The Use of Context-Free Grammar and Gamification to Reduce User Story Ambiguity and Raise Ambiguity Awareness: A Proposal

Anis R. Amna [1]

[1] *Ghent University, Faculty of Economics and Business Administration, Ghent, Belgium*

Abstract

User story ambiguity has been identified as a serious challenge in Requirements Engineering (RE) due to various linguistic and contextual constraints. These challenges have been addressed by research employing mainly Natural Language Programming (NLP) techniques and Conceptual Modeling. However, existing studies have only limitedly investigated the impact of personal and organizational elements on user story ambiguity and user awareness of ambiguity issues. Therefore, this research aims to mitigate user story ambiguity by developing a knowledge repository for user story writing from different domains. To do so, we aim to investigate the role of cognitive elements to enhance user story clarity. Technically, the use of context-free grammar will be investigated to maintain the quality of the knowledge repository. Finally, for raising awareness of ambiguity issues, gamification concepts (e.g., levels/tags, points) will be applied to monitor user progress.

Keywords

User story, ambiguity, reuse user story, cognitive elements, context-free grammar, gamification

1. Introduction

As a widely known RE artifact, the user story is perceived as a compelling yet challenging artifact in Agile Software Development (ASD). Despite their effectiveness as requirements documentation instruments, user story creation and/or review have also been recognized as challenging processes due to implicit ambiguity. This ambiguity emerges as a result of differences in linguistics expressions and context understanding.

Regarding linguistic issues, user story expression is prone to misunderstanding because one word might have several related meanings regarding specific contexts. For example, the term “key account” might be perceived by marketing as valuable customers, while procurement comprehends the term as an important supplier. *Distributed cognition (DCog)* theory explained that these various understandings emerge due to information processing in people’s minds and associated environment where the knowledge have been acquired and distributed. The personal and organizational aspects that influence knowledge understanding consists of knowledge, skills, and experiences, which is referred as *socio-cultural cognition* [1]. Furthermore, Harris [2] described that the DCog may unconsciously prompts different perceptions in a particular field (i.e., business, IT) through human and machine interactions. This phenomenon has been shown by Wautelet et al. [3] and Jia et al. [4], who demonstrated a

In: F.B. Aydemir, C. Gralha, S. Abualhaija, T. Breaux, M. Daneva, N. Ernst, A. Ferrari, X. Franch, S. Ghanavati, E. Groen, R. Guizzardi, J. Guo, A. Herrmann, J. Horkoff, P. Mennig, E. Paja, A. Perini, N. Seyff, A. Susi, A. Vogelsang (eds.): Joint Proceedings of REFSQ-2021 Workshops, OpenRE, Posters and Tools Track, and Doctoral Symposium, Essen, Germany, 12-04-2021

EMAIL: *AnisRahmawati.Amna@UGent.be*

significant involvement of socio-cultural cognition (i.e., knowledge background, experience) in creating a shared understanding of requirements between user and developer communities.

Besides apparently emanating in requirements elicitation and documentation activities, the literature study we performed indicates that ambiguity also affects requirements analysis and negotiation activities. Regarding requirements elicitation and documentation, ambiguity issues have typically been resolved by employing linguistic techniques. The solutions were generally concerned with NLP algorithms detecting user story similarity [5, 6] and inconsistency [7, 8]. Studies focusing on requirements analysis and negotiation have approached ambiguity in terms of a lack of shared understanding of requirements between user and developer communities. Proposed solutions have generally mapped user stories to (graphical) conceptual models [9, 10]. However, our literature study revealed that the number of studies addressing ambiguity issues related to requirements documentation with user stories is relatively small. Furthermore, this research topic is also immature, as almost half of the proposed solutions have not been empirically verified.

Apart from these gaps, the studies emphasizing linguistic solutions have barely considered human factors (e.g., knowledge, skills, experience) in the search for a solution to overcome user story ambiguity. Therefore, we aim to extend the state of the art by exploring personal and organizational factors that lead to user story ambiguity during requirements elicitation and documentation. Our intention is to investigate how a focus on these factors might help creating solutions for mitigating user story ambiguity and raising awareness of user story ambiguity. To do so, we propose the development of a knowledge repository of user stories, based on context-free grammar for duplication and similarity detection, augmented with a gamification concept (e.g., levels/tags, points) for promoting individual cognitive skills and organizational culture to enhance a shared understanding between user and developer communities.

