



Adopting Agile Software Development Combined with User-Centered Design and Lean Startup: A Systematic Literature Review on Maturity Models

Maximilian Zorzetti¹(✉), Cassiano Moralles¹, Larissa Salerno¹, Eliana Pereira², Sabrina Marczak¹, and Ricardo Bastos¹

¹ MunDDoS Research Group – PPGCC, School of Technology, Pontifical Catholic University of Rio Grande Do Sul, Porto Alegre, RS, Brazil

{maximilian.zorzetti,cassiano.mora,larissa.salerno}@acad.pucrs.br,
{sabrina.marczak,bastos}@pucrs.br

² Instituto Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
eliana.pereira@restinga.ifrs.edu.br

Abstract. The use of Agile in the software development industry in the past two decades revealed that it is lackluster in some aspects, such as in guaranteeing user involvement and assuring that the right software is being built. There are reports that combining Agile with Lean Startup and User-Centered Design (UCD) helps in overcoming these shortcomings while also yielding several other benefits. However, there is not much documentation on how to use this “combined approach” and adapting existing organizations to use it is a challenge in of itself, in which the use of an instrument to guide or assess such transformations is typically pivotal to their success. As such, in this paper we seek to identify maturity models that assess the use of Agile, Lean Startup, and UCD. We conducted a systematic literature review of maturity models for these three methods published between 2001 and 2020. We characterized the maturity models and determined how they see maturity, how they are applied, and how they were evaluated. As an extended version of a previous paper, we augmented our analysis criteria and further classified the models in how they interpret maturity and what strategy they suggest when undergoing an improvement process, in addition to providing new insight on various aspects of the models. We found 35 maturity models, of which 23 were for Agile, 5 for Lean thinking, 5 for UCD, and 2 for Agile and UCD combined. No models for the combination of the three methods were found (nor for Lean Startup), as expected due to the novelty of the approach. Existing models mostly focus on practice adoption and acquiring continuous improvement capabilities, and are typically developed with a specific context in mind. We also note a lack of proper evaluation procedures being conducted on the majority of models, which could be due to the lack of well-established maturity model development methods and guidelines.

Keywords: Systematic literature review · Maturity model · Agile software development · Lean startup · User-Centered Design

1 Introduction

Agile is extensively used by organizations today [26] as it serves as a powerful and adaptive alternative to the rigid and wasteful software development approaches that were previously used, e.g., waterfall. However, there are some issues with Agile that indicate that it might not be enough by itself—such as lack of user involvement [62] and clear identification of added value [35]. Recent industry cases [22, 66] 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 [57] focuses on adding value to business stakeholders through strategic experimentation, while UCD [43] 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) [48], 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 [39] and provide guidance towards improving software engineering processes [18]. Maturity models are widely used in several domains as a means to improve something (e.g., processes or products) and are typically suggest “levels” of maturity to be achieved, as is the case for the well-known Capability Maturity Model Integration (CMMI) [12].

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 review following existing best practices for systematic reviews on the software engineering domain. We report on several aspects of maturity models found both in academic and gray literature. As an extended version of a previous paper [78], our study provides new analysis criteria and novel insight on currently available maturity models.

The remainder of this paper is organized as follows: Sect. 2 discusses the use of a combined approach of Agile, Lean Startup, and UCD in software development; Sect. 3 discusses related work; Sect. 4 explains how the systematic literature review was conducted and outlines the research questions; Sect. 5 presents our findings; Sect. 6 deliberates on our analysis; and Sect. 7 deliberates on this study and considers future work.

2 Agile, Lean Startup, and User-Centered Design

The Agile movement dates back to 2001 with the introduction of the Agile Manifesto [5], 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 [62]. 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 [75] and that is drawing the attention of academics [13, 75] and industry practitioners [22, 66].

This combined approach tackles business-level issues with Lean Startup, an entrepreneurship method 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, enabling organizations to pivot away from ideas that data suggests to be unfruitful and persevere on the ones most likely to succeed [57]. This idea pulls heavily from traditional Lean values, by way of reducing development waste on new products or features that do not earn back enough to warrant being successful. Although not specifically a software development method, studies have reported on it being a great driving force for software development [15, 76], although embracing its continuous experimentation practices requires proper technological capabilities (e.g., continuous deployment) and organizational support (e.g., culture) [37].

To ensure that the software not only meets business demands but also the users’, the combined approach enlists the use of UCD to enable developers to understand the user’s real needs and create improved software with better usability and user satisfaction [59]. 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 [43] at varying degrees of intensity, from consultation of their needs to having them actively participate in the design process [1]. Integrating UCD and software development can help developers with the difficult practice of involving customers, and the wider concern of how to integrate human-computer-interaction concerns with software engineering [9, 61]. As it stands, UCD has evolved into an umbrella term for similar approaches, thus encompassing terms like Design Thinking and Human-Centered Design.

One successful example of the combined approach is fashion retailer Nordstrom’s Discovery by Design. Grossman-Kahn and Rosensweig [22] report on the evolution of the Nordstrom Innovation Lab, a Nordstrom initiative 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. The team with its “iterative mindset, relentless focus on the needs of the customer, and bias towards rapid experimentation, prototyping and testing” [22] emerged as a powerful and dynamic asset for Nordstrom.

In academia, Moralles et al. [41] 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 [16], backed by an agile maturing framework [17], 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 Related Work

There has been several studies contemplating the integration of Agile Software Development with UCD [61]. Adding Lean Startup to this “method combo” is rather of a novelty given the time frame of the three approaches. Current literature encompasses studies contemplating several aspects of the approach itself (e.g., benefits, challenges, or use of experimentation) [66,67,74] and also studies that propose models to using the combined approach with varying degrees of abstraction [13,22,75]. The combined approach can be very different from typical agile development as it requires a certain degree of developer empowerment that larger organizations might not be used to, making adaptation efforts difficult and the use of maturity models and enticing choice.

Maturity models for agile development has been an interesting research subject ever since the rise of Agile: studies on such maturity models can be traced back to the early 2000s. Leppänen [36] reports that maturity models for Agile have varying levels of maturity, and Nurdiani et al. [44] compares the practice adoption order proposed in existing maturity models with that of industry experts. Ozcan-Top and Demirörs [46] evaluated the strengths and weaknesses of Agile maturity models and frameworks from a process assessment and improvement perspective. Fontana et al. [19] conducted a systematic literature review on Agile maturity models and delineated a classification criteria for how maturity can be defined based on the analysis of the identified models. Henriques and Tanner [25] bring to light that maturity models lack research providing them validation. Pereira and Serrano [52] analyze the main development and evaluation methods for IT maturity models.

4 Systematic Review Protocol

This study was conducted as a systematic literature review based on guidelines for conducting systematic literature reviews in software engineering [33]. 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. The protocol for the systematic literature review is documented next.

