Using process mining in agile software development methodologies: a systematic mapping study

Michael Arias
Computer Science Department
School of Engineering
Pontificia Universidad Católica de Chile
Santiago, Chile
m.arias@uc.cl

Maíra R. Marques
Computer Science Department
Universidad de Chile
Santiago, Chile
mmarques@dcc.uchile.cl

Eric Rojas
Computer Science Department
School of Engineering
Pontificia Universidad Católica de Chile
Santiago, Chile
eric.rojas@uc.cl

***Abstract*—The structure and the way in which organizations manage their projects have evolved. Agile software development has emerged as an alternative to manage projects management processes more efficiently. Process mining allows the analysis of project historical information and proposing improvements for agile processes. A systematic mapping study (SMS) was conducted to classify the proposed approaches in agile development methodologies that uses process mining. A total of 502 studies were identified, and finally 6 studies were selected and analyzed according to distinct aspects. Conference proceedings is the most common venue. There is a concentration of approaches published that comes from Asia and Europe. Disco tool is the most frequently used tool. Meanwhile, the process discovery being the most relevant process mining type used by researchers in this research area. There are two evaluation methods reported as being used: case study and running example, where Scrum is the most frequently methodology used. To the best of our knowledge, this is the first research that has been conducted to generate a SMS in this research area.**

***Index Terms*—Process mining, project management, agile development methodology, agile method, systematic mapping study**

I. Introduction

Agile software development methodologies have become an essential aspect in a world in which organizations need to adapt to a dynamic business environment. The constant changes in which software development environments are exposed have made traditional development methodologies too cumbersome; it is hard to meet the fast changing requirements and short life-cycles needed by companies [18] in software development. In order to face this changing environment, agile software development methodologies were proposed, and the research community has devoted a great deal of attention to them [9]. These methodologies are based on iterative development, prototyping, and the use of templates, where solutions are evolving through collaboration of teammembers and the adoption of fundamental principles of the Agile Manifesto [5].

Within software development there are processes that are executed to generate software systems. According to [28], two types of processes can be distinguished: well-structured routine processes (a predefined control flow), and agile processes (flexible control flow). Analyzing the latter, and the analysis of the execution data (log files), represents a challenge nowadays for practitioners and researchers because it is not always possible to characterize and understand the process adequately [6].

In the literature, there are two well-know disciplines that are concerned with business processes: Business Process Management (BPM) and Process Mining (PM). On the one hand, Business Process Management enables companies to carry out phases for modeling, executing and analyzing their business processes [10], and it is considered a process-oriented discipline. On the other hand, Process Mining focus on extracting useful and sometimes, unexpected knowledge from large amount of data; normally stored in event logs available in many companies information systems [32], and it is considered a data-oriented discipline [31]. Previously, this discipline has been applied in multiple fields such as healthcare [1], [25], marketing [21], Intranets [7], education [19], and resource management [3].

With the rising of technology use and, in particular, the use of information systems in software development

life-cycle, Process Mining can provide several techniques and tools, which could be used for effective software analysis at run-time [27]. This task can lead to different types of analysis and to the identification of potential improvements in software development processes. Due to the importance that software development processes has, and the contribution that Process Mining could make, we conducted a Systematic Mapping Study (SMS) [23]. The aim of this study is to identify and determine the number of research articles that have been published reporting the use of Process Mining in agile software development.

To the best of our knowledge, there is no report publicly available that systematizes and classifies the use of Process Mining within agile software development. We also want to map some important aspects of this area: publishing vehicles used to report research on this area; what kind of research is being done (research type), what are the agile methodologies reported as being used; the geographical distribution of researchers of the area, and Process Mining techniques being used and reported tools being used. As such, this study provides an initial assessment for researchers and practitioners interested in evaluating the amount of evidence that exists in this research area and synthesize the obtained results.

The rest of this paper is organized as follows. Section II explains the need to perform an SMS and present the objectives and methodology followed to conduct the study. Section III presents the results obtained. In Section IV, the analysis and discussion are presented, and also the validity procedures of the study. Finally, Section V concludes the paper and define future research topics.

II. Objectives and Methodology

A. Objectives

Two objectives have been identified to guide this study:

- Obtain an overview of the state of the art on how process mining has been applied with agile methodologies.
- Identify characteristics of the identified studies where process mining and agile methodologies has been applied: the main publication venues, geographical areas where the study was carried out, what process mining tools and techniques were applied, what type of analysis was executed, what agile methodologies have been used in these processes, and, what type of research this studies are based on.

