, this can affect the results of experiments as subjects are aware of their participation to some degree [18]. The need to acquire consent can also raise pressure for allowing customers to withdraw from the experiment or to erase their usage data on demand [17]. Due to the high costs of building this capability and the need for protecting their competitive advantage, companies can be reluctant to allow their customers and competitors to become aware of the details of their experiments.

Much of the research on the ethics of using human subjects in experiments has been conducted in behavioural and social sciences and especially in medical research where traditions

mandate ethical reviews for performing experiments [22]. Even though many of the same essential principles apply for experimentation in the commercial domain, it is not so clear whether companies and their employees follow internal or external codes of conduct, such as the Belmont Report [1], if they leave defining ethics to user and customer agreements, or if there is a common understanding of experimentation ethics.

Vinson and Singer [21] identified four key principles as guidelines for conducting empirical studies involving human subjects: 1) Subjects must give informed consent to their participation; this implicitly includes the requirement of notifying users in order to allow them to give consent. 2) Before conducting experiments, it is important to assess whether the benefits outweigh the harm, risks, and efforts, and whether the obtained user data will really be trustworthy, whether the experiment results can be used for decision making, and whether the time spent on experiments is worth spending. 3) Experimenters must take all possible measures to maintain confidentiality. 4) The experiment should have value in order to motivate subjects to expose themselves to the risks. How practitioners in software companies understand the ethical aspects of involving users, and the impact on the company, can be studied in the light of these guidelines. Taking into consideration the possible reluctance of companies to make their customers aware of experiments, as discussed above, we focused our investigation on the companies' ethical perspectives of informing customers and on the general assessment of experimentation (i.e., Vinson and Singer's first two guidelines).

### III. RESEARCH APPROACH

To gain more understanding of how people working in different roles in software companies perceive issues related to involving users in experimentation, we planned and conducted a survey. In this section, we describe the research design and execution.

#### A. Research Question

The goal of this study was to explore the ethical attitudes and perceptions towards involving customers and users in experimentation. We pose the following research question:

*RQ: How do practitioners in software companies view ethical concerns related to notifying and involving users in experimentation?*

#### B. Research Design

To address the research question, we used an online questionnaire to collect data from industry participants. The items regarding ethics were part of a larger survey investigating how companies involve users in their software development. Apart from a background section collecting demographic data, the survey included sections on user involvement and experimentation practices, and attitudes and ethical concerns towards involving users in experimentation. Only data from the background and ethics sections are considered in this paper.

The background section was tailored so that practitioner roles matched actual titles or roles in each company. We used

the term “customer” and “user” involvement interchangeably in the survey in order to refer to the primary user relevant in each company context (see Section III-C). For instance, consultancy companies tended to prefer the term “customer”. We maintained a mapping between company-specific job titles and the role categories we considered most prominent in software companies: “developers” (persons performing technical duties, such as programmers, architects, and testers), “UX designers” (e.g. persons involved with planning user interfaces, usability, and visual design of user interfaces), “managers” (e.g. team or product managers), and “other” (e.g. office admins, sales). At this stage, we considered such a coarse-grained division to be appropriate, given the lack of prior work on the subject.

The ethics section asked respondents to rate their agreement with 14 statements covering ethical aspects and perceptions of involving users in experimentation (see Table I). The first seven statements concerned notifying users about their involvement in experiments, and the other half concerned aspects of involving users in experiments in general. The statements were partly organised hierarchically and intentionally overlapping: general statements (e.g. “Users do not need to know they are involved”) were followed by more specific statements (e.g. “Users can be involved in an experiment without their knowledge if we let them know afterwards”). This was done to allow respondents to express general attitudes as well as exceptions under special conditions. Respondents rated the statements on a five-point scale ranging from “strongly agree” to “strongly disagree”. The option “I don't know” was also given and counted as a missing answer.

#### C. Study Sample and Execution

We recruited participants from four software companies operating in Finland. We aimed for a diverse set, with different sizes and domains. The companies are A) a division of a very large telecommunications network company, B) a large information security company, C) a medium-sized company providing a user interface development toolkit, and D) a large digital consultancy providing software development services. Contact persons at each company helped us tailor the background section to fit their company contexts. They also helped distribute a call for participation as well as reminders to a suitable set of company practitioners. Data was collected for two weeks in each company during November 2016 to April 2017. Participants were distributed over the companies' offices in Europe and the United States.

