Highlights

Making Ethics Practical: User Stories as a Way of Implementing Ethical Consideration in Software Engineering

Erika Halme, Marianna Jantunen, Ville Vakkuri, Kai-Kristian Kemell, Pekka Abrahamsson

- The paper presents the concept of Ethical User Stories (EUS) as a novel practice for implementing ethics in software engineering.

- The concept of EUS is validated through 689 user stories produced in different contexts, including industry projects.

- By building on an established practice (user stories), ethics can be made more approachable for developers, whereas it has usually been challenging to incorporate ethics into SE.

Making Ethics Practical: User Stories as a Way of Implementing Ethical Consideration in Software Engineering

Erika Halme[a], Marianna Jantunen[a], Ville Vakkuri[b], Kai-Kristian Kemell[c], Pekka Abrahamsson[d]

[a]*Faculty of Information Technology, University of Jyvaskyla, Mattilanniemi 2, Jyvaskyla, 40100, Finland*
[b]*School of Marketing and Communication, University of Vaasa, Wolffintie 32 FI-65200 Vaasa PL 700, Vaasa, FI-65101, Finland*
[c]*Department of Computer Science, University of Helsinki, Pietari Kalmin katu 5, Helsinki, 00560, Finland*
[d]*Faculty of Information Technology and Communication Sciences, Tampere University, Kanslerinrinne 1, Tampere, 33014, Finland*

Abstract

Context: Shortcomings of AI/ML systems have recently brought ethics into the spotlight in Software Engineering (SE) in the form of AI ethics. However, actually implementing ethics into practice remains a challenge in both AI ethics and SE at large. Translating abstract ethical principles into requirements and features is difficult and lacks established processes, as well as practices and methods.

Objective: In this study, we explored user stories as a way of implementing ethics in SE. Initially, we simply investigated whether user stories could be utilized for this purpose. After we began to consider this possible, we began to develop the concept of Ethical User Stories (EUS) as a specific practice for this purpose.

Method: We utilized a Design Science Research (DSR) approach to first and explore the use of user stories in implementing ethics, and then to develop the concept of Ethical User Stories (EUS). This process featured three DSR phases through which the concept of EUS was iteratively developed with empirical data.

Results: Over three DSR iterations, we studied 689 user stories produced in different contexts including both student and industry settings. Based on the data, we defined the concept of EUS and provided empirical validation

for it.

Conclusions: The concept of EUS provides a novel way of tackling ethics in SE. This paper presents the concept in-depth, along with practical suggestions for utilizing EUS.

---

1. Introduction

Recently, the shortcomings of existing AI/ML systems have highlighted various ethical issues that have resulted in calls for ethical consideration in AI/ML. Many of these failures have made global headlines and have prompted widespread public and academic discussion on AI ethics. As a result, numerous guidelines for ethical AI have been proposed, including ones from national and supranational actors such as the EU, alongside scientific ones and ones devised by companies developing AI [1].

However, as with ethics in SE in general, bringing this ethical discussion into practice is a recurring challenge in AI ethics as well. For example, in terms of professional ethics, the ACM guidelines for professional conduct for software professionals [2] are a widely known set of ethical guidelines in the area. Yet, McNamara et al. [3] argue that it has had little impact out on the field, and Gogoll et al. [4] posit that this is the case for Codes of Conduct in SE in general due to their abstract nature. This has also been the case in AI ethics, where various sets of guidelines [1, 5] for ethical AI have been introduced, and whose impact in practice has similarly been called into question [6, 7]. This has also served as the primary motivation behind this paper. While such guidelines with their ethical principles communicate important issues to address, they arguably do not provide the tools for doing so.

AI systems are ultimately still software, although they have been described to be probabilistic as opposed to the deterministic nature of conventional software [8]. To this end, various existing papers have highlighted the importance of SE in ML development [9, 10]. While ML systems contain unique ML components and processes related to them not seen in conventional software systems, these ML components are ultimately just a part of

2

the larger system. Yet, in AI ethics, SE as a whole remains a gap. A myriad of conceptual papers exist, alongside various papers focusing on technical ML solutions related to AI ethics, while papers focused on AI ethics from the point of view of SE remain few and far between. Currently, the conceptual discussion on AI ethics and the resulting principles and guidelines seem to have had little impact on SE practice in the area [7]. This is also highlighted in various conceptual papers discussing the challenges of bringing AI ethics into practice (e.g., [6]).

To facilitate the development of ethical AI systems, we consequently wish to further shift the focus towards SE in AI ethics and the *development* of these systems. In this paper, we propose a novel way of implementing ethics (and AI ethics in particular): Ethical User Stories (EUS). As the name implies, EUS build on traditional User Stories (US), which are an established agile requirements engineering practice in SE [11]. Consequently, the idea of EUS is to help turn ethics into tangible requirements to be implemented into systems, as one way of implementing ethics in practice. While AI ethics has motivated this research, this approach to implementing ethics is not exclusive to AI ethics at all. AI ethics is simply one of the more topical ethics-related issues in SE currently, both in academic literature and out on the field.

The contribution of this paper is two-fold. First, we present the concept of EUS as a novel theoretical concept. We then empirically demonstrate EUS through a two-phase study featuring, altogether, 435 user stories. The empirical portion of the paper provides insights into how EUS function in practice. The specific research questions we tackle in this paper are: **(RQ1)** What are Ethical User Stories?, and **(RQ2)** What are the traits/elements/attributes that frame the EUS concept?

The rest of this paper is structured as follows. Section 2 presents the theoretical background of this study (ethics in SE, user stories, AI ethics). Section 3 presents the concept of Ethical User Stories. Section 4 presents the research approach used to formulate the concept of EUS and validate it in practice through empirical data. Section 5 presents our findings from the empirical studies. Section 6 discusses the theoretical and practical implications of our findings and their limitations. Section 7 concludes the paper.

2. Background

In this section, we discuss the theoretical background of this paper. In Section 2.1, we discuss ethics in SE in general, taking on a more general point

3

of view to our specific topic. In Section 2.2, we discuss the specific ethical context that we focus on in this paper, AI ethics. In Section 2.3, we discuss User Stories (US) on a general level.

2.1. Incorporating Ethics in SE

Incorporating ethics in SE has been a challenge for decades, and became more widely known topic in the 1970s and 1980s. Software and software systems influence individuals, as well as create new contexts, which warrants ethical consideration [12]. Some have argued that they create entirely new ethical issues instead of simply providing existing ones with new application contexts [13]. There have been two traditional approaches to tackling these issues: education (of software engineers) [12, 13], and raising the level of professionalism in the field (e.g., the works of Gotterbarn et al. [2] on the ACM code of ethics).

Over the decades, a number of practical issues have prompted discussion related to ethics in SE or, more generally, IT. Some examples of such issues include: piracy (especially in the early 2000s) (e.g., [14]), security policy compliance in organizations [15], green IT (e.g., [16]), social media use (addiction etc.) (e.g., [17]), ethics in big data [18] and more generally data privacy issues, and, most recently, the downright boom in AI ethics (which we discuss in Section 2.2). Specific problem contexts such as these have provided lenses through which ethics has been approached in SE. Instead of more general, far-reaching attempts to incorporate ethics in SE, it could be argued that much of the discussion on ethics in SE has been focused on specific problems in specific problem contexts such as these.

One example of a more general approach to ethics in software development is Value-Sensitive Design (VSD) [19, 20]. VSD is a design approach that focuses on incorporating human values in software systems and advocates tackling ethical issues through design choices. It has achieved an established status in research. Although it still struggles with practical adoption among software developers, it has remained prominent largely among the design community. Over the years, VSD has received its share of criticism. For example, framing and prioritizing values, alignment with existing methods, and industry adaptation have all been recognized as major challenges for VSD, and therefore incorporating ethics [21].

Whereas in the past ethics has often remained on the sidelines, especially in the eyes of practitioners, the rise of AI ethics [22] has arguably resulted in new-found interest in ethical aspects in SE in recent years. For example,

4

various widely known conferences have recently invited papers on ethical aspects in SE [1]. AI ethics in particular has now become a prominent topic across IT disciplines.

2.2. AI Ethics

Though this paper ultimately discusses a way of implementing ethics that could arguably work in SE overall, our specific application context is that of AI ethics. In part, this is due to our own existing research interest in AI ethics, which has served as one motivation for this paper as well.