1.1. Relevance

Requirements ambiguity has been studied from linguistics [11] and contextual [12] perspectives. Compared to the contextual approach, the linguistics approach has drawn smaller attention from researchers. Given that situation, our study aims to develop a comprehensive solution to mitigate user story ambiguity by elaborating socio-cultural cognition that have generally been neglected in these studies. Our solution will take advantage of the DCog theory to investigate which socio-cultural cognition aspects influencing user story writing. We will also exploit NLP techniques to elaborate these finding by constructing a knowledge repository comprised of word glossaries and ontology as a standard terminology for avoiding semantic ambiguity. Finally, a gamification concept will be introduced in the shape of points and levels to observe the adequacy of the solution to improve user awareness for mitigating ambiguity while writing user stories.

2. Research plan

Our research plan comprises of four sections: research problems and questions, research methodology, proposed solution, and novelty of the solution

2.1. Research problems and questions

In order to identify the main problems in user story ambiguity, we conducted a systematic literature review in 2020 (in progress). From the literature study, we extracted 38 documents addressing user story ambiguity for the last 20 years. The study then identified four shortcomings or research gaps:

1. Numerous studies highlighting user story ambiguity have promoted the use of conceptual models, while a smaller number of the studies have proposed a linguistics-based approach. Especially, the latter approach is interesting for avoiding, rather than detecting and remedying, ambiguity in user stories.

2. Only a few studies have addressed socio-cultural cognition (e.g., knowledge background, experience) as a factor impacting the shared understanding of user stories.
3. The existing studies have mainly focused on providing solutions for user story ambiguity for one particular business or societal sector (e.g., e-commerce, university)
4. The existing solutions were limitedly verified and/or applied in a real-world setting.

Given this analysis of state of the art, our research questions have been formulated as follows:

RQ1: To what extent have socio-cultural cognition been investigated in the search for solutions for mitigating user story ambiguity?

RQ2: How do socio-cultural cognition improve/impede a shared understanding of user stories?

RQ3: How can the sharing of contextualized knowledge of user stories help in writing unambiguous user stories and improve the awareness of potential ambiguity issues?

Hence, our study strives to leverage mechanisms related to distributed cognition (i.e., the contextualized knowledge) to reduce ambiguity and misunderstanding of user stories. In order to do so, socio-cultural cognition aspects related to individual and organizational elements will be explored to improve user story clarity and promote a shared understanding of requirements contexts. Semantic ambiguity will be detected and reduced with a knowledge repository based on a glossary, ontology, and context-free grammar. The result will be evaluated through a case study involving a medium-sized enterprise with Scrum practitioners.

2.2. Research methodology

Our research methodology will be implemented in sequential steps as follows. The steps consist of a literature review, data collection and exploratory study, solution development, and validation.

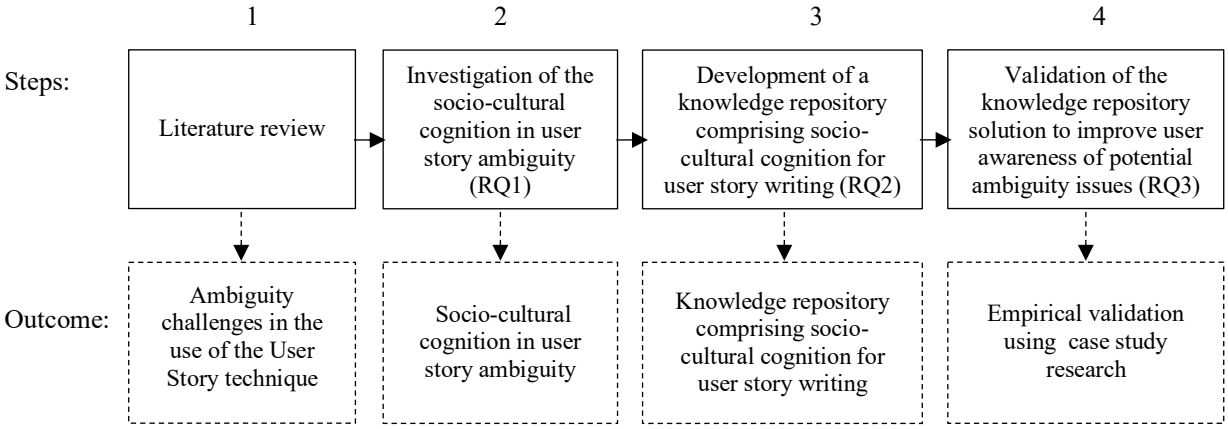

Figure 1: Research methodology

Step 1: Literature review

A literature review is essential to understand to which extent user story ambiguity has been investigated, which solutions have been proposed, and to what extent they have been validated academically or evaluated in practice [13]. Due to the absence of a literature review focusing on user story research, our research started by conducting such a secondary study. The study classified the RE issues on user stories for the last 20 years, the proposed solutions to resolve ambiguity, and validation of the solutions. From this literature review, we obtained insights under what circumstances user story ambiguity has emerged and which solutions have been proposed and validated or evaluated.

