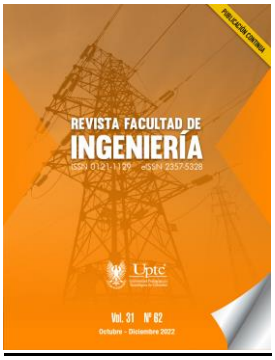
Revista Facultad de Ingeniería
Journal Homepage: https://revistas.uptc.edu.co/index.php/ingenieria



Metrics Model to Complement the Evaluation of DevOps in Software Companies

Carlos-Eduardo Orozco-Garcés[1]
César-Jesús Pardo-Calvache[2]
Elizabeth Suescún-Monsalve[3]

Received: May 08, 2022 **Accepted:** September 29, 2022 **Published:** October 04, 2022

Citation: C.-E. Orozco-Garcés, C.-J. Pardo-Calvache, E. Suescún-Monsalve, "Metrics Model to Support the Evaluation of DevOps in Software Companies", *Revista Facultad de Ingeniería*, vol. 31 (62), e14766, 2022. https://doi.org/10.19053/01211129.v31.n62.2022.14766

Abstract

This article presents a model to complement the evaluation of DevOps in software companies. It was designed by harmonizing the elements of the DevOps process identified through a systematic mapping of the literature and aimed to know the state of the art of methodological solutions and tools to evaluate DevOps in the industry.

[1] M. Sc. Universidad del Cauca (Popayán-Cauca, Colombia). carlosorozco@unicauca.edu.co. ORCID: 0000-0003-2615-3294
[2] Ph. D. Universidad del Cauca (Popayán-Cauca, Colombia). cpardo@unicauca.edu.co. ORCID: 0000-0002-6907-2905
[3] Ph. D. Universidad EAFIT (Medellín-Antioquia, Colombia). esuescu1@eafit.edu.co. ORCID: 0000-0001-7872-7638

The process elements were identified, compared, and integrated into a common process structure that was used to establish a total of 11 metrics using the Goal-Question-Metric approach. The model was evaluated by a focus group of expert DevOps professionals. They determined that the model is clear, easy to apply, and provides valuable information to companies to improve their DevOps practices.
Keywords: assessment; DevOps; evaluation; GQM; Metrics.

Modelo de métricas para complementar la evaluación de DevOps en empresas de software

Resumen

En este artículo se presenta un modelo de métricas para apoyar la evaluación de las prácticas, dimensiones y valores propuestos para la implementación de DevOps en empresas de software. El modelo de métricas fue diseñado haciendo una armonización de los elementos del proceso DevOps identificados en un mapeo sistemático de la literatura. Este se realizó con el propósito de conocer el estado del arte de las soluciones metodológicas y herramientas para evaluar DevOps en la industria. Los elementos del proceso fueron identificados, comparados e integrados en una estructura común que permitió definir un total de 11 métricas usando el enfoque *Goal-Question-Metric*. Posteriormente, el modelo de métricas fue evaluado por un grupo focal de profesionales expertos en DevOps, quienes determinaron que el modelo es claro, fácil de aplicar y aporta valor a las empresas para la mejora de sus prácticas de DevOps.
Palabras clave: DevOps; evaluación; GQM; Métricas.

Modelo de medições para complementar a avaliação de DevOps em empresas de software

Resumo

Este artigo apresenta um modelo de métricas para apoiar a avaliação das práticas, dimensões e valores propostos para avaliar a implementação de DevOps em empresas de software. O modelo de métricas foi desenhado por meio da harmonização dos elementos do processo DevOps identificados por meio de um

mapeamento sistemático da literatura, realizado com a finalidade de conhecer o estado da arte em relação à definição de soluções metodológicas e ferramentas para avaliar DevOps na indústria. Os elementos do processo foram identificados, comparados e integrados em uma estrutura de processo comum que foi usada para definir um total de 11 métricas através da aplicação da abordagem *Goal-Question-Metric*. Posteriormente, o modelo de métricas foi avaliado por um grupo focal formado por profissionais especialistas em DevOps, que determinaram que o modelo é claro, de fácil aplicação e fornece informações valiosas para as empresas melhorarem suas práticas relacionadas ao DevOps.

Palavras-chave: Avaliação; Desenvolvimento e Operações; DevOps;MSL.

I. INTRODUCTION

Currently, software development companies face challenges to deploy solutions with high quality standards in short time intervals [1]. To achieve this, companies seek to improve their processes by implementing approaches and/or frameworks that allow them to enhance the quality of their products [1]. In this sense, proposals related to the software product implementation life cycle (Dev) that can be classified as traditional and agile have been made. Some of the most used traditional solutions are CMMI [2], RUP [3], Cascade model [4], Spiral model [5], and Rapid Application Design (RAD) [6]. Some common agile solutions are Scrum [7], Lean Software [8], Test Driven Development (TDD) [9], Extreme Programming (XP) [10], [11], Crystal Clear [12], Adaptive Software Development [13], and Dynamic Systems Development Method [14]. Moreover, hybrid solutions that take advantage of both approaches have been proposed, e.g., Scrum & XP [15], Scrumban [16], and Scrum & CMMI [17]. However, software companies have also paid special attention to the processes related to operations management in Information Technology (Ops), which are applied to establish strategies that allow defining and implementing a set of best practices to guarantee the stability and reliability of the solutions in productive environments. Software development life cycle management brings multiple benefits to companies including continuously reducing development, integration, and deployment times; delegating repetitive tasks to automated processes; reducing errors caused by human intervention [18], [19], among others. To achieve this, solutions related to operations management such as ITIL [20], COBIT [21], the ISO/IEC 20000 standard [22], and ISO/IEC 27000 standard [23] have been proposed. Debois [24] introduced the term DevOps in 2009 with the aim of integrating the best practices proposed for development and operations (Dev and Ops). Over the years, DevOps has proven to bring multiple benefits related to the improvement of activities of the projects' life cycle, especially in productivity, quality, and competitiveness of software development companies [25], [26]. In general, DevOps focuses on defining practices that allow enhancing tasks related to continuous integration [27], change management [28], automated tests [29], continuous deployment [30], continuous maintenance [31], among others. According

