Maturity Models for Agile, Lean Startup, and User-Centered Design in Software Engineering: A Combined Systematic Literature Mapping

Maximilian Zorzetti[1], Matheus Vaccaro[1], Cassiano Moralles[1], Bruna Prauchner[1], Ingrid Signoretti[1], Eliana Pereira[2], Larissa Salerno[1], Ricardo Bastos[1] and Sabrina Marczak[1]

[1]*School of Technology, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Brazil*

[2]*Instituto Federal do Rio Grande do Sul, Porto Alegre, Brazil*

Keywords: Software Engineering, Maturity Model, Agile, Lean Startup, Lean, User-Centered Design.

Abstract: In a bid to reduce the risk accompanied by innovation, IT companies have been trying to boost their Agile development practices by combining Lean Startup and User-Centered Design (UCD) with their existing work processes. Undergoing this transformation in large enterprises can be a difficult challenge without an instrument to help in conducting the adoption and assessment of this novel development approach. In this paper we seek to identify maturity models that assess the use of Agile, Lean Startup, and UCD; characterize these maturity models; and see how they are applied and evaluated. We conducted a systematic literature mapping of maturity models published between 2001 and 2020 taking existing systematic review guidelines into account; and we analyzed the models using an adapted maturity model classification criteria. There are 35 maturity models, of which 23 are maturity models for Agile, 5 for Lean thinking, 5 for User-Centered Design, and 2 for Agile and UCD combined. We found that agile models have been published fairly consistently throughout the years (2001–2020), while Lean thinking and UCD models have mostly been published in the last decade, which might be related to the somewhat recent use of Design Thinking and Lean Startup in software engineering. However, there are no maturity models for a combined use of Agile, Lean Startup, and UCD. We believe that this is the case due to the approach's infancy, as it is seeing success among industry practitioners.

1 INTRODUCTION

Agile is extensively used by organizations today (Hoda, 2017) as it serves as a powerful and adaptive alternative to the rigid and wasteful software development approaches of the past. However, there are some issues with Agile that indicate that it might not be enough by itself—such as lack of user involvement (Schön et al., 2017) and clear identification of added value (Kuusinen et al., 2017). Recent industry cases (Grossman-Kahn and Rosensweig, 2012; Signoretti et al., 2019) show that a combined use of Agile, Lean Startup, and User-Centered Design (UCD) can be a way to overcome the aforementioned issues: Lean Startup (Ries, 2011) focuses on adding value to business stakeholders through strategic experimentation, while UCD (Norman, 2002) puts the user at the center of the discussion to foster empathy.

Adopting such a combined approach can lead to several organizational challenges of different nature, such as cultural (e.g., trust), structural (e.g., roles), and technical (e.g., techniques), which are aggravated when dealing with large enterprises as new large-scale issues arise (e.g., inter-team coordination) (Paasivaara et al., 2018), making instruments to guide and assess the transformation essential in these cases. An example of such instruments are maturity models, which can gauge the transformation in a not overly expensive and time-consuming manner (Maier et al., 2012). We aim to show what is the current state of the art in maturity models for a software development approach composed of Agile, Lean Startup, and UCD pillars through a systematic literature mapping.

The remainder of this paper is divided as follows: Section 2 discusses the use of a combined approach of Agile, Lean Startup, and UCD in software development; Section 3 explains how the systematic literature mapping was conducted and highlights the research questions; Section 4 presents our findings, including data extracted from the mapping; Section 5 deliberates on this study, examines threats to its validity, and considers future work.

Zorzetti, M., Vaccaro, M., Moralles, C., Prauchner, B., Signoretti, I., Pereira, E., Salerno, L., Bastos, R. and Marczak, S.
Maturity Models for Agile, Lean Startup, and User-Centered Design in Software Engineering: A Combined Systematic Literature Mapping.
DOI: 10.5220/0009582401450156
In *Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS 2020) - Volume 2*, pages 145-156
ISBN: 978-989-758-423-7

2 BACKGROUND

The Agile movement dates back to 2001 with the introduction of the Agile Manifesto (Beck et al., 2001), a result of the then-current wasteful and rigid software development culture and work processes. The extensive use of agile in the past two decades has brought to light some of its weaknesses, such as difficulty in increasing user involvement (Schön et al., 2017). A development method composed of Agile, Lean Startup, and UCD is a novel approach that has been argued as a way to overcome such weaknesses (Ximenes et al., 2015) and that is drawing the attention of academics (Ximenes et al., 2015; Dobrigkeit et al., 2019) and industry practitioners (Grossman-Kahn and Rosensweig, 2012; Signoretti et al., 2019).

Lean Startup is an entrepreneurship methodology that focuses on developing a business plan iteratively through the use of a "build-measure-learn" loop, where business hypotheses are evaluated through carefully planned and efficient experiments that gather useful customer feedback, which feeds into the strategic decision process that leads to the next loop (Ries, 2011). UCD consists of a set of procedures, processes, and techniques that focus on setting the user as the center of the design space or development process (Norman, 2002), enabling developers to understand the user's real needs and create improved software with better usability and user satisfaction (Salah et al., 2015).

Grossman-Kahn and Rosensweig (2012) report on the evolution of the Nordstrom Innovation Lab, an initiative of the fashion retailer Nordstrom to rapidly and cheaply test novel concepts internally. Each iteration of the lab improved upon the shortcomings of the former, turning what started as an isolated agile development team into an acclaimed innovation team with its own development methodology which encapsulates Design Thinking, Lean Startup, and Agile.

Moralles et al. (2019) conducted an empirical study to compare Extreme Programming (XP), Lean, and UCD concepts identified through literature reviews with what was being used in practice by two software development teams that use a development methodology that encompasses the three methods. Their findings suggest that both teams use a complementary subset of concepts from each pillar, in addition to techniques and roles not found in the literature. Their study motivated us to seek maturity models that propose the combination of the three aforementioned pillars. Maturity models, which can be prescriptive or descriptive, aim to offer guidance on practices that are relevant to master. The Agile Compass (Fontana et al., 2015), backed by an agile maturing framework (Fontana et al., 2015), is an example of a checklist-based agile maturity model which introduces the category of outcomes an agile team should seek as it matures with regards to the use of practices. Such models can be of help to bring awareness to newcomers to the combined use of Agile, Lean Startup, and UCD.