AI ethics is a long-standing area of research far pre-dating the more recent advances in ML in the past decade. These recent advances, however, have certainly served to notably accelerate the discussion in the area. [22] Aside from the surge of research interest in the area, this discussion has made its way into the public sphere as well. Various incidents related to AI/ML systems have made the news, resulting in an increase of public awareness of these issues as well. Building on the lengthy history of AI ethics and its current topics of interest, AI ethics as a concept has now been defined and applied by various initiatives [7, 23] to establish rules, principles, guidelines, and standards [24, 25]. In fact, interest in AI ethics has surged to the point where multiple papers simply reviewing the various existing guidelines on AI ethics have been published recently [1, 26], highlighting the number of such guidelines publicly available, in addition to any internal AI ethics guidelines utilized by organizations that are not available for public review. Such guidelines also include ones published by particularly notable parties such as the IEEE [27], and even the EU [28]. These guidelines are typically built around various *principles* that describe what an ethical AI system should be like. Bringing these principles into practice, however, remains a key challenge in the area.

In existing papers [7, 29, 30], we have studied the state of practice in AI ethics and have argued, based on empirical data, that little is currently done in the way of implementing AI ethics out on the field. In particular, the numerous ethical guidelines created to help implement AI ethics (Jobin et al. [24] and Hagendorff [26] review these guidelines in their papers) seem to have had little impact on practice. The numerous tools related to AI ethics

---

[1]Some examples include ICSE2023 (https://conf.researchr.org/track/icse-2023/icse-2023-technical-track) and XP2022 (https://www.agilealliance.org/xp2022/call-for-submissions/call-for-research-papers/)

discussed by Morley et al. [31] seem to also not see much use, based on the studies looking at the current state of practice. While technical ML tools can help tackle ethical issues such as bias, it seems that the motivation to use these tools may often be missing. Moreover, such tools, while helpful and practical for what they are intended to do, are narrow in scope. Far fewer solutions exist for taking into account AI ethics on a project level in SE.

As one potential solution, we have proposed the ECCOLA method for implementing AI ethics in SE [32], and Siqueira De Cerqueira et al. [33] build on ECCOLA with their RE4AI guide. However, we feel that there is still much work to be done in bringing AI ethics into practice in SE. There is still a lack of empirical research in AI ethics overall [34]. Even though the implementation of AI ethics in practice remains a challenge, there is now demand for ways to do so. Policy-makers and companies have begun to hold an interest in AI ethics. Perhaps in part due to the interest in AI ethics, ethics in SE more generally speaking has recently been in the spotlight as well.

There are many open questions that still need addressing in bringing AI ethics into practice. This includes the question of *who* should be involved in implementing AI ethics is also an open question. According to Canca [35] simpler ethical issues could be addressed by developers alone, while more complex issues could necessitate the presence of an AI ethics expert. Hallamaa & Kallikoski [36] propose that ethical issues could be approached primarily from the point of view of safety and risks, focusing on favorable outcomes and avoiding unwanted effects in order to make the discussion more practical and less theoretical. This discussion is still nascent, and this paper presents one more SE practice-oriented contribution aimed at tackling the challenge of incorporating AI ethics and ethics in general into SE.

2.3. User Stories

User Stories are popular tools [37, 38, 39] in agile software development, and specifically in requirements engineering, for facilitating communication between the customer and development team in relation to requirements [40] [41]. User stories act as a bridge between requirements and software features. The idea of a user story is to provide context by describing how a system would be used by a user, and in doing so, to provide guidance for developing the system. From the point of view of requirement engineering, user stories can be linked to the entire requirement engineering process from requirement

6

elicitation and documentation to requirement validation and management [42].

There are varied ideas of what a user story should look like. However, they are typically short sentences, written on an index card, a post-it note, or on a user story template expressed via natural language that emphasize the goals of the system and its usage [43]. Jeffries [41] describes USs using a triple C component scheme, introduced by Wake [44], where the *Cards* are a physical medium, *conversation* as the discussion surrounding the cards, and *confirmation* for tests that verify them [44].

USs are currently often formulated on the basis of a 2001 template by Connextra ² that progresses as follows: "As a [user], I want to [capability], so that [receive benefit]" [40], or some variation of this three-part template. Lucassen [11] proposes a new way of conceptualizing Cohn's [40] user stories by approaching it through four core elements: format, role, means, and end (e.g., means-end as opposed to Cohn's reason-goal).

In this paper, we turn to Cohn's [40] description of user stories, using the aforementioned three-part-template. This is a common way of approaching user stories in the industry, and we wish to use established practices given that our topic otherwise is novel, to make it easier to approach (as further discussed in Section 3, and as is recommended in [32]) Cohn [40] describes user story writing as a story creation process and emphasize that the customer or user should be involved in the process from start to finish, contrary to traditional requirements engineering practices in which the user is typically only involved at the start and at the end of the process.

3. Conceptual Model - Ethical User Stories

In this section, we discuss the concept of Ethical User Stories (EUS). In Section 3.1, we discuss the theoretical justification behind the concept of EUS and our motivation for pursuing it (i.e., the 'why'). In Section 3.2, we present the concept of EUS and explain how EUS are utilized in practice (i.e., the 'what' and 'how').

This conceptual model is based on the empirical studies discussed later in this paper in Sections 4 and 5. In the rest of this paper, we describe the process that ultimately led to the formulation of the concept described

---

7

here, through multiple sets of data. This concept is the primary contribution of this paper. However, through the empirical data we also provide some interesting observations and lessons learned to suggest further practical and theoretical implications, in addition to using it to provide support for this concept.

3.1. Theoretical Justification and Motivation

Our motivation for exploring the idea of EUS originally stems from our past studies in AI ethics, as we have discussed in Section 1 and Section 2.2. These is currently a strong demand for ethical systems in the field of AI, following global public discussion on the topic. This is highlighted by the large number of guidelines being proposed for ethical AI [1]. There is also a more widespread interest in ethics in SE in general as well in recent years. E.g., privacy is arguably a current mega-trend among users.

However, implementing ethics in practice is often challenging due to the abstract nature of the principles and values it is conceptualized through. For example, AI ethics is a field where bridging the gap between theoretical discussion and practice remains an on-going challenge (as we discuss in [7, 29] among others). This theoretical discussion in the area has resulted in a set of recurring ethical principles (transparency, fairness, etc.), which highlight various potential ethical issues [1, 6]. However, these principles continue to provide little guidance in terms of tackling these issues in practice. In an existing study, we proposed a method for implementing AI ethics in order to help bridge this gap [32]. Yet much more work is still arguably needed in this regard, as empirical studies in AI ethics in SE remain scarce [34]. Thus, we propose the concept of EUS as one way of approaching ethics in practice in SE.

Implementing ethics in practice in SE is a challenge on a more general as well and is hardly limited to the specific area of AI ethics. AI ethics has simply been the most prominent driver of ethical consideration in SE in recent years. Ethical consideration is not a natural part of SE. Developers do not normally take ethics into consideration, and ethics is not present in SE practices and methods typically used in SE. Even when pushing developers to take into account ethical guidelines (the ACM Code of Ethics [2]), McNamara et al. [3] saw little change in their ways of working. In part, this also highlights that it is difficult to approach ethics through abstract guidelines and principles, as is also the case in AI ethics [6, 7].

8

Converting abstract principles into practice is challenging. Even if relevant values are recognized through principles, if such guidelines do not provide actionable ways of implementing them, the responsibility of doing so lies entirely on the developers and the organization developing the system at hand [7].

We argue that, in order for ethical principles to get implemented, ethics needs to become a part of the SE process [32]. It is ultimately the developers, and the development team more generally, whose responsibility it is to somehow implement these ethical issues into the system being developed. Though the demand for ethics may come from the customer or other project stakeholders, the technical implementation lies on the developers. Thus, ethics needs to be present in SE practice, either through novel practices or existing ones. On this note, Zuber et al. [45], calling for well-integrated ethical deliberation in software development, suggest that agile processes are particularly suitable for integrating them in the field.

With the concept of EUS, we approach ethics through Requirements Engineering (RE). In SE, requirements drive development. Requirements describe what the system should do and are a tool for formulating features to be implemented in practice, based on customer or other stakeholder requirements. If we wish to see ethical issues being taken into account in development, we argue that they should be converted into requirements. In this regard, we have turned to user stories, which are a well-established RE practice for formulating requirements (as we have discussed in Section 2.3).

We have opted to utilize user stories for several reasons. First, user stories have been utilized in the past to introduce new elements into SE (e.g., quality [46]), and have also been explored as an avenue for implementing ethical consideration in SE [47]. Secondly, user stories being a widespread practice in RE is also beneficial in that it is a familiar practice for many developers and organizations. By utilizing an existing, widely known practice, we wish to make the implementation of ethics in this manner more approachable for development teams. To this end, Zuber et al. [45] also argue that the prevalence of Agile development eases the integration of ethical thinking through familiar methods, and the structures in Agile development empower developers and thus foster ethical thinking. Thirdly, devising user stories requires discussion among developers and other stakeholders involved in devising them. Ethics is not straightforward and involves deliberation. Discussion and deliberation are consequently important when looking to determine how to approach ethical issues in a given (SE) context. Finally,

9

writing stories to illustrate requirements and features can help make ethics more approachable for developers by illustrating what types of situations can involve ethical issues in practice.