A description of the process implemented to produce this SMS is provided below.

Table. I: Structure Proposed by [17] to Create Research Questions

Criteria	Description
Population	Agile software development methodologies that uses process mining.
Intervention	Techniques, tools and analysis applied.
Comparison	For comparison we are comparing what was already being done so far, related with the usage of process mining in agile software development methodologies
Outcome	Describe the state of the art of what was done so far about process mining in agile software development methodologies.
Context	Describe the domain of use, in our case the domain where process mining is used within agile software development teams.

B. Methodology

A systematic mapping study following Kitchenham's et al. guidelines [17] was carried out in this work; in their work they states that systematic mapping studies "are designed to provide a wide overview of a research area, to establish if research evidence exists on a topic and provide an indication of the quantity of the evidence". As a first step a set of research questions was defined (Section II-C) the ones that drive this mapping. With eyes on these research questions (RQs) the search string was defined for selecting the primary studies from relevant academic databases (Section II-D). To finish the preparation of this study, a set of inclusion and exclusion criteria were created, to select the primary studies (Section II-E). At the end, a backward snowballing was employed on the initial set of papers. Backward snowballing are normally used as a way to extend the literature review and to minimize possible bias created by the defined search string [34].

C. Definition of Research questions

Systematic Mapping Studies are normally done to: evaluate what was already done in a specific area of research, what are the main themes of the area; and also to evaluate how evolved a specific area of research is. The latter is the main objective of this work to elucidate how evolved is the area of process mining agile development. But this statement is too broad, to be able to tackle with it, six research questions were created. Kitchenham et al. [17] guidelines suggest a procedure to define research questions, which have the structure: Population, Intervention, Comparison, Outcome and Context; also known as PICOC structure. Table I shows the structure used.

The proposed research questions were:

- RQ1 - What are the main venues of publication of the area?
- RQ2 - What are the origin of the works of the area (geographical area)?
- RQ3 - What process mining techniques and tools are used?
- RQ4 - Which type of process mining analysis is applied?
- RQ5 - What are the agile methodologies reported being used?
- RQ6 - What are the research types being used in publication of the area?

D. Search strategy

Based on the research questions created, keywords of the area were identified. Table II shows the words used and related synonyms that were considered in our search.

Table. II: Search String

Term	Keywords	Synonyms
A	Process Mining	"process mining"
B	Agile method	"agile method", "agile methodology", "agile development"

The study corresponds to the first exploratory analysis of using process mining in agile, so only a general "process mining" keyword was used to identify papers. Related terms such as "business process management" or "business process intelligence" were excluded to limit the study to very specific process mining publications.

The string **"process mining" AND ("agile method" OR"agile methodology" OR"agile development")** was used to search in titles, abstracts and keywords. There was no limitation related to year of publication in the search performed. The search for primary studies was done on the following digital libraries: ACM Digital Library[1], IEEE Xplore Digital Library[2], ScienceDirect[3], Scopus[4], Springer Link[5], Wiley[6] and Web of Knowledge[7].

E. Inclusion/exclusion criteria

The criteria used to select the papers that we followed were the ones below:

- Inclusion criteria:
  - English, peer-reviewed published articles in journals, conferences and workshops
  - Work published until March 2018
  - Work published in English, Spanish or Portuguese
  - Work published that uses any agile development process with any type or methodology of mining process
- Exclusion criteria:
  - The paper is not peer reviewed
  - The paper is not not available online
  - The paper does not use any sort of mining process to analyze their agile software process development

Table III presents the breakdown of the amount of articles retrieved from each database during the search phase of the mapping process.

Table. III: Retrieved Papers by Database

Database	Search results
IEEE Xplore	-
ACM Digital Library	46
Science Direct	7
Scopus	42
Springer	30
Wiley	246
Web of Knowledge	272
Total studies	643
Excluded (merged)	141
Studies for phase data extraction	502

F. Extraction Procedure

The procedure followed used in this work can also be seen in Figure 1:

- Phase 1:
  1) Paper extraction: In the case of work that appears in more than one venue or type of publication, we included the latest version, preferring journal articles over conference papers, and conference papers over workshop papers.
  2) Paper database: a repository with all the papers was created by one of the authors and all authors used only this repository as data source for the papers, avoiding different versions of the paper being read by any of the authors.
  3) Paper selection step:
     a) All the authors read titles and abstracts of all papers selected to enter phase 1 (502),