3 RESEARCH METHOD

This study was conducted as a systematic literature mapping based on guidelines for conducting systematic literature mappings in software engineering (Petersen et al., 2015). Our first effort on mapping maturity models for a combined approach of the three aforementioned pillars found zero results, so we expanded our effort into 7 systematic literature reviews (SR) about maturity models for Agile, Lean Startup, UCD, and their intersections: Agile combined with Lean Startup; Agile combined with UCD; Lean Startup combined with UCD; and Agile, Lean Startup, and UCD combined (each is hereinafter referred to as a *search context*). The goal of these SRs is to identify and assess primary and secondary studies regarding the use, structure, and evaluation of maturity models for the three pillars.

3.1 Research Questions

All SRs address the same research questions, each related to their respective search context.

RQ1. What maturity models are available?

RQ2. How are these maturity models characterized?

RQ3. How are these maturity models applied and evaluated?

3.2 Search

As suggested by Kitchenham (Kitchenham and Charters, 2007), we used the PICO criteria to guide the formulation of our search string.

Population: Primary and secondary studies related to their respective search context.

Intervention: Maturity models related to their respective search context.

Comparison: This criterion does not apply to our RQs because the goal of this study is not to compare the identified maturity models.

Outcomes: Understanding of use, structure, and evaluation of identified maturity models.

All SRs followed the same search process. We retrieved studies from electronic databases that met the following source selection criteria:

- Databases that include journal articles, conference, and workshop papers related to their respective SR context;
- Databases with an advanced search mechanism that allows filtering of the results by keywords that address the research questions; and
- Databases that provide access to full papers written in English.

Based on these criteria, we selected the following databases: ACM Digital Library, IEEExplore, Science Direct, Scopus, and Springer Database. We adapted the search string (Equation 1) for each database based on the search functionality offered by the given database. Each search string consisted of two parts—S1 and S2—defined as follows:

- S1 is a string composed of keywords related to maturity models, namely: maturity model, capability model, self assessment, health check, and team assessment; and
- S2 is a string composed of keywords related to the search context of each SR. Table 1 presents the keywords used.

As Lean Startup is the newest of the three pillars, we chose to broaden its search context by including other Lean thinking schools, such as Lean UX.

Equation 1. Search criteria boolean expression.

$$S1 \text{ AND } S2 \quad (1)$$

Afterwards, inclusion and exclusion criteria were applied by a varying number of researchers for each SR on the retrieved studies in two distinct rounds, as explained in Section 3.3. The first round consisted of title and abstract inspection to triage the candidate studies based on the inclusion and exclusion criteria. The second round consisted of a thorough inspection with full text reading to further filter the studies and to perform the data extraction procedure (Section 3.5).

3.3 Study Selection

To determine whether a study should be selected, all SRs applied the following selection criteria.

Table 1: Keywords used in the search string of each SR.

ID	SR	Keywords
K1	Agile	"Agile" OR "Agile Method*" OR "Agile Development" OR "Agile Software Development" OR "Agile Practice" OR "Test Driven Development" OR "Test-driven Development" OR "Behavior-driven Development" OR "Behavior Driven Development" OR "Behaviour-driven Development" OR "Behaviour Driven Development" OR "Extreme Programming" OR "Scrum" OR "Kanban"
K2	Lean	"Lean Startup" OR "Lean Start-Up" OR "Lean UX" OR "Lean User Experience" OR "Lean Software" OR "Lean Development"
K3	UCD	"Design Thinking" OR "*Centered Design" OR "*Centred Design" OR "User Experience" OR "Usability" OR "Human Computer Interaction" OR "Computer-Human Interaction" OR "Human Factor" OR "User Interface"
K4	Agile and Lean	K1 AND K2
K5	Agile and UCD	K1 AND K3
K6	Lean and UCD	K2 AND K3
K7	Agile, Lean, and UCD	K1 AND K2 AND K3

Inclusion Criteria: (I1) the study presents a maturity model for its SR context; (I2) the study is written in English; (I3) the study is fully written in electronic format; (I4) the study was retrieved from a conference, workshop, or journal.

Exclusion Criteria: (E1) the study does not present a maturity model for its SR context; (E2) the study is an extended abstract or editorial paper; (E3) the study is duplicated.

We only searched for studies published between 2001 and 2020. We chose 2001 as the lower bound as

Table 2: Quality criteria for study selection.

Criteria	Response grading
C1	{1, 0.5, 0} (Yes, Moderately, No)
C2	{1, 0.5, 0} (Yes, Moderately, No)
C3	{1, 0.5, 0} (Yes, Moderately, No)
C4	(>80% = 1), (<20% = 0), (in-between = 0.5)

it is the publication date of the Agile Manifesto (Beck et al., 2001). Additionally, we performed a manual, informal search on the internet and considered gray literature studies, as these concern very current issues which might have not yet been covered in academic literature (Kiteley and Stogdon, 2013).

3.4 Quality Assessment

We used a set of quality criteria proposed by Guyatt et al. (2008)—later used by Dybå and Dingsøyr (2008) in software engineering—to assess the methodological quality of the studies selected for review, as they cover thoroughness, trustworthiness, and significance of the studies (Inayat et al., 2015). The criteria are based on four quality assessment questions:

(C1) Is the research objective clearly defined?
(C2) Is the research context well addressed?
(C3) Are the findings clearly stated?
(C4) Based on the findings, how valuable is the research?

We graded the selected studies on each criterion using an ordinal scale instead of a dichotomous scale to obtain a more accurate assessment (Inayat et al., 2015). Table 2 shows the grading scale for each criterion. When there was not an agreement on a study's grade, we had meetings to discuss the issue until we agreed upon a single grade.