3.2. What is an Ethical User Story and How to Devise Them

Ethical User Stories are a practice that enables ethically motivated features to be implemented into a software product. They are user stories intended to formalize and help tackle ethical issues in SE, by using an ethical framework as the lens for doing so. EUS help bring ethics into SE practice by utilizing an existing RE practice (user stories) to convert ethics into tangible requirements.
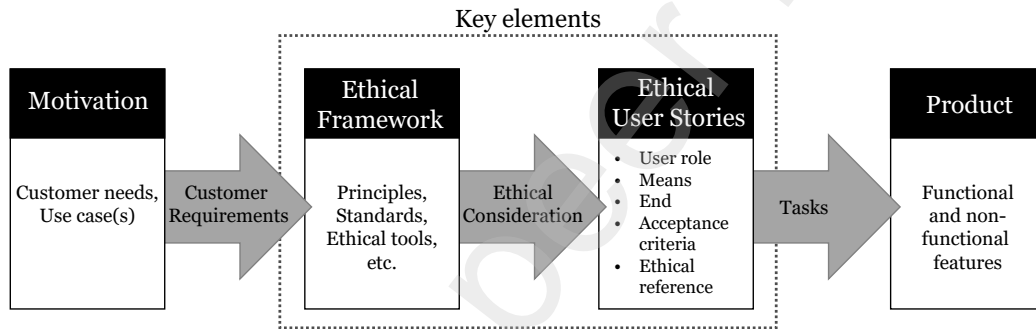

Figure 1: Conceptual model of Ethical User Stories

Figure 1 describes the process of devising EUS. In practice, EUS are highly similar to traditional user stories, as seen in Figure 2. The main differences come from the process used to devise them (Figure 1, as well as the contents of the stories (which are related to ethical issues). Below, we cover this process in more detail by discussing the individual elements (i.e., (1) motivation, (2), ethical framework, (3) (ethical) user stories, (4) product)).

Motivation. A motivation to develop an ethical system is required as a starting point for devising EUS. This motivation, though, can also have a practical impact on the development. For example, we have seen, as a part of our research endeavors, industrial cases where it was the (external) *customer* who was concerned with developing an ethical AI system when commissioning one. Thus, the customer was the one pushing the development team(s) to take into account ethical issues and tackling these issues in practice was something the stakeholders had to handle among themselves. In this

10

fashion, the stakeholders need to determine, together, what they wish to achieve in terms of ethics by formulating requirements for the system. To this end, motivation also involves considering *what* (kind of system) is being developed, and this has a direct impact on what type of ethical framework should/could be selected for the project.

Ethical framework. An ethical framework is required to define relevant ethical issues in any given project context. An ethical framework provides the lens through which ethics is approached in the project. This ethical framework can be any ethical tool: a method, a set of guidelines, or even one as abstract as a an ethical theory. The purpose of the framework is to help determine what is 'ethical' in the given context, and to provide direction on which ethical issues to take into account. The framework should be as closely related to the given context as possible. E.g., one should use an AI ethics framework when working with AI/ML systems, as opposed to a more generic one.

(Ethical) user stories. As artefacts, EUS closely resemble traditional user stories. They contain all the elements typically present in user stories, e.g., the typical three-part-template "As a [user], I want to [capability], so that [receive benefit]". We only propose the addition of an ethics-related component, *ethical reference*, that links the user story to the framework at hand. The ethical reference indicates what part(s) of the ethical framework(s) the particular user story is related to. E.g., if using an AI ethics related framework, one might document which AI ethics principle a given EUS is related to. Producing a paper trail and documenting trade-offs is an important part of ethical consideration [32].

Additionally, in terms of their contents, EUS should tackle the potential ethical issues identified based on the framework of choice. We explore this in more detail in the empirical portion of this paper in Section 5.

Product. Based on the EUS, functional and non-functional requirements are determined, which are then implemented into the product. As a result, a system that takes ethical issues into account is (ideally) developed. However, benchmarking whether a system is ethical or not cannot be done through EUS alone. Such benchmarking is context-dependent and requires a framework, and overall presents various challenges, e.g., in AI ethics. We consider this an interesting topic for future research, as we highlight later in this paper.

Having introduced the concept of EUS and covered the creation of EUS as a process, we now elaborate further on EUS as an artefact through a practical example of an EUS. Figure 3 showcases an EUS written as a part

11

Project:

| Project week: | Date: |

| Priority: | Time estimate: | Ethical reference: |

Story title:

User Story description:

   As a *< type of user >*

   want to *< perform type of task >*

   so that I can *< achieve some goal >*

Acceptance Criteria:
What needs to be done, or
when can the user story be considered done?

Figure 2: Ethical User Story

of a real industrial case. In this case, an AI/ML-enabled smart terminal system was being developed for use in harbor environments for passenger flow management and estimation. The customer was concerned with making the system ethical and ECCOLA introduced as a tool for facilitating ethical consideration in the project. ECCOLA was used to direct the creation of (ethical) user stories to formulate requirements for this system.

4. Research Design

While the purpose of this paper is to introduce, in-depth, the concept of EUS as a novel way of implementing ethics in practice in SE, we consider it relevant to also present some empirical validation for the concept. In doing
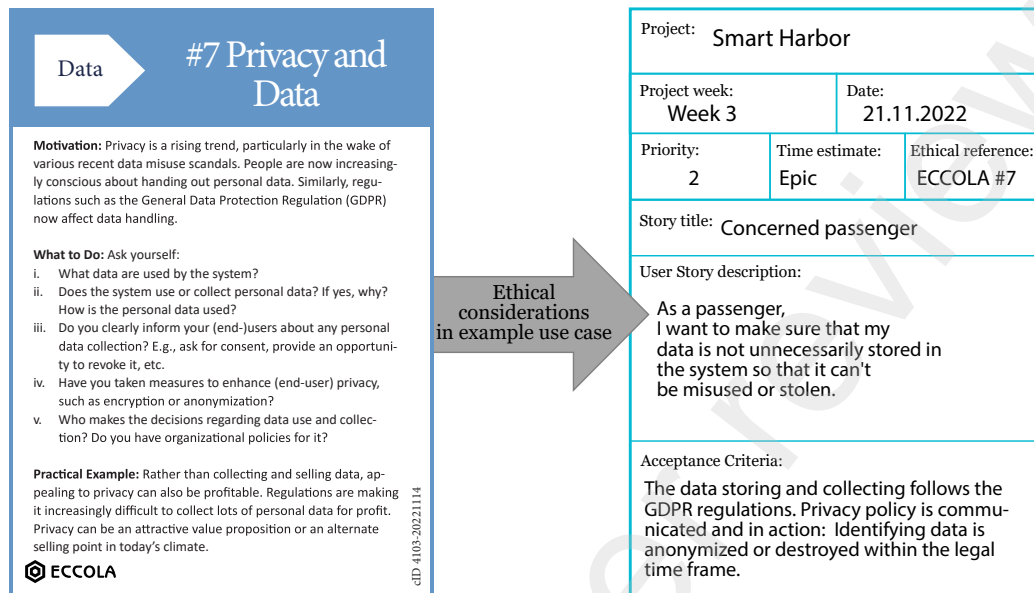
12

Figure 3: An EUS example from an industrial project, using ECCOLA as framework.

so, we also further demonstrate how EUS can be written and what they could look like. For this empirical validation, we have conducted a holistic multiple case study, utilizing different sources of data. The general research approach we have used is Design Science Research (DSR). In this section, we describe this DSR process and its three iterations (referred to as phases).

In Figure 4, we illustrate this research process and its phases thus far. In the process, we cite existing studies we have published on the topic where applicable. The empirical validation and application of EUS is still an ongoing process, and we continue to explore the use of EUS in different contexts, including industrial ones. In this paper, we wish to present the concept of EUS in-depth. I.e., what are EUS, and how do you write them? We also discuss results and lessons learned from these existing papers in relation to this point of view.