4.1 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 do these maturity models envision maturity?
- RQ4.** How are these maturity models applied?
- RQ5.** How are these maturity models evaluated?

4.2 Search

As suggested by Kitchenham [33], 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 SR 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, IEEEExplore, Science Direct, Scopus, and Springer Database. We adapted the search string (Eq. 1) for each database based on the search functionality offered by each. 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

Table 1. Keywords used in the search string of each SR [78].

| 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 |

- 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.

$$S1 \text{ AND } S2 \tag{1}$$

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 Sect. 4.3. 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 (Sect. 4.5).

4.3 Study Selection

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

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 it is the publication date of the Agile Manifesto [5]. 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 [34].

4.4 Quality Assessment

We used a set of quality criteria proposed by Guyatt et al. [23]—later used by Dybå and Dingsøyr [14] in software engineering—to assess the methodological quality of the studies selected for review, as they cover thoroughness, trustworthiness, and significance of the studies [28]. 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 scored the selected studies on each criterion using an ordinal scale instead of a dichotomous scale to obtain a more accurate assessment [28]. Table 2 shows the scoring scale for each criterion. When there was not an agreement on a study's score, we had meetings to discuss the issue until we agreed upon the same score.

Table 2. Quality criteria scoring scheme [78].

| Criteria | Response scoring |
|----------|--|
| 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) |

4.5 Data Extraction and Analysis

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

- D1** – Study identification (RQ1);
- D2** – Publication source and year (RQ1);
- D3** – Audience: the expected users of the model (RQ2);
- D4** – Aim: if the model determines necessary improvements for its use case (analysis) or if it presents best practices for comparison (benchmarking) (RQ2);
- D5** – Scope: if the model is generic or limited to a specific method (RQ2);
- D6** – Strategy: whether the model suggests a “big bang” or “gradual” approach to its adoption process (RQ2);
- D7** – Maturity levels: if the model has specified quantifiable standards (levels) of maturity and has definitions for each (RQ3);
- D8** – Maturity class: whether the model’s characterization of maturity fits into “practices adoption”, “continuous improvement”, “sustaining approach”, “project performance”, or “highly productive teams”; (RQ3)
- D9** – Administration mechanism: if the model has defined a mechanism to apply the model (RQ4).
- D10** – Evaluation: if the model was evaluated, such as by having it applied in a real context (RQ5); and
- D11** – Evaluation type: whether the model’s evaluation process (if any) can be regarded as an “author evaluation”, “domain expert evaluation”, or “practical setting evaluation” (RQ5);

Most of these 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.

New to this study are items D2, D3, D6, D8, and D11. We enhance our existing evaluation analysis (D10) by categorizing existing evaluation methods as defined by Helgesson, Höst, and Weyns [24] and using Salah and Cairn’s [58] nomenclature: “author evaluation”, an evaluation performed by the model’s authors to assess the model’s processes regarding its intended use or to compare it to similar models; “domain expert evaluation”, an evaluation performed

by domain experts external to the model’s development; and “practical setting evaluation”, an evaluation that involves applying the model in a practical setting.

We also make use of the maturity definitions outlined by Fontana et al. [19] and analyze each model’s characterization of maturity, which can resolve into the following categories: “practices adoption”, maturity is increased when new practices are adopted; “continuous improvement”, maturity is similar to CMMI-DEV’s, i.e., mature organizations/teams focus on process improvement; “sustaining approach”, maturity implies organizations/teams with lean processes and adherence to agile values; “project performance”, maturity is a way to obtain results; and “highly productive teams”, maturity is related directly to team productivity.

Additionally, we draw on the work of Julian, Noble, and Anslow [30], who argue that having a strategy to adopting a new approach is crucial for its success, and categorize the suggested strategies of existing models as either “big bang”, in which a set of practices are adopted all at once by-the-book for teams then to learn and modify, or “gradual”, in which practices are gradually adopted and adapted alongside in-place non-agile practices in what is a typically longer transitioning period.

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 [19] that focused on Agile maturity models as our search resolved into a superset of the models it identified.

5 Results

This section summarizes the results of each SR. Table 3 presents the results of the search process in the electronic databases selected in Sect. 4.2. We analyze the studies in light of our research questions based on the data extracted using the procedure in Sect. 4.5 next. Entries marked as “—” on the following tables stand for “unspecified”.

Table 3. Number of identified studies during the distinct rounds of our systematic search for maturity models [78].

| 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 |

Table 4. Selected maturity model studies and their sources.

| Context | Study | Year | Source | Source type |
|---------------|-------|--|--|-------------|
| Agile | [42] | 2001 | EUROMICRO | Conference |
| | [38] | 2005 | International Software Process Workshop | Conference |
| | [65] | 2007 | Innovations in Systems and Software Engineering | Journal |
| | [49] | 2007 | Agile Development Conference | Conference |
| | [56] | 2008 | Journal of Systems and Software | Journal |
| | [50] | 2009 | Int'l. Journal of Software Engineering | Journal |
| | [51] | 2009 | Journal of Software | Journal |
| | [27] | 2009 | Web publishing (ThoughtWorks) | Grey Lit |
| | [6] | 2010 | Hawaii Int'l. Conference on System Sciences | Conference |
| | [54] | 2010 | Web publishing (personal blog) | Grey Lit |
| | [77] | 2011 | Int'l. Conference on Software Engineering Advances | Conference |
| | [8] | 2011 | Int'l. Conference on Product Focused Software Development and Process Improvement | Conference |
| | [40] | 2012 | Hawaii Int'l. Conference on System Sciences | Conference |
| | [70] | 2013 | Agile Development Conference | Conference |
| | [20] | 2014 | Int'l. Conference on Agile Processes in Software Engineering and Extreme Programming | Conference |
| | [68] | 2014 | Int'l. Conference on the Quality of Information and Communications Technology | Conference |
| | [47] | 2014 | Int'l. Conference on Process Improvement and Capability Determination | Conference |
| | [69] | 2015 | Iberian Conference on Information Systems and Technologies | Conference |
| | [16] | 2015 | Journal of Systems and Software | Journal |
| | [72] | 2015 | EUROMICRO | Conference |
| [3] | 2016 | Web publishing (CrossTalk) | Grey Lit | |
| [71] | 2018 | Telecommunications Forum | Conference | |
| [64] | 2020 | Int'l. Journal of Information Management | Journal | |
| Lean | [29] | 2007 | Int'l. Federation for Information Processing | Book Chap. |
| | [31] | 2012 | Euromicro | Conference |
| | [11] | 2012 | International Journal of Advanced Manufacturing Technology | Journal |
| | [63] | 2015 | Int'l Conference on Industrial Engineering, Management Science and Applications | Conference |
| | [2] | 2018 | IEEE Transactions on Software Engineering | Journal |
| UCD | [73] | 2009 | Int'l Conference on Human Centered Design | Conference |
| | [10] | 2014 | Int'l. Conference of Design, User Experience, and Usability | Conference |
| | [32] | 2016 | Symposium on Applied Computing | Conference |
| | [45] | 2017 | Int'l. Conference on Software Process Improvement and Capability Determination | Conference |
| | [55] | 2018 | Int'l Conference on Human Computer Interaction | Conference |
| Agile and UCD | [53] | 2014 | Agile Development Conference | Conference |
| | [60] | 2016 | Int'l. Conference on Software Process Improvement and Capability Determination | Conference |