Step 2: Investigation of the socio-cultural cognition in user story ambiguity (RQ1)

In the second step, an exploratory study will be conducted to investigate which socio-cultural cognition factors have been studied in user story ambiguity. We choose ERP projects as a growing number of companies implement, extend, adapt, or customize ERP solutions [14]. Furthermore, ERP

adoption has encountered many challenges due to requirements complexity, features misfit, and others [14, 15].

Our research study aims to apprehend individual and organizational determinants to a shared understanding of user story context. Hence, the data collection will be focused on three categories of items: (1) business or societal sector, (2) user profiles and organizational culture, (3) participant qualification. For the business or societal sector, the information will be collected from manufacturing, information technology, and financial services. The information should be presented in user stories or text-based requirements, which has similar components to user stories: *as a [role], I want [goal] so that [reason]*.

Regarding determinants, we intend to conduct an exploratory study to verify which factors of user profile and organizational culture have impacted user story ambiguity related to a shared understanding. As a starting point, we refer to the study of Wautelet et al. [3] and Jia et al. [4], who have confirmed knowledge background as an important factor to a shared understanding. Subsequently, we attempt to investigate working experience (i.e., entry-level, intermediate, mid-level, senior) and managerial responsibility level (i.e., operational, middle manager, senior manager, executive) as we posit that those factors might significantly influence a common concept of user story context. In respect of organizational culture, we determine to explore the organizational structure, communication culture, and employee turnover because these factors have been committed as a crucial elements to ASD adoption [16] (see Table 1). Information on user-profiles and organizational culture will be carefully analyzed to understand the impact of those elements on word selection and grammatical expression in user story writing and/or review.

Finally, we select Scrum practitioners as participants considering that the method is the most widely practiced in ASD [17]. Additionally, we require five years of experience and a geographically distributed team.

Table 1.

Personal and Organization determinants

Qualifications type	Element
Personal	Educational background
	Working experience
	Responsibility
Organizational	Organization structure
	Communication culture
	Employee turnover

Table 2.

Respondent Qualifications

Qualifications	Agile practitioners
ASD method	Scrum
Experience	>=5 years
Organizational Structure	Distributed
Project Management Culture	Pair programming

Step 3: Development of a knowledge repository comprising socio-cultural cognition aspects for user story writing (RQ2)

Next, the requirements from different domains will be structured to ontology, while terminology will be arranged in a glossary. For a terminology having different interpretations across domains (i.e., homonyms), the meanings will be classified accordingly to support different interpretations for the various objects/actions. Furthermore, synonyms will be recorded to facilitate alternative phrases to refer to the same object/action (i.e., polysemy). Context-free grammar will be adopted as a classifier to verify sentence similarity, predict subsequent phrases, and provide alternatives for user story sentences.

With respect to socio-cultural cognition factors, suppose there is evidence of personal and organizational determinants to word choice, the element will be carefully considered a branch of the ontology. Next, corresponding words will be examined to distill the glossary. The recommendation for

user story writing will be provided according to a user profile and project domain that have been recorded into our knowledge repository. User awareness of user story ambiguity will be assessed by comparing the sentence grammar with the standard terminology, and the result will then be converted into points. User improvement of producing unambiguous user stories will be evaluated based on standard vocabulary from the knowledge repository. The shared understanding of user story context will be validated by specifying user stories into product backlogs.

Step 4: Validation of the knowledge repository solution to improve user awareness of potential ambiguity issues (RQ3)

The effectiveness of our solution will be evaluated using case study research. The study will observe the solution performance to improve a shared understanding of user story context and raise user awareness of user story ambiguity. We determine accuracy, efficiency, and consistency as key variables to measure the adequacy of our solution in practice. The validation will be performed by measuring user acceptance and duration to formulate unambiguous user stories with the help of our knowledge repository.