[1]http://dl.acm.org
[2]http://ieee.org/ieeexplore
[3]http://www.sciencedirect.com
[4]http://www.scopus.com
[5]http://link.springer.com
[6]http://onlinelibrary.wiley.com
[7]http://apps.webofknowledge.com

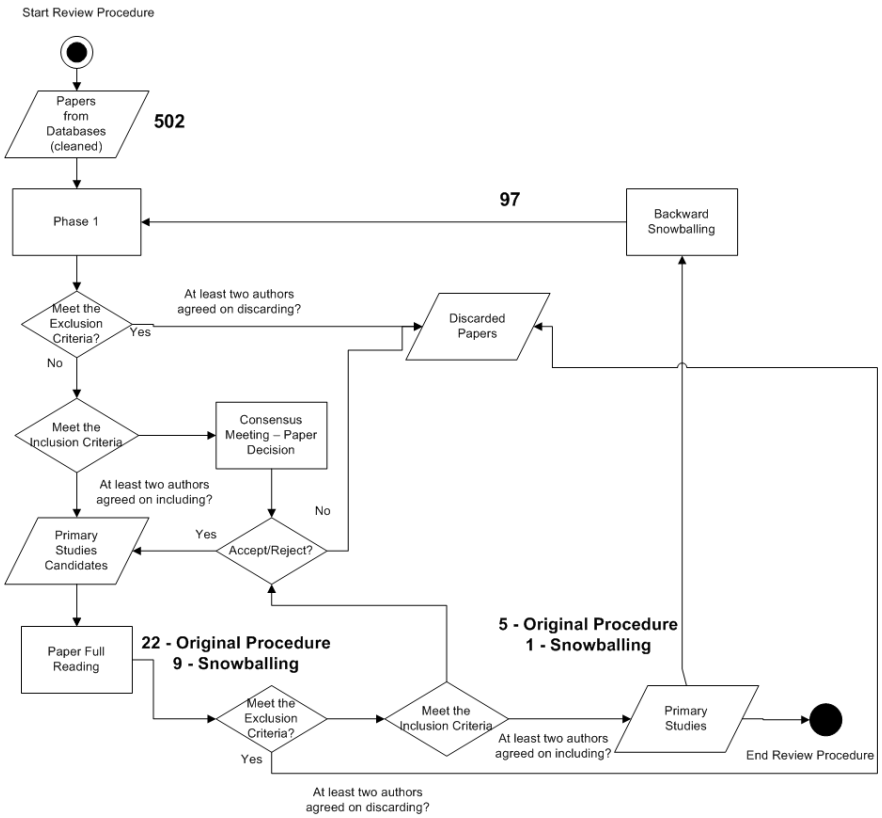

Figure. 1: SMS - Methodology

the papers should meet all the inclusion criteria items and does not meet any of the exclusion criteria items to continue to phase 2 of the work.

b) After all the three authors read and decided which papers would continue to phase 2 and which one would not continue. All the papers selected to continue were gather in a spreadsheet. The papers that were considered to continue, were the ones where at least two authors considered relevant to phase 2 selection. The ones that were considered by only one of the authors were discussed between all the authors to reach an agreement if the paper really fit the criteria of inclusion/exclusion.

c) A final library of 22 was created with the papers to enter the next phase.

4) Phase 2:

a) An online spreadsheet with all the papers and the six research questions was created.

b) All authors read the 22 papers in full, and each one decided separately if the paper really meet the inclusion and does not meet the exclusion criteria.

c) All authors extracted the data of papers which one considered that meet the selection criteria, and filled the spreadsheet with the answers.

d) Once all primary studies were processed by the authors, the extracted data was compared, disagreements between the authors were discussed until a final consensus was made. At the end of the phase 2, 5 papers were selected.

From now on, the selected papers will be called primary studies.

5) Backward Snowballing:

a) All the citations mentioned in the primary studies selected were gathered, a set of 97 papers in total.

b) These 97 papers passed the full extension

Table. IV: Venues

Primary Study	Year	Type of Venue	Venue
[20]	2014	Conference	International Conference on Software Engineering
[15]	2014	Conference	International Conference on Mining Software Repositories (MSR)
[14]	2014	Conference	International Conference on Software Engineering
[27]	2014	Conference	International Conference on Software and System Process
[6]	2016	Conference	International Conference on the Quality of Information and Communications Technology (QUATIC)
[12]	2017	Conference	International Conference on Software Process Improvement and Capability Determination

of the already mentioned Phase 1 and Phase 2.

c) At the end of this backward snowballing, 1 paper was added to the primary study final list.

