Evaluation of Human DevOps with users

The usability of Human DevOps, which is recognized as an essential software product quality characteristic [1], was examined through two complementary approaches combining quantitative and qualitative methods. Qualitative feedback was complemented with quantitative evidence collected during the same validation phase, using established benchmarks to ensure meaningful interpretation. Two distinct samples were employed in separate evaluations: a group of expert evaluators and a cohort of final-year Computer Engineering students.

To enhance the validity of both evaluations, data source triangulation was performed. In this regard, we propose integrating the Technology Acceptance Model (TAM) [2] with ISO 9241-11 usability principles and objective performance metrics. TAM provides insights into perceived usefulness (PU) and perceived ease of use (PEOU), which are critical predictors of user acceptance, while ISO 9241-11 offers a complementary perspective by assessing usability through effectiveness (goal achievement), efficiency (resource expenditure such as time and effort), and satisfaction (user experience quality) in a defined context of use. The two frameworks are complementary: PEOU aligns closely with efficiency and satisfaction, as ease of use reduces effort and frustration, while PU corresponds to effectiveness, as perceived utility implies successful goal attainment. Consequently, a mixed evaluation approach can be implemented, combining TAM-based quantitative measures (e.g., Likert-scale surveys) with ISO 9241-11 objective indicators (e.g., task completion time, success rate, and error frequency), thereby reducing bias and enhancing the robustness of the analysis.

The quantitative measures aligned with ISO 9241-11 employed in this study are as follows: for effectiveness, the percentage of tasks completed without error; for efficiency, the time required to complete a task; and for satisfaction, the System Usability Scale (SUS), a widely adopted standardized questionnaire designed to assess user satisfaction and perceived usability of interactive systems. These metrics have been compared against established benchmarks to ensure meaningful interpretation.

The following research question was formulated: RQ. *What is the users' acceptance of Human DevOps according to the constructs of the TAM and the performance metrics defined by ISO 9241-11?*

The assessment performed by experts was conducted by a team of 19 usability experts who thoroughly examined the tool's capabilities and interface. Six participants are faculty members from the Department of Computer Science at the University of Murcia, with substantial expertise in usability evaluation. Among them, two have delivered undergraduate courses focused on usability. Furthermore, they possess demonstrable proficiency in DevOps practices, which reinforces their qualification as experts in the specific tool examined in this study. The remaining 13 participants are professionals from the software industry, each with between two and ten years of practical experience in software development, thereby forming a diverse and well-qualified expert group. The second assessment was conducted with 31 final-year Computer Engineering students (after excluding two outliers) enrolled in an elective course entitled User Interfaces, which is part of the Software Engineering specialization within the Computer Engineering degree program. This specialization ensures that students graduate with advanced knowledge in areas such as interface design, software development, software quality, DevOps, and usability principles, making them a technically qualified group to evaluate the Human DevOps tool from an end-user perspective.

A scenario was presented involving a software development company undertaking the modernization of its software architecture through the adoption of microservices. The participant assumed the role of DevOps team leader and was tasked with analyzing the results provided by the Human DevOps system. This analysis was conducted using both the interactive diagram and the list of recommendations, following the assessment of human factors and the capture of sociotechnical dynamics affecting performance. These dynamics had been monitored through daily surveys distributed via the Slack application.

Prior to interacting with the application, participants received instructions by viewing a demonstration video detailing the use of Human DevOps. The video explained the process of response collection via Slack, back-end data processing, and front-end dashboard analysis. Subsequently, the participant was required to interact with a deployed instance of the application to complete two designated tasks.

Upon completion of the tasks and the recording of time measurements, participants were asked to complete an online questionnaire. This questionnaire comprised 16 items based on the TAM, 10 items measuring user satisfaction according to the SUS, and 10 items assessing effectiveness and efficiency in accordance with ISO 9241-11.

The document provided to participants, which described the steps to be followed, is available at the following URL: https://github.com/investigaciongiis/human-devops/raw/refs/heads/main/resources/usability-evaluation/TasksStatement.pdf

1. TAM Evaluation

Participants filled out a five-point Likert scale questionnaire to measure the PEOU, PU, attitude towards use (AU), and intention to use (ITU) of the system by answering 16 items adapted from the TAM framework.