The rest of this section is split into three subsections. Each subsection covers one phase, and thus there are three subsections, one for each of the three phases seen in Figure 4.
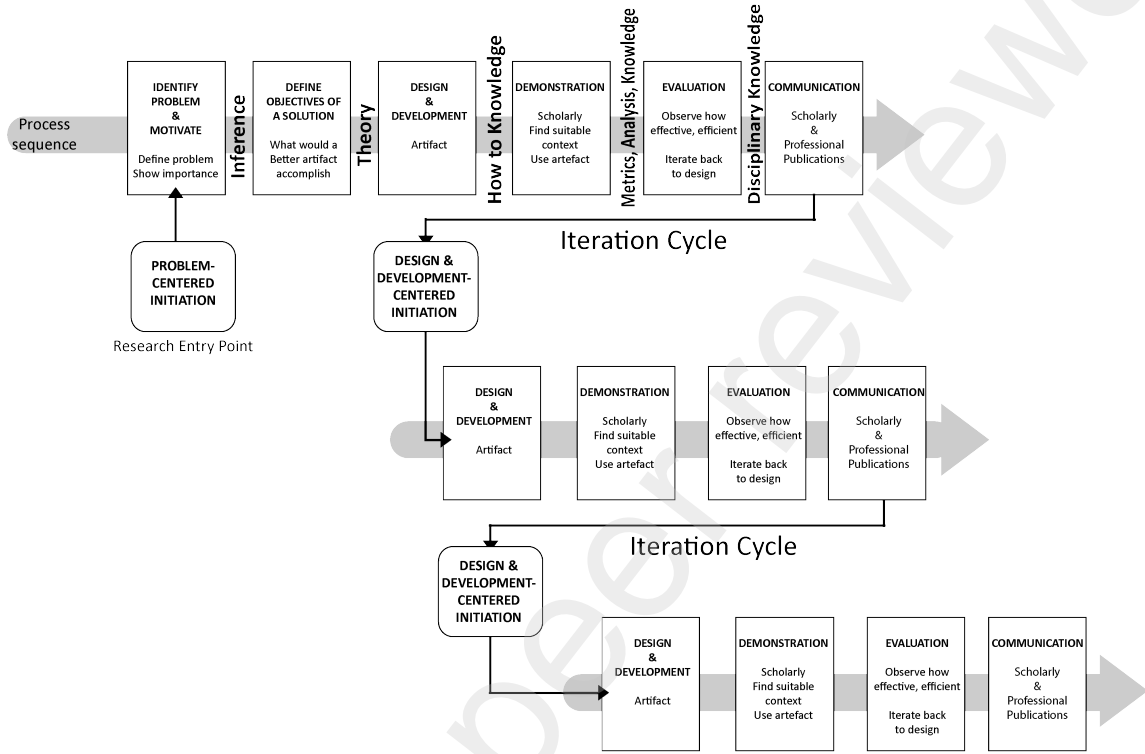
13

Figure 4: EUSdevProcess

4.1. Phase 1

Identify problem & motivate. We have extensively discussed the problem context and motivation in this paper in Sections 1, 2, and 3. In brief, the lack of ethical consideration in AI/ML development is an on-going problem that is becoming increasingly relevant the more ubiquitous such systems become. While this has been our primary motivation for this study, incorporating ethics into SE more generally is also a recurring challenge acknowledged in extant research.

Define objectives of solution. This has also been covered in this paper, particularly in section 3. We aim to create an artefact to help incorporate ethics into SE processes. The artefact should be approachable and familiar to software developers, and thus we turned to existing SE practices (user stories) for inspiration.

Design & development. We have discussed the concept of EUS extensively in Section 3. However, at this stage of the process, going into Phase 1,

14

we were still simply experimenting with the idea of utilizing user stories to implement ethics in SE and had no clear artefact in mind for what an ethical user story should be like. Unsure of whether this was plausible in the first place, we conducted an initial, exploratory study into the topic, which we discuss next.

Demonstration. We first began to explore the utilization of user stories to implement ethics in SE in 2021. Data for phase was collected through a practical, project-based course on Information Systems (IS). The participating students were master level students of IS and Computer Science (CS). These students formed 15 project teams in total with 3-5 members per team, and, during the course, worked as teams on a real-world case given by a client company.

The project duration was six weeks. During this time, the students received five assignments, one each week after the first week's introductory lecture. These assignments comprised two parts: non-technical and technical. The non-technical part was the focus of this study and formed the basis of data collection. User stories were discussed during the lectures to familiarize the teams with the practice of producing them. Then, during the following weeks, the teams produced user stories and worked on implementing them. The first ones were produced for the second week, and, in the following weeks, each week of the project featured user story revision, creation of new user stories if/when applicable, and a check to see if user stories were implemented into the product.

The purpose of this study was to explore how ethical consideration could impact user story creation. Thus, we introduced an ethical framework into the process. To this end, these 15 teams were split into two groups: (1) the control group, that received generic instructions on how to write user stories, but no ethical framework, and (2) the ECCOLA group, who received an ethical framework to utilize in the form of some cards from the ECCOLA method [32] in addition to instructions on how to write user stories. There were six teams in the control group and nine teams in the ECCOLA group. The purpose of this split was to see what impact the use of an ethical framework had on the user stories being produced. Altogether, the teams produced 298 user stories during the project [3].

---

[3]These are available online in an external repository: https://doi.org/10.6084/m9.figshare.14210753

15

Evaluation. These 298 user stories were analyzed using the INVEST model. The INVEST model [48] is a tool for evaluating the quality of user stories through a scoring system. The name of the tool is an acronym that contains the list of attributes to be scored. i.e., according to INVEST, user stories should be: [I]ndependent, [N]egotiable, [V]aluable to purchasers/users, [E]stimable, [S]mall, and [T]estable. This scoring was done by three teachers and/or researchers. The purpose of the INVEST evaluation was to understand whether user stories dealing with ethical issues could be of high quality as well, as it has been argued [40] that user stories have to be of high quality to get implemented at all.

In addition to utilizing INVEST as a framework for analysing the data from Phase 1, we also looked at the contents of the user stories more generally through a grounded theory approach, which was used to code the data. This was done in order to make more qualitative observations about the data as well, given the exploratory nature of the topic.

Communication. We published initial results about the use of user stories to implement ethics in 2021 in reference [49]. This existing paper further details the study described above, including the coding process and its results. In Section 5, we discuss further lessons learned from this paper that relate to the concept of EUS and their utilization in practice.

4.2. Phase 2

Design & development. Based on the data from Phase 1, we began to consider EUS an actual possibility for implementing ethics, prompting us to conduct further studies on the topic. The DSR process thus continued iteratively, as we returned to the *design & development* stage of the process. The *problem & motivation* and the *objectives of the solution* remained the same as in phase 1. In Phase 1, we had the pre-conception that an EUS was simply a human-centric user story. Going into Phase 2, we became more interested in the effects of the ethical framework being introduced into the evaluation process. We began to consider ethical user stories closely linked with the ethical framework used to devise them to the point where it seemed like a defining aspect. I.e., in order to determine whether a user story is an EUS or not, it needs to contain elements from an ethical framework. Still working to explore and define the concept of EUS, we continued to be interested in understanding how EUS differ from conventional user stories in Phase 2.

16

Demonstration. The demonstration of Phase 2 was carried out in a somewhat similar manner to that of Phase 1. Master level CS and IS students were tasked with producing (E)US, and were once again split into a control group who utilized no ethical framework, and the ECCOLA [32] group that utilized one. These were students from Phase 1 who were already familiar with writing user stories. The main differences compared to Phase 1 were as follows. First, the (E)US were produced by individuals rather than teams. Secondly, the (E)US described requirements for a hypothetical system rather than for a real-world case. Thirdly, we asked the students to evaluate how much time they spent on devising the user stories. Fourthly, the user stories were produced independently at one's own pace, rather than in a course project environment. In this fashion, 137 user stories were produced.

In Phase 2, we were interested in any observations related to EUS as an artefact or a process, and as mentioned earlier, particularly in relation to the ethical framework and the contents of the (E)US. Additionally, we explored a time aspect: would it take longer to devise EUS compared to traditional user stories? This would be relevant going into industrial use, and indeed, the purpose of Phase 2 was to also provide further validation for the concept of EUS before moving on to industrial evaluation.

Evaluation. In Phase 2, we focused on the *contents* of the 137 (E)US. The user stories were evaluated on their contents to determine whether they focused on ethical issues related to the system or not. We compared such user stories to other user stories to further explore the idea of EUS.

The evaluation was carried out in two stages. In the first stage, three separate evaluators evaluated all 137 user stories on a scale of 0 to 2 (where 0 = ethical component not present, 1 = ethical component present, and 2 = ethical component unclear). If two evaluators agreed on a score of (be it 0, 1, or 2), this was the final score for the particular user story. On the other hand, US with no such score consensus were discussed among the evaluators in order to reach a consensus, while reflecting on what it was that caused the conflicting scoring.