## IV. RESULTS

In this section, we first give an overview of the survey respondents and then outline the data analysis and its results. For this, we use the median (denoted  $\bar{x}$ ) of the overall responses to evaluate the respondents' tendencies against the whole data set and in subsets split by their roles in each company. In addition, we use correlation analysis to examine relationships of variables in the dataset.

TABLE I  
SURVEY STATEMENTS.

Code	Statement
<i>How much do you agree with the following statements regarding notifying users about experiments? Please answer according to your personal beliefs.</i>	
S1	Users do not need to know they are involved
S2	If we collect personal information, users need to be notified
S3	If no laws are being broken, users do not need to be notified
S4	Users can be involved in an experiment without their knowledge if we let them know afterwards
S5	Users should always be notified when they are being involved in an experiment
S6	It is ok not to disclose all the experiment details to users involved
S7	It is ok to intentionally deceive or mislead the user if experiment results depend on it
<i>How much do you agree with the following statements about involving users in experiments? Please answer according to your personal beliefs.</i>	
S8	I cannot trust that the experiment results will be correct
S9	Involving users in experiments is time-consuming
S10	Our company does not have the needed technical infrastructure to run experiments
S11	Users would not like to be part of experiments
S12	Users have to be convinced of the benefit before taking part in an experiment
S13	Experiments give users false expectations
S14	Experiments reveal secrets about the product strategy

(Response options: 1: completely disagree, 2: disagree, 3: neither agree nor disagree, 4: agree, 5: completely agree.)

### A. Overview of the data set

We obtained a total of 130 responses, which was 17% of the invited participants. For each company, the answer rate was: A: 15%, B: 32%, C: 16%, and D: 17%. Table II summarizes the number of responses per company and role.

The majority of respondents are software developers, with an emphasized share in companies A (57%) and D (66%). Company B has more responses for managers (75%) because the contact person was mostly interested in the attitudes of managers. Company C has a more even role distribution.

Cleaning and analysis of the data was performed mainly with a standard spreadsheet and the R language. After cleaning, there were 103–128 responses for different variables which we used for descriptive statistics. Pairwise complete data points were used for correlations.

### B. Ethical attitudes regarding notifying users

Table I shows the statements (S1–7) about notifying users about their participation in experiments. Figure 1 displays the distribution of the responses. We observe that there is a general agreement that users should be notified of their involvement in experiments (S1,  $\bar{x}=2$ ). A strong agreement that

users should be notified if personal information is collected (S2,  $\bar{x}=5$ ) exists, and the same was found to apply to the case when no laws are broken (S3,  $\bar{x}=2$ ). However, in cases where notification was delivered after the experiment and when the wording was “always notify” (S4 and S5,  $\bar{x}=3$ ), respondents tend to neither agree nor disagree. Most respondents agreed that companies should be allowed *not* to disclose all details of their experiments (S6,  $\bar{x}=4$ ), yet they perceived that deceiving or misleading subjects of experiments is unethical (S7,  $\bar{x}=2$ ).

Figure 3 displays the response distribution for statements for each role. We can observe that among the four roles, UX designers agree the least on S1 (14%) while managers disagree the most on S1 (78%). In addition, we find that UX designers are mostly undecided on S3 (45% in comparison to 20% agree and 35% disagree). On S4, managers disagree the most (68%), while the rest of the roles are mostly undecided. Responses on S5 reveal that managers tend to agree the most (65%) in comparison to the other roles, who are mostly undecided.

We examined the correlations between statements, and found a strong negative correlation between S1 and S5 (-0.66), and a lesser negative correlation between S4 and S5 (-0.56), and S3 and S5 (-0.53). Respondents who agree with one of these statements are likely to disagree with the other. Positive correlations were found between S1 and S3 (0.55), S1 and S4 (0.46), and S3 and S4 (0.48).

We also looked at patterns across the companies, yet recognizing that it may not be feasible to make inferences from them. While doing so, one observation was made: a majority of the respondents in Company A clearly agreed that users should always be notified (S5,  $\bar{x}=4$ ), whereas those in Company D clearly disagreed with the same statement (S5,  $\bar{x}=2$ ). As the distribution of the roles did not directly suggest the reasons for such differences, the contexts of the companies would have to be investigated carefully in order to make further interpretations.

TABLE II  
NUMBER OF SURVEY RESPONDENTS PER COMPANY AND ROLE.  
PERCENTAGES REPRESENT THE RATIO WITHIN THE COMPANY.