Table 5. Selected maturity model studies and their respective quality scores [78].

| Context | Study | C1 | C2 | C3 | C4 | Avg. |
|---------------|-------|-----|-----|-----|-------|-------|
| Agile | [42] | 1.0 | 1.0 | 0.5 | 0.5 | 0.75 |
| | [38] | 1.0 | 0.5 | 0.5 | 0.5 | 0.625 |
| | [65] | 1.0 | 0.5 | 0.5 | 1.0 | 0.75 |
| | [49] | 1.0 | 0.5 | 1.0 | 1.0 | 0.875 |
| | [56] | 0.5 | 0.5 | 1.0 | 0.5 | 0.625 |
| | [50] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | [51] | 0.5 | 1.0 | 0.5 | 1.0 | 0.75 |
| | [27] | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | [6] | 0.5 | 0.5 | 0.5 | 1.0 | 0.625 |
| | [54] | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | [77] | 1.0 | 0.5 | 1.0 | 1.0 | 0.875 |
| | [8] | 0.5 | 1.0 | 0.5 | 1.0 | 0.75 |
| | [40] | 1.0 | 1.0 | 1.0 | 0.5 | 0.875 |
| | [70] | 1.0 | 1.0 | 0.5 | 1.0 | 0.875 |
| | [20] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | [68] | 1.0 | 1.0 | 0.5 | 0.5 | 0.75 |
| | [47] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | [69] | 1.0 | 1.0 | 1.0 | 0.5 | 0.875 |
| | [16] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | [72] | 1.0 | 0.5 | 1.0 | 1.0 | 0.875 |
| [3] | 0.5 | 0.5 | 0.5 | 0.0 | 0.375 | |
| [71] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | |
| [64] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | |
| Lean | [29] | 1.0 | 1.0 | 0.5 | 0.5 | 0.75 |
| | [31] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | [11] | 1.0 | 1.0 | 0.5 | 0.5 | 0.75 |
| | [63] | 1.0 | 1.0 | 0.0 | 0.5 | 0.875 |
| | [2] | 1.0 | 1.0 | 0.5 | 0.0 | 0.625 |
| UCD | [73] | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | [10] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | [32] | 1.0 | 1.0 | 0.5 | 0.5 | 0.75 |
| | [45] | 1.0 | 1.0 | 0.5 | 1.0 | 0.875 |
| | [55] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| Agile and UCD | [53] | 1.0 | 0.5 | 0.5 | 0.5 | 0.625 |
| | [60] | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

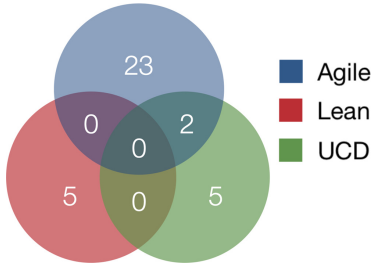


Fig. 1. Venn diagram of maturity models for Agile, Lean, and UCD (adapted [78]).

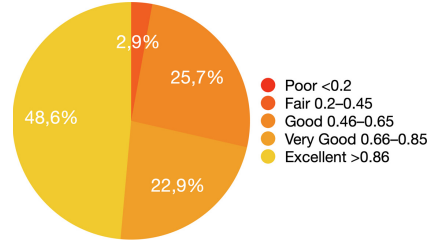


Fig. 2. Normalized quality score distribution for maturity model studies.

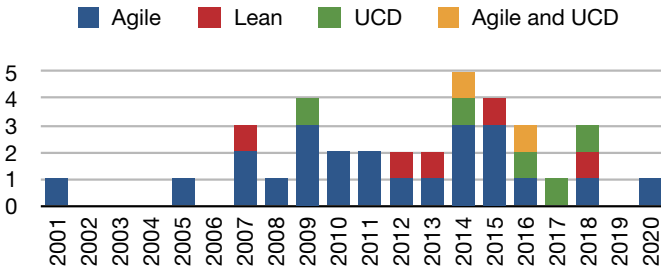


Fig. 3. Publication frequency of maturity models [78].

5.1 RQ1. What Maturity Models Are Available?

As mentioned in Table 3, our systematic literature review 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 [50, 56, 65, 77] and 2 gray literature studies [3, 54] for a total of 35 studies. Table 4 shows the selected maturity models studies and their publication year and venue/source. Our initial objective was to identify maturity models for a combined approach of Agile, UCD, and Lean Startup. This search, however, proved fruitless. 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. The higher number of maturity models for Agile is expected, as it is the most dominant approach to software engineering worldwide. Of the existing Lean studies, we point out that none concern the use of Lean Startup.

We assessed the quality of the papers as per systematic literature review guidelines [33] (see Table 5). The papers scored approximately 0.78 on average, with at most a 0.03 score difference between identified categories (Agile, UCD, Lean Startup, and Agile with UCD). Studies with a low score (0.45 and below)

tended to be short studies with a low page count. Figure 2 shows the studies' scores in a normalized fashion.

Figure 3 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.

5.2 RQ2. How Are These Maturity Models Characterized?

Table 6 shows an overview of the maturity models. While earlier agile maturity models focused on XP [6, 38, 42] due to it being one of the first agile methods to become popular [4] (and also harboring a strong influence from CMMI), the majority of the maturity models are generic, focusing on the general idea and values of the method they adhere to. Although most of them come from a more general need for maturity models, some originate from very specific demands, such as Lui and Chan's model [38], which was specifically developed to help Chinese companies dealing with commercial-off-the-shelf software; or Peres' [53], which focuses on integrating Agile and UCD on CMMI level two compliant organizations.

We identified some benchmark-based maturity models (two for Agile [6, 49] and two for Lean [29, 31]) that were based on CMMI-DEV [12] and adapted from industry models to be used in a generic fashion. The remaining models have analysis as their focus, evaluating the situation of teams and organizations and making use of the notion that each organization has its own unique context and characteristics, where comparisons can be inadequate and faced with resistance [17].

The Lean maturity models were mostly influenced by the manufacturing industry and propose gradually evolving circumstances, focusing on a sustainable adoption of their method of choice [2, 11, 31, 63], with the exception of Cil and Turkan [11] with its analytic network process approach. Whilst Julian et al. [30] identifies two paths to minimize the effects of adopting "big bang" and "gradual" strategies, Lui and Chan [38] interestingly considers both approaches.