With respect to excluded articles (91), these were excluded because they did not meet the established inclusion criteria, and the authors did not made evaluations on what was the exact reason why each paper was not considered, so there was no statistical data about the excluded papers.

III. Results

This section presents the results generated from having conducted the SMS. Here are the answers to the research questions defined in Section II-C.

A. RQ1 - What are the main venues of publication of the area?

Analyzing the data of the primary studies it is possible to see the venues of publication: three of them (50%) were presented in a specific conference venue, International Conference on Mining Software Repositories (MSR). The other three were presented in different conferences. No journals or book have been published regarding this topic. The details of each of the venues of publication for the studies is presented in Table IV.

B. RQ2 - What are the origin of the works of the area (geographical area)?

The search on the literature reports two main geographical regions in which the published studies of process mining with agile development has been carried out (check Figure 2).

The greatest concentration of these studies is in Asia (around 43%), followed only by Europe (around 43%), and Oceania with only one study involved (around 14%). There are no publications regarding Africa or America.

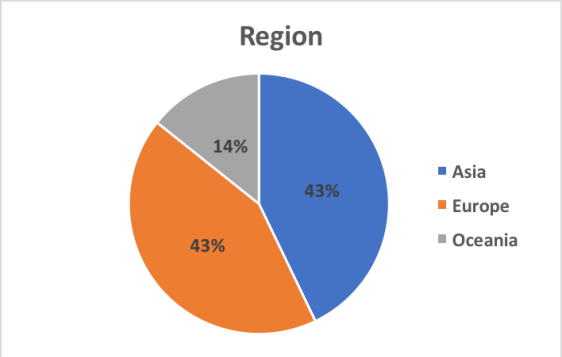

Figure. 2: Region of Origin of Primary Studies Authors

Regarding the specific countries in these regions, the countries where the authors were performing the research are: India with 3 studies [14], [15], [20], followed by Russia and Netherlands [27], Turkey and Australia [12] and Portugal [6], all with one study (check figure 3).

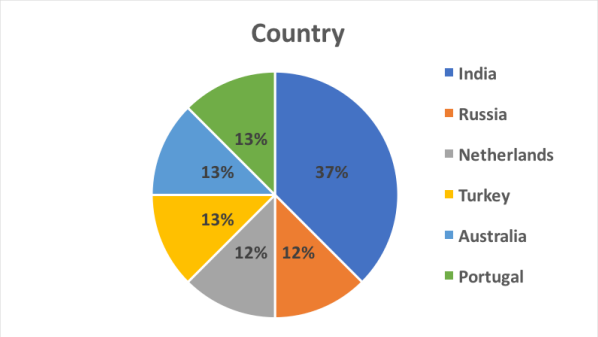

Figure. 3: Country of Origin of Primary Studies Authors

C. RQ3 - What process mining techniques and tools are used?

The main techniques or algorithms used in the primary studies are outlined in Table V. The main technique used in several studies is Disco Miner algorithm [8] used in 5 studies (83%). The main functionality of this algorithm is the automated discovery of process maps by interpreting the sequences of activities in the imported event log [13]. Next, the performance analysis technique was used in 3 different studies (33%). For

[8]https://fluxicon.com/disco/

this type of analysis, a process model is discovered and the time information is annotated in order to diagnose performance problems; i.e.; finding bottlenecks and the usage temporal information for predictions and recommendations. The other reported techniques are mainly focused in organizational mining analysis using metrics such as Handover of Work, Subcontracting and Working Together. The organizational analysis focuses on examining information about resources with the aim of studying how resources are related and involved in the process; i.e.; uncovering organizational knowledge, such as organizational structures and social networks, enables process owners to understand organizational structures and improve business processes [30].

Table. V: Mining techniques reported being used

Mining technique	Primary Study	Tool
Disco Miner Algorithm	[12], [14], [15], [20], [27]	Disco
Disco Performance Analysis	[14], [15], [27]	Disco
Handover of Work	[14], [15]	ProM
Subcontracting	[14], [15]	ProM
Working Together	[14], [15]	ProM
Joint Activities	[14], [15]	ProM
Not Reported	[6]	Eclipse

Regarding the used tools (software applications), we found that Disco is the most frequently used tool (83%), followed by ProM [9] (33%) and Eclipse [10] (17%) (check Figure 4).