1.1. Experts

The initial perceptions of the experts suggest that 'Human DevOps' has significant potential for market acceptance due to its ease of use and practical utility. Table 1 shows that the application received high praise for its intuitive interface, which ensures smooth and accessible navigation, features that are crucial for rapid user adoption. Moreover, the application's functionality was acknowledged as a fundamental determinant in comprehending the well-being of professionals and facilitating the prompt resolution of any arising issues. Notice that in the assessment of perceived usefulness, the values vary depending on the question. While questions that explore the tool's utility for evaluating and improving individuals' psychological well-being receive ratings with a median of 5, when the question concerns the impact, the tool will have on productivity and project quality, the median is 4. These discrepancies can be explained through well-established theories on perceived usefulness [3]. TAM posits that individuals' intention to use a system is strongly influenced by their belief that it will enhance task performance. When the tool's primary function, improving psychological well-being, aligns closely with users' personal needs, the perceived utility becomes concrete, leading to higher ratings. In contrast, the impact on productivity and project quality is more contingent on organizational factors, making it less tangible and thus rated lower [4].

The experts were asked to identify the functionalities they considered most valuable, as well as those they perceived to be lacking. The Human DevOps tool has been positively evaluated by reviewers for several key functionalities that enhance its practical value. First, its seamless integration with Slack facilitates incorporation into existing workflows without disrupting established communication channels. Second, the tool's data collection process is minimally intrusive, requiring little effort from employees while still yielding valuable insights that would otherwise be difficult to obtain. Third, the team leader interface is simple with a minimalist design, contributing to a high level of usability. Fourth, the representation of data by using graphs is both visually appealing and informative, aiding in the interpretation of complex team dynamics. Fifth, the tool's recommendation list serves as a useful foundation for decision-making processes aimed at improving team members' well-being and performance.

The experts also proposed several enhancements to further improve the functionality and user experience of the Human DevOps tool. First, to increase employee engagement, gamification of the questionnaire, along with simple feedback mechanisms such as emotive icons, could serve as motivational incentives. On the other hand, while the graph view effectively illustrates relationships among factors, it could be complemented by a parallel tabular or list view to better evidence quantitative data, such as factor percentages sorted in descending order. A bidirectional linkage between the graph and the table or list, using linked highlighting, would enhance interactivity and data exploration. Moreover, if the number of responses varies across factors, node size in the graph could reflect response volume to improve interpretability. Although tooltips currently provide detailed information, implementing semantic zooming could allow users to view factor names and percentages dynamically as they zoom in, with spatial filtering enhancing clarity. Finally, tracking the implementation of recommended actions over time would be valuable. For instance, enabling access to a recommended measure and allowing it to be marked as completed to observe their continued relevance and impact in the future. Subsequent reports could then indicate whether the action continues to be recommended and whether the recommendation rate has changed, thereby providing evidence as to whether the measures have been effective.

Table 1.

Descriptive statistics of the item scores from the TAM evaluation questionnaire on experts' perceptions of Human DevOps "M": Mean; "SD": Standard deviation; "Md": Median; "p": P values returned by the Wilcoxon Signed Rank test when comparing the sample median to the neutral midpoint value of 3.

Evaluation questionnaire item		**M**	**SD**	**Md**	**p**
1.Do you consider that the tool requires little training to use it effectively?	PEOU	4,58	0,837	5,00	<0,001
2.How would you rate the ease of navigation within the tool's interface?	PEOU	4,74	0,561	5,00	<0,001
3.To what extent do you believe that using this tool increases your productivity?	PU	3,74	1,045	4,00	0,013
4.Does using the tool improve the quality of the projects you work on?	PU	3,74	0,991	4,00	0,01
5.I would use 'Human DevOps' if I needed to assess the psychological and emotional well-being of development and operations teams.	PU	4,32	1,108	5,00	0,001
6.I would use 'Human DevOps' to improve team well-being in software engineering environments.	PU	4,37	1,116	5,00	<0,001
7.Would you recommend 'Human DevOps' to other colleagues in the industry?	PU	4,26	0,991	5,00	<0,001
8.If it were available, I would like to use it in the company where I work.	ATU	4,11	1,196	5,00	0,004
9.Overall, I found the application intuitive.	PEOU	4,63	0,830	5,00	<0,001
10.I can easily identify the actions that need to be taken and the negative human factors present in the DevOps team.	PEOU	4,68	0,671	5,00	<0,001