In the second stage, for further rigour, two additional evaluators worked independently to score the same 137 user stories. Once finished, all five evaluators went over the scores together and discussed any unclear cases to come to a final conclusion for each individual user story. In this second stage, the final scores assigned to each user story in the first stage were compared to the evaluations of the two individual evaluators of stage 2, so that, like in the first round, two out of three scores had to match. The purpose of this

17

second stage was to further understand what types of elements or contents in the user stories caused disagreements in the scores.

Communication. The results of this phase have not been communicated in an existing study. We discuss them in this paper in Section 5. Most importantly, the results were looked at among the research team and motivated us to move on with utilizing the concept of EUS in an industrial setting in Phase 3.

4.3. Phase 3

Design & development. Based on the results of Phases 1 and 2, we began to study EUS in an industrial context. At this stage, we officially added the *ethical reference* as a component to the EUS concept based on Phase 2. The purpose of the ethical framework is to provide a definition for what is 'ethical' in the given context when looking at the contents of a user story. Moreover, the ethical reference serves to directly link a user story to a component of the framework, providing transparency to the process (which is considered pivotal in AI ethics for example [1].

Demonstration. In Phase 3, we began to utilize EUS in an industrial context after exploring it in a university course setting in Phases 1 and 2. This industrial setting was a project between ten industrial partners and five research institutes, which was led by an R&D company. The project was a maritime industry project that focused on utilizing AI/ML in a smart terminal context through three use cases: (1) ship turnaround, (2) truck traffic, and (3) passenger flow. The project also concerned with AI ethics and looked at ways to implement it into such solutions. To this end, we explored the implementation of AI ethics through EUS in all three use cases. As was the case in Phases 1 and 2, we utilized the ECCOLA method [32] as the ethical framework here. Otherwise, the process was split into two stages.

In the first stage, three virtual workshops focused on devising EUS were organized, one for each use case. In these workshops, the use case owner(s) were asked to choose the themes from the ethical framework (i.e., the EC-COLA cards) that were best suited for their use case. The selected cards were then placed on a digital whiteboard and the workshop participants were divided into groups. These groups were divided into breakout rooms and each room contained industrial partners and one researcher familiar with the EC-COLA method and AI ethics in order to facilitate the discussion if/where needed. The participants also received instructions on how to use the ethical framework beforehand. The cards were used to direct and facilitate ethical

18

discussion from the point of view of the project and the system(s) being developed. The participants made 367 notes based on the discussion and based on their previous experiences with the project. From the point of view of the project, the purpose of the workshops was to address any further ethics-related issues that had not already been addressed earlier in the project.

In the second stage, 8 researchers from the University of Jyväskylä (who worked as developers in the project) began to devise EUS based on the workshop notes from the first stage. This work was done over the course of multiple online meeting sessions where the participants worked on a shared spreadsheet containing the notes while being on a voice/video call. This continued until all the notes had been exhausted by converting them into user stories. Altogether, 254 EUS were devised. These user stories were also cross-referenced among the researchers to ensure the quality of the user stories, as quality is important for user stories [40].

Evaluation. In Phase 3, we focused on two aspects in our evaluation (1) the process of devising EUS (this time in a real-world industrial setting with various stakeholders), and (2) what do real-world EUS look like? While the students in Phase 1 also devised EUS for a real case project, this was a fully industrial setting with no student data. In terms of the process, we looked at, e.g., how the notes translated into user stories, and whether the presence of a familiar ethical framework could guarantee that a user story became an EUS, etc. with an explorative approach. In terms of more formal data analysis, we utilized a grounded theory approach to code the 254 user stories, looking at recurring characteristics of the user stories. Additionally, we also looked at each user story from the point of view of the framework (ECCOLA) and its themes (cards) to determine whether a single EUS could overlap with multiple themes of the framework, etc.

Communication. In this paper, we make further observations about the concept and process of devising EUS in this paper based on this industrial data (367 workshop notes and 254 EUS). We have published some preliminary results based on a subset of this data in an existing paper [50]. In the existing paper, we only looked at 125 user stories from a single use case (passenger flow).

5. Results and Lessons Learned

In this section, we discuss our results from the three DSR phases detailed in Section 4. Sections 4.1 to 4.3 cover one phase each, while Section 4.4

19

provides a brief look at the current state of this on-going research endeavor and provides some additional observations. d

5.1. Phase 1 Results

In Phase 1, we began to explore the concept of EUS. At this stage, our plan was to simply see whether user stories could be used as an approach to implementing ethics. We did not plan on creating a separate concept for this initially.

As we analyzed the 298 user stories, the idea of EUS began to emerge. We argued in the existing paper published on this data that it served as proof of concept for EUS [49]. We found that some of the user stories contained two elements we considered to be relevant for an EUS: (1) high quality, and (2) human-centricity. Moreover, such user stories were also successfully implemented into the system being worked on in the project. Based on these observations, we considered user stories a potential way of bringing ethical consideration into the SE process. It seemed that EUS could be written with the same skill-set as traditional user stories, and even without an ethical prompt.

While analyzing the data of Phase 1, we placed emphasis on evaluating the quality of the user stories. Cohn [40] argues that quality is key in having user stories actually get implemented. Thus, if an EUS is not of high quality, it will (likely) not get implemented, and thus the implementation of ethics in this manner is likely to fail as well. Based on our analysis of the data of Phase 1 (presented in more detail in [49]), it seemed to be possible to write EUS that are of high quality, tackling this concern. In brief, by scoring the user stories using the INVEST model, we identified many user stories that we considered ethical that were also of high quality on the INVEST scale. Thus, EUS as user stories did not seem to be impractical.

However, at this point our idea of EUS was arguably still fuzzy. We considered EUS, for which we did not have a definition yet, to be human-centric user stories of high quality, as we attempted to look at the concept from a very general point of view. It was not until later that we began to consider the link between EUS and an ethical framework essential (as we discussed in Section 3) in defining the concept. Nonetheless, at this stage we considered our findings encouraging in that implementing ethics through user stories seemed plausible, leading us to continue on this path.

Finding 1: *Implementing ethics through user stories is possible.*

Finding 2: *Ethical user stories can be, and need to be, of high quality.*

20

5.2. Phase 2 Results

In Phase 2, we placed more emphasis on the fit between the ethical framework and the user stories. In our analysis of 137 user stories in Phase 2, we looked at the the user stories through the lens of the ethical framework being used (i.e., ECCOLA [32]). As the user stories themselves did not involve any direct references to the framework, we went over the user stories, scoring them in the manner discussed in Section 4.2 (where 0 = no ethical component, 1 = ethical component, 2 = unclear). Here, the 'ethical component' was considered to be a link to the framework; did the user story deal with the contents of the framework? In doing so, we were also able to look at the differences between user stories scored 0 and 1, as well as what types of user stories received a score of 2.

Based on this analysis, 54 out of the 137 user stories total were considered EUS (i.e., were scored 1). Looking at these 54 EUS, we determined three elements that (at least) comprised an EUS: (1) motivation, (2) presence of ethical components (according to the framework), and (3) end-user focus. However, in Phase 3, we began to doubt whether an end-user focus was in fact required at all, as we discuss in Section 5.3. Whereas (2) is self-explanatory, motivation, here, refers to a user story being motivated by an ethical issue. I.e., even if the user story discussed a straightforward technical feature, it could be an EUS if this feature was motivated by an ethical concern.

Finding 3: *EUS should be motivated by ethical consideration. Even a technical user story can be an EUS if it is motivated by ethical consideration.*

Finding 4: *EUS should contain an ethical reference that denotes what ethical framework and what part of that framework the user story relates to.*

Additionally, in Phase 2, we looked at EUS from a resource point of view. I.e., how long does it to take to write an EUS compared to a traditional user story? For this purpose, the time spent on writing user stories was self-reported by the participants. Here, we looked at the points of data where this time was reported on the level of individual user stories. 21 participants reported time estimates for 102 individual user stories.

On average, based on these 102 user stories, a user story took 18 minutes to create. The shortest time spent on an individual user story was 6 minutes, while the longest time was 37 minutes. However, variances in the time spent were between individuals rather than types of user stories: some participants simply took longer than others to devise user stories, and vice versa. Based on this data, there seemed to be no difference in time spent on EUS compared to traditional user stories.

21

Finding 5: *EUS do not seem to take longer to devise than traditional user stories.*

We considered this an interesting observation going into industrial settings. Arguably, resources are always important from the point of view of companies. It is far more likely for ethical consideration to receive some attention the less additional resources doing so takes from the organization. As we have discussed in Section 3, this was one reason we opted to look at existing SE practices here as opposed to introducing entirely new ones: a familiar practice should take less effort to utilize.

5.3. Phase 3 Results

Going into industrial evaluation in Phase 3, we reflected on our experiences and the data from the previous 2 cycles. In the design & development stage of Phase 3, based on our experiences in Phase 2, we determined that an ethical framework was necessary for defining EUS.