Disco is considered a fairly complete commercial solution, focused on the end user. It offers a series of algorithms that are easy-to-use. It is worth to highlight its capabilities for processes discovery, performance analysis, application of filters to different analysis, and it has a good interface to visualize the results generated. ProM is an open-source framework widely used by the process mining community. ProM allows to standardize the development of analysis techniques through the use of plug-ins. It also offers different types of algorithms to execute discovery, conformance, organizational and performance analysis. Meanwhile, Eclipse is an integrated development environment (IDE) that can be adapted to develop particular solutions for different types of process analysis.

D. RQ4 - Which type of process mining analysis is applied?

Van der Aalst [31], identified three types of process mining: discovery, conformance and enhancement. The discovery process is done through the analysis of the log (record of events), which has as objective to obtain a model of the process as a result. The conformance checking analysis compares an existing process model against an event record (Log) of the same process, in order to evaluate if there are differences between the process model (ideal process) and the information contained in the event log (real process). With respect to the enhancement analysis, the objective is to extend or improve an existing process model using the actual (and used) information about the process.

In addition, it is possible to execute process mining analysis according to four process perspectives: control-flow, organizational, case and performance. Control-flow focuses on the order in which activities are executed. Organizational perspective describes who are involved in the performance of the activities and how they are related to each other. Case perspective allows to analyze the characteristics that identify each case of the process. Finally, the performance perspective seeks to analyze the timing of the activities within the use of process.

Of these, five have been previously used to study agile software development as can be seen in Table VI. All primary studies have applied at least one of these types of analysis. The discovery type has been applied in all six identified studies, being the most relevant type of process mining in the analyzed studies. Following, performance analysis with three studies and organizational analysis with two studies reporting its use, and the remaining conformance checking and enhancement have only one primary study that reports its use.

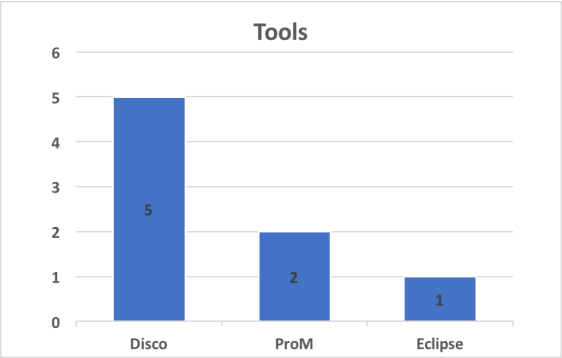

Figure. 4: Tools Reported being used - Primary Studies

E. RQ5 - What are the agile methodologies reported being used?

Several Agile methodologies have been applied in the past years in software development [9], including Agile Unified Process [11], Extreme Programming (XP) [4],

[9]http://www.promtools.org
[10]https://www.eclipse.org/

Table. VI: Process Mining Approach

Approach	Primary Study
Discovery	[6], [12], [14], [15], [20], [27]
Conformance	[14]
Enhancement	[6]
Organizational	[14], [15]
Performance	[14], [15], [27]

Scrum [29], Feature-driven Development [22], Crystal Methodology [8], Kanban Method [2] and others.

Among these methodologies, only two have been explicitly used in the primary studies: Scrum and XP. Scrum within software development processes was reported in three studies, while XP software development processes was used in only one study. Additional studies did not indicate explicitly or did not indicate at all the agile methodology they followed, just indicated the specific practices such as issue tracking or peer review coding. For more details on these studies check table VII.

F. RQ6 - What are the research types and evaluation methods being used in publication of the area?

Inspired by Wieringa et al. [33], we considered three research types: proposal of solution, validation research and evaluation research (see Table VIII). Additionally, based on the classification proposed by Prat et al. [24], we considered the following evaluation methods to assess primary studies: running example, simulation, case study, and several case studies (see description of this methods in Table IX).

The distribution of primary studies in terms of evaluated method can be seen in table VII. It should be noted that four reports evaluate the method using a case study while two studies used running examples.

IV. Analysis and Discussion

We have compiled this systematic mapping study in order to understand the state of research about the use of process mining techniques within agile software development. The analysis has been undertaken considering six research questions that had the goal of highlighting key aspects of the topic. A discussion of the obtained results is presented below.

After carrying the protocol of Kitchenham [17] to perform this mapping study, we are able to provide an initial classification of the literature published in the area of process mining agile software development, something that had not been done until now. Through the study conducted, it is possible to see that, this research area can be considered incipient, yet. We managed to identify only six primary studies, which have been published in the period between 2014 and April 2018, evidencing that the interest of researchers and practitioners to propose approaches in this area is just beginning.

With respect to publishing vehicles, conference proceedings is the only type of venue identified so far.