to the global survey report on the state of agility in 2021 [32], 75% of the participants mentioned that a transformation towards a culture supported by DevOps brings multiple benefits for companies in terms of reduced effort, cost, and time. However, adopting DevOps in software companies is not a simple task [33], to minimize the risk of error in its adoption, they must establish mechanisms that allow quantifying how it is applied in their projects and identify improvement opportunities to fine-tune their practices and improve their internal processes [34]. The efforts and proposals related to the evaluation of DevOps in software companies were identified with a systematic mapping of the literature carried out in [35]. Two mechanisms were used to define methodological solutions (models, metrics, certification standards) and tools developed by active players in the industry that seek to assess DevOps in multiple ways. However, the results show a high degree of heterogeneity in the proposed solutions, since there is no consensus in the definitions, relationships, and concepts related to DevOps [36]. In consequence, the solutions identified in the literature were proposed in accordance with a set of values, principles, activities, roles, practices, and tasks considered relevant by each author. Although the analyzed solutions follow the same objective: “assess the degree of DevOps capacity, maturity and/or competence”, they have different perceptions, scopes and, in some cases, they are ambiguous. Likewise, the solutions described in [35] establish "what" to do; however, they do not define "how" to implement the proposed practices, which can cause confusion when applying DevOps in software companies. Besides, there are studies related to the evaluation of DevOps in companies of different sizes, most of them focus on large and medium-sized companies and leave aside small and micro software companies. According to the digital transformation report of the Economic Commission for Latin America and the Caribbean (CEPAL) in 2021 [37], they correspond to approximately 99% of the legally constituted companies in Latin America and have gradually become active industry players looking to apply DevOps in their projects.

Hence, there are solutions and tools to evaluate DevOps; however, each author suggests his own terminology, evaluation criteria, concepts, practices, and process elements. It results in a high degree of heterogeneity that can generate confusion,

inconsistency, and terminological conflicts during the adoption of DevOps practices. This article presents a metrics model defined following the Goal, Question, Metric (GQM) approach [38], and aims to complement the evaluation of DevOps. The model organizes its elements around four dimensions: people, culture, technology, and processes and aims to define what and how to evaluate DevOps compliance in the software industry. The paper is structured as follows: Section 2 analyzes the state of the art of solutions to evaluate DevOps in software companies; Section 3 presents a metrics model to evaluate DevOps according to the practices, dimensions, and values found, analyzed, and harmonized from the literature; Section 4 describes the protocol to form a focus group as an evaluation method. Finally, Section 5 presents the conclusions and future work.

II. Materials and Methods

A. Background

After executing a systematic mapping of the literature (SML), as reported in [35], it was analyzed to identify the solutions proposed by different authors in relation to the definition of processes, models, techniques and/or or tools to evaluate DevOps in software companies. Three types of studies were identified: (i) exploratory studies, (ii) methodological solutions, and (iii) tools. The results obtained are presented below.

1) Exploratory Studies. In [39], an exploratory study was carried out to analyze different tools to evaluate DevOps in small and medium software companies. In [11], [36], [40], [41], SML were made to identify the process elements that must be considered to certify that a company applies DevOps appropriately. In [42]–[45], studies were conducted to know the use of maturity models to evaluate DevOps.

2) Methodological Solutions. In [46]–[49], metrics to evaluate the construction, integration, and continuous deployment practices in software companies are proposed; [50], [51] propose competency models; [42]–[45], [52]–[57], maturity models; [50], a model to evaluate DevOps collaboration; [57], a DevOps evaluation model based on the Scrum Maturity Method (SMM); [58], a method to certify the use

of best DevOps practices; [59], a model to evaluate development, security, and operations (DevSecOps); and [60], a standard to adopt DevOps in software companies.

3) Tools. [39], [50], [61] mentioned the following tools: DevOps Maturity Assessment [62], Microsoft DevOps Self-Assessment [63], IBM DevOps Self-Assessment [64], and IVI's DevOps Assessment [65]. However, the tools presented in the studies were not assessed exhaustively. To expand the knowledge on the definition of tools to evaluate DevOps, an exploratory study was carried out based on the methodology proposed in [66]. As a result, 13 tools were identified and are presented in Table 1. The tools analysis considers accessibility (A1): free to access, trial period, or paid; evaluation method (A2): surveys, frameworks, consulting, or another mechanism; and objective or scope of the evaluation (A3): the tool performs an evaluation of the process, practices, activities, tasks, or other aspects/element. In relation to accessibility (A1), it was observed that 7 tools (54%) ([62]–[64], [67]–[70]) are free, 5 tools (38.4%) ([65], [71]–[74]) are paid, and 1 tool (7.6%) [75] offers a 30 day trial period. Regarding the evaluation method (A2), different mechanisms were observed: 6 tools (46.2%) ([62], [63], [67]–[70]) evaluate DevOps through of surveys, 5 tools (38.4%) ([71]–[75]) evaluate DevOps through consulting processes, and 2 tools (15.4%) ([64], [65]) evaluate DevOps through methodological guides and frameworks. In relation to the objective or scope of the evaluation (A3), 6 tools (46.2%) ([65], [71]–[75]) evaluate DevOps according to the set of principles, values, activities and roles applied by a company; 5 tools (38.4%) ([62], [64], [67]–[69]) evaluate continuous integration and deployment practices; and 2 tools (15.4%) ([63], [70]) evaluate DevOps according to compliance with the Culture, Automation, Lean, Measurement and Shared Use principles.

Table 1. Tools to evaluate DevOps.

No	Ref.	Company	A1	A2	A3
1	[62]	ATOS	Free	Survey	Practices
2	[63]	Microsoft	Free	Survey	Principles
3	[67]	Infostretch	Free	Survey	Practices
4	[68]	InCycle	Free	Survey	Practices
5	[64]	IBM	Free	Guide	Practices
6	[69]	Xmatters	Free	Survey	Practices

7	[70]	Attlasian	Free	Survey	Principles
8	[65]	IVI's	Payment	Framework	Process
9	[71]	Veritis	Payment	Consulting	Process
10	[72]	Boxboat	Payment	Consulting	Process
11	[73]	Humanitec	Payment	Consulting	Process
12	[74]	Attlasian	Payment	Consulting	Process
13	[75]	Eficode	Free demo	Consulting	Process

B. Protocol to Harmonize DevOps Process Elements

It was necessary to carry out a harmonization process that allowed identifying the elements to define a generic model to evaluate DevOps. Nevertheless, each model and tool has its own structure, concepts, and characteristics. To establish a homogeneous solution, HPROCESS was used to harmonize the models [76] with the following activities: identification, carried out during the SML; homogenization; comparison; and integration.