5.3 RQ3. How Do These Maturity Models Envision Maturity?

Maturity can be seen as a state to be achieved by teams or organizations, as such, maturity models mostly quantify this by defining maturity levels. Models usually define four to six levels, most commonly defaulting to five (see Table 7). Some authors do not use levels to quantify maturity, however, such as Fontana et al. [17], who argue that "maturity is obtained when results are accomplished in various aspects of software development" and propose an outcome-based model; Cil and Turkan [11], who propose a model based on analytic network process;

Table 6. Overview of data extracted from selected maturity model studies (adapted [78]).

| Context | Study | Adoption process | Aim | Scope | Evaluation | Maturity level | Audience | Admin. Mech. |
|---------------|----------|------------------|--------------|-------------------|------------|-------------------|---------------------------|--------------|
| Agile | [42] | Big bang | Analysis | XP | Yes | Yes | Organizations adopting XP | Partial |
| | [38] | Big bang/gradual | Analysis | XP | No | Yes | Inexperienced XP Teams | No |
| | [65] | Gradual | Analysis | Generic | Yes | Yes | Agile Teams | Yes |
| | [49] | Gradual | Benchmarking | Generic | Yes | Yes | Agile Teams | Yes |
| | [56] | Gradual | Analysis | Generic | Yes | Yes | Agile Teams | Yes |
| | [50] | Gradual | Analysis | Generic | Yes | Yes | Agile Teams | Yes |
| | [51] | Gradual | Analysis | Generic | Yes | Yes | Agile Teams | Yes |
| | [27] | Gradual | Analysis | Generic | Yes | Yes | Agile Teams | Yes |
| | [6] | Big bang | Benchmarking | XP | Yes | Yes | XP Teams | Partial |
| | [54] | Gradual | Analysis | Generic | No | Yes | Any organization | No |
| | [77] | Gradual | Analysis | Scrum | Yes | Yes | Scrum Teams | Partial |
| | [8] | Gradual | Analysis | Generic | No | Yes | Agile Teams | No |
| | [40] | — | Analysis | Generic | No | Yes | Agile Teams | Yes |
| | [70] | Gradual | Analysis | Generic | No | Yes | Agile Teams | Yes |
| | [20] | — | Analysis | Generic | No | Yes | Agile Teams | No |
| | [68] | Gradual | Analysis | Quality Assurance | Yes | Yes | Agile and CMMI | No |
| | [47] | Gradual | Analysis | Generic | Yes | Yes | Implementing Agile | Yes |
| | [69] | Gradual | Analysis | Generic | No | Yes | Agile and CMMI | No |
| | [16] | — | Analysis | Generic | No | No | Agile Teams | Yes |
| | [72] | Big bang | Analysis | Generic | Yes | Yes | Scaling Agile | Yes |
| [3] | Gradual | Analysis | Generic | No | Yes | Agile Teams | No | |
| [71] | Big bang | Analysis | Generic | No | Yes | Distributed teams | Partial | |
| [64] | Gradual | Analysis | Generic | No | Yes | No evidence | No | |
| Lean | [29] | — | Benchmarking | Generic | No | Yes | Lean Sustainability | No |
| | [31] | Gradual | Benchmarking | Generic | Yes | Yes | Soft. Dev. Teams | Yes |
| | [11] | — | Analysis | Generic | Yes | No | Lean Transformation | Yes |
| | [63] | Gradual | Analysis | Generic | No | No | Lean Implementation | No |
| | [2] | Gradual | Analysis | Kaizen | Yes | No | Soft. Dev. Teams | Yes |
| UCD | [73] | Gradual | Analysis | Generic | No | Yes | All Organization | No |
| | [10] | Gradual | Analysis | Generic | No | Yes | Soft. Dev. Teams | No |
| | [32] | Gradual | Analysis | Generic | Yes | Yes | Soft. Dev. Teams | Yes |
| | [45] | — | Analysis | Generic | No | No | Soft. Dev. Teams | Yes |
| | [55] | Gradual | Analysis | Generic | Yes | Yes | Soft. Dev. Teams | Yes |
| Agile and UCD | [53] | Big bang | Analysis | Generic | Yes | Yes | Scrum Teams | Yes |
| | [60] | Gradual | Analysis | Generic | Yes | Yes | Agile Teams | Yes |

and Schröders and Cruz-Machados [63], whose model individually assesses and quantitatively rates several criteria pertaining to leadership, culture, knowledge, and process.

5.4 RQ4. How Are These Maturity Models Applied?

Table 8 shows the administration mechanisms of the selected maturity models. Administration mechanisms tend to be simple, not deviating much from instruments similar to questionnaires or checklists, attesting to the inexpensive and somewhat swift quality maturity models are known for. Of note is the work of Patel and Ramachandran [51], which has the support of a web-based tool; and Cil and Turkan [11], which employs the analytical network process to resolve assessments. Almost half of the maturity models, however, provide no mechanism at all, perhaps overstepping on simplicity.

We notice a necessity for more studies and/or discussions on how administration mechanisms must be designed or used in a way that is adequate for the specific needs for which their corresponding maturity model was designed for. As a mechanism sensitive to the context it is applied on, it most likely needs to be highly dynamic features to properly assess and identify improvement needs and to perform follow-up measurements. Only a few maturity models present in-depth metrics, typically analysis-focused ones [11], but we speculate that using both a quantitative and qualitative approach to analysis might be a better path to proper assessment.

5.5 RQ5. How Are These Maturity Models Evaluated?

Table 9 shows how each maturity model was evaluated. Of the models that were in fact evaluated, most of them underwent either a domain-expert or practical setting evaluation. Given that author evaluations seem to be the cheapest alternative and that the majority of the models were not evaluated at all, this perhaps suggests that author evaluations are seen as not worthwhile by academics, or that there is a lack of industry involvement during the development of these artifacts. Indeed, only a few models seem to be actively used in the software industry and none had follow-up studies, revealing a concerning detachment between academia and industry with regard to maturity models.