Finding 6: *An ethical framework is required to determine whether a user story is an EUS.*

In Phase 3, we looked at 254 EUS devised as a part of an industrial project, as discussed in Section 4.3. Based on Phase 2, we now introduced a direct link to the ethical framework in the EUS in the form of the *ethical reference* component. I.e., each user story should contain information on what part of what ethical framework it is inspired by. Each of the 254 user stories contained such a component and these were a part of the analysis in Phase 3. Overall in Phase 3, we further continued to look at EUS in order to better characterize the concept, this time in an industrial setting.

Given the real-world industrial setting, the process of devising EUS was also approached from a practical point of view in Phase 3. I.e., which stakeholders should be involved in the process and how could it be carried out? In this case, the process was organized in the manner described in Section 4.3. Based on our experiences, it needs to be highlighted that the motivation of the relevant stakeholders to implement ethics is crucial, as much of ethical consideration is based on discussion among stakeholders. The work done with different stakeholders (e.g., use case owners) during the workshops was what enabled the developers to write EUS. In this process, the role of the ethical framework was also crucial as it enabled the participants to have shared understanding of ethics in the given context, and served as a way of directing the discussion towards relevant issues.

22

Finding 7: *Motivation to consider ethical issues is not only important on the level of individual user stories, but on the level of the entire project.*

Finding 8: *Ethical consideration involves discussion, and the choice of an ethical framework is important in defining relevant issues to discuss.*

Moving from the process to the EUS themselves, we can make some observations based on the grounded theory analysis. The introduction of the ethical reference made the resulting user stories directly linked to the ethical framework. Each user story was written based on an element of the framework and converserly could be traced back to it through the direct reference.

However, as we analyzed the user stories, it seemed that a single user story could involve elements from different components of the framework. In this case, we could, in some cases, assign multiple ECCOLA cards to a single user story, even though it was motivated by just one. Though a potentially interesting observation, it is more related to the framework and AI ethics principles than the concept of EUS.

Moreover, during the process of writing these 254 EUS, we characterized three types of EUS: (1) human-centric EUS (e.g., personnal training, user needs, system goals, regulations), (2) EUS with a technical focus (e.g., testing, back-ups, logistics), and (3) EUS focused on data and information (e.g., data privacy, data handling). Thus, contradicting with our initial ideas from Phase 2, a user story could be an EUS without being human-centric, as long as it ethically motivated and can be linked to the ethical framework. I.e., EUS can be very feature-oriented and technical as well, largely externally resembling a conventional user story.

Finding 9: *Human-centricity is not required of an EUS, although the motivation behind the EUS still needs to have such a focus, even if the contents are technical in nature.*

5.4. Results Summary & Current Situation

In Table 1, we summarize our findings from the DSR process we have presented in this section. Together, these findings have resulted in the EUS concept we present in Section 3. Some of the findings also provide some practical insights into how EUS could or should be utilized.

At this stage, we feel that the EUS concept itself has been validated and is mature enough to be presented through this paper. Our research on EUS, however, continues, with a focus on their use in practice in industrial settings. We feel that there is still work to do on how EUS. In particular, we are still interested on understanding how EUS are prioritized and actually

23

#	Finding Description
1	Implementing ethics through user stories is possible.
2	Ethical user stories can be, and need to be, of high quality.
3	EUS should be motivated by ethical consideration. Even a technical user story can be an EUS if it is motivated by ethical consideration.
4	EUS should contain an ethical reference that denotes what ethical framework and what part of that framework the user story relates to.
5	EUS do not seem to take longer to devise than traditional user stories.
6	An ethical framework is required to determine whether a user story is an EUS.
7	Motivation to consider ethical issues is not only important on the level of individual user stories, but on the level of the entire project.
8	Ethical consideration involves discussion, and the choice of an ethical framework is important in defining relevant issues to discuss.
9	Human-centricity is not required of an EUS, although the motivation behind the EUS still needs to have such a focus, even if the contents are technical in nature.

Table 1: Findings

24

implemented in industrial settings. We further discuss this in relation to future research suggestions in Section 8.

6. Discussion

In this paper, we began to tackle a practical challenge in the field of AI ethics: bringing ethics into practice in SE. AI ethics is a field where practical implementation is an on-going challenge [7]. Despite various guidelines being proposed to help implement ethics [1], utilizing these in practice is difficult [7], and existing tools are largely either ML techniques of narrow scope or guidelines [31]. Taking ethics into account is also a challenge in SE overall [4, 47]. AI ethics is simply the most prominent field in this regard currently.

Based on our results, we argue that Ethical User Stories (EUS) (see Section 3) present one way of bringing ethics into practice in SE. By leveraging an existing SE practice with minor changes, we are able to use a familiar practice that is easily understood by developers, making it more approachable. In our DSR process, we have seen that EUS can be produced like any other user story, even if the individuals producing them do not have a notable background in ethics. Overall, we have demonstrated that user stories can be utilized to convert ethical consideration into tangible requirements through the concept of EUS. This concept has been developed iteratively through three phases of DSR. In doing so, we have provided a contribution that may help in bridging the prominent gap in AI ethics discussed in various extant studies (e.g., [6, 7]). This contributions also extends to implementing ethics in SE in general, as EUS are not a practice specific to AI ethics.

While we have introduced the practice of EUS in this paper, the processes for utilizing it and for implementing ethical consideration into SE overall are still an open question. The workshop oriented approach for industrial settings that we utilized in Section 5.3 involved both developers and AI ethics experts and comprised of two stages (the initial workshop with both parties present and the later sessions for converting the workshop notes into user stories among the AI ethics experts). This is but one way of devising EUS. For example, Canca [35] suggests that simple ethical issues could be tackled by developers alone, while more complex ones could necessitate the presence of AI ethics experts.

Seeing as all organizations arguably do not have AI ethics experts at their disposal, utilizing practices (and processes) that developers alone could potentially utilize successfully could make AI ethics more approachable to

25

a wider range of organizations. Utilizing familiar agile practices helps in this regard [45], which has been the case here as we build on an existing practice: user stories. In Phases 1 and 2 of the empirical portion of this paper EUS were successfully devised without the involvement of AI ethics experts. We consider a suitable ethical framework to be particular important in this regard, as a complex and abstract framework is likely to be more difficult to utilize without the presence of ethics experts with domain knowledge.

The choice of an ethical framework is a notable issue in and of itself. First, finding and selecting a suitable framework can be difficult without involving experts (and possibly even for experts). Secondly, there is a lack of established ethical frameworks both in AI ethics [6] and SE in general, which is the root of this problem. Nonetheless, as ethics is easily a vague and broad topic without a framework to guide discussion, selecting any framework is arguably better than trying to rely ethical assumptions and personal convictions.

Let us return to the two research questions (RQs) we outlined at the start of this paper: (1) what are ethical user stories?, and (2) what are the traits/elements/attributes that frame the EUS concept? Through the DSR process (Sections 4 and 5) utilized to define and validate the concept, we have answered both of these research questions in Section 3. In brief, an ethical user story is *a user story related to ethical issues or themes present in a given ethical framework, which transforms them into features of a software product.*

We answer the second RQ through the process of devising EUS (Figure 1), which highlights the elements needed to devise EUS. First (in addition to the EUS themselves that are implemented into the product), a motivation to implement ethics, on both project-level and on the level of individual user stories, is needed. Secondly, an ethical framework is required to define what is 'ethical' in the given context, and to direct ethical consideration and discussion towards relevant issues. Thirdly, the resulting user stories need to include references to this ethical framework, and their contents should reflect those of the framework in order to be considered EUS.

6.1. Practical Implications

The EUS present an approachable way of bringing ethical consideration into SE through a familiar practice. The concept is presented in detail Section 3. Here, we present practical insights we have gained from our own experiences with EUS this far over the three DSR phases described in this paper.

26

Select a suitable ethical framework. The ethical framework is key in writing EUS. It (1) defines what 'ethics' means in the given context, (2) provides you with an idea on what kind of ethical issues you should be tackling. While an ethical framework can even be a classic ethical theory, it should be something more context-specific when possible. For example, if you are interested in AI ethics, it would ideally be something related to AI ethics (e.g., guidelines [1], or the ECCOLA method [32] we have used in this paper). One more general-purpose ethical framework could be the RESOLVEDD strategy [51, 52], or the recent IEEE standard for addressing ethical concerns during systems design [53].

Discuss, discuss, discuss. As ethics remains a novel topic in the context of SE, it necessitates consideration, which as achieved through discussion. Instead of trying to directly go to tangible requirements, it would be beneficial to have a more open discussion related to the system and its use context. Use the ethical framework to direct the discussion: what issues pointed out by the framework seem relevant for *your* development context? Moreover, discussion is at the very core of ethics in general, as there are seldom any 'right' answers.