	Roles				Total
	Developers	Managers	UX	Others	
Company A	20 (57%)	10 (29%)	2 (0.6%)	3 (0.9%)	<b>35</b>
Company B	1 (13%)	6 (75%)	-	1 (13%)	<b>8</b>
Company C	6 (29%)	4 (19%)	4 (19%)	7 (33%)	<b>21</b>
Company D	44 (66%)	3 (0.5%)	16 (24%)	3 (0.5%)	<b>66</b>
<b>Total</b>	<b>71</b>	<b>23</b>	<b>22</b>	<b>14</b>	<b>130</b>

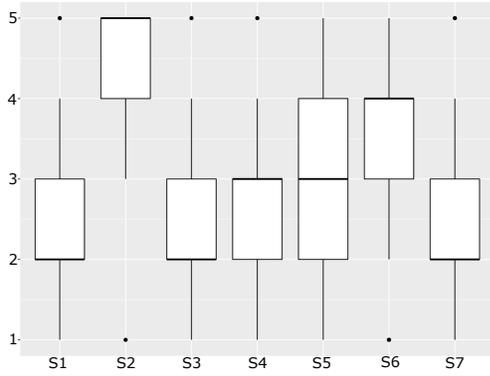


Fig. 1. Distribution of all responses to statements 1–7. 1: Strongly disagree, 2: Disagree, 3: Neither disagree nor agree, 4: Agree, 5: Strongly agree.

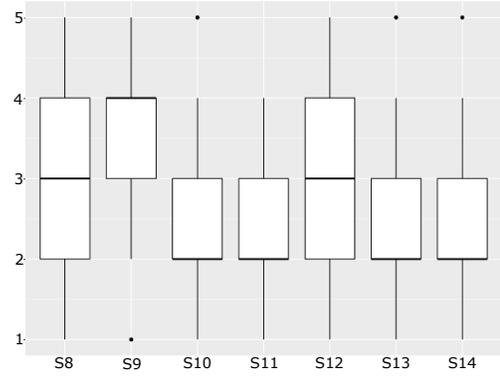


Fig. 2. Distribution of all responses to statements 8–14. 1: Strongly disagree, 2: Disagree, 3: Neither disagree nor agree, 4: Agree, 5: Strongly agree.

### C. Perceptions of involving users in experimentation

When answering questions S8–14 (Table I), respondents expressed their general perceptions on involving users in experiments. A summary of the response distribution is displayed in Figure 2 where we can observe that the respondents are undecided about the trustworthiness of experiment results (S8,  $\bar{x}=3$ ). Overall, involving users in experiments was considered time-consuming (S9,  $\bar{x}=4$ ) even though the infrastructure for experimentation would be in place in their companies (S10,  $\bar{x}=2$ ). In general, respondents thought that users would be favourable to take part in experiments (S11,  $\bar{x}=2$ ) but they were undecided on whether users would need convincing of the benefits of their participation first (S12,  $\bar{x}=3$ ). Respondents did not think that experiments give users false expectations (S13,  $\bar{x}=2$ ), nor that they would reveal secrets about the company’s product strategy (S14,  $\bar{x}=2$ ).

Figure 3 reveals that developers agreed and disagreed equally on S8 (35.9%) while a majority of the managers and the “other” role disagreed (50% and 52.4% respectively). A majority of the respondents in all roles agreed on S9, with developers having the most agreement (61%) followed by UX designers (57%). Regarding the technical infrastructure (S10), all roles except UX designers agreed; the latter were more evenly distributed on the scale. It also appears that all roles except “other” mostly disagreed on S11 with UX designers having the highest disagreement rate (94.7%). Developers and UX designers mostly disagreed on S12 (42.4% and 47.6% respectively), with managers and others mostly agreeing on S12 (70% and 53.8% respectively). Overall, we observe that developers and managers disagreed most on S13 (61.7% and 65% respectively) and all the roles strongly disagreed on S14. Two noteworthy correlations emerged in this section: between S8 and S11 (0.44) and S11 and S13 (0.43).

## V. DISCUSSION

In this section, we discuss the results and answer the research question. We divide the answer into practitioners’ views on 1) user notification and 2) involvement, and consider what these mean for ethical guidelines.