Table 7. Maturity class and levels of maturity models.

| Context | Study | Maturity class | Maturity levels |
|---------------|------------------------|---|---|
| Agile | [42] | Practices adoption, project performance | 4 levels, based on existence and adherence to a structured process |
| | [38] | Practices adoption | 4 levels, based on learning |
| | [65] | Practices adoption, sustaining approach | 5 levels, based on existence and adherence to a structured process |
| | [49] | Highly productive teams | 5 levels, based on people |
| | [56] | Sustaining approach | 6 levels, based on existence and adherence to a structured process |
| | [50] | Project performance | 5 levels, based on existence and adherence to a structured process |
| | [51] | Practices adoption | 4 levels, based on existence and adherence to a structured process |
| | [27] | Continuous improvement | 5 levels, based on existence and adherence to a structured process |
| | [6] | Highly productive teams | 6 levels, based on existence and adherence to a structured process |
| | [54] | Project performance | 5 levels, based on existence and adherence to a structured process (practices) |
| | [77] | Continuous improvement | 5 levels, based on existence and adherence to a structured process (Scrum practices) |
| | [8] | Continuous improvement | — |
| | [40] | Sustaining approach | — |
| | [70] | Practices adoption, continuous improvement | — |
| | [20] | Practices adoption | 3 levels, but they mention adherence to a process should not be measured |
| | [68] | Practices adoption | 6 levels, based on existence and adherence to a structured process |
| | [47] | Practices adoption, continuous improvement | 4 levels, based on existence and adherence to a structured process (practices) |
| | [69] | Continuous improvement | 5 levels, based on existence and adherence to a structured process |
| | [16] | — | — |
| [72] | Sustaining approach | 5 levels, based on existence and adherence to a structured process | |
| [3] | Practices adoption | — | |
| [71] | Continuous improvement | — | |
| [64] | Practices adoption | 6 levels, based on existence and adherence to a structured process | |
| Lean | [29] | Practices adoption, highly productive teams | 5 levels, based on existence and adherence to a structured lean culture and process |
| | [31] | Practices adoption | 5 levels, based on existence and adherence to a structured lean culture and process (focus on products) |
| | [11] | — | — |
| | [63] | Highly productive teams, sustaining approach | — |
| | [2] | Practices adoption, project performance, highly productive teams | 3 levels, based on team productivity and team experience |
| UCD | [73] | Practices adoption | 5 levels, based on existence and adherence to a structured user experience process (practices) |
| | [10] | Continuous improvement | 5 levels, based on existence and adherence to a structured user experience process |
| | [32] | Continuous improvement | 3 levels, based on existence and adherence to a structured usability process |
| | [45] | — | — |
| | [55] | Continuous improvement | 5 levels, based on integration to a structured UCD process |
| Agile and UCD | [53] | Practices adoption, continuous improvement | — |
| | [60] | Practices adoption, continuous improvement, highly productive teams | — |

Table 8. Administration mechanisms of maturity models [78].

| Context | Study | Administration mechanism |
|---------------|--|--|
| Agile | [42] | Partial, provides a list of items to be observed by an XP tracker but leaves the definition of an assessment method open-ended |
| | [38] | – |
| | [65] | Uses a Goal-Question-Indicator-Metric approach to measure readiness for practice adoption; provides a 4-step process for organizations to adopt agile |
| | [49] | Reports a case that used user stories based on agile goals |
| | [56] | Provides a custom analytical tool that evaluates agile methods through four distinct perspectives |
| | [50] | Uses questionnaires that are distributed to a project’s development team and any other associated personnel; and a roadmap for software process improvement |
| | [51] | 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 |
| | [27] | Provides a somewhat superficial plan-do-check-act cycle to roll out improvements throughout an organization |
| | [6] | 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 |
| | [54] | – |
| | [77] | Partial, uses an undisclosed checklist of Scrum practices for each maturity level |
| | [8] | – |
| | [40] | Reports a case that used a survey |
| | [70] | Collects data on indicators and through a series of computations resolves into numeric scores for strategies, principles, and objectives |
| | [20] | – |
| | [68] | – |
| | [47] | Uses a questionnaire about specific practices and generic agile practices |
| | [69] | – |
| | [16] | Provides a checklist that helps teams to identify which outcomes they have attained |
| | [72] | Provides indicators to be assessed in assessment meetings |
| [3] | – | |
| [71] | Partial, describes two methods to interpret maturity parameters that are evaluated using discrete scoring, but does not specify how to obtain the latter | |
| [64] | – | |
| Lean | [29] | – |
| | [31] | Provides assessment items for the lean practices in each of its process areas |
| | [11] | Uses the analytical network process |
| | [63] | – |
| | [2] | Specifies a process area for evaluation methods |
| UCD | [73] | – |
| | [10] | – |
| | [32] | Uses a questionnaire to assess organizational strategic usability |
| | [45] | Uses a questionnaire for peer and self-assessment |
| | [55] | Uses U+A SPICE, a mechanism adapted from ISO/IEC 15504 |
| Agile and UCD | [53] | Synchronizes the iterative development cycle with its process improvement suggestions; provides a list of relevant measurements |
| | [60] | 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 |

Table 9. Evaluations performed on maturity models (adapted [78]).

| Context | Study | Evaluation type | Evaluation |
|---------------|-------|------------------|---|
| Agile | [42] | Practical | 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 |
| | [38] | – | – |
| | [65] | Expert | Evaluated through questionnaires answered by 28 members of the agile community |
| | [49] | Practical | Evaluated through 20 teams using it for over six months |
| | [56] | Practical | Evaluated through 2 industry case studies |
| | [50] | Expert | Evaluated through a discussion with 3 different organizations |
| | [51] | Expert | Evaluated through case study in three companies |
| | [27] | – | – |
| | [6] | Practical | Evaluated through a case study in a multinational communication company |
| | [54] | – | – |
| | [77] | Expert/practical | Evaluated through action research, which incorporated interviews with Scrum, Agile, and CMMI experts |
| | [8] | – | – |
| | [40] | – | – |
| | [70] | – | – |
| | [20] | – | – |
| | [68] | Expert | Evaluated through a survey based on the opinion of experts |
| | [47] | Practical | Evaluated through an exploratory case study in a government organization |
| | [69] | – | – |
| | [16] | – | – |
| | [72] | Practical | Evaluated through a case study in a large organization |
| [3] | – | – | |
| [71] | – | – | |
| [64] | – | – | |
| Lean | [29] | – | – |
| | [31] | Author | Evaluated by comparison |
| | [11] | Practical | Evaluated through a case study in a company |
| | [63] | – | – |
| | [2] | Expert | Evaluated through a survey |
| UCD | [73] | – | – |
| | [10] | – | – |
| | [32] | Expert/practical | Evaluated survey between novices and experts |
| | [45] | Expert/practical | Evaluated through a survey with experts and a case study |
| | [55] | – | – |
| Agile and UCD | [53] | Expert | Evaluated by a panel of experts |
| | [60] | Expert | Evaluated by domain experts |

6 Discussion

A large sum of the maturity models resort to the simple course of action of practice adoption as their main focus to improving maturity, though many also

see the competence of being able to continuously improve as the key takeaway of maturity. We highlight that Lean maturity models generally consider the culture and behavior domains, while UCD models focus on development teams trying to adhere to UCD during their work process; outlining the disparity of what maturity even is among different studies. This “confusion” of maturity is not without reason (after all, different models try to solve different things), but makes for a difficult time in trying to establish guidelines on proper maturity model development [7]. Several maturity models are developed to solve the problems and fulfill the specific needs of certain contexts and/or problem domains identified by academia or requested by the industry, even though most of them make use of generic methods. Furthermore, many go without a suitable evaluation procedure and none were demonstrated being applied on real teams or organizations in one or more follow-up studies, casting some doubt on the actual capabilities of some models, despite the fact that our quality assessment resolved into generally positive scores. This could be due to the lack of established guides on maturity model development [52], which would reflect on the quality of the models themselves [21], the inaccessibility of domain-experts in some academia circles, or the higher costs of conducting a proper evaluation procedure with industry partners.