3.5 Data Extraction and Classification

We performed a full text reading of each study to identify, categorize, and analyze the following items:

- Study identification (RQ1);
- Aim: if the model determines necessary improvements for its use case (analysis) or if it presents best practices for comparison (benchmarking) (RQ2);
- Scope: if the model is generic or domain-specific (limited to a determined method) (RQ2);
- Evaluation: if the model was evaluated, such as by having it applied in a real context (RQ3);
- Maturity levels: if the model has defined quantifiable levels of maturity (RQ2);
- Maturity description[1]: if the model has definitions for what constitutes different standards of maturity (RQ2); and
- Administration mechanism: if the model has defined a mechanism to apply the model (RQ3).

Aside from the first, the items were adapted from the guidelines for developing maturity grids by Maier, Moultrie, and Clarkson (2012). Although the guidelines concern maturity grids, we found them adequate to fulfill the needs of our study. We chose guideline elements that facilitate the categorization of maturity models. Each researcher received an equal amount of studies to extract data from and apply the study selection criteria again. We made use of the data found in a similar literature review study (Fontana et al., 2018) that focused on Agile maturity models as our search resolved into a superset of the models it identified.

4 RESULTS

This section summarizes the results of each SR. Table 3 presents the results of the search process in the electronic databases selected in Section 3.2. Table 4 shows the selected studies categorized by search context, along with their quality grading as defined in Section 3.4. We analyze the studies in light of our research questions based on the data extracted (Table 5, 6, and 7) using the procedure in Section 3.5 next.

Table 3: Number of identified studies during the distinct rounds of our systematic search for maturity models.

Search context	Retrieved	Round 1		Round 2	
		Excl.	Incl.	Excl.	Incl.
Agile, Lean, and UCD	82	77	5	5	0
Agile and Lean	152	144	8	8	0
Agile and UCD	77	72	5	3	2
Lean and UCD	78	73	5	5	0
Agile	2188	2095	93	76	17
Lean	231	207	24	19	5
UCD	3194	3142	52	47	5
Total	**5920**	**5810**	**192**	**163**	**29**

[1]Defined as "cell texts" in (Maier et al., 2012)

Table 4: Selected maturity model studies and their respective quality gradings.

Context	Study	C1	C2	C3	C4	Avg.
Agile	(Nawrocki et al., 2001)	1.0	1.0	0.5	0.5	0.75
	(Lui and Chan, 2006)	1.0	0.5	0.5	0.5	0.625
	(Sidky et al., 2007)	1.0	0.5	0.5	1.0	0.75
	(Packlick, 2007)	1.0	0.5	1.0	1.0	0.875
	(Qumer and Henderson-Sellers, 2008)	0.5	0.5	1.0	0.5	0.625
	(Patel and Ramachandran, 2009a)	1.0	1.0	1.0	1.0	1.0
	(Patel and Ramachandran, 2009b)	0.5	1.0	0.5	1.0	0.75
	(Humble and Russell, 2009)	0.5	0.5	0.5	0.5	0.5
	(Benefield, 2010)	0.5	0.5	0.5	1.0	0.625
	(Proulx, 2010)	0.5	0.5	0.5	0.5	0.5
	(Yin et al., 2011)	1.0	0.5	1.0	1.0	0.875
	(Buglione, 2011)	0.5	1.0	0.5	1.0	0.75
	(Medappa and Bhattacharya, 2012)	1.0	1.0	1.0	0.5	0.875
	(Soundararajan et al., 2013)	1.0	1.0	0.5	1.0	0.875
	(Fontana et al., 2014)	1.0	1.0	1.0	1.0	1.0
	(Silva et al., 2014)	1.0	1.0	0.5	0.5	0.75
	(Özcan Top and Demirörs, 2014)	1.0	1.0	1.0	1.0	1.0
	(Soares and Meira, 2015)	1.0	1.0	1.0	0.5	0.875
	(Fontana et al., 2015)	1.0	1.0	1.0	1.0	1.0
	(Stojanov et al., 2015)	1.0	0.5	1.0	1.0	0.875
	(Ambler and Lines, 2016)	0.5	0.5	0.5	0.0	0.375
	(Stanisavljevic et al., 2018)	1.0	1.0	1.0	1.0	1.0
	(Shukla and Sushil, 2020)	1.0	1.0	1.0	1.0	1.0
Lean	(Jørgensen et al., 2007)	1.0	1.0	0.5	0.5	0.75
	(Karvonen et al., 2012)	1.0	1.0	1.0	1.0	1.0
	(Cil and Turkan, 2013)	1.0	1.0	0.5	0.5	0.75
	(Schröders and Cruz-Machado, 2015)	1.0	1.0	0.0	0.5	0.875
	(Al-Baik and Miller, 2019)	1.0	1.0	0.5	0.0	0.625
UCD	(Van Tyne, 2009)	0.5	0.5	0.5	0.5	0.5
	(Chapman and Plewes, 2014)	1.0	1.0	1.0	1.0	1.0
	(Kieffer and Vanderdonckt, 2016)	1.0	1.0	0.5	0.5	0.75
	(Ogunyemi et al., 2017)	1.0	1.0	0.5	1.0	0.875
	(Quintal and Macías, 2018)	1.0	1.0	1.0	1.0	1.0
Agile and UCD	(Peres et al., 2014)	1.0	0.5	0.5	0.5	0.625
	(Salah et al., 2016)	1.0	1.0	1.0	1.0	1.0

4.1 RQ1. What Maturity Models Are Available?

As mentioned in Table 3, our systematic literature mapping identified a total of 29 studies establishing maturity models for Agile, Lean Startup, UCD, and their intersections. From our manual search, we selected an additional 4 academic studies (Sidky et al., 2007; Qumer and Henderson-Sellers, 2008; Patel and Ramachandran, 2009a; Yin et al., 2011) and 2 gray literature studies (Ambler and Lines, 2016; Proulx, 2010) for a total of 35 studies. The higher number of maturity models for Agile is expected, as it is the most dominant approach to software engineering worldwide. There are few maturity models for intersections of the pillars—only 2 for a combined use of Agile and UCD—with a notable absence of models for all three pillars combined. Figure 1 shows the number of maturity models for each category using a Venn diagram. Of the existing Lean studies, we point out that none concern the use of Lean Startup.