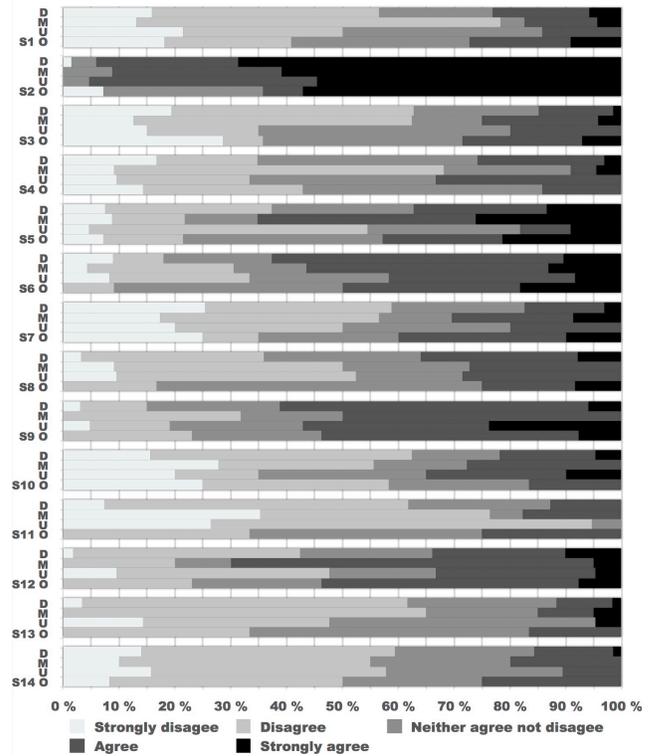


Fig. 3. Relative response distribution of selected statements among roles. (D: developers, M: managers, U: UX designers, O: others).

### A. Views on notifying users

Our analysis shows both positive and negative correlations between certain statements. There are positive correlations between users not needing to know that they are involved in an experiment in general (S1), not needing to notify unless laws are being broken (S3), and notifying after users are involved (S4), as well as between the two latter ones. This group of statements indicates a relaxed attitude towards notifying users. Conversely, the aforementioned statements (S1, S3, and S4) are negatively correlated to “always” notifying (S5). The strong wording of S5 may have influenced the results, but the negative correlation

also makes sense given the sentiment of the other statements. We posit that there are two opposite attitudes towards notifying users: one of always wanting to notify and one which permits exceptions to the rule, e.g. because notification can come afterwards (S4), or because it is considered sufficient to follow legal requirements (S3). Guidelines should explain when and how withholding notification is ethically justified.

Developers' responses indicate a preference for users being aware that they are part of an experiment. However, they tend to accept withholding experiment details (S6). Managers are strongest in agreeing that "users should always be notified" (S5). Intentionally deceiving or misleading the user (S7) was ranked highest among UX designers, yet showing indecisiveness on average; the same was true for users needing to know they are involved in general (S1).

Collecting personal information stands out as a particularly contentious issue: there is strong overall agreement regarding notifying users if their personal information is collected (S2). This applies even in situations where it would not be legally required (S3). However, managers perhaps tend to view the company-customer relationship from a business point of view and be more concerned about legal aspects. UX designers, on the other hand, are often more familiar with the notion of experiment-driven development and user involvement, as their work includes user-centric design, usability tests, and related techniques. Their experience might influence their attitude that sometimes users might need to be unaware of the experiment or be misinformed about the details. Guidelines should elaborate on what personal information is and how its collection and retention should be handled in different situations.

A distinguishable observation specific to Company D is that it had the highest overall disagreement with "Users should always be notified when they are being involved in an experiment" (S5) and agreement with "Users do not need to know they are involved" (S1). This is a strong sign that even though the population is diverse, respondents in that company have an understanding of user involvement in experimentation which acknowledges that different kinds of experiments can be conducted with users, and that they come with different ethical concerns regarding notification. This finding is preliminary and should be confirmed, particularly by looking into the companies' contexts and other demographic information. If confirmed, it could indicate that guidelines need to take the maturity of experimentation in companies into account.

### *B. Views on involving users*

The correlations between trusting experiment results (S8), users wanting to participate, (S11), and giving false expectations (S13) lead us to posit that there are two major attitudes towards user involvement through experimentation: support and opposition. Overall, the results indicate that the respondents believe users would want to participate in experiments (S11), the experiments would not give false expectations (S13) nor reveal product strategies (S14) (e.g., about new product features) and the participants considered their companies to have sufficient means to conduct experiments (S10). On the

other hand, participants also believe that involving users in experiments is a time-consuming task (S9) and experiment results might not always be correct (S8).