1) Homogenization Method. This method compares the general information of each solution and tool in a common structure that shows the characteristics of each study in relation to the rest [77]. It was defined from the process elements established in the PrMO ontology [78]. The characterization is available at https://bit.ly/3QDJOT9.

2) Comparison Method. The comparison was made by applying the set of activities proposed by MaMethod [79] adapted to compare the dimensions, values, and practices identified in the homogenization stage through the following activities: (i) analyze the solutions, (ii) design the comparison, and (iii) make the comparison. To do it, it was necessary to establish a base model that was crossed with all the solutions through a matrix that relates the set of practices, dimensions, and values proposed by each solution. The base model was chosen considering as selection criteria C1: the solution is generic; C2: the solution has a clearly defined set of dimensions, values, and practices; and C3: the solution was peer-reviewed by experts. After the analysis, it was determined that the reference model proposed in [80] meets all the criteria. The base model was compared with 23 solutions and 3 tools. The details of all the comparisons can be consulted at https://bit.ly/3c4nzaa.

3) Integration Method. IMethod [81] was applied to carry out the integration. It proposes five (5) activities: design, define an integration criterion, execute it, analyze the results, and present the integrated model. After the integration, 12 practices were considered fundamental and 6 complementary. 4 dimensions and 4 values were obtained, which represent the state of knowledge related to all the solutions. Table 2 summarizes the practices, dimensions, and values resulting from the integration process.

The detail of the results can be consulted at https://bit.ly/3dItx0M. Finally, an activity was conducted to identify the relationship between practices, dimensions, and values. It can be consulted at https://bit.ly/3T05Q45.

Table 2. Integrated process elements.

Elements	Acronym	Name
Fundamental practices (PF)	**IC**	Continuous integration
	EC	Continuous delivery
	PC	Continuous testing
	GR	Requirement management
	GD	Data Management
	SS	Security supervision
	DE	Strategic direction
	GC	Configuration management
	MC	Continuous monitoring and observability
	ED	Education around DevOps
	RC	Continuous feedback
	MCu	Culture measurement
Complementary practices (PC)	**DC**	Continuous deployment
	Ico	Infrastructure as Code
	GA	Privilege access management
	AC	Continuous learning
	ExC	Continuous experimentation
	SL	Laboral satisfaction
Dimensions	**HERR**	Tools
	PROC	Processes
	CULT	Culture
	PERS	People
Values	**AUT**	Automatization
	COL	Collaboration
	MED	Measurement
	COM	Communication

Note: The acronyms are in Spanish.

III. Results and Discussion

The goal of the metrics in software engineering is to identify the essential parameters present in the projects [82]. The harmonization process allowed to obtain 12 fundamental practices, 6 complementary practices, 4 dimensions, and 4 values. The model follows a hierarchical structure in which the values are the aspects that must be considered to ensure that DevOps culture is applied properly, the dimensions describe each of the activities required to implement the values proposed for DevOps effectively, and the practices represent what must be applied to comply with each of the dimensions.

A. Purpose of the Model

The metrics model seeks to evaluate the implementation of DevOps based on a set of questions that allow knowing the degree of compliance of DevOps practices, dimensions, and values. The metrics model aims to support the evaluation made by a consultant, through a set of clearly defined metrics, of the implementation of DevOps in software companies. It also aims to identify areas to be improved with respect to the mechanisms used by companies to adopt and/or apply DevOps. The metrics model was defined following the guidelines described by the GQM approach [38]: a conceptual level (Goal), an operational level (Question), and a quantitative level (Metric). At the conceptual level, the dimensions, practices and values proposed by DevOps were identified. At the operational level, the questions associated with each DevOps practice were defined according to a set of goals associated with each practice. Finally, at the quantitative level, a set of metrics that enable knowing the degree of implementation of DevOps practices, dimensions, and values were defined.

B. Goals

Initially, a set of goals related to each practice defined at the harmonization stage was defined. As a result, 42 goals and 63 questions related to fundamental practices were set; and 19 goals and 29 questions related to complementary practices. Table 3 shows the goals related to the continuous deployment (DC) practice. The rest of the objectives can be consulted at https://bit.ly/3SZQSuZ.

Table 3. Integrated process elements.

Practice acronym	Goal	Goal acronym
DC	Define a pipeline for continuous deployment	O1_DC
	Define container and virtualization practices	O2_DC
	Reduce dependencies between team members (Dev-Ops-QA)	O3_DC

Note: The acronyms are in Spanish.

C. Questions

Each goal is associated with one or more questions that relate the aspects to be evaluated quantitatively. The questions use a nominal scale with two possible values (YES: 100%, NO: 0%). They were defined following the criteria proposed in [83], which seeks to avoid ambiguous questions, vague terms, cognitive overload, among others. Table 4 presents the questions associated with DC. The rest of the questions can be consulted at https://bit.ly/3ChM1zi.

Table 4. Proposed questions for complementary practices

Practice acronym	Question	Goal acronym
DC	**P1_CD.** Have clear procedures been defined to guarantee the integrity of an artifact to be deployed?	O1_DC
	P2_CD. Have tools been defined to automate the deployment of artifacts?	O2_DC
	P3_CD. Has a direction process been defined to ensure the deployment of new artifacts?	O3_DC
	P4_CD. Does the deployment process require manual approval?	O4_DC

Note: The acronyms are in Spanish.

A questionnaire-type evaluation instrument was designed with two possible answers (“YES”, “NO”). The template used to answer the questions can be found at https://bit.ly/3LDgfik. The answer to each question is given according to the following criteria: “YES”; (i) collection of opinions about each role involved in the practices or (ii) consistent historical records that evidence compliance. "NO"; (i) if the company does not present evidence of compliance with the practice.