This could be seen as a first attempt to present and build the area; as this research area reaches a higher level of maturity, it is expected that other research vehicles can be publishing research more frequently, for example, higher impact journals.

With respect to geographic analysis, there is a clear cluster of researchers in Asia and Europe. Specifically India, with three primary studies, the country with most publications in this area. This concentration reflects the interest that the area has awakened in this country considered as a great software outsourcing provider [16], which can serve as a reference and incentive for other researchers from other regions.

The Disco tool is the one that has been most commonly used in the primary studies selected. The techniques to perform the process discovery and performance analysis (Disco Miner and Disco performance analysis) have been preferred. Additionally, the analysis through the organizational perspective has allowed to analyze the information within different metrics. It should be noted that there is a wide variety of tools (eg. Celonis [11] or Lana [12]) that can be used to perform different types of analysis.

Regarding the types of process mining applied in the primary studies, we have identified that process discover is the most used technique. Moreover, other types of analysis, such as organizational metrics have been used. This organizational process perspective have a relevant interest in the research community [35].

In the case of the agile methodologies reported as used, Scrum is the most used methodology (3 studies), followed by XP (1 study). This question allows to validate Scrum as the most used methodology from the perspective of software development [26]. However, it would be relevant to be able to compare the results when using process mining along with other agile methodologies, such as Kanban or Crystal.

Finally, regarding research types analysis, it was reported that five of the six primary studies use evaluation research. In addition, four studies use case study as evaluation method. This trend can be explained by the possible associations of methodologies that can be used with case studies, since it is possible to use synthetic

[11]https://www.celonis.com
[12]https://lana-labs.com/

Table. VII: Agile Methodology, Research Type and Evaluation Method

Primary Study	Agile Methodology	Research Type	Evaluation Method
[20]	XP	Evaluation research	Case Study
[15]	Does not specify a methodology, just practices (issue tracking, version control system, peer coding review)	Evaluation research	Case Study
[14]	Not Specified	Evaluation research	Running example
[27]	Scrum	Evaluation research	Case Study
[6]	Scrum	Proposal of solution	Running example
[12]	Scrum	Evaluation research	Case Study

Table. VIII: Research types classification based on [33]

Research Type	Description
Proposal of Solution	*"An innovative solution for a new problem or significant extension to an existing technique."*
Validation Research	*"Solution that has not yet been implemented in practice. May involve aspects such as prototyping, simulation, experiments, mathematical systematic analysis and mathematical proof of properties."*
Evaluation Research	*"Involve a problem assessment considering an implemented solution, using for instance, case studies, field studies and field experiments."*

Table. IX: Evaluation methods classification based on [24]

Evaluation method	Description
Running Example	Uses a hypothetical execution of a business process.
Simulation	Execute the experimental/simulation using synthetic data.
Case Study	Implement a case study using real-life data.
Several Case Studies	Implement two or more case studies using real-life data.

data for experimentation, functional prototype, among other techniques, besides the fact that a case study can be made based on the use of real data.

Based on the identified studies and the resulting analysis, several challenges have been identified. Is it a research field that should be expanded with more studies, including data from multiple projects and/or unstructured processes, development settings, methodologies and geographical locations. This expansion should result in a clearer and broader understanding of how process mining can impact the software development process in a more broader way. Another important challenge is to go deeper in the understanding of how developers, and other involved resources, interact in each project, and how this can be studied from the process mining perspective. Understanding the interaction between actors and their roles in this type of processes is a critical factor to success.

Additional to the challenges identified, important typical issues presented while developing software with agile methodologies must be addressed. These issues must be studied to see if process mining can provide guidance in possible solutions or improvements. Such issues include, and are not limited to: understand non-standardized software development process, explore existing software support to manage agile projects, reduce rework for next iterations or sprints, analyze collaboration patterns, make more accurate and realist effort estimation for each iteration/sprint, identify particular development practices (e.g. presence of loops), best possible task resource allocation and agile team formation, among others.

As can be seen with the previous analyzes, the study of the application of process mining with agile methodologies is just beginning and can be considered as an opportunity for the development of new methodologies, techniques, methods, tools and their application in new research.

A. Validity Procedures

Since a systematic literature was performed, the corresponding validity procedures established for this type of review mentioned by Kitchenham [17] and Wohlin [34] were used. We relied on the triangulation of decisions made by three different authors, to avoid bias in inclusion and exclusion criteria. We also performed our search in

the most used online databases in computer science. And to minimize bias on the construction of the search string, a backward snowballing was performed to overcome this bias.