11.Using Human DevOps would help me make better decisions about the team's well-being.	ATU	4,47	1,020	5,00	<0,001
12.Human DevOps would be a valuable resource for managing the psychological aspects of software development teams.	ATU	4,26	1,147	5,00	0,002
13.I would use Human DevOps to monitor and improve team dynamics over time.	IU	4,32	1,108	5,00	0,001
14.I would use Human DevOps as part of my regular workflow to support team health and performance.	IU	4,32	1,204	5,00	0,001
15.I will likely suggest adopting Human DevOps in future team projects.	IU	4,05	1,078	4,00	0,003
16.I see myself becoming a regular user of Human DevOps if it proves to be effective.	IU	4,00	1,247	4,00	0,012

Table 2 presents a summary of the psychometric properties of the constructs evaluated in the study for experts. Table 2 includes, for each construct, the items used, the mean and standard deviation of the responses, the median value, and the corresponding Cronbach's α coefficient. The reliability column provides a qualitative interpretation of the internal consistency based on the Cronbach's α value.

Table 2.

Psychometric properties of the constructs from the TAM evaluation questionnaire on experts' perceptions of Human DevOps. "M": Mean; "SD": Standard deviations; "Md": Median.

Construct	Items	M	SD	Md	Cronbach's α	Reliability
PEOU	P1, P2, P9, P10	4,66	0,18	4,71	0,94	Excellent
PU	P3–P7	4,08	0,45	4,05	0,96	Excellent
ATU	P8, P11, P12	4,28	0,33	4,26	0,93	Excellent
BI	P13–P16	4,17	0,46	4,26	0,95	Excellent
Overall TAM	—	4,29	0,48	4,29	0,98	Excellent

1.2. Final-year Computer Engineering students

The evaluations collected from Computer Engineering students are slightly lower than those obtained from experts, as shown in Table 3. This can be explained by differences in experience and expectations. Experts typically have a deeper understanding of the tool's intended purpose and its potential benefits within organizational contexts, which may lead to more favourable assessments. Students, on the other hand, might lack practical exposure to real-world software development environments, making it harder for them to fully appreciate the tool's strategic value.

The highest ratings for PEOU align with the principle that usability is immediately observable and requires minimal domain-specific knowledge. According to the TAM, ease of use strongly influences initial impressions, especially among novice users. Favourable ratings for PU and ATU suggest that students recognize some potential benefits, though these perceptions may be moderated by their limited experience. The lowest ratings for IU may reflect a slight lack of clarity regarding the tool's practical application in real-world scenarios. In particular, lower scores on the items *"I will likely suggest adopting Human DevOps in future team projects"* ($p = 0.095$) and *"I see myself becoming a regular user of Human DevOps if it proves to be effective"* ($p = 0.374$) can be attributed to the contextual characteristics of the sample. The participants were final-year Computer Engineering students enrolled in an academic setting, where the use of

Human DevOps was limited to an evaluative exercise rather than real-world project integration. According to the TAM, clarity and purpose are critical for user acceptance. When users, particularly students, do not fully understand how the tool will be integrated into organizational workflows or how it contributes to measurable outcomes, their confidence in its intended use diminishes. Additionally, expectancy-value theory [5] suggests that individuals evaluate tools based on the perceived likelihood of achieving valued outcomes. If the connection between the tool and its intended organizational impact is not explicit, users may assign lower ratings. This effect can be amplified among students who lack practical experience in organizational contexts, making it harder for them to envision the tool's strategic role.

Table 3.

Descriptive statistics of the item scores from the TAM evaluation questionnaire on Computer Engineering students' perceptions of Human DevOps "M": Mean; "SD": Standard deviations; "Md": Median; "p": P values returned by the Wilcoxon Signed Rank test when comparing the sample median to the neutral midpoint value of 3.