The accuracy of the repository will be evaluated by inviting Agile practitioners to write/review user stories from the given scenario using our knowledge repository. Then, they will be requested to specify those stories into product backlogs. The differences between the backlogs and the expected features will be compared to predict accuracy.

The evaluation of solution efficiency will be performed by measuring time spent by the participants to fulfill the expected requirements of the user stories. If the time to write/review user stories supporting the knowledge repository is less than manual writing/review, our solution will be considered sufficient.

The solution consistency will be verified through the prevalence of the sentences to precisely interpret user needs in the form of the product backlog. To do so, user performance during the reference periods (i.e., points and levels) will be measured.

2.3. Proposed solution

Our solution will be proposed by incorporating socio-cultural cognition as a result of personal behavior (related to user profiles) and the organizational cultures of their working environment as components of an integrated solution to mitigate user story ambiguity. These factors have been carefully selected based on empirical studies confirming the positive influence of those factors to improve a shared understanding of user stories [4], [6]. The factors will be verified through an exploratory study to finalize what cognitive factors have influenced word selection in writing and/or reviewing user stories. The factor will then be inserted as an ontology branch, while the corresponding word will be installed in the parallel glossary.

User story duplication and similarity will be avoided by adopting context-free grammar in the knowledge repository. Finally, user activity will be awarded into a point, and the awareness of user story ambiguity will be monitored through level every time they work using our repository.

2.4. Novelty of the solution

Our research focuses on mitigating user story ambiguity by means of a knowledge repository, which is developed by adopting DCog theory as a theoretical framework to establish socio-cultural cognition aspects in user story writing. Owing to socio-cultural cognition, we determine user profile and organizational culture as the solution attributes because these factors have been limitedly studied in user story research.

The knowledge repository will be developed by reusing user stories on the basis of ERP systems from different business domains. The standard terminology will be extracted from the business domain by utilizing context-free grammar. Then, the terminology will be classified using an ontology. Furthermore, the gamification concept (i.e., points, levels) will also be adopted by virtue of reward and level to raise user awareness of user story ambiguity.

3. Research Methods

In order to provide a solution to mitigate user story ambiguity, our study follows the design science research method by developing a knowledge repository. In order to do so, we will contact software companies to participate in our research. We intentionally set some minimum requirements for our participants (see Table 2.) as we expect diverse working experiences and practices. Then, an online survey will be conducted to collect user profiles and organizational culture data from our participants (see Table 1). User profiles will be collected by applying a purposive sample technique, targeting Product Owners, Business Analysts, System Analysts, and Developers. This information will be useful to initiate a corresponding glossary with regard to word selection and grammatical expression.

Organizational culture data will be compiled from an in-depth interview that will be conducted with Scrum Masters and Managers (e.g., software team managers, product line managers, COO, CIO) as such profiles might best comprehended organizational culture. Additionally, requirements documentation (i.e., user stories) will be solicited from participants as a base for our glossary and ontology. Therefore, the participants are required to supply documentation of some ERP projects from various business sectors. Afterward, the knowledge repository will be developed by reusing user stories, user profiles, and organizational culture parallel with the business domain. Context-free grammar will be used to avoid user story similarity and duplication by detecting similar words or meanings of a new user story. The new user story could be added as new knowledge by the Scrum team after getting authorization from the Scrum Master. Eventually, a case study will be conducted by observing the knowledge repository performance in terms of accuracy, efficiency, and consistency. The accuracy will be assessed by comparing the recommended user stories with expected requirements from users, while efficiency will be measured by calculating time to write and/or review user stories with the repository. The consistency will be evaluated by observing user experience working with our knowledge repository. Furthermore, the improvement of user awareness to user story ambiguity will be monitored through a gamification concept (i.e., levels/tags, points) installed as a feature of our knowledge repository

4. Progress

Concerning the research plan that has been presented in Section 2, the first stage, a systematic literature review, has nearly been completed (see Figure 1). This proposal is structured based on research gaps that have been found during our literature study.

From selected papers addressing user story ambiguity, our study indicated that the ambiguity emerged due to natural language and contextual understanding of the user story. Despite less mature, existing solutions focusing on ambiguity have mostly neglected socio-cultural cognition aspects as crucial elements influencing ambiguity. Our study discloses that linguistics-related solutions have typically been proposed to overcome ambiguity during requirements elicitation and documentation, while the (graphical) conceptual model has generally promoted requirements analysis and negotiation to enhance a shared understanding between user and developer communities.