V. Conclusions and future work

The application of process mining in agile methodologies allows software development experts to understand the followed process when developing software. Not only process mining provides insight on how the process is being executed task by task, but also provides insight on the performance and the organizational aspects of it.

This article provides a systematic mapping study about the main approaches used to apply process mining with agile software development. After applying a well defined review protocol, the scope of this SMS covers 6 primary studies. The study includes the main tools and techniques applied, the executed process mining analysis, the research type of each study, a breakdown analysis by geographical area and on applied agile methodologies, and finally, the venue of publication.

The results reveal that the application of process mining techniques in agile software development area is an emerging topic. This systematic mapping study now can serve as a reference and quick guide for future researchers in the field, so they can understand and get a basis of the fundamentals of how process mining was applied into agile software development process and how this can provide benefits to improve future development iterations.

Also, we hope the results of this work can inspire research studies into other key aspects regarding agile methodologies, such as coordination of work between teams in agile environments, discover collaboration patterns between team members, and explore project behavior in different software development scale (e.g. small, medium or large), among others aspects..

A. Future Work

For future work, and knowing the state of the art of the application of process mining with agile methodologies, the group will try to establish formal mechanisms to apply this in a more formal way. Methodologies, specific techniques, and case studies may be executed in the future. Regarding the specific content of the review, in the future this may be extended with newer studies and with more detail.

Acknowledgment

The authors thanks the Universidad de Costa Rica Grant Professor Fellowships.

References

[1] Camilo Alvarez, Eric Rojas, Michael Arias, Jorge Munoz-Gama, Marcos Sepúlveda, Valeria Herskovic, and Daniel Capurro. Discovering role interaction models in the emergency room using process mining. *Journal of biomedical informatics*, 2017.

[2] David J Anderson. *Kanban: successful evolutionary change for your technology business*. Blue Hole Press, 2010.

[3] Michael Arias, Rodrigo Saavedra, Maira R Marques, Jorge Munoz-Gama, and Marcos Sepúlveda. Human resource allocation in business process management and process mining: A systematic mapping study. *Management Decision*, 56(2):376–405, 2018.

[4] Kent Beck. *Extreme programming explained: embrace change*. Addison-Wesley professional, 2000.

[5] Kent Beck, Mike Beedle, Arie Van Bennekum, Alistair Cockburn, Ward Cunningham, Martin Fowler, James Grenning, Jim Highsmith, Andrew Hunt, Ron Jeffries, et al. Manifesto for agile software development. 2001.

[6] João Caldeira and Fernando Brito e Abreu. Software development process mining: Discovery, conformance checking and enhancement. In *10th International Conference on the Quality of Information and Communications Technology, QUATIC 2016, Lisbon, Portugal, September 6-9, 2016*, pages 254–259, 2016.

[7] Michael Arias Chaves and Eric Rojas Córdoba. Deciphering event logs in sharepoint server: A methodology based on process mining. In *Computing Conference (CLEI), 2014 XL Latin American*, pages 1–12. IEEE, 2014.

[8] Alistair Cockburn. *Crystal Clear: A Human-Powered Methodology for Small Teams: A Human-Powered Methodology for Small Teams*. Pearson Education, 2004.

[9] Torgeir Dingsøyr, Sridhar P. Nerur, Venugopal Balijepally, and Nils Brede Moe. A decade of agile methodologies: Towards explaining agile software development. *Journal of Systems and Software*, 85(6):1213–1221, 2012.

[10] Marlon Dumas, Marcello La Rosa, Jan Mendling, and Hajo A. Reijers. *Fundamentals of Business Process Management*. Springer, 2013.

[11] Charles Edeki. Agile unified process. *International Journal of Computer Science*, 1(3), 2013.

[12] Sezen Erdem and Onur Demirörs. An exploratory study on usage of process mining in agile software development. In *Software Process Improvement and Capability Determination - 17th International Conference, SPICE 2017, Palma de Mallorca, Spain, October 4-5, 2017, Proceedings*, pages 187–196, 2017.

[13] Christian W. Günther and Anne Rozinat. Disco: Discover your processes. In *Proceedings of the Demonstration Track of the 10th International Conference on Business Process Management (BPM 2012), Tallinn, Estonia, September 4, 2012*, pages 40–44, 2012.

[14] Monika Gupta. Nirikshan: process mining software repositories to identify inefficiencies, imperfections, and enhance existing process capabilities. In *Companion Proceedings of the 36th International Conference on Software Engineering*, pages 658–661. ACM, 2014.