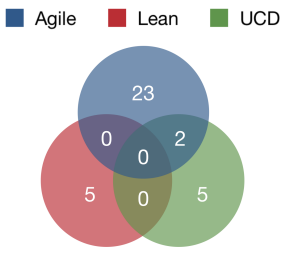

Figure 1: Venn diagram of maturity models for Agile, Lean, and UCD.

Table 5: Overview of data extracted from selected maturity model studies.

Context	Study	Aim	Scope	Evaluation	Maturity Level	Maturity Desc.	Admin. Mech.
Agile	(Nawrocki et al., 2001)	Analysis	XP	Yes	Yes	Yes	Partial
	(Lui and Chan, 2006)	Analysis	XP	No	Yes	No	No
	(Sidky et al., 2007)	Analysis	Generic	Yes	Yes	Yes	Yes
	(Packlick, 2007)	Benchmarking	Generic	Yes	Yes	Yes	Yes
	(Qumer and Henderson-Sellers, 2008)	Analysis	Generic	Yes	Yes	No	Yes
	(Patel and Ramachandran, 2009a)	Analysis	Generic	Yes	Yes	No	Yes
	(Patel and Ramachandran, 2009b)	Analysis	Generic	No	Yes	Yes	Yes
	(Humble and Russell, 2009)	Analysis	Generic	No	Yes	Yes	Yes
	(Benefield, 2010)	Benchmarking	XP	Yes	Yes	No	Partial
	(Proulx, 2010)	Analysis	Generic	No	Yes	No	No
	(Yin et al., 2011)	Analysis	Scrum	Yes	Yes	No	Partial
	(Buglione, 2011)	Analysis	Generic	No	Yes	Yes	No
	(Medappa and Bhattacharya, 2012)	Analysis	Generic	No	Yes	Yes	Yes
	(Soundararajan et al., 2013)	Analysis	Generic	No	Yes	Yes	Yes
	(Fontana et al., 2014)	Analysis	Generic	No	Yes	No	No
	(Silva et al., 2014)	Analysis	Quality Assurance	Yes	Yes	Yes	No
	(Özcan Top and Demirörs, 2014)	Analysis	Generic	Yes	Yes	No	Yes
	(Soares and Meira, 2015)	Analysis	Generic	No	Yes	Yes	No
	(Fontana et al., 2015)	Analysis	Generic	No	No	Yes	Yes
	(Stojanov et al., 2015)	Analysis	Generic	Yes	Yes	Yes	Yes
	(Ambler and Lines, 2016)	Analysis	Generic	No	Yes	Yes	No
	(Stanisavljevic et al., 2018)	Analysis	Generic	No	Yes	No	Partial
	(Shukla and Sushil, 2020)	Analysis	Generic	No	Yes	Yes	No
Lean	(Jørgensen et al., 2007)	Benchmarking	Generic	No	Yes	No	No
	(Karvonen et al., 2012)	Benchmarking	Generic	No	Yes	Yes	Yes
	(Cil and Turkan, 2013)	Analysis	Generic	No	No	Yes	Yes
	(Schröders and Cruz-Machado, 2015)	Analysis	Generic	No	No	Yes	No
	(Al-Baik and Miller, 2019)	Analysis	Kaizen	Yes	No	No	Yes
UCD	(Van Tyne, 2009)	Analysis	Generic	No	Yes	No	No
	(Chapman and Plewes, 2014)	Analysis	Generic	No	Yes	Yes	No
	(Kieffer and Vanderdonckt, 2016)	Analysis	Generic	No	Yes	Yes	Yes
	(Ogunyemi et al., 2017)	Analysis	Generic	No	No	No	Yes
	(Quintal and Macías, 2018)	Analysis	Generic	No	Yes	Yes	Yes
Agile and UCD	(Peres et al., 2014)	Analysis	Generic	Yes	Yes	Yes	Yes
	(Salah et al., 2016)	Analysis	Generic	No	Yes	Yes	Yes

Table 6: Evaluations performed on maturity models.

Context	Study	Evaluation
Agile	(Nawrocki et al., 2001)	Evaluated by 5 project teams composed of 6 students each in a university. Teams were asked to organize their work according to the model and apply as many XP practices as possible
	(Lui and Chan, 2006)	—
	(Sidky et al., 2007)	Evaluated through questionnaires answered by 28 members of the agile community
	(Packlick, 2007)	Evaluated through 20 teams using it for over six months
	(Qumer and Henderson-Sellers, 2008)	Evaluated through 2 industry case studies
	(Patel and Ramachandran, 2009a)	Evaluated through a discussion with 3 different organizations
	(Patel and Ramachandran, 2009b)	—
	(Humble and Russell, 2009)	—
	(Benefield, 2010)	Evaluated through a case study in a multinational communication company
	(Proulx, 2010)	—
	(Yin et al., 2011)	Evaluated through action research, which incorporated interviews with Scrum, Agile, and CMMI experts
	(Buglione, 2011)	—
	(Medappa and Bhattacharya, 2012)	—
	(Soundararajan et al., 2013)	—
	(Fontana et al., 2014)	—
	(Silva et al., 2014)	Evaluated through a survey based on the opinion of experts
	(Özcan Top and Demirörs, 2014)	Evaluated through an exploratory case study in a government organization
	(Soares and Meira, 2015)	—
	(Fontana et al., 2015)	—
	(Stojanov et al., 2015)	Evaluated through a case study in a large organization
	(Ambler and Lines, 2016)	—
	(Stanisavljevic et al., 2018)	—
	(Shukla and Sushil, 2020)	—
Lean	(Jørgensen et al., 2007)	—
	(Karvonen et al., 2012)	—
	(Cil and Turkan, 2013)	—
	(Schröders and Cruz-Machado, 2015)	—
	(Al-Baik and Miller, 2019)	Evaluated through a survey
UCD	(Van Tyne, 2009)	—
	(Chapman and Plewes, 2014)	—
	(Kieffer and Vanderdonckt, 2016)	—
	(Ogunyemi et al., 2017)	—
	(Quintal and Macías, 2018)	—
Agile and UCD	(Peres et al., 2014)	Evaluated by a panel of experts
	(Salah et al., 2016)	—

Table 7: Administration mechanisms of maturity models.