Evaluation questionnaire item		M	SD	Md	p
1.Do you consider that the tool requires little training to use it effectively?	PEOU	4.39	0.955	5.00	<0.001
2.How would you rate the ease of navigation within the tool's interface?	PEOU	4.65	0.709	5.00	<0.001
3.To what extent do you believe that using this tool increases your productivity?	PU	3.62	0.882	4.00	0.002
4.Does using the tool improve the quality of the projects you work on?	PU	3.81	1.078	4.00	0.002
5.I would use 'Human DevOps' if I needed to assess the psychological and emotional well-being of development and operations teams.	PU	4.00	0.894	4.00	<0.001
6.I would use 'Human DevOps' to improve team well-being in software engineering environments.	PU	4.07	0.892	4.00	<0.001
7.Would you recommend 'Human DevOps' to other colleagues in the industry?	PU	3.58	1.089	4.00	0.009
8.If it were available, I would like to use it in the company where I work.	ATU	3.68	1.045	4.00	0.003
9.Overall, I found the application intuitive.	PEOU	4.29	1.070	5.00	<0.001
10.I can easily identify the actions that need to be taken and the negative human factors present in the DevOps team.	PEOU	4.55	0.723	5.00	<0.001
11.Using Human DevOps would help me make better decisions about the team's well-being.	ATU	3.94	1.031	4.00	<0.001
12.Human DevOps would be a valuable resource for managing the psychological aspects of software development teams.	ATU	4.00	1.007	4.00	<0.001
13.I would use Human DevOps to monitor and improve team dynamics over time.	IU	3.90	1.136	4.00	<0.001
14.I would use Human DevOps as part of my regular workflow to support team health and performance.	IU	3.61	1.022	4.00	0.004
15.I will likely suggest adopting Human DevOps in future team projects.	IU	3.36	1.112	3.00	0.095
16.I see myself becoming a regular user of Human DevOps if it proves to be effective.	IU	3.35	1.257	3.00	0.374

Table 4 presents a summary of the psychometric properties of the constructs evaluated in the study for experts. Table 4 includes, for each construct, the items used, the mean and standard deviation of the responses, the median value, and the corresponding Cronbach's α coefficient. The reliability column provides a qualitative interpretation of the internal consistency based on the Cronbach's α value.

Table 4.

Psychometric properties of the constructs from the TAM evaluation questionnaire on students' perceptions of Human DevOps. "M": Mean; "SD": Standard deviations; "Md": Median.

Construct	Items	M	SD	Md	Cronbach's α	Reliability
PEOU)	P1, P2, P9, P10	4,47	0,44	4,48	0,78	Acceptable
PU	P3–P7	3,81	0,63	3,81	0,88	Good
ATU	P8, P11, P12	3,87	0,44	3,94	0,84	Good
BI	P13–P16	3,52	0,67	3,56	0,85	Good
Overall TAM	—	3,81	1,14	3,81	0,87	Good

2. ISO 9241-11 Evaluation

2.1. Satisfaction measured using the SUS

According to ISO 9241-11, user satisfaction is a key component of usability, reflecting the degree to which users find the system pleasant and acceptable to use. In this study, satisfaction was measured using the SUS, a widely adopted and validated instrument for assessing perceived usability. The benchmark applied corresponds to the global average SUS score of 68, which represents the 50th percentile based on an extensive meta-analysis of 446 studies and over 5,000 individual responses [6]. Our research question is: *Does the observed SUS score significantly exceed the established benchmark of 68?*

Two samples of 19 experts and 31 students were considered. For the expert group, the SUS score obtained from the 19 participants was 87.368, surpassing the established benchmark. Based on the Sauro–Lewis curved grading scale for interpreting SUS scores, this result is categorized as "A+ Outstanding" [6,7]. Moreover, Cronbach's α was calculated for the SUS questionnaire, resulting in a value of 0.90. This indicates excellent internal reliability. To assess whether the observed mean SUS score significantly differed from the benchmark, a one-sample t-test was conducted. The analysis revealed that the observed mean SUS score was significantly greater than the benchmark, $t(18) = 4.843$, $p = 0.00007$. For students, the SUS score obtained from the 31 participants was 80.967, which exceeds the established benchmark. According to the Sauro–Lewis curved grading scale interpretation of SUS scores, this value is classified as 'A Excellent'. In addition, Cronbach's α was calculated for the SUS questionnaire, resulting in a value of 0.80. This indicates good internal reliability. To determine whether the observed mean SUS score significantly differed from this benchmark, a one-sample t-test was applied. Observed mean SUS score was statistically significantly higher than the benchmark, $t(30) = 5.085$, $p = 0.00001$.