Make notes. Notes leave a paper trail (and it is important to leave a paper trail in ethics [7, 32]), but more importantly, the notes serve as a basis for ethical user stories. Notes are a low effort way of bridging the gap between the ethical framework and user stories. They help you formulate your ideas, while also making them visible for others taking part in the discussion. Notes can also be iteratively combined and modified as the discussion progresses, and multiple notes can converge into one user story (as we have discussed in Section 4.3).

Write ethical user stories. Your notes serve as a natural starting point for devising EUS. The EUS themselves largely resemble traditional user stories: the difference is that they tackle ethical issues you have highlighted in your discussion. The only practical difference is that they should contain a reference to your ethical framework (e.g., which ECCOLA [32] card you used for that user story, if you used ECCOLA as the framework). User story quality (e.g., as measured with the INVEST model [48]) is important for ethical user stories practice, quality in user stories usually translates into practical, actionable user stories. Even though you are dealing with ethical issues, your user stories should be something your developers can work with.

27

6.2. Limitations

As we have utilized different types of data collection and analysis approaches in the three DSR phases of this study, there are also various limitations that should be acknowledged. We would like to emphasize, however, that the main contribution of this paper is the concept/practice of EUS. The role of the empirical studies has been to validate the feasibility of this novel approach to ethics in SE through its use in practice. Thus, while the evaluations have their shortcomings, we argue that they nonetheless have served, at very minimum, as a proof-of-concept for EUS.

Perhaps the most prominent limitation stems from the way we have utilize ethical frameworks to devise EUS in this paper (the process of which we have discussed in Section 3). In this regard, the limitation is that, in all three phases of the study, we have utilized the same ethical framework: ECCOLA. There were two reasons behind this: (1) all the empirical contexts were related to AI/ML and we were not aware of any other suitable framework for AI ethics in particular (seeing as, e.g., AI ethics guidelines are considered to be poorly suited for practical use [6]), (2) we developed ECCOLA and we were interested in potentially further studying the use of ECCOLA in the process as well. Nonetheless, the framework plays a notable role in devising EUS, as we discuss in Section 3, and thus it would be of interest to explore how different frameworks might affect the process and the resulting EUS. For example, if the ethical framework is difficult to understand, the process of devising EUS based on it is likely to also require more effort, and may result in lower quality user stories. We highlight this as a future research suggestion in Section 7.

Evaluating the quality of user stories presents another limitation. In Phase 1 of the DSR process (Section 5.1), we utilized the INVEST model to evaluate the quality of the user stories in this phase. We selected INVEST due to it being a tool promoted by the Agile Alliance [48], although ultimately it is but one of different frameworks for doing so. It nonetheless provided us with some metric of evaluating whether EUS could be 'good', or practical, as many of the quality attributes of INVEST are related to the user story being practical.

Moreover, during the process, we reflected on the relevance of concept validity in this context. I.e., how to measure what is ethical, and to that end, how to determine the presence of ethical consideration in a user story. Ultimately, based on the DSR process, we chose to tackle this issue by introducing the use of an ethical framework as a requirement for devising EUS.

28

The ethical framework is now used to determine what is 'ethical' in a given context, and adherence to the framework needs to be evaluated while evaluating the ethical consideration in a user story. While, arguably, it is possible to write a user story that contains ethical consideration even without actively using an ethical framework to do so, we consider the systematic implementation of ethics to require intent, which actualized through the use of an ethical framework. As ethics tends to be a vague concept, a framework provides the means to shared understanding of ethics in a given context.

Finally, in this paper, we have not studied how the user stories have ultimately impacted the systems being developed in the different DSR phases (the user stories of Phases 1 and 3 were written for real-world systems being developed). We have relied on the generally accepted assumption that high quality user stories are more likely to get implemented, and in this regard, we have confirmed that EUS can (and should) be of high quality. This is something we wish to focus on in future studies on the topic, while the purpose of this paper was to establish the concept and practice of EUS and to provide initial empirical support for it.

7. Conclusions

The recent surge of interest towards AI ethics has also highlighted challenges in implementing ethics in practice in SE. In AI ethics in particular, bringing the various ethical principles discussed in various existing works [1] into practice remains a key challenge in the area [6, 7, 32]. Making ethics a part of SE practice is a larger challenge outside the specific context of AI ethics as well [3], however, and this recent interest in AI ethics seems to have resulted in a heightened interest towards ethical issues in SE at large.

In this paper, we have discussed the implementation of ethics in SE through the use of user stories. We have developed and presented the concept of Ethical User Stories (EUS) (Section 3) through Design Science Research (DSR) (Section 4). EUS is a novel practice that builds on an existing SE practice, i.e., user stories, to provide a way of incorporating ethical consideration as a part of existing SE processes. In building on an established, existing practice, our aim has been to make ethics more approachable for software developers in particular.

The concept has been developed over three DSR iterations. In the process, we have studied 689 user stories produced in different settings, using different data collection and analysis approaches to do so. In addition to using this

29

data to build the concept of EUS, we have highlighted various observations and lessons learned related to the use of EUS in Section 5. At present, EUS provide one potential way of implementing ethics in SE, which we have provided some empirical support for.

Future Research

Further empirical studies into ethics in SE are still sorely needed, especially in the field of AI ethics. Existing studies highlight various problems with bringing abstract, high-level ethical principles into practice in SE (e.g., [3, 6, 7]). However, few empirical studies in general exist in AI ethics [34], and thus we, too, recommend further empirical studies in the area of AI ethics. To this end, we urge further studies into how ethics could be implemented in practice outside the specific area of AI ethics as well. In particular, we feel that further practices and methods that could make ethics a part of SE are needed. Ethical consideration, at present, is arguably not something that is conventionally seen in SE, and is likely to feel distant to developers. This is something that suitable practices and methods could help address.

We also recommend further studies into utilizing user stories as one such practice. This paper simply presents a starting point for doing so, with some empirical support. First, we would recommend utilizing different types of ethical frameworks in doing so. E.g., would AI ethics guidelines, which are often considered impractical [6], function as an ethical framework in devising EUS? Secondly, we personally have plans to continue exploring EUS by looking at how they are ultimately converted into features once they have been written and are being utilized in development.

References

[1] A. Jobin, M. Ienca, E. Vayena, The global landscape of ai ethics guidelines, Nature Machine Intelligence 1 (9) (2019) 389–399.

[2] D. W. Gotterbarn, B. Brinkman, C. Flick, M. S. Kirkpatrick, K. Miller, K. Vazansky, M. J. Wolf, Acm code of ethics and professional conduct (2018).
URL https://www.acm.org/code-of-ethics

[3] A. McNamara, J. Smith, E. Murphy-Hill, Does acm's code of ethics change ethical decision making in software development?, in: Proceed-

30

ings of the 2018 26th ACM ESEC/FSE, ESEC/FSE 2018, ACM, New York, NY, USA, 2018, pp. 729–733. doi:10.1145/3236024.3264833.

[4] J. Gogoll, N. Zuber, S. Kacianka, T. Greger, A. Pretschner, J. Nida-Rümelin, Ethics in the software development process: from codes of conduct to ethical deliberation, Philosophy & Technology (2021) 1–24.

[5] B. Attard-Frost, A. D. los Ríos, D. R. Walters, The ethics of ai business practices: a review of 47 ai ethics guidelines, AI and Ethics (2022).

[6] B. Mittelstadt, Principles alone cannot guarantee ethical ai, Nature Machine Intelligence 1 (11) (2019) 501–507.

[7] V. Vakkuri, K. Kemell, J. Kultanen, P. Abrahamsson, The current state of industrial practice in artificial intelligence ethics, IEEE Software 37 (4) (2020) 50–57.

[8] V. Vakkuri, K.-K. Kemell, J. Kultanen, P. Abrahamsson, The current state of industrial practice in artificial intelligence ethics, IEEE Software 37 (4) (2020) 50–57.

[9] S. Martínez-Fernández, J. Bogner, X. Franch, M. Oriol, J. Siebert, A. Trendowicz, A. M. Vollmer, S. Wagner, Software engineering for ai-based systems: A survey, ACM Trans. Softw. Eng. Methodol. 31 (2) (apr 2022). doi:10.1145/3487043.
URL https://doi.org/10.1145/3487043

[10] G. Giray, A software engineering perspective on engineering machine learning systems: State of the art and challenges, Journal of Systems and Software 180 (2021) 111031. doi:https://doi.org/10.1016/j.jss.2021.111031.
URL https://www.sciencedirect.com/science/article/pii/S016412122100128X

[11] G. Lucassen, F. Dalpiaz, J. M. E. Van Der Werf, S. Brinkkemper, Forging high-quality user stories: towards a discipline for agile requirements, in: 2015 IEEE 23rd international requirements engineering conference (RE), IEEE, 2015, pp. 126–135.