The combined approach of Agile, UCD, and Lean Startup seems most promising still, and it and its variants are being used by organizations worldwide. As larger institutions move to use it, the hardships of large-scale adoption will become more apparent, highlighting the need of maturity models to support them and likely instigating research on the topic. The maturity models reported in this paper garner a lot of knowledge that could help in developing a proper model for the combined approach, even if Lean Startup was not directly addressed by them.

6.1 Threats to Validity

As with any systematic literature 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.

The systematic literature review was conducted by one PhD, one PhD candidate, three graduate students, and two undergraduate students (all from IT-related education); with guidance from two senior researchers.

7 Conclusion

This paper reports on a systematic literature review of maturity models for Agile, Lean Startup, UCD, and their intersections in a software engineering

context following existing systematic review guidelines [33]. 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 [14, 23, 28]. Then, we analyzed and categorized the maturity models using criteria adapted from maturity grid guidelines [39] and other studies [24, 58].

The absence of maturity models for the combined approach of the three pillars is likely due to its infancy. Research on the use of Lean Startup in software development is not as extensive as the other two pillars, which also already have a subject area specific to researching integration efforts between the two, i.e., Agile User-Centered Design Integration (AUCDI). Additionally, all three pillars lack a widely accepted theoretical basis that properly defines each pillar, leading back to issues like “what is a mature agile team?” and what issues should each pillar tackle individually or together; making integration efforts difficult. The yet unexplained inner workings of the combined approach make the development of a maturity model for it a daunting task.

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 [22]. Lean Startup deals heavily with continuous experimentation, a practice that is very much intertwined with the method’s somewhat risk-tolerant mindset, perhaps making future maturity models for it focus on cultural concerns, much like the reported Lean models of this study.

The identified maturity models show a worrisome trend in evaluation procedures: about half of the studies did not report on evaluating their maturity model. Although many studies lack a sound theoretical basis and methodology [21], Pereira and Serrano [52] report that several maturity model development guidelines have been created, but that authors choose to follow their own method instead, which could be the cause of this trend. Nevertheless, the lack of evaluation shows an alarming disconnect from the industry, which is where the maturity models are to be applied in the first place.

For future work, the development of a maturity model for the combined approach of the three pillars is evident, although a better understanding of how the pillars interact should be attained first, even if it has been suggested that cultural and mindset factors should be one of the focus points of such a model [22]. The combined approach is a promising take on software development [22, 66], albeit an understudied one. As practices from Lean Startup continue to be adopted by the software industry, we hope to see improved interest in this subject.

Acknowledgments. We acknowledge that this research is sponsored by Dell Brazil using incentives of the Brazilian Informatics Law (Law no. 8.248, year 1991).

References

1. Abras, C., Maloney-krichmar, D., Preece, J.: User-centered design. In: Encyclopedia of Human-Computer Interaction. SAGE Publications Inc, Bainbridge Island (2004)
2. Al-Baik, O., Miller, J.: Integrative double kaizen loop (IDKL): towards a culture of continuous learning and sustainable improvements for software organizations. *IEEE Trans. Softw. Eng.* **45**(12), 1189–1210 (2019)
3. Ambler, S., Lines, M.: The disciplined agile framework: a pragmatic approach to agile maturity. *J. Defense Softw. Eng. - CrossTalk* **29**, 25–31 (2016)
4. Beck, K.: Embracing change with extreme programming. *Computer* **32**(10), 70–77 (1999)
5. Beck, K., et al.: Manifesto for agile software development (2001)
6. Benefield, R.: Seven dimensions of agile maturity in the global enterprise: a case study. In: Hawaii International Conference on System Sciences, Koloa, Kauai, HI, USA, pp. 1–7. IEEE Computer Society, December 2010
7. de Bruin, T., Freeze, R., Kulkarni, U., Rosemann, M.: Understanding the main phases of developing a maturity assessment model. In: Australasian Conference on Information Systems, January 2005
8. Buglione, L.: Light maturity models (LMM): an agile application. In: Proceedings of the International Conference on Product Focused Software Development and Process Improvement, Torre Canne, Brindisi, Italy, pp. 57–61. ACM (2011)
9. Chamberlain, S., Sharp, H., Maiden, N.: Towards a framework for integrating agile development and user-centred design. In: Abrahamsson, P., Marchesi, M., Succi, G. (eds.) XP 2006. LNCS, vol. 4044, pp. 143–153. Springer, Heidelberg (2006). https://doi.org/10.1007/11774129_15
10. Chapman, L., Plewes, S.: A UX maturity model: effective introduction of UX into organizations. In: Marcus, A. (ed.) DUXU 2014. LNCS, vol. 8520, pp. 12–22. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07638-6_2
11. Cil, I., Turkan, Y.S.: An ANP-based assessment model for lean enterprise transformation. *Int. J. Adv. Manuf. Technol.* **64**(5–8), 1113–1130 (2013)
12. CMMI Product Team: CMMI for Development, 1.3 edn. Software Engineering Institute, Carnegie Mellon University (2010)
13. Dobrigkeit, F., de Paula, D., Uflacker, M.: InnoDev: a software development methodology integrating design thinking, scrum and lean startup. In: Meinel, C., Leifer, L. (eds.) Design Thinking Research. UI, pp. 199–227. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-97082-0_11
14. Dybå, T., Dingsøy, T.: Empirical studies of agile software development: a systematic review. *Inf. Softw. Technol.* **50**(9), 833–859 (2008)
15. Fagerholm, F., Guinea, A.S., Mäenpää, H., Münch, J.: The RIGHT model for continuous experimentation. *J. Syst. Softw.* **123**, 292–305 (2017)
16. Fontana, R.M., Reinehr, S., Malucelli, A.: Agile compass: a tool for identifying maturity in agile software-development teams. *IEEE Softw.* **32**(6), 20–23 (2015)
17. Fontana, R., Meyer Jr., V., Reinehr, S., Malucelli, A.: Progressive outcomes: a framework for maturing in agile software development. *J. Syst. Softw.* **102**, 88–108 (2015)
18. Fontana, R.M.: Maturity in agile software development. Ph.D. thesis, Pontifical Catholic University of Paraná, Curitiba, Brazil (2015)