Context	Study	Administration Mechanism
Agile	(Nawrocki et al., 2001)	Partial, provides a list of items to be observed by a XP tracker but leaves the definition of an assessment method open-ended
	(Lui and Chan, 2006)	—
	(Sidky et al., 2007)	Uses a Goal-Question-Indicator-Metric approach to measure readiness for practice adoption; provides a 4-step process for organizations to adopt agile
	(Packlick, 2007)	Reports a case that used user stories based on agile goals
	(Qumer and Henderson-Sellers, 2008)	Provides a custom analytical tool that evaluates agile methods through four distinct perspectives
	(Patel and Ramachandran, 2009a)	Uses questionnaires that are distributed to a project's development team and any other associated personnel; and a roadmap for software process improvement
	(Patel and Ramachandran, 2009b)	Proposes the use of user stories for each desired maturity level; has a web-based tool to assess organizational suitability to use story card-based requirements engineering and agile practices
	(Humble and Russell, 2009)	Provides a somewhat superficial plan-do-check-act cycle to roll out improvements throughout an organization
	(Benefield, 2010)	Partial, uses an undisclosed list of required measures and evidence to determine maturity levels that is to be used by a third party and as a self-assessment tool to allow for complementary views
	(Proulx, 2010)	—
	(Yin et al., 2011)	Partial, uses an undisclosed checklist of Scrum practices for each maturity level
	(Buglione, 2011)	—
	(Medappa and Bhattacharya, 2012)	Reports a case that used a survey
	(Soundararajan et al., 2013)	Collects data on indicators and through a series of computations resolves into numeric scores for strategies, principles, and objectives
	(Fontana et al., 2014)	—
	(Silva et al., 2014)	—
	(Özcan Top and Demirörs, 2014)	Uses a questionnaire about specific practices and generic agile practices
	(Soares and Meira, 2015)	—
	(Fontana et al., 2015)	Provides a checklist that helps teams to identify which outcomes they have attained
	(Stojanov et al., 2015)	Provides indicators to be assessed in assessment meetings
	(Ambler and Lines, 2016)	—
	(Stanisavljevic et al., 2018)	Partial, describes two methods to interpret maturity parameters that are evaluated using discrete scoring, but does not specify how to obtain the latter
	(Shukla and Sushil, 2020)	—
Lean	(Jørgensen et al., 2007)	—
	(Karvonen et al., 2012)	Provides assessment items for the lean practices in each of its process areas
	(Cil and Turkan, 2013)	Uses the Analytical Network Process
	(Schröders and Cruz-Machado, 2015)	—
	(Al-Baik and Miller, 2019)	Specifies a process area for evaluation methods
UCD	(Van Tyne, 2009)	—
	(Chapman and Plewes, 2014)	—
	(Kieffer and Vanderdonckt, 2016)	Uses a questionnaire to assess organizational strategic usability
	(Ogunyemi et al., 2017)	Uses a questionnaire for peer and self-assessment
	(Quintal and Macías, 2018)	Uses U+A SPICE, a mechanism adapted from ISO/IEC 15504
Agile and UCD	(Peres et al., 2014)	Synchronizes the iterative development cycle with its process improvement suggestions; provides a list of relevant measurements
	(Salah et al., 2016)	Provides a performance scale to rate organizational performance and an assessment procedure that contains a sheet template for information recording, maturity scores for comparison, guidelines, and other benchmarks

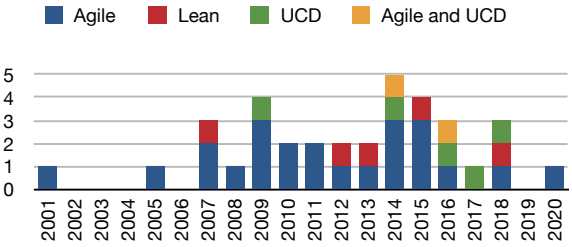

Figure 2: Publication frequency of maturity models.

Figure 2 shows the publication frequency of the maturity model studies on a stacked bar chart. Agile maturity models see a fairly consistent publication rate throughout the years. Most Lean maturity models and all UCD ones were published in the past ten years, likely due to the rising popularity of Design Thinking and Lean Startup in software engineering.

4.2 RQ2. How Are These Maturity Models Characterized?

We plotted the extracted data (aim, scope, maturity levels, and maturity description) on stacked bar charts. Figure 3 divides the maturity models into either benchmarking or analysis types. Figure 4 categorizes the maturity models as generic (e.g., for generic Agile practices (Sidky et al., 2007)) or specific (for Scrum (Yin et al., 2011), for XP (Nawrocki et al., 2001; Lui and Chan, 2006; Benefield, 2010), for Quality Assurance (Silva et al., 2014), and for Kaizen practices (Al-Baik and Miller, 2019)). Figure 5 shows how many maturity models have defined maturity levels or have no evidence of having done so. Figure 6 shows the percentage of maturity models that have descriptions for maturity standards.

4.3 RQ3. How Are These Maturity Models Applied and Evaluated?

We plotted summarized data on model evaluations and administration mechanisms on stacked bar charts (details are shown in Table 6 and 7). Figure 7 show how many maturity models were evaluated in some way. Figure 8 shows how many maturity models have instruments or procedures to applying them in their target context. A "partial evidence" value was chosen when the study does not disclose the tool.