Developers perceive that experiments would not give false expectations to the users (S13) nor reveal the product strategy (S14). They also indicate that users would like to take part in experiments (S11). As developers represent a large portion of the dataset (55%), they affect the overall results strongly. Altogether, even though they acknowledge that it might be time-consuming (S9), developers show positive attitudes towards user involvement in experimentation.

Managers, on the other hand, tend to think differently: that users should first be convinced of the benefit of the experiment before taking part (S12). This could perhaps be because they are more cautious about the privacy and legal concerns of their companies and customers. Another reason might be that their knowledge and experience about experimentation could be at a more theoretical level compared to developers and UX designers, who must implement experiments in practice. Therefore, it might be difficult for them to anticipate situations such as when the experiment subjects change their behaviour due to awareness of being involved. During the survey design, some managers stated concerns that experimentation could give users false expectations and reveal secrets about their product strategy, and wanted to know if the employees of their companies could have the same concerns. As results for S13 and S14 portray clearly, these concerns were not supported in any of the roles or companies in our data set.

Contrary to developers and managers, UX designers more strongly think experimentation results will be correct (S8) and users would want to take part in experiments, yet think involving users is time-consuming (S9). Also, they portray hesitance on experiments giving false expectations to users (S13). These could again be due to experience with designing experiments and running them – as anyone who has seriously tried will testify, experiment design is no trivial task, and unforeseen things can happen during execution. These differences between roles should be taken into account in setting guidelines and when managing cross-functional teams conducting experiments.

### *C. Limitations*

Differences in the operational domains and organizational structures of the companies might have contributed to different understandings about experimentation and user involvement. This creates a threat to the construct validity by increasing the risk of misunderstandings of the questions of our survey. To mitigate this, we discussed the survey, its goals and central concepts with company contacts, piloted the survey first with 10 persons and iteratively improved it twice before the actual data collection. In addition, as the current study was mainly descriptive, internal validity, with its focus on causal relationships, was not considered relevant at this stage.

Transferring the results to other contexts should consider contextual differences. With respect to generalizability, we did not attempt to estimate the statistical power of our results, as on one hand, our aim was to gain initial insight into an emerging

phenomenon. Despite the limited number of responses and lack of analysis of all variables in the data set, the phenomenon was well examined through the systematic analysis of the data and therefore, we consider the results to be grounded in actual practice. Future work will address the threat carefully by analysing the whole data set and comparing company contexts.

## VI. CONCLUSION

We investigated software practitioners' views on ethical issues regarding notifying and involving users in experimentation using a survey approach. We obtained responses from 130 participants in four companies. We discovered ethical aspects which were shared by all the participants, e.g., that users should always be notified if personal information is collected, as well as role-specific attitudes, e.g., managers' cautious approach towards users first having to be convinced of the benefit of experiments before participation.

The aggregated results show two opposite ethical attitudes towards notifying users: one of always wanting to notify the users and one portraying exceptions, such as notification afterwards, yet always preferring notification if personal information is collected. Overall, the latter group is larger. In addition, we observe positive perceptions from all the roles: users are believed to want to take part in experiments and experiments are believed not to give false expectations to users. On the other hand, all participants tended to agree that involving users in experimentation can be a time-consuming task.

Differences across practitioner roles can be explained by their field of work and experience. For instance, managers are observed to be more cautious and think strongly that users always should be notified and that they need to be convinced of the benefit before taking part in experimentation. On the other hand, UX designers consider it acceptable to sometimes mislead the users during an experiment if the validity of the results depend on it. In addition, some aspects were likely affected by the companies' own experimentation culture, as some previous work has indicated [8], [7].

Actual behaviour may differ from stated opinions, and in practice, people might not even notice when they are making decisions with an ethical dimension. In part, ethics guidelines aim to raise awareness of such issues. Our findings contribute to the understanding of how practitioners view ethics in software experiments, and to the preparation of ethical guidelines and codes of conduct for experiments involving users. In particular, analysis of the user notification and involvement statements point to customized guidelines for the roles.

The ethical perspectives presented in this paper are part of a larger ongoing research effort concerning involving users in experimentation. This includes aspects such as where in the software development activities companies involve users and what methods the companies use. In future work, we aim to conduct a broader analysis on ethics, as well as involve more companies in the study. Examining whether perceptions differ between national cultures could also be of importance.