19. Fontana, R.M., Albuquerque, R., Luz, R., Moises, A.C., Malucelli, A., Reinehr, S.: Maturity models for agile software development: what are they? In: Larrucea, X., Santamaria, I., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2018. CCIS, vol. 896, pp. 3–14. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97925-0_1
20. Fontana, R.M., Reinehr, S., Malucelli, A.: Maturing in agile: what is it about? In: Cantone, G., Marchesi, M. (eds.) XP 2014. LNBIP, vol. 179, pp. 94–109. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-06862-6_7
21. García-Mireles, G., Moraga, M., Garcia, F.: Development of maturity models: a systematic literature review. In: International Conference on Evaluation & Assessment in Software Engineering, Ciudad Real, Spain, vol. 2012, pp. 279–283, January 2012
22. Grossman-Kahn, B., Rosensweig, R.: Skip the silver bullet: driving innovation through small bets and diverse practices. *Leading Through Design*, p. 815 (2012)
23. Guyatt, G., Rennie, D., Meade, M., Cook, D.: *Users' Guide to the Medical Literature Essentials of Evidence-Based Clinical Practice*, 2nd edn., vol. 270, no. 21. McGraw-Hill, New York (2008)
24. Helgesson, Y.Y.L., Höst, M., Weyns, K.: A review of methods for evaluation of maturity models for process improvement. *J. Softw. Evol. Process* **24**(4), 436–454 (2012)
25. Henriques, V., Tanner, M.: A systematic literature review of agile and maturity model research. *Interdiscip. J. Inf. Knowl. Manag.* **12**, 53–73 (2017)
26. Hoda, R., Salleh, N., Grundy, J., Tee, H.M.: Systematic literature reviews in agile software development: a tertiary study. *Inf. Softw. Technol.* **85**, 60–70 (2017)
27. Humble, J., Russell, R.: *The agile maturity model applied to building and releasing software* (2009)
28. Inayat, I., Salim, S.S., Marczak, S., Daneva, M., Shamshirband, S.: A systematic literature review on agile requirements engineering practices and challenges. *Comput. Hum. Behav.* **51**, 915–929 (2015). *Computing for Human Learning, Behaviour and Collaboration in the Social and Mobile Networks Era*
29. Jørgensen, F., Matthiesen, R., Nielsen, J., Johansen, J.: Lean maturity, lean sustainability. In: Olhager, J., Persson, F. (eds.) APMS 2007. ITIFIP, vol. 246, pp. 371–378. Springer, Boston, MA (2007). https://doi.org/10.1007/978-0-387-74157-4_44
30. Julian, B., Noble, J., Anslow, C.: Agile practices in practice: towards a theory of agile adoption and process evolution. In: Kruchten, P., Fraser, S., Coallier, F. (eds.) XP 2019. LNBIP, vol. 355, pp. 3–18. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-19034-7_1
31. Karvonen, T., Rodriguez, P., Kuvaja, P., Mikkonen, K., Oivo, M.: Adapting the lean enterprise self-assessment tool for the software development domain. In: Euromicro Conference on Software Engineering and Advanced Applications, Cesme, Izmir, Turkey, pp. 266–273, September 2012
32. Kieffer, S., Vanderdonckt, J.: STRATUS: a questionnaire for strategic usability assessment. In: Proceedings of the 31st Annual ACM Symposium on Applied Computing, Pisa, Italy, pp. 205–212. ACM (2016)
33. Kitchenham, B., Charters, S.: *Guidelines for performing systematic literature reviews in software engineering* (2007)
34. Kiteley, R., Stogdon, C.: *Literature Reviews in Social Work*. Sage, Thousand Oaks (2013)

35. Kuusinen, K., et al.: Knowledge sharing in a large agile organisation: a survey study. In: International Conference on Agile Software Development, Cologne, Germany, pp. 135–150 (2017)
36. Leppänen, M.: A comparative analysis of agile maturity models. In: Pooley, R., Coady, J., Schneider, C., Linger, H., Barry, C., Lang, M. (eds.) *Information Systems Development*, pp. 329–343. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-4951-5_27
37. Lindgren, E., Münch, J.: Raising the odds of success: the current state of experimentation in product development. *Inf. Softw. Technol.* **77**, 80–91 (2016)
38. Lui, K.M., Chan, K.C.C.: A road map for implementing eXtreme programming. In: Li, M., Boehm, B., Osterweil, L.J. (eds.) *SPW 2005. LNCS*, vol. 3840, pp. 474–481. Springer, Heidelberg (2006). https://doi.org/10.1007/11608035_38
39. Maier, A.M., Moultrie, J., Clarkson, P.J.: Assessing organizational capabilities: reviewing and guiding the development of maturity grids. *IEEE Trans. Eng. Manag.* **59**(1), 138–159 (2012)
40. Medappa, P., Bhattacharya, S.: Towards a framework for assessing agility. In: *Hawaii International Conference on System Sciences*, pp. 5329–5338, January 2012
41. Morales, C., et al.: 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., Rocha, C. (eds.) *WBMA 2019. CCIS*, vol. 1106, pp. 25–40. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-36701-5_3
42. Nawrocki, J., Walter, B., Wojciechowski, A.: Toward maturity model for extreme programming. In: *Proceedings EUROMICRO Conference 2001: A Net Odyssey*, Warsaw, Poland, pp. 233–239, September 2001
43. Norman, D.A., Draper, S.W.: *User Centered System Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associates Inc., New Jersey (1986)
44. Nurdiani, I., Börstler, J., Fricker, S., Petersen, K., Chatzipetrou, P.: Understanding the order of agile practice introduction: comparing agile maturity models and practitioners' experience. *J. Syst. Softw.* (2019). <https://doi.org/10.1016/j.jss.2019.05.035>
45. Ogunyemi, A., Lamas, D., Stage, J., Lárusdóttir, M.: Assessment model for HCI practice maturity in small and medium sized software development companies. In: Mas, A., Mesquida, A., O'Connor, R.V., Rout, T., Dorling, A. (eds.) *SPICE 2017. CCIS*, vol. 770, pp. 55–69. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67383-7_5
46. Ozcan-Top, O., Demirörs, O.: Assessment of agile maturity models: a multiple case study. In: Woronowicz, T., Rout, T., O'Connor, R.V., Dorling, A. (eds.) *SPICE 2013. CCIS*, vol. 349, pp. 130–141. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38833-0_12
47. Top, Ö.Ö., Demirörs, O.: Assessing software agility: an exploratory case study. In: Mitasiunas, A., Rout, T., O'Connor, R.V., Dorling, A. (eds.) *SPICE 2014. CCIS*, vol. 477, pp. 202–213. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-13036-1_18
48. Paasivaara, M., et al.: Large-scale agile transformation at Ericsson: a case study. *Empirical Softw. Eng.* **23**, 2550–2596 (2018). <https://doi.org/10.1007/s10664-017-9555-8>
49. Packlick, J.: The agile maturity map a goal oriented approach to agile improvement. In: *AGILE 2007*, Washington, DC, USA, pp. 266–271. IEEE Computer Society (2007)