5 CONCLUSION

This paper reports on a systematic literature mapping of maturity models for Agile, Lean Startup, UCD,

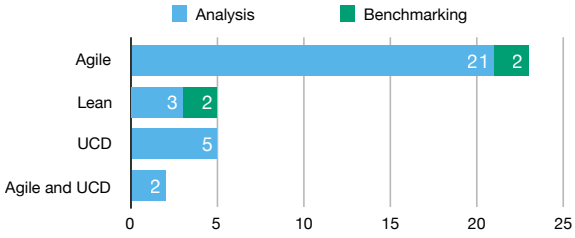

Figure 3: Distribution of maturity model aims.

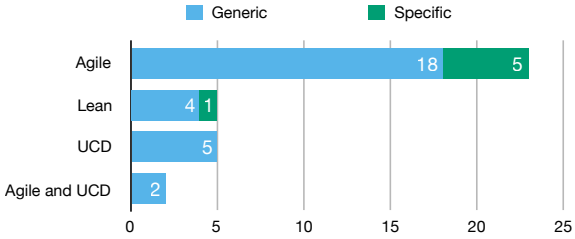

Figure 4: Distribution of maturity model scopes.

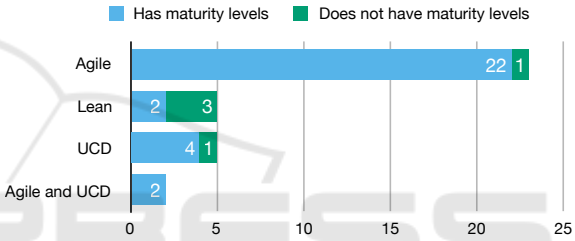

Figure 5: Distribution of maturity level definition.

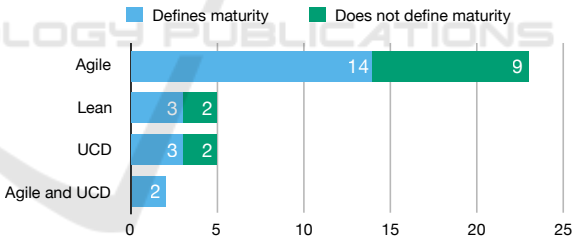

Figure 6: Distribution of maturity description definition.

and their intersections in a software engineering context. We found a total of 35 maturity models, but none were of a combined approach of the three pillars. The methodological quality of the maturity model studies was evaluated using previously established criteria (Guyatt et al., 2008; Dybå and Dingsøyr, 2008; Inayat et al., 2015). Then, we categorized the maturity models using criteria adapted from maturity grid guidelines (Maier et al., 2012) and plotted the resulting data on stacked bar charts.

Although we found some maturity models for Lean thinking, none were specifically for Lean Startup, which seems to be a major driving force behind the combined approach of Agile, Lean Startup, and UCD (Grossman-Kahn and Rosensweig, 2012).

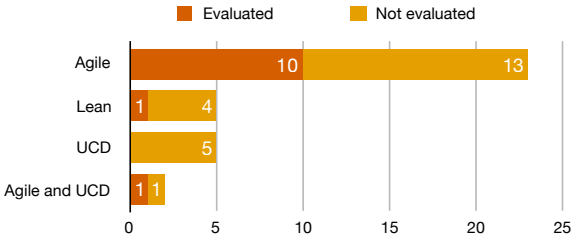

Figure 7: Distribution of evaluations.

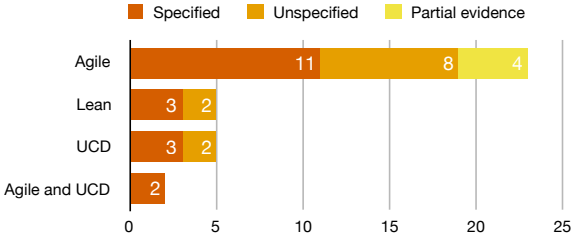

Figure 8: Distribution of administration mechanism definition.

We found no maturity models for a combined approach of the three pillars. This absence of combined models could be attributed to the lack of widely accepted theoretical bases for Agile, Lean Startup, and UCD and of documentation on how to develop theoretically sound maturity models (de Bruin et al., 2005); leading to many models not being developed with a sound methodology (García-Mireles et al., 2012), making combination efforts difficult.

As with any systematic review, most threats to validity concern study selection bias and inaccuracy during data extraction. We carried out procedures to reduce such threats, but our protocol is prone to faults:

- The first round of inclusion and exclusion criteria was applied only once by multiple researchers (no study was evaluated more than once);
- The studies that participated in the second round of inclusion and exclusion criteria were assessed by two researchers, but no metric to rate inter-rater agreement among the researchers was calculated;
- Data extraction results obtained from a researcher were not checked by another; and
- No snowball search of any kind was executed.

Despite the lack of research on the topic, there is a clear interest on the approach on behalf of the industry (Grossman-Kahn and Rosensweig, 2012; Ximenes et al., 2015; Dobrigkeit et al., 2019; Moralles et al., 2019; Signoretti et al., 2019). We believe a rise in demand on this research topic is imminent. For future work on this topic, the development of a maturity model that assesses a combined approach of the three pillars is evident, but perhaps it needs a proper theoretical foundation laid down before it is made, so as to avoid common maturity model pitfalls.

ACKNOWLEDGMENTS

We acknowledge that this research is sponsored by Dell Brazil using incentives of the Brazilian Informatics Law (Law no. 8.2.48, year 1991).

REFERENCES

Al-Baik, O. and Miller, J. (2019). Integrative double kaizen loop (idkl): Towards a culture of continuous learning and sustainable improvements for software organizations. *IEEE Transactions on Software Engineering*, 45(12):1189–1210.

Ambler, S. and Lines, M. (2016). The disciplined agile framework: A pragmatic approach to agile maturity. *The Journal of Defense Software Engineering – CrossTalk*, 29:25–31.