D. Metrics

Table 5 shows the scale to assess the degree of implementation (gi) of practices, dimensions, and values. It was defined following the formalism proposed in [80]. The metrics definition process was carried out by assigning weights to each practice, dimension, and value by applying the linear weighting method [84], and metrics were defined using the GQM [38].

Table 5. Scale to measure the implementation level.

Acronym	Implementation level	Scale
NA	Not achieved	0% ≤ gi ≤ 15%
PA	Partially achieved	15% < gi ≤ 50%
AA	Widely achieved	50% < gi ≤ 85%
CA	Completely achieved	85% < gi ≤ 100%

Note: The acronyms are in Spanish.

1) Metrics to Evaluate Practices. As a result of the weighting process [84], it was identified that each practice has an associated weighted percentage $(\%PPA)$. The combined weighted percentage $(\%PPC)$ corresponds to the weight associated with all practices during the total evaluation. Table 6 shows the weights of each fundamental and complementary practice.

Table 6. Integrated process elements.

Practices	Acronym	%PPA	%PPC	Total
Fundamental practices (PF)	**IC**	8,5%	70%	100%
	EC	8,4%		
	PC	8,4%		
	GR	8,3%		
	GD	8,3%		
	SS	8,3%		
	DE	8,3%		
	GC	8,3%		
	MC	8,3%		
	ED	8,3%		
	RC	8,3%		
	MCu	8,3%		
Complementary practices (PC)	**DC**	18%	30%	
	ICo	16%		
	GA	16%		
	AC	16%		
	ExC	16%		
	SL	18%		

Note: The acronyms are in Spanish.

Table 7 describes the metrics that relate the degree of individual, combined, weighted and total implementation of the fundamental and complementary practices.

Table 7. Metrics to assess the degree of implementation of practices.

Id	Scale [%]	Equation	Variables
$\%PFI$	$[0, 100]$	$\frac{1}{n}\sum_{i}^{n} \%P_i$	$\boldsymbol{n}$: Number of questions; $\%\boldsymbol{P_i}$: Question percentage (YES: 100%, NO: 0%).
$\%PFIP$	$[0, \%PPA]$	$\%PFI * \%PPA$	$\%\boldsymbol{PFI}$: Percentage of individual compliance with a fundamental practice; **%PPA:** Associated weighted percentage (see Table 6).
$\%PFT$	$[0, 100]$	$\sum_{i}^{n} \%PFIP_i$	$\boldsymbol{n}$: Number of fundamental practices; $\%\boldsymbol{PFIP_i}$: Percentage of weighted compliance with a fundamental practice.
$\%PCI$	$[0, 100]$	$\frac{1}{n}\sum_{i}^{n} \%P_i$	$\boldsymbol{n}$: Number of questions associated with the practice; $\%\boldsymbol{P_i}$: Percentage obtained by a specific question (YES: 100%, NO: 0%).
$\%PCIP$	$[0, \%PPA]$	$\%PCI * \%PPA$	$\%\boldsymbol{PCI}$: Percentage of individual compliance with a complementary practice; **%PPA:** Associated weighted percentage (see Table 11).
$\%PCT$	$[0, 100]$	$\sum_{i}^{n} \%PCIP_i$	$\boldsymbol{n}$: Number of practices; $\%\boldsymbol{PCIP_i}$: Percentage of weighted compliance with a complementary practice.
$\%PTP$	$[0, 100]$	$(\%PPCF * \%PFT) + (\%PPCC * \%PCT)$	$\%\boldsymbol{PPCF}$: 70% (see Table 6); $\%\boldsymbol{PPCC}$: 30% (see Table 6); $\%\boldsymbol{PFT}$: percentage of compliance with fundamental practices; $\%\boldsymbol{PCT}$: percentage of compliance with complementary practices.

2) Metrics to evaluate dimensions. A total of 4 dimensions were obtained: (i) tools, (ii) processes, (iii) culture and (iv) people. Each dimension has a set of practices associated with it and a weighted percentage ($\%PPAD = 25\%$). Table 8 shows the metrics to evaluate the degree of implementation of dimensions.

Table 8. Metrics to assess the degree of implementation of dimensions.

Id	Scale [%]	Equation	Variables
$\%PCDI$	$[0, 100]$	$\frac{\%PPCF}{n}\sum_{i}^{n} \%PFI_i + \frac{\%PPCC}{m}\sum_{j}^{m} \%PCI_j$	$\boldsymbol{n}$: Number of fundamental practices; $\boldsymbol{m}$: Number of complementary practices; $\%\boldsymbol{PPCF}$: 70% (see Table 6); $\%\boldsymbol{PPCC}$: 30% (see Table 6); $\%\boldsymbol{PFI_i}$: Percentage of individual compliance with a specific fundamental practice (see Table 7); $\%\boldsymbol{PCI_i}$: Percentage of individual compliance with a specific complementary practice (see Table 7).
$\%PCDT$	$[0, 100]$	$\sum_{i}^{n} \%PPAD_i * \%PCDI_i$	$\boldsymbol{n}$: Number of dimensions; $\%\boldsymbol{PPAD_i}$: Weighted percentage associated with a specific dimension; $\%\boldsymbol{PCDI_i}$: Percentage of individual compliance with a dimension.

DOI: https://doi.org/10.19053/01211129.v31.n62.2022.14766

3) Metrics to evaluate values. A total of 4 values were obtained: (i) automation, (ii) collaboration, (iii) measurement and (iv) communication. Each value has a set of associated dimensions and a weighted percentage ($\%PPAV = 25\%$). Table 9 presents the metrics to evaluate the degree of implementation of the values.

Table 9. Metrics to assess the degree of implementation of values.

Id	Scale [%]	Equation	Variables
$\%PCVI$	$[0, 100]$	$\frac{1}{n}\sum_{i}^{n} \%PCDI_i$	$\mathbf{n}$: Number of dimensions associated with the value; $\%\mathbf{PCDI_i}$: Percentage of compliance with a dimension associated with a value (see Table 8).

4) Metrics to evaluate DevOps. Table 10 presents the metric to know the implementation degree of DevOps in a software development company.