[15] Monika Gupta, Ashish Sureka, and Srinivas Padmanabhuni. Process Mining Multiple Repositories for Software Defect Resolution from Control and Organizational Perspective. In *Proceedings of the 11th Working Conference on Mining Software Repositories*, pages 122–131, New York, NY, USA, 2014. ACM.

[16] Siffat Ullah Khan, Mahmood Niazi, and Rashid Ahmad. Barriers in the selection of offshore software development outsourcing vendors: An exploratory study using a systematic literature review. *Information & Software Technology*, 53(7):693–706, 2011.

[17] Barbara Kitchenham and Stuart Charters. Guidelines for performing Systematic Literature Reviews in Software Engineering.

Technical Report EBSE 2007-01, Keele University and Durham University, July 2007.

[18] Jeffrey A Livermore. Factors that impact implementing an agile software development methodology. In *SoutheastCon, 2007. Proceedings. IEEE*, pages 82–86. IEEE, 2007.

[19] Jorge J Maldonado, René Palta, Jorge Vázquez, Jorge L Bermeo, Mar Pérez-Sanagustín, and Jorge Munoz-Gama. Exploring differences in how learners navigate in moocs based on self-regulated learning and learning styles: A process mining approach. In *Computing Conference (CLEI), 2016 XLII Latin American*, pages 1–12. IEEE, 2016.

[20] Megha Mittal and Ashish Sureka. Process mining software repositories from student projects in an undergraduate software engineering course. In *Companion Proceedings of the 36th International Conference on Software Engineering*, pages 344–353. ACM, 2014.

[21] Anibal Silva Osses, Luiz Quelves Da Silva, Bernardita Fernandez Cobo, Michael Arias, Eric Rojas, Jorge Munoz-Gama, and Marcos Sepulveda Fernadez. Business process analysis in advertising: An extension to a methodology based on process mining projects. In *Computer Science Society (SCCC), 2016 35th International Conference of the Chilean*, pages 1–12. IEEE, 2016.

[22] Steve R Palmer and Mac Felsing. *A practical guide to feature-driven development*. Pearson Education, 2001.

[23] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering, EASE 2008, University of Bari, Italy, 26-27 June 2008*, pages 71–80, 2008.

[24] Nicolas Prat, Isabelle Comyn-Wattiau, and Jacky Akoka. A taxonomy of evaluation methods for information systems artifacts. *Journal of Management Information Systems*, 32(3):229–267, 2015.

[25] Eric Rojas, Jorge Munoz-Gama, Marcos Sepúlveda, and Daniel Capurro. Process mining in healthcare: A literature review. *Journal of biomedical informatics*, 61:224–236, 2016.

[26] Kenneth S Rubin. *Essential Scrum: A practical guide to the most popular Agile process*. Addison-Wesley, 2012.

[27] Vladimir A. Rubin, Irina A. Lomazova, and Wil M. P. van der Aalst. Agile development with software process mining. In *International Conference on Software and Systems Process 2014, ICSSP '14, Nanjing, China - May 26 - 28, 2014*, pages 70–74, 2014.

[28] Stefan Schönig, Cristina Cabanillas, Stefan Jablonski, and Jan Mendling. Mining the organisational perspective in agile business processes. In *Enterprise, Business-Process and Information Systems Modeling - 16th International Conference, BPMDS 2015, 20th International Conference, EMMSAD 2015, Held at CAiSE 2015, Stockholm, Sweden, June 8-9, 2015, Proceedings*, pages 37–52, 2015.

[29] Ken Schwaber. Scrum development process. In *Business object design and implementation*, pages 117–134. Springer, 1997.

[30] Minseok Song and Wil MP Van der Aalst. Towards comprehensive support for organizational mining. *Decision Support Systems*, 46(1):300–317, 2008.

[31] Wil M. P. van der Aalst. *Process Mining - Data Science in Action*. Springer, 2016.

[32] Wil M. P. van der Aalst and A. J. M. M. Weijters. Process mining: a research agenda. *Computers in Industry*, 53(3):231–244, 2004.

[33] Roel Wieringa, Neil Maiden, Nancy Mead, and Colette Rolland. Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. *Requirements Engineering*, 11(1):102–107, 2006.

[34] Claes Wohlin. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, page 38. ACM, 2014.

[35] Weidong Zhao and Xudong Zhao. Process mining from the organizational perspective. *Advances in Intelligent Systems and Computing*, pages 701–708, 2014.