[12] J. H. Moor, What is computer ethics?, Metaphilosophy 16 (4) (1985) 266–275.

31

[13] D. G. Johnson, Computer ethics, Prentice-Hall, Inc., 1994.

[14] T. T. Moores, J. C.-J. Chang, Ethical decision making in software piracy: Initial development and test of a four-component model, MIS Quarterly 30 (1) (2006) 167–180.
URL http://www.jstor.org/stable/25148722

[15] D. Chatterjee, S. Chatterjee, S. Chatterjee, An empirical investigation of the influence of organizational virtues on information technology security policy compliance, in: ACIS 2021 proceedings, 2021.

[16] T. Butler, M. Daly, R. Hackney, Socio-technical transitions towards environmental sustainability through green ict, in: SIGGreen Pre-ICIS 2015 Workshop, 2015, pp. 1–12.

[17] J. R. Drake, D. Hall, J. B. Becton, C. Posey, Job applicants' information privacy protection responses: Using socia media for candidate screening, Transactions on Human-Computer Interaction 8 (4) (2016) 160.

[18] L. Perreault, Big data and privacy: control and awareness aspects, in: CONF-IRM, 2015.
URL https://aisel.aisnet.org/confirm2015/15

[19] B. Friedman, Value-sensitive design, interactions 3 (6) (1996) 16–23.

[20] B. Friedman, D. G. Hendry, Value sensitive design: Shaping technology with moral imagination, Mit Press, 2019.

[21] B. Friedman, M. Harbers, D. G. Hendry, J. van den Hoven, C. Jonker, N. Logler, Eight grand challenges for value sensitive design from the 2016 lorentz workshop, Ethics and Information Technology 23 (2021) 5–16.

[22] J. Borenstein, F. S. Grodzinsky, A. Howard, K. W. Miller, M. J. Wolf, Ai ethics: A long history and a recent burst of attention, Computer 54 (1) (2021) 96–102.

[23] V. Dignum, Ethics in artificial intelligence: introduction to the special issue (2018).

[24] A. Jobin, M. Ienca, E. Vayena, The global landscape of ai ethics guidelines, Nature Machine Intelligence 1 (9) (2019) 389–399.

[25] H. AI, High-level expert group on artificial intelligence (2019).

[26] T. Hagendorff, The ethics of ai ethics: An evaluation of guidelines, Minds and Machines (2020) 1–22.

[27] T. I. G. I. on Ethics of Autonomous, I. Systems, Ethically aligned design: Prioritizing human wellbeing with autonomous and intelligent systems (2019).
URL https://standards.ieee.org/content/dam/ieee-standards/standards/web/docume
$undefinedutm_source$ $=$ $undefinedutm_campaign$ $=$ $undefinedutm_content = undefinedutm_term = undefined$

[28] Ethics guidelines for trustworthy ai (2019).
URL https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trus

[29] V. Vakkuri, K. Kemell, J. Kultanen, M. T. Siponen, P. Abrahamsson, Ethically aligned design of autonomous systems: Industry viewpoint and an empirical study, EJBO - Electronic Journal of Business Ethics and Organization Studies 27 (2022).

[30] V. Vakkuri, K.-K. Kemell, M. Jantunen, P. Abrahamsson, "this is just a prototype": How ethics are ignored in software startup-like environments, in: V. Stray, R. Hoda, M. Paasivaara, P. Kruchten (Eds.), Agile Processes in Software Engineering and Extreme Programming, Springer International Publishing, Cham, 2020, pp. 195–210.

[31] J. Morley, L. Floridi, L. Kinsey, A. Elhalal, From what to how: an initial review of publicly available ai ethics tools, methods and research to translate principles into practices, Science and engineering ethics 26 (4) (2020) 2141–2168.

[32] V. Vakkuri, K.-K. Kemell, M. Jantunen, E. Halme, P. Abrahamsson, Eccola — a method for implementing ethically aligned ai systems, Journal of Systems and Software 182 (2021) 111067. doi:https://doi.org/10.1016/j.jss.2021.111067.
URL https://www.sciencedirect.com/science/article/pii/S0164121221001643

[33] J. A. Siqueira De Cerqueira, A. Pinheiro De Azevedo, H. Acco Tives, E. Dias Canedo, Guide for artificial intelligence ethical requirements elicitation-re4ai ethical guide, in: 55th Hawaii International Conference on System Sciences, 2022.

33

[34] M. Sloane, J. Zakrzewski, German ai start-ups and "ai ethics": Using a social practice lens for assessing and implementing socio-technical innovation, in: 2022 ACM Conference on Fairness, Accountability, and Transparency, FAccT '22, Association for Computing Machinery, New York, NY, USA, 2022, p. 935–947.

[35] C. Canca, Operationalizing ai ethics principles, Communications of the ACM 63 (12) (2020) 18–21.

[36] J. Hallamaa, T. Kalliokoski, Ai ethics as applied ethics, Frontiers in Computer Science (2022).

[37] S. Dimitrijević, J. Jovanović, V. Devedžić, A comparative study of software tools for user story management, Information and Software Technology 57 (2015) 352–368.

[38] I. K. Raharjana, D. Siahaan, C. Fatichah, User stories and natural language processing: A systematic literature review, IEEE Access 9 (2021) 53811–53826.

[39] W. Hudson, User stories don't help users: Introducing persona stories, interactions 20 (6) (2013) 50–53.

[40] M. Cohn, User stories applied: For agile software development, Addison-Wesley Professional, 2004.

[41] R. Jeffries, A. Anderson, C. Hendrickson, Extreme programming installed, Addison-Wesley Professional, 2001.

[42] A. De Lucia, A. Qusef, Requirements engineering in agile software development, Journal of emerging technologies in web intelligence 2 (3) (2010) 212–220.

[43] X. Wang, L. Zhao, Y. Wang, J. Sun, The role of requirements engineering practices in agile development: an empirical study, in: Requirements engineering, Springer, 2014, pp. 195–209.

[44] B. Wake, Invest in good stories, and smart tasks (Aug 2003).
URL https://xp123.com/articles/invest-in-good-stories-and-smart-tasks/

[45] N. Zuber, J. Gogoll, S. Kacianka, A. Pretschner, J. Nida-Rümelin, Empowered and embedded: ethics and agile processes, Humanities and Social Sciences Communications 9 (1) (2022) 1–13.

[46] G. Lucassen, F. Dalpiaz, J. M. E. M. v. d. Werf, S. Brinkkemper, The use and effectiveness of user stories in practice, in: M. Daneva, O. Pastor (Eds.), Requirements Engineering: Foundation for Software Quality, Springer International Publishing, Cham, 2016, pp. 205–222.

[47] P. Kamthan, N. Shahmir, On ethically-sensitive user story engineering, in: Proceedings of the 4th International Conference on Computer Science and Software Engineering, CSSE '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 71–79. doi:10.1145/3494885.3494899.
URL https://doi.org/10.1145/3494885.3494899

[48] What does invest stand for? (2021).
URL https://www.agilealliance.org/glossary/invest/

[49] E. Halme, V. Vakkuri, J. Kultanen, M. Jantunen, K.-K. Kemell, R. Rousi, P. Abrahamsson, How to write ethical user stories? impacts of the eccola method, in: International Conference on Agile Software Development, Springer, Cham, 2021, pp. 36–52.

[50] E. Halme, M. Agbese, J. Antikainen, H.-K. Alanen, M. Jantunen, A. A. Khan, K.-K. Kemell, V. Vakkuri, P. Abrahamsson, Ethical user stories : Industrial study, in: Joint Proceedings of REFSQ-2022 Workshops, Doctoral Symposium, and Posters & Tools Track co-located with the 28th International Conference on Requirements Engineering: Foundation for Software Quality (REFSQ 2022), CEUR Workshop Proceedings, 3122, 2022.

[51] R. S. Pfeiffer, R. P. Forsberg, Ethics on the Job: Cases and Strategies, Wadsworth Publishing Company, 1993.

[52] V. Vakkuri, K. Kemell, P. Abrahamsson, Ethically aligned design: An empirical evaluation of the resolvedd-strategy in software and systems development context, arXiv preprint arXiv:1905.06417 (2019). arXiv:1905.06417.
URL http://arxiv.org/abs/1905.06417

[53] IEEE, IEEE Standard Model Process for Addressing Ethical Concerns during System Design, Standard, IEEE (Sep. 2021).

36