Table 10. Total degree of implementation of DevOps (own elaboration)

Id	Scale [%]	Equation	Variables
$\%PCDev$	$[0, 100]$	$\sum_{i}^{n} \%PPAV_i * \%PCVI_i$	**Variables:** $\mathbf{n}$: Number of values; $\%\mathbf{PPAV_i}$: Weighted percentage associated with a specific value;$\%\mathbf{PCVI_i}$: Percentage of individual compliance with a value (see Table 9).

IV. Evaluation

A. Focus Group Protocol

The procedure to form the focus group followed the guidelines defined in [85], which proposes 5 phases: (i) planning, (ii) recruitment, (iii) moderation, (iv) analysis and report of results, and (v) limitations. To conduct the focus group, a questionnaire that aimed to assess the suitability, completeness, ease of understanding, and applicability of the metrics model was designed.

1) Planning. During this phase, the general goal of the focus group and the research objective were defined. Subsequently, the materials and procedures necessary to carry out the discussion session were identified. The general goal of the research was oriented to know the perceptions, opinions, and suggestions made by the

participants of the focus group. The research goals had the purpose of identifying possible improvement actions or suggestions made by DevOps experts about the degree of acceptance or rejection of suitability, completeness, ease of comprehension, and applicability of the model in software development companies. The materials were a questionnaire, a work agenda, a protocol structure, and a proposal to be evaluated.

2) Recruitment. The research group defined the profile of the attendees with the aim of choosing people with the necessary experience and knowledge about DevOps. As a result, the participants must meet the following criteria: be an active professional in the industry or academic environment, have knowledge and experience in the definition or application of agile approaches, and have at least one year of experience working with DevOps. Considering that criteria, 15 potential participants were invited, out of which 14 were accepted.

3) Moderation. The debate session took an hour and a half, and the agenda was: (i) thanks to the participants for attending; (ii) presentation of the goals for the focus group; (iii) presentation of the metrics model; (iv) discussion of the observations and suggestions identified by each of the participants; and (v) completion of an online form to know the opinion of each participant. The activities were coordinated by a moderator, who ensured that the interventions of the participants were within the objectives and scope of the focus group, and a rapporteur who recorded the perceptions, suggestions, and comments of each participant. At the end of the discussion session, the participants were asked to fill a form answering 17 questions defined according to the levels of conformity proposed in the Likert scale [86] including 5 possible values: (1) Very bad, very dissatisfied; (2) Bad, little satisfied; (3) Good, sufficient, adequate, somewhat satisfied; (4) Fairly good, adequate, satisfied; and (5) Very good, very adequate, very satisfied. Additionally, there were two open questions that allowed the participants to propose adjustments to the process and make additional comments. The relationship between each question and the following criteria is presented: comprehensibility; applicability; suitability; and completeness. According to the distribution of the questions, questions P1-P3 evaluate the comprehensibility of the proposal; P4-P5 and P16, its applicability; P5-

P14, its suitability; P8-P17, its completeness; and open questions P18 and P19 evaluate all aspects. Table 12 presents the details of the questions.

Table 12. Focus group results.

Id	Discrete scale Question
P1	Do you consider that the practices proposed in the model are clear and easy to understand?
P2	Do you consider that the dimensions proposed in the model are clear and easy to understand?
P3	Do you consider that the values proposed in the model are clear and easy to understand?
P4	According to your experience, do you consider that the evaluation of practices, dimensions, and values is adequate and allows to identify valuable aspects for software development companies?
P5	Do you consider that the weighting defined for the fundamental and complementary practices is adequate?
P6	Do you consider that the weighting defined for the dimensions is adequate?
P7	Do you consider that the weighting defined for the values is adequate?
P8	Based on your experience, do you consider that the metrics have sufficient mathematical rigor?
P9	Do you consider that the metrics defined to evaluate the degree of adoption of fundamental practices are adequate?
P10	Do you consider that the metrics to evaluate the degree of adoption of complementary practices are adequate?
P11	Do you consider that the metric to calculate the degree of total adoption of practices is adequate?
P12	Do you consider that the metrics to calculate the degree of adoption of dimensions are adequate?
P13	Do you consider that the metric to calculate the degree of adoption of values is adequate?
P14	Do you consider that the metric to calculate the degree of adoption of DevOps is adequate?
P15	Do you consider that the metrics proposed in the model are sufficient to guarantee a complete evaluation of DevOps?
P16	Do you consider that the results obtained after applying the metrics model will allow a company to identify aspects to be improved in its processes?
P17	Based on your experience, do you consider that the proposed metrics model can be applied in companies with limited infrastructure, capital, personnel, or size?
Id	***Open questions***
P18	Would you add, delete, or modify elements in the metrics model?
P19	Do you have any additional comments about the proposed metrics model?

Note: The acronyms are in Spanish.

4) Analysis and Report of Results. According to the results, the participants had a positive perception about the practices, dimensions, and values proposed in the model. Besides, they considered that the practices are sufficient and necessary to guarantee the evaluation of DevOps and that the proposed dimensions and values are coherent because they approach DevOps including aspects related to soft skills such as communication, cooperation, transparency, and teamwork. A high degree of agreement was observed regarding the applicability of the proposal. The participants stated that the elements in the model provide value to companies and open opportunities for improvement after an evaluation. They also stated that the defined weights are consistent and adequate according to the distribution of fundamental and complementary practices. In addition, a high degree of agreement was observed regarding the suitability of the proposal. According to the participants, the model has a solid mathematical basis according to the goals proposed for each practice. In this sense, the participants stated that the metrics can offer a result that allows companies to identify possible aspects to be improved. Finally, it was possible to observe a favorable opinion regarding aspects related to mathematical rigor and the usefulness of the proposed metrics; however, the ones related to the applicability of the proposal in small and medium-sized companies were identified. They were considered and applied to refine a new version of the proposal. The detail of the improvement actions can be consulted at the following link: https://bit.ly/3pw659y.