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., and Thomas, D. (2001). Manifesto for agile software development.

Benefield, R. (2010). Seven dimensions of agile maturity in the global enterprise: A case study. In *Hawaii International Conference on System Sciences*, pages 1–7, Koloa, Kauai, HI, USA. IEEE Computer Society.

Buglione, L. (2011). Light maturity models (lmm): An agile application. In *Proceedings of the International Conference on Product Focused Software Development and Process Improvement*, pages 57–61, Torre Canne, Brindisi, Italy. ACM.

Chapman, L. and Plewes, S. (2014). A ux maturity model: Effective introduction of ux into organizations. In *International Conference of Design, User Experience, and Usability*, pages 12–22, Heraklion, Crete, Greece. Springer.

Cil, I. and Turkan, Y. S. (2013). An anp-based assessment model for lean enterprise transformation. *The International Journal of Advanced Manufacturing Technology*, 64(5-8):1113–1130.

de Bruin, T., Freeze, R., Kulkarni, U., and Rosemann, M. (2005). Understanding the main phases of developing a maturity assessment model. *Australasian Conference on Information Systems*.

Dobrigkeit, F., de Paula, D., and Uflacker, M. (2019). *InnoDev: A Software Development Methodology Integrating Design Thinking, Scrum and Lean Startup: Looking Further: Design Thinking Beyond Solution-Fixation*, pages 199–227. Springer International Publishing.

Dybå, T. and Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic review. *Information and Software Technology*, 50(9):833 – 859.

Fontana, R., Jr, V., Reinehr, S., and Malucelli, A. (2015). Progressive outcomes: A framework for maturing in agile software development. *Journal of Systems and Software*, 102:88–108.

Fontana, R. M., Albuquerque, R., Luz, R., Moises, A. C., Malucelli, A., and Reinehr, S. S. (2018). Maturity models for agile software development: What are they? In Larrucea, X., Santamaria, I., O'Connor, R. V., and Messnarz, R., editors, *Proceedings of the European Systems, Software and Services Process Improvement*, volume 896 of *Communications in Computer and Information Science*, pages 3–14, Bilbao, Spain. Springer.

Fontana, R. M., Reinehr, S., and Malucelli, A. (2015). Agile compass: A tool for identifying maturity in agile software-development teams. *IEEE Software*, 32(6):20–23.

Fontana, R. M., Reinehr, S. S., and Malucelli, A. (2014). Maturing in agile: What is it about? In Cantone, G. and Marchesi, M., editors, *International Conference on Agile Processes in Software Engineering and Extreme Programming*, volume 179 of *Lecture Notes in Business Information Processing*, pages 94–109, Rome, Italy. Springer.

García-Mireles, G., Moraga, M., and Garcia, F. (2012). Development of maturity models: A systematic literature review. In *International Conference on Evaluation & Assessment in Software Engineering*, volume 2012, pages 279–283, Ciudad Real, Spain.

Grossman-Kahn, B. and Rosensweig, R. (2012). Skip the silver bullet: driving innovation through small bets and diverse practices. *Leading Through Design*, page 815.

Guyatt, G., Rennie, D., Meade, M., and Cook, D. (2008). *Users' guide to the medical literature essentials of evidence-based clinical practice*, volume 270 (21). Mc Graw Hill, 2nd edition.

Hoda, R.; Salleh, N. G. J. T. H. M. (2017). Systematic literature reviews in agile software development: A tertiary study. *Information and Software Technology*, 85:60–70.

Humble, J. and Russell, R. (2009). The agile maturity model applied to building and releasing software.

Inayat, I., Salim, S. S., Marczak, S., Daneva, M., and Shamshirband, S. (2015). A systematic literature review on agile requirements engineering practices and challenges. *Computers in Human Behavior*, 51:915–929. Computing for Human Learning, Behaviour and Collaboration in the Social and Mobile Networks Era.

Jørgensen, F., Matthiesen, R. V., Nielsen, J., and Johansen, J. (2007). *Lean Maturity, Lean Sustainability*, volume 246, pages 371–378. Springer.

Karvonen, T., Rodriguez, P., Kuvaja, P., Mikkonen, K., and Oivo, M. (2012). Adapting the lean enterprise self-assessment tool for the software development domain. In *Euromicro Conference on Software Engineering and Advanced Applications*, pages 266–273, Cesme, Izmir, Turkey.

Kieffer, S. and Vanderdonckt, J. (2016). Stratus: A questionnaire for strategic usability assessment. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, pages 205–212, Pisa, Italy. ACM.

Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering.

Kiteley, R. and Stogdon, C. (2013). *Literature reviews in social work*. Sage.

Kuusinen, K. et al. (2017). Knowledge sharing in a large agile organisation: A survey study. In *International Conf on Agile Software Development*, pages 135–150, Cologne, Germany.

Lui, K. M. and Chan, K. C. C. (2006). A road map for implementing extreme programming. In *Software Process Workshop: Unifying the Software Process Spectrum*, pages 474–481, Beijing, China. Springer Berlin Heidelberg.

Maier, A. M., Moultrie, J., and Clarkson, P. J. (2012). Assessing organizational capabilities: Reviewing and guiding the development of maturity grids. *IEEE Transactions on Engineering Management*, 59(1):138–159.

Medappa, P. and Bhattacharya, S. (2012). Towards a framework for assessing agility. *Hawaii International Conference on System Sciences*, 0:5329–5338.

Moralles, C., Vaccaro, M., Zorzetti, M., Pereira, E., Trindade, C., Prauchner, B., Marczak, S., and Bastos, R. (2019). On the mapping of underlying concepts of a combined use of lean and user-centered design with agile development: The case study of the transformation process of an it company. In Meirelles, P., Nelson, M. A., and Rocha, C., editors, *Agile Methods*, pages 25–40, Belo Horizonte, Brazil. Springer International Publishing.