At this stage, we are in the second stage of our research study. Therefore, we attempt to contact software companies in Belgium to join our research study. Scrum practitioners will be called participants, considering that the Scrum method has been widely adopted. The participants are required to have a minimum of five years of working experience, besides working in distributed teams. On top of that, we expect to conduct an exploratory study to collect user profiles and organizational culture data as a basis of our proposed solution. We also aim to collect user stories or other similar user requirements representation from ERP projects to construct glossary and ontology components for our knowledge repository.

[1] M. Cole and Y. Engeström, "A cultural-historical approach to distributed cognition," in *Distributed Cognitions: Psychological and Educational Considerations*, G. Salomon Ed.: Cambridge University Press, 1997, ch. 1, pp. 1-46.

[2] D. Harris, "Distributed cognition in flight operations," in *Engineering Psychology and Cognitive Ergonomics. Applications and Services. EPCE 2013.* , vol. 8020 of Lecture Notes in Computer Science, D. Harris, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 125-133.

[3] Y. Wautelet, M. Velghe, S. Heng, S. Poelmans, and M. Kolp, "On modelers ability to build a visual diagram from a user story set: A goal-oriented approach," in *Lecture Notes in Computer Science. Requirements Engineering: Foundation for Software Quality. REFSQ 2018*, vol. 10753, H. J. Kamsties E., Dalpiaz F., Ed.: Springer, Cham, 2018, pp. 209-226.

[4] J. Jia, X. Yang, R. Zhang, and X. Liu, "Understanding software developers' cognition in agile requirements engineering," *Science of Computer Programming 178,* pp. 1-19, 2019, doi: 10.1016/j.scico.2019.03.005.

[5] R. Barbosa, A. E. A. Silva, and R. Moraes, "Use of Similarity Measure to Suggest the Existence of Duplicate User Stories in the Srum Process," presented at the 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshop (DSN-W), 2016.

[6] E. Dilorenzo *et al.*, "Enabling the Reuse of Software Development Assets Through a Taxonomy for User Stories," *IEEE Access,* vol. 8, pp. 107285-107300, 2020, doi: 10.1109/access.2020.2996951.

[7] M. S. Murtazina and T. V. Avdeenko, "An ontology-based approach to support for requirements traceability in agile development," *Procedia Computer Science 150,* pp. 628-635, 2019, doi: 10.1016/j.procs.2019.02.044.

[8] P. L. de Souza, A. F. do Prado, W. L. de Souza, S. M. dos Santos Forghieri Pereira, and L. F. Pires, "Improving Agile Software Development with Domain Ontologies," *Information Technology-New Generations 738,* pp. 267-274, 2018, doi: 10.1007/978-3-319-77028-4_37.

[9] L. B. De Araujo and F. L. Siqueira, "Using i∗ with scrum: An initial proposal," iStar, 2016, vol. 1674, pp. 19-24.

[10] M. Elallaoui, K. Nafil, and R. Touahni, "Automatic Transformation of User Stories into UML Use Case Diagrams using NLP Techniques," *Procedia Computer Science,* vol. 130, pp. 42-49, 2018, doi: 10.1016/j.procs.2018.04.010.

[11] D. M. Berry and E. Kamsties, "Ambiguity in requirements specification," *Perspectives on Software Requirements,* pp. 7-44, 2004, doi: 10.1007/978-1-4615-0465-8_2.

[12] M. Trkman, J. Mendling, P. Trkman, and M. Krisper, "Impact of the conceptual model's representation format on identifying and understanding user stories," *Information and Software Technology,* vol. 116, 2019, doi: 10.1016/j.infsof.2019.08.001.

[13] T. Dybå and T. Dingsøyr, "Empirical studies of agile software development: A systematic review," *Information and Software Technology,* vol. 50, no. 9-10, pp. 833-859, 2008, doi: 10.1016/j.infsof.2008.01.006.

[14] G. Panorama Consulting, "The 2020 ERP Report," 2020. [Online]. Available: https://www.panorama-consulting.com/resource-center/2020-erp-report/

[15] S. Sancar Gozukara, B. Tekinerdogan, and C. Catal, "Obstacles of on-premise Enterprise Resource Planning Systems and solution directions," *Journal of Computer Information Systems,* doi: 10.1080/08874417.2020.1739579.

[16] M. Gupta, J. F. George, and W. Xia, "Relationships between IT department culture and agile software development practices: An empirical investigation," *International Journal of Information Management,* vol. 44, no. September 2018, pp. 13-24, 2019, doi: 10.1016/j.ijinfomgt.2018.09.006.