B. Limitations

Each limitation found during the focus group and the solutions they applied are presented below. Although all the participants met the selection criteria, they did not have the same level of knowledge and experience with DevOps; the metric model was sent to all participants three weeks in advance to guarantee that all participants were aware of the context of the proposal. According to [85], the focus group should have at least 6 participants; therefore, 15 people were invited to reduce the possibility of not reaching the minimum number of attendees. At the beginning of the session, it was possible to observe that the participation was low, this was corrected

by the rapporteur and the moderator, who encouraged them to participate by asking questions or animate them to express their comments. Due to the number of participants, some of the comments made during the discussion were outside the scope of the proposed evaluation objectives; it was decided to clarify each comment quickly to continue with the discussion. The focus group was carried out following biosafety protocols to avoid crowds: the session was held remotely and permission was requested to record it and analyze the observations and comments that could have been omitted during the session.

V. Conclusions

The metrics model was the result of several stages executed in a structured and organized manner: (i) a SML on the evaluation of DevOps in software companies, (ii) the harmonization of the methodological solutions and tools identified in the SML, and (iii) the definition of a metrics model applying the GQM approach that allowed establishing the set of process elements to evaluate DevOps. The harmonization of the solutions and tools identified in the systematic mapping allowed a much broader and clearer picture of the set of practices, dimensions and values associated with DevOps through an organized, clear, and generic structure. The metrics model proposed in this article provides support to expert DevOps professionals and consultants who seek to assess the degree of implementation of DevOps practices, dimensions, and values. As a result, a company can quickly understand the degree of DevOps implementation at general and specific levels. The evaluation of the proposal through a focus group allowed to confirm that the model is consistent and defines a set of clear metrics that evaluate vital aspects to the application of DevOps. Likewise, the focus group allowed to receive feedback thanks to the recommendations of software engineering experts with experience in the definition, adoption, and application of DevOps processes, and to identify aspects to be improved. Those aspects were analyzed to obtain a refined version of the proposal. Finally, the future work gaps that are currently being addressed include the execution of multiple case studies to evaluate the metrics model in operational environments, the construction of a tool to automate the application of the metrics, and the

execution of additional exploratory studies to identify new proposals that can be integrated into the model.

ACKNOWLEDGEMENTS

Ph.D. César Pardo and MSc. Carlos Orozco appreciates the contribution of Universidad del Cauca, where they work as Associate Professor and postgraduate student, respectively.

AUTHOR'S CONTRIBUTION

Carlos-Eduardo Orozco-Garces: Investigation, Formal Analysis, Methodology, Writing-original draft.

César-Jesús Pardo-Calvache: Supervision, Methodology, Validation, Writing-Review and Editing.

Elizabeth Suescun-Monsalve: Supervision, Methodology, Validation.

REFERENCES

[1] H. Conradi, A. Fuggetta, "Improving software process improvement," *IEEE Software*, vol. 19, no. 4, pp. 92–99, 2002. https://doi.org/10.1109/MS.2002.1020295
[2] CMMI Institute, *Capability maturity model integration for development*, 2018.
[3] Rational Software, "Rational Unified Process," *Best Practices for Software Development Teams*, 2020.
[4] W. W. Royce, "Managing the development of large software systems: concepts and techniques," in *Proceedings of 9th International Conference on Software Engineering*, 1987, pp. 328–338.
[5] B. W. Boehm, "A spiral model of software development and enhancement," *Computer (Long Beach Calif)*, vol. 21, no. 5, pp. 61–72, 1988. https://doi.org/10.1109/2.59
[6] J. Martin, *Rapid application development*. Macmillan Publishing Co., Inc., 1991.
[7] K. Schwaber, J. Sutherland, *The scrum guide the definitive guide to scrum: The rules of the game*, 2017.
[8] M. Poppendieck, T. Poppendieck, *Lean Software Development: An Agile Toolkit*. Addison-Wesley, 2003.
[9] K. Beck, *Test Driven Development: By Example*, 1st ed. Addison-Wesley Professional, 2002.
[10] K. Beck, E. Gamma, *Extreme Programming Explained: Embrace Change*, 2000.
[11] J. Guerrero, C. Certuche, K. Zúñiga, C. Pardo, "What is there about DevOps? Preliminary Findings from a Systematic Mapping Study," in *JIISIC 2019*, 2019.
[12] A. Cockburn, *Crystal clear: A human-powered methodology for small teams: A human-powered methodology for small teams*, Pearson Education, 2004.
[13] J. Highsmith, *Adaptive software development: a collaborative approach to managing complex systems*, Addison-Wesley, 2013.
[14] J. Stapleton, *DSDM, dynamic systems development method: the method in practice*, Cambridge University Press, 1997.
[15] H. Kniberg, *Scrum and XP from the Trenches*, 2015.
[16] C. Ladas, *Scrumban-essays on kanban systems for lean software development*, 2009.
[17] J. Sutherland, C. R. Jakobsen, K. Johnson, "Scrum and CMMI level 5: The magic potion for code warriors," in *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, 2008, pp. 466–466. https://doi.org/10.1109/hicss.2008.384