Nawrocki, J., Walter, B., and Wojciechowski, A. (2001). Toward maturity model for extreme programming. In *Proceedings EUROMICRO Conference. 2001: A Net Odyssey*, pages 233–239, Warsaw, Poland.

Norman, D. A. (2002). *The Design of Everyday Things*. Basic Books, New York, USA.

Ogunyemi, A., Lamas, D., Stage, J., and Lárusdóttir, M. (2017). Assessment model for hci practice maturity in small and medium sized software development companies. In *International Conference on Software Process Improvement and Capability Determination*, pages 55–69, Palma de Mallorca, Spain. Springer.

Özcan Top, Ö. and Demirörs, O. (2014). Assessing software agility: An exploratory case study. In Mitasiunas, A., Rout, T., O'Connor, R. V., and Dorling, A., editors, *International Conference on Software Process Improvement and Capability Determination*, pages 202–213, Vilnius, Lithuania. Springer International Publishing.

Paasivaara, M. et al. (2018). Large-scale agile transformation at ericsson: A case study. *Empirical Software Engineering*, 23:2550–2596.

Packlick, J. (2007). The agile maturity map a goal oriented approach to agile improvement. In *AGILE 2007*, pages 266–271, Washington, DC, USA. IEEE Computer Society.

Patel, C. and Ramachandran, M. (2009a). Agile maturity model (amm): A software process improvement

framework for agile software development practices. *International Journal of Software Engineering*, 2.

Patel, C. and Ramachandran, M. (2009b). Story card maturity model (SMM): A process improvement framework for agile requirements engineering practices. *Journal of Software*, 4(5):422–435.

Peres, A. L., Silva, T. S. D., Silva, F. S., Soares, F. F., Carvalho, C. R. M. D., and Meira, S. R. D. L. (2014). Agileux model: Towards a reference model on integrating ux in developing software using agile methodologies. In *Agile Conference*, pages 61–63, Orlando, FL, USA.

Petersen, K., Vakkalanka, S., and Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64:1–18.

Proulx, M. (2010). Yet another agile maturity model (amm) – the 5 levels of maturity.

Quintal, C. and Macías, J. A. (2018). A capability maturity proposal for usability and accessibility centered processes. In *Proceedings of the International Conference on Human Computer Interaction*, pages 15:1–15:8, Palma, Spain. ACM.

Qumer, A. and Henderson-Sellers, B. (2008). A framework to support the evaluation, adoption and improvement of agile methods in practice. *Journal of Systems and Software*, 81(11):1899–1919.

Ries, E. (2011). *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses*. Crown Business.

Salah, D., Paige, R., and Cairns, P. (2015). Patterns for integrating agile development processes and user centred design. In *Proceedings of the 20th European Conference on Pattern Languages of Programs*, page 19, Kaufbeuren, Germany. ACM.

Salah, D., Paige, R., and Cairns, P. (2016). A maturity model for integrating agile processes and user centred design. In Clarke, P. M., O'Connor, R. V., Rout, T., and Dorling, A., editors, *Software Process Improvement and Capability Determination*, pages 109–122, Dublin, Ireland. Springer International Publishing.

Schön et al. (2017). Key challenges in agile requirements engineering. In *International Conference on Agile Software Development*, pages 37–51, Cologne, Germany. Springer.

Schröders, T. and Cruz-Machado, V. (2015). Assessing lean implementation. In *Industrial Engineering, Management Science and Applications*, pages 803–811. Springer.

Shukla, S. K. and Sushil (2020). Evaluating the practices of flexibility maturity for the software product and service organizations. *International Journal of Information Management*, 50:71–89.

Sidky, A. S., Arthur, J. D., and Bohner, S. A. (2007). A disciplined approach to adopting agile practices: the agile adoption framework. *Innovations in Systems and Software Engineering*, 3:203–216.

Signoretti, I., Marczak, S., Salerno, L., d. Lara, A., and Bastos, R. (2019). Boosting agile by using user-centered design and lean startup: A case study of the adoption of the combined approach in software development. In *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pages 1–6, Porto de Galinhas, Recife, Brazil.

Silva, F. S., Soares, F. S. F., Peres, A. L., de Azevedo, I. M., Pinto, P. P., and de Lemos Meira, S. R. (2014). A reference model for agile quality assurance: Combining agile methodologies and maturity models. In *International Conference on the Quality of Information and Communications Technology*, pages 139–144, Guimaraes, Portugal. IEEE Computer Society.

Soares, F. and Meira, S. (2015). An agile strategy for implementing cmmi project management practices in software organizations. In *Iberian Conference on Information Systems and Technologies*, pages 1–4, Águeda, Portugal.

Soundararajan, S., Balci, O., and Arthur, J. D. (2013). Assessing an organization's capability to effectively implement its selected agile method(s): An objectives, principles, strategies approach. In *2013 Agile Conference*, pages 22–31, Leuven, Belgium.

Stanisavljevic, Z., Walter, B., Vukasovic, M., Todosijevic, A., Labedzki, M., and Wolski, M. (2018). Geant software maturity model. In *Telecommunications Forum*, pages 420–425, Belgrade, Serbia.

Stojanov, I., Turetken, O., and Trienekens, J. J. M. (2015). A maturity model for scaling agile development. In *Euromicro Conference on Software Engineering and Advanced Applications*, pages 446–453, Madeira, Portugal.

Van Tyne, S. (2009). Corporate user-experience maturity model. In *International Conference on Human Centered Design*, pages 635–639, San Diego, CA, USA. Springer.

Ximenes, B. H., Alves, I. N., and Araújo, C. C. (2015). Software project management combining agile, lean startup and design thinking. In Marcus, A., editor, *Design, User Experience, and Usability: Design Discourse*, pages 356–367, Los Angeles, CA, USA. Springer International Publishing.

Yin, A., de Lemos Figueiredo, S., and Mira da Silva, M. (2011). Scrum maturity model. In *International Conference on Software Engineering Advances*, Barcelona, Spain. IARIA.