50. Patel, C., Ramachandran, M.: Agile maturity model (AMM): a software process improvement framework for agile software development practices. *Int. J. Softw. Eng.* **2**, 3–28 (2009)
51. Patel, C., Ramachandran, M.: Story card maturity model (SMM): a process improvement framework for agile requirements engineering practices. *J. Softw.* **4**(5), 422–435 (2009)
52. Pereira, R., Serrano, J.: A review of methods used on it maturity models development: a systematic literature review and a critical analysis. *J. Inf. Technol.* **35**, 161–178 (2020)
53. Peres, A.L., Silva, T.S.D., Silva, F.S., Soares, F.F., Carvalho, C.R.M.D., Meira, S.R.D.L.: AGILEUX model: towards a reference model on integrating UX in developing software using agile methodologies. In: Agile Conference, Orlando, FL, USA, pp. 61–63, July 2014
54. Proulx, M.: Yet another agile maturity model (AMM) - the 5 levels of maturity (2010)
55. Quintal, C., Macías, J.A.: A capability maturity proposal for usability and accessibility centered processes. In: Proceedings of the International Conference on Human Computer Interaction, Palma, Spain, pp. 15:1–15:8. ACM (2018)
56. Qumer, A., Henderson-Sellers, B.: A framework to support the evaluation, adoption and improvement of agile methods in practice. *J. Syst. Softw.* **81**(11), 1899–1919 (2008)
57. Ries, E.: *The Lean Startup: How Today’s Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses*. Crown Business, New York (2011)
58. Salah, D., Paige, R., Cairns, P.: An evaluation template for expert review of maturity models. In: Jedlitschka, A., Kuvaja, P., Kuhrmann, M., Männistö, T., Münch, J., Raatikainen, M. (eds.) PROFES 2014. LNCS, vol. 8892, pp. 318–321. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-13835-0_31
59. Salah, D., Paige, R., Cairns, P.: Patterns for integrating agile development processes and user centered design. In: Proceedings of the 20th European Conference on Pattern Languages of Programs, Kaufbeuren, Germany, p. 19. ACM (2015)
60. Salah, D., Paige, R., Cairns, P.: A maturity model for integrating agile processes and user centered design. In: Clarke, P.M., O’Connor, R.V., Rout, T., Dorling, A. (eds.) SPICE 2016. CCIS, vol. 609, pp. 109–122. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-38980-6_9
61. Salah, D., Paige, R.F., Cairns, P.: A systematic literature review for agile development processes and user centred design integration. In: Proceedings of the International Conference on Evaluation and Assessment in Software Engineering. Association for Computing Machinery, New York (2014)
62. Schön, E.-M., Winter, D., Escalona, M.J., Thomaschewski, J.: Key challenges in agile requirements engineering. In: Baumeister, H., Lichter, H., Riebisch, M. (eds.) XP 2017. LNBIP, vol. 283, pp. 37–51. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57633-6_3
63. Schröders, T., Cruz-Machado, V.: Assessing lean implementation. In: Gen, M., Kim, K.J., Huang, X., Hiroshi, Y. (eds.) Industrial Engineering, Management Science and Applications 2015. LNEE, vol. 349, pp. 803–811. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-47200-2_84
64. Shukla, S.K., Sushil: Evaluating the practices of flexibility maturity for the software product and service organizations. *Int. J. Inf. Manag.* **50**, 71–89 (2020)
65. Sidky, A.S., Arthur, J.D., Bohner, S.A.: A disciplined approach to adopting agile practices: the agile adoption framework. *Innov. Syst. Softw. Eng.* **3**, 203–216 (2007)

66. Signoretti, I., Marczak, S., Salerno, L., d. Lara, A., Bastos, R.: 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), Porto de Galinhas, Recife, Brazil, pp. 1–6, September 2019
67. Signoretti, I., Salerno, L., Marczak, S., Bastos, R.: Combining user-centered design and lean startup with agile software development: a case study of two agile teams. In: Stray, V., Hoda, R., Paasivaara, M., Kruchten, P. (eds.) XP 2020. LNBP, vol. 383, pp. 39–55. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49392-9_3
68. Silva, F.S., Soares, F.S.F., Peres, A.L., de Azevedo, I.M., Pinto, P.P., de Lemos Meira, S.R.: A reference model for agile quality assurance: combining agile methodologies and maturity models. In: International Conference on the Quality of Information and Communications Technology, Guimaraes, Portugal, pp. 139–144. IEEE Computer Society (2014)
69. Soares, F., Meira, S.: An agile strategy for implementing CMMI project management practices in software organizations. In: Iberian Conference on Information Systems and Technologies, Águeda, Portugal, pp. 1–4, June 2015
70. Soundararajan, S., Balci, O., Arthur, J.D.: Assessing an organization’s capability to effectively implement its selected agile method(s): an objectives, principles, strategies approach. In: 2013 Agile Conference, Leuven, Belgium, pp. 22–31, August 2013
71. Stanisavljevic, Z., Walter, B., Vukasovic, M., Todosijevic, A., Labedzki, M., Wolski, M.: GEANT software maturity model. In: Telecommunications Forum, Belgrade, Serbia, pp. 420–425, November 2018
72. Stojanov, I., Turetken, O., Trienekens, J.J.M.: A maturity model for scaling agile development. In: Euromicro Conference on Software Engineering and Advanced Applications, Madeira, Portugal, pp. 446–453, August 2015
73. Tyne, S.: Corporate user-experience maturity model. In: Kurosu, M. (ed.) HCD 2009. LNCS, vol. 5619, pp. 635–639. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-02806-9_74
74. Vargas, B.P., Signoretti, I., Zorzetti, M., Marczak, S., Bastos, R.: On the understanding of experimentation usage in light of lean startup in software development context. In: Proceedings of the Evaluation and Assessment in Software Engineering, Trondheim, Norway, pp. 330–335. ACM (2020)
75. Ximenes, B.H., Alves, I.N., Araújo, C.C.: Software project management combining agile, lean startup and design thinking. In: Marcus, A. (ed.) DUXU 2015. LNCS, vol. 9186, pp. 356–367. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20886-2_34
76. Yaman, S.G., et al.: Introducing continuous experimentation in large software-intensive product and service organisations. *J. Syst. Softw.* **133**, 195–211 (2017)
77. Yin, A., de Lemos Figueiredo, S., Mira da Silva, M.: Scrum maturity model. In: International Conference on Software Engineering Advances. IARIA, Barcelona, Spain, January 2011
78. Zorzetti, M., et al.: Maturity models for agile, lean startup, and user-centered design in software engineering: a combined systematic literature mapping. In: Proceedings of the 22nd International Conference on Enterprise Information Systems, vol. 2, pp. 145–156. SciTePress, Prague (2020)