[18] B. de França, H. Jeronimo, G. Travassos, *Characterizing DevOps by Hearing Multiple Voices*. New York, NY: Association for Computing Machinery, 2016. https://doi.org/10.1145/2973839.2973845
[19] M. Sanchez Gordon, R. Colomo Palacios, "Characterizing DevOps Culture: A Systematic Literature Review," in *International Conference on Software Process Improvement and Capability Determination. SPICE 2018*, 2018, pp. 3–15. https://doi.org/10.1007/978-3-030-00623-5_1
[20] A. Hochstein, R. Zarnekow, W. Brenner, "ITIL as common practice reference model for IT service management: formal assessment and implications for practice," *EEE*, pp. 704-710, 2005. https://doi.org/10.1109/eee.2005.86
[21] J. Young, G. Ridley, P. Carroll, "COBIT and Its Utilization: A Framework from the Literature," *HICSS*, 2014.
[22] ISO/IEC, *Calidad de los servicios TI*, 2019.
[23] ISO/IEC, *Sistemas de Gestión la Seguridad de la Información*, 2013.
[24] P. Debois, *Devopsdays - Organizing Guide*, 2009.
[25] M. Virmani, "Understanding Devops & Bridging The Gap From Continuous Integration To Continuous Delivery," in *International Conference on Innovation and Computational Technology*, 2015, pp. 78–82. https://doi.org/10.1109/intech.2015.7173368
[26] S. S. Samarawickrama, I. Perera, "Continuous scrum: A framework to enhance scrum with DevOps," in *International Conference on Advances in ICT for Emerging Regions*, Sep. 2017, pp. 1–7. https://doi.org/10.1109/icter.2017.8257808
[27] M. Shahin, M. A. Babar, L. Zhu, "Continuous integration, delivery and deployment: a systematic review on approaches, tools, challenges and practices," *IEEE Access*, vol. 5, pp. 3909–3943, 2017. https://doi.org/10.1109/access.2017.2685629
[28] C. Orozco, C. Pardo, S. Vásquez, H. Ordoñez, E. Suescún, "An agile process to support software configuration management," *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, vol. 2020, no. E32, 2020.
[29] J. Michelsen, "Dysfunction Junction: A Pragmatic Guide to Getting Started with DevOps," *CA Technologies*, p. 26, 2014.
[30] E. Diel, S. Marczak, D. S. Cruzes, "Communication Challenges and Strategies in Distributed DevOps," in *11th International Conference on Global Software Engineering*, 2016, pp. 24–28. https://doi.org/10.1109/ICGSE.2016.28
[31] M. Soni, "End to End Automation on Cloud with Build Pipeline: The Case for DevOps in Insurance Industry, Continuous Integration, Continuous Testing, and Continuous Delivery," in *IEEE International Conference on Cloud Computational in Emerging Markets*, 2015, pp. 85–89. https://doi.org/10.1109/CCEM.2015.29
[32] Digital.ai, *15th Annual State of Agile Report*, 2021. https://bit.ly/3Lgw4KE
[33] J. Wettinger, V. Andrikopoulos, F. Leymann, "Automated Capturing and Systematic Usage of DevOps Knowledge for Cloud Applications," in *International Conference on Cloud Engineering*, 2015, pp. 60–65.
[34] F. Erich, C. Amrit, M. Daneva, *Report: DevOps Literature Review*, 2014. https://doi.org/10.13140/2.1.5125.1201
[35] C.-E. Orozco-Garcés, C.-J. Pardo-Calvache, Y.-H. Salazar-Mondragón, "What is There About DevOps Assessment? A Systematic Mapping," *Revista Facultad de Ingeniería*, vol. 31, no. 59, pp. e13896, 2022.
[36] J. Guerrero, K. Zúniga, C. Certuche, C. Pardo, "A systematic mapping study about DevOps," *Journal de Ciencia e Ingeniería*, vol. 12, no. 1, pp. 48–62, 2020. https://doi.org/10.46571/jci.2020.1.5
[37] M. Dini, N. Gligo, A. Patiño, *Transformación digital de las mipymes: elementos para el diseño de políticas*, 2021.
[38] V. R. Basili, *Software modeling and measurement: the Goal/Question/Metric paradigm*, 1992.
[39] M. Muñoz, J. Mejia, B. Corona, J. A. Calvo-Manzano, T. San Feliu, J. Miramontes, "Analysis of Tools for Assessing the Implementation and Use of Agile Methodologies in SMEs," in *International Conference on Software Process Improvement and Capability Determination*, 2016, pp. 123–134. https://doi.org/10.1007/978-3-319-38980-6_10
[40] A. Mishra, Z. Otaiwi, "DevOps and software quality: A systematic mapping," *Computer Science Review*, vol. 38, e100308, 2020. https://doi.org/10.1016/j.cosrev.2020.100308
[41] G. Rong, H. Zhang, D. Shao, "CMMI guided process improvement for DevOps projects: an exploratory case study," in *Proceeding of International Conference on Software System*, 2016, pp. 76–85. https://doi.org/10.1145/2904354.2904372
[42] M. Gasparaite, S. Ragaisis, *Comparison of devops maturity models*, 2019.