[17] Digital.ai, "14th annual state of Agile report," 2020. [Online]. Available: https://explore.digital.ai/state-of-agile/14th-annual-state-of-agile-report

5. References

[1] M. Cole and Y. Engeström, "A cultural-historical approach to distributed cognition," in *Distributed Cognitions: Psychological and Educational Considerations*, G. Salomon Ed.: Cambridge University Press, 1997, ch. 1, pp. 1-46.

[2] D. Harris, "Distributed cognition in flight operations," in *Engineering Psychology and Cognitive Ergonomics. Applications and Services. EPCE 2013.* , vol. 8020 of Lecture Notes in Computer Science, D. Harris, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 125-133.

[3] Y. Wautelet, M. Velghe, S. Heng, S. Poelmans, and M. Kolp, "On modelers ability to build a visual diagram from a user story set: A goal-oriented approach," in *Lecture Notes in Computer Science. Requirements Engineering: Foundation for Software Quality. REFSQ 2018*, vol. 10753, H. J. Kamsties E., Dalpiaz F., Ed.: Springer, Cham, 2018, pp. 209-226.

[4] J. Jia, X. Yang, R. Zhang, and X. Liu, "Understanding software developers' cognition in agile requirements engineering," *Science of Computer Programming 178,* pp. 1-19, 2019, doi: 10.1016/j.scico.2019.03.005.

[5] R. Barbosa, A. E. A. Silva, and R. Moraes, "Use of Similarity Measure to Suggest the Existence of Duplicate User Stories in the Srum Process," presented at the 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshop (DSN-W), 2016.

[6] E. Dilorenzo *et al.*, "Enabling the Reuse of Software Development Assets Through a Taxonomy for User Stories," *IEEE Access,* vol. 8, pp. 107285-107300, 2020, doi: 10.1109/access.2020.2996951.

[7] M. S. Murtazina and T. V. Avdeenko, "An ontology-based approach to support for requirements traceability in agile development," *Procedia Computer Science 150,* pp. 628-635, 2019, doi: 10.1016/j.procs.2019.02.044.

[8] P. L. de Souza, A. F. do Prado, W. L. de Souza, S. M. dos Santos Forghieri Pereira, and L. F. Pires, "Improving Agile Software Development with Domain Ontologies," *Information Technology-New Generations 738,* pp. 267-274, 2018, doi: 10.1007/978-3-319-77028-4_37.

[9] L. B. De Araujo and F. L. Siqueira, "Using i∗ with scrum: An initial proposal," iStar, 2016, vol. 1674, pp. 19-24.

[10] M. Elallaoui, K. Nafil, and R. Touahni, "Automatic Transformation of User Stories into UML Use Case Diagrams using NLP Techniques," *Procedia Computer Science,* vol. 130, pp. 42-49, 2018, doi: 10.1016/j.procs.2018.04.010.

[11] D. M. Berry and E. Kamsties, "Ambiguity in requirements specification," *Perspectives on Software Requirements,* pp. 7-44, 2004, doi: 10.1007/978-1-4615-0465-8_2.

[12] M. Trkman, J. Mendling, P. Trkman, and M. Krisper, "Impact of the conceptual model's representation format on identifying and understanding user stories," *Information and Software Technology,* vol. 116, 2019, doi: 10.1016/j.infsof.2019.08.001.

[13] T. Dybå and T. Dingsøyr, "Empirical studies of agile software development: A systematic review," *Information and Software Technology,* vol. 50, no. 9-10, pp. 833-859, 2008, doi: 10.1016/j.infsof.2008.01.006.

[14] G. Panorama Consulting, "The 2020 ERP Report," 2020. [Online]. Available: https://www.panorama-consulting.com/resource-center/2020-erp-report/

[15] S. Sancar Gozukara, B. Tekinerdogan, and C. Catal, "Obstacles of on-premise Enterprise Resource Planning Systems and solution directions," *Journal of Computer Information Systems,* doi: 10.1080/08874417.2020.1739579.

[16] M. Gupta, J. F. George, and W. Xia, "Relationships between IT department culture and agile software development practices: An empirical investigation," *International Journal of Information Management,* vol. 44, no. September 2018, pp. 13-24, 2019, doi: 10.1016/j.ijinfomgt.2018.09.006.

[17] Digital.ai, "14th annual state of Agile report," 2020. [Online]. Available: https://explore.digital.ai/state-of-agile/14th-annual-state-of-agile-report