## REFERENCES

- [1] T. L. Beauchamp. The belmont report. *The Oxford textbook of clinical research ethics*, pages 21–28, 2008.
- [2] G. Birchley, R. Huxtable, M. Murtagh, R. ter Meulen, P. Flach, and R. Gooberman-Hill. Smart homes, private homes? an empirical study of technology researchers' perceptions of ethical issues in developing smart-home health technologies. *BMC medical ethics*, 18(1):23, 2017.
- [3] J. Constine. The Morality Of A/B Testing. <https://techcrunch.com/2014/06/29/ethics-in-a-data-driven-world/>, Jun 2014. Accessed: 2017-06-23.
- [4] F. Fagerholm, A. S. Guinea, H. Mäenpää, and J. Münch. Building Blocks for Continuous Experimentation. In *1st International Workshop on Rapid Continuous Software Engineering*, pages 26–35. ACM, 2014.
- [5] Fagerholm, F. and Guinea, A. S. and Mäenpää, H. and Münch, J. The RIGHT model for Continuous Experimentation. *Journal of Systems and Software*, 123:292–305, 2017.
- [6] M. S. Featherman and P. A. Pavlou. Predicting e-services adoption: a perceived risk facets perspective. *International Journal of Human-Computer Studies*, 59(4):451 – 474, 2003. Zhang and Dillon Special Issue on HCI and MIS.
- [7] O.. C Ferrell and J. Fraedrich. *Business ethics: Ethical decision making & cases*. Nelson Education, 2015.
- [8] O.C. Ferrell, M. D. Hartline, and S. W. McDaniel. Codes of ethics among corporate research departments, marketing research firms, and data subcontractors: An examination of a three-communities metaphor. *Journal of Business Ethics*, 17(5):503–516, 1998.
- [9] L. Floridi. The informational nature of personal identity. *Minds and machines*, 21(4):549, 2011.
- [10] C. Goodwin. Privacy: Recognition of a consumer right. *Journal of Public Policy and Marketing*, 10(1):149–166, 1991.
- [11] J. Grimmelmann. The law and ethics of experiments on social media users. *Colorado Technology Law Journal*, 13:219–272, 2015.
- [12] J. Jouhki, E. Lauk, M. Penttinen, N. Sormanen, and T. Uskali. Facebook's emotional contagion experiment as a challenge to research ethics. *Media and Communication*, 4(4):75–85, 2016.
- [13] R. Kohavi, A. Deng, B. Frasca, R. Longbotham, T. Walker, and Y. Xu. Trustworthy online controlled experiments: Five puzzling outcomes explained. In *18th ACM SIGKDD, KDD '12*, pages 786–794, New York, NY, USA, 2012. ACM.
- [14] R. Kohavi, A. Deng, B. Frasca, T. Walker, Y. Xu, and N. Pohlmann. Online controlled experiments at large scale. In *19th ACM SIGKDD, KDD '13*, pages 1168–1176, New York, NY, USA, 2013. ACM.
- [15] R. Kohavi, R. Longbotham, D. Sommerfield, and R. M. Henne. Controlled experiments on the web: survey and practical guide. *Data mining and knowledge discovery*, 18(1):140–181, 2009.
- [16] H. H. Olsson, H. Alahyari, and J. Bosch. Climbing the “stairway to heaven”—a multiple-case study exploring barriers in the transition from agile development towards continuous deployment of software. In *38th EUROMICRO SEEA*, pages 392–399. IEEE, 2012.
- [17] A. Schuster and D. Berrar. The omnipresent computing menace to information society. *Journal of Advanced Computational Intelligence*, 15(7), 2011.
- [18] W. R. Shadish, T. D. Cook, and D.T. Campbell. *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning, 2002.
- [19] P. W. Taylor. *Principles of ethics: An introduction*. Dickenson Pub. Co., 1975.
- [20] M. Tichy, J. Bosch, M. Goedicke, and B. Fitzgerald. 2nd international workshop on rapid continuous software engineering (rcose 2015). In *37th International Conference on Software Engineering-Volume 2*, pages 993–994. IEEE Press, 2015.
- [21] N. G. Vinson and J. Singer. A practical guide to ethical research involving humans. In *Guide to Advanced Empirical Software Engineering*, pages 229–256. Springer, 2008.
- [22] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén. *Experimentation in software engineering*. Springer Science & Business Media, 2012.
- [23] S. G. Yaman, F. Fagerholm, M. Munezero, J. Münch, M. Aaltola, C. Palmu, and T. Männistö. Transitioning towards continuous experimentation in a large software product and service development organisation—a case study. In *Product-Focused Software Process Improvement: 17th International Conference*, pages 344–359. Springer, 2016.