[43] M. Zarour, N. Alhammad, M. Alenezi, K. Alsarayrah, “A research on DevOps maturity models,” *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, pp. 4854–4862, 2019.
[44] C. Marnewick, J. Langerman, “DevOps and Organisational Performance: The Fallacy of Chasing Maturity,” *IEEE Software,* vol. 38, no. 5, pp. 48-55, 2021. https://doi.org/10.1109/MS.2020.3023298
[45] R. Feijter, R. Vliet, E. Jagroep, S. Overbeek, S. Brinkkemper, “Towards the adoption of DevOps in software product organizations: A Maturity Model Approach,” *Technical Report Series*, no. UU-CS-2017-009. UU BETA ICS Departement Informatica, 2017.
[46] L. König, A. Steffens, “Towards a quality model for devops,” in *Continuous Software Engineering & Full-scale Software Engineering*, vol. 37, pp. 37-42, 2018.
[47] S. Kruis, *Designing a metrics model for DevOps at Philips IT*, Master Thesis, University of Technology, Eindhoven, 2014.
[48] L. Prates, J. Faustino, M. Silva, R. Pereira, “Devsecops metrics,” in *EuroSymposium on Systems Analysis and Design*, 2019, pp. 77–90. https://doi.org/10.1007/978-3-030-29608-7_7
[49] P. Batra, A. Jatain, “Measurement Based Performance Evaluation of DevOps,” in *International Conference on Computational Performance Evaluation*, 2020, pp. 757–760. https://doi.org/10.1109/compe49325.2020.9200149
[50] P. Rittgen, S. Cronholm, H. Göbel, “Towards a Model for Assessing Collaboration Capability Between Development and Operations,” in *European Conference on Software Process Improvement*, 2019, pp. 111–122. https://doi.org/10.1007/978-3-030-28005-5_9
[51] T. Masombuka, E. Mnkandla, “A DevOps collaboration culture acceptance model,” in *Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists*, 2018, pp. 279–285. https://doi.org/10.1145/3278681.3278714
[52] J. M. Radstaak, “Developing a DevOps maturity model: a validated model to evaluate the maturity of DevOps in organizations,” Master Thesis, University of Twente, 2019.
[53] D. Teixeira, R. Pereira, T. Henriques, M. M. da Silva, J. Faustino, M. Silva, “A maturity model for DevOps,” *International Journal of Agile Systems and Management*, vol. 13, no. 4, pp. 464–511, 2020. https://doi.org/10.1504/ijasm.2020.10034553
[54] T. Neubrand, T. Haendler, *Development of a GQM-based Technique for Assessing DevOps Maturity*, 2020.
[55] R. de Feijter, S. Overbeek, R. van Vliet, E. Jagroep, S. Brinkkemper, “DevOps competences and maturity for software producing organizations,” in *Enterprise, Business-Process and Information Systems Modeling*, 2018, pp. 244–259. https://doi.org/10.1007/978-3-319-91704-7_16
[56] T. Seppä-Lassila, A. Järvi, S. Hyrynsalmi, *An assessment of DevOps maturity in a software project*, Master Thesis, University of Turku, 2017.
[57] R. Costa, R. Rodrigues, A. C. S. Dutra, “Application of Scrum Maturity Model in SoftDesign Company,” in *Brazilian Workshop on Agile Methods*, 2016, pp. 39–49. https://doi.org/10.1007/978-3-319-55907-0_4
[58] M. Anisetti, C. A. Ardagna, F. Gaudenzi, E. Damiani, “A Continuous Certification Methodology for DevOps,” in *Proceedings of the 11th International Conference on Management of Digital EcoSystems*, 2019, pp. 205–212. https://doi.org/10.1145/3297662.3365827
[59] N. Tomas, J. Li, H. Huang, “An empirical study on culture, automation, measurement, and sharing of devsecops,” in *International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, 2019, pp. 1–8. https://doi.org/10.1109/cybersecpods.2019.8884935
[60] IEEE, “IEEE Standard for DevOps: Building Reliable and Secure Systems Including Application Build, Package, and Deployment,” *IEEE Std 2675-2021*, vol. 1, pp. 1–91, 2021. https://doi.org/10.1109/IEEESTD.2021.9415476
[61] O. E. Adalı, Ö. Özcan-Top, O. Demirörs, “Evaluation of agility assessment tools: a multiple case study,” in *International Conference on Software Process Improvement and Capability Determination.*, 2016, pp. 135–149. https://doi.org/10.1007/978-3-319-38980-6_11
[62] ATOS, *DevOps Maturity Assessment*, 2020. https://bit.ly/3uTbPve
[63] Microsoft, *Microsoft DevOps Self-Assessment*, 2021. https://bit.ly/2RZCHLz
[64] IBM, *IBM DevOps Practice Self Assesment*, 2021. https://ibm.co/3w2bWEW
[65] IVI, *IVI’s DevOps Assessment*, 2021. https://bit.ly/3w9LGZd
[66] B. Kitchenham, S. Linkman, D. Law, “DESMET: a methodology for evaluating software engineering methods and tools,” *Computing & Control Engineering Journal*, vol. 8, no. 3, pp. 120–126, 1997. https://doi.org/10.1049/cce:19970304

[67] Infostretch, *Infostretch DevOps Self-Assessment*, 2020. https://bit.ly/3fh4krm
[68] InCycle, *InCycle Evaluacion de devops*, 2020. https://bit.ly/2RqYQCl
[69] Xmatters, *DevOps Maturity Survey Report*, 2021. https://bit.ly/33N8iCD
[70] Atlassian, *DevOps Maturity model*, 2021. https://bit.ly/2Rq1o3N
[71] Veritis, *Veritis*, 2021. https://bit.ly/3yhZhQ0
[72] Boxboat, *Boxboat*, 2021. https://bit.ly/3yqsDMm
[73] Humanitec, *DevOps Assessment*, 2021. https://humanitec.com/devops-assessment
[74] Atlassian, *DevOps Assessment*, 2021. https://bit.ly/3fChUpB
[75] Eficode, *Eficode DevOps Assesment*, 2021. https://bit.ly/3omPkfD
[76] C. Pardo, F. J. Pino, F. García, M. Piattini, M. T. Baldassarre, "A process for driving the harmonization of models," in *Proceedings of the 11th International Conference on Product Focused Software*, 2010, pp. 51–54. https://doi.org/10.1145/1961258.1961271
[77] C. Pardo, F. J. Pino, F. Garcia, "Towards an integrated management system (IMS), harmonizing the ISO/IEC 27001 and ISO/IEC 20000-2 standards," *International Journal of Software Engineering and Its Applications*, vol. 10, no. 9, pp. 217–230, 2016.
[78] C. Pardo, F. García, F. J. Pino, M. Piattini, M. T. Baldassarre, "PrMO: An Ontology of Process-reference Models," in *XVII Jornadas de Ingeniería del Software y Bases de Datos*, 2012.
[79] F. J. Pino, M. T. Baldassarre, M. Piattini, G. Visaggio, "Harmonizing maturity levels from CMMI-DEV and ISO/IEC 15504," *Journal of Software Maintenance and Evolution: Research and Practice*, vol. 22, no. 4, pp. 279–296, 2010. https://doi.org/10.1002/spip.437
[80] J. Guerrero, *DevOps Model - Modelo de referencia para la adopción de DevOps en empresas de desarrollo de software*, Grade Thesis, Universidad del Cauca, Colombia, 2021.
[81] C. Pardo, F. Garcia, M. Piattini, F. J. Pino, S. Lemus, M. T. Baldassarre, "Integrating multiple models for definition of IT governance model for banking ITGSM," *International Business Management*, vol. 10, no. 19, pp. 4644–4653, 2016.
[82] N. Fenton, J. Bieman, *Software metrics: a rigorous and practical approach*. CRC press, 2014. https://doi.org/10.1201/b17461
[83] Universidad de Cantabria, *Tema 2. Métodos de Valoración: Cuestionarios*, 2022, https://ocw.unican.es/mod/page/view.php?id=498#6
[84] M. F. Triola, *Probabilidad y estadística*. Pearson educación, 2004.
[85] J. Kontio, J. Bragge, L. Lehtola, "The focus group method as an empirical tool in software engineering," in *Guide to advanced empirical software engineering*, Springer, 2008, pp. 93–116. https://doi.org/10.1007/978-1-84800-044-5_4
[86] R. Likert, "A technique for the measurement of attitudes," *Archives of psychology*, vol. 22 no. 140, pp. 55, 1932.