Therefore, we can be approximately 99.993% confident for experts, and 99.999% confident for students, that Human DevOps has an average SUS score greater than the industry average of 68, thus reflecting an exceptionally high level of user satisfaction.

2.2. Effectiveness measured using the percentage of tasks completed without error

Effectiveness, as defined by ISO 9241, was assessed using the metric "percentage of tasks completed without error." This metric represents the proportion of tasks that users completed without committing any errors, thereby serving as an indicator of the system's capacity to enable accurate and dependable task execution. By concentrating on error-free performance, it provides a quantitative foundation for assessing the usability dimension of effectiveness within the framework of human-computer interaction. To interpret this measure, a benchmark was established, specifically, a completion rate exceeding 70% was set as the minimum acceptable threshold [8]. Our research question is: *Does the Human DevOps tool enable users to achieve an effectiveness level that meets or exceeds the benchmark of 70% error-free task completion as defined by ISO 9241?*

The primary objective was to determine whether the observed success rate in the usability test significantly surpassed this predefined target. For statistical comparison, appropriate tests were applied based on sample size and the nature of the hypothesis. Since the sample size was small, the one-sample binomial test was employed to calculate the exact probability under the binomial distribution

An exact binomial test was run on a sample of 31 students and 19 experts to determine if a greater proportion of participants completed each task without committing any errors compared to those who did. Of the 19 experts, 19 (100%) in Task 1 and 19 (100%) in Task 2, carried out the work without committing any errors The p-value obtained was $p = 0.001$ for Task 1 and $p = 0.001$ for Task 2, indicating that the probability that the true (population) completion rate exceeds 70% is 99.999% for Task 1 and 99.999% for Task 2. Of the 31 students, 26 (83.87%) in Task 1 and 28 (90.32%) in Task 2, carried out the work without committing any errors. The p-value obtained was $p = 0.038$ for Task 1 and $p = 0.005$ for Task 2, indicating that the probability that the true (population) completion rate exceeds 70% is 99.962% for Task 1 and 99.995% for Task 2. This outcome suggests that Task 1 and Task 2 demonstrate a higher level of effectiveness compared to the benchmark.

2.3. Efficiency measured using the time required to complete a task

Efficiency, as defined by ISO 9241, was measured using the metric "time required to complete a task". This metric captures the duration taken by users to successfully finish a given task, thereby providing a quantitative indicator of the system's ability to support rapid and streamlined interaction. This measure reflects the extent to which the system minimizes the resources, particularly temporal effort, needed to achieve user goals, which is a core aspect of the usability dimension of efficiency within the ISO 9241 framework. The benchmark employed was set at twice the completion time achieved by an expert under optimal conditions.The benchmark times reported for Task 1 (21 seconds) and Task 2 (14 seconds) were obtained from a controlled experimental session involving 5 domain experts using the Human DevOps application. To minimize individual variability, the average completion time was calculated across 10 executions performed by these experts under standardized conditions. Our research question is: *Does the Human DevOps tool enable users to achieve an efficiency level such that at least 95% of users can complete Task 1 and Task 2 in less than twice the time taken by an expert (benchmark)?*

To determine whether the observed mean completion time significantly differed from this benchmark, a one-sample t-test was applied. For small samples, the geometric mean is considered the most appropriate estimator of the average task completion time. In the case of the expert sample, with GM = 13,534 seconds for Task 1, and GM = 13,453 seconds for Task 2, the time to complete tasks was statistically significantly lower than the benchmark, $t(18) = 6{,}701$, $p < 0.001$ for Task 1, and $t(18) = 4.685$, $p < 0.001$ for Task 2. Therefore, the probability of observing an average completion time of 13,534 seconds for Task 1 and 13,453 seconds for Task 2, assuming the true population mean exceeds 42 seconds in Task 1, and 28 seconds in Task 2, is lower than 0.001% and 0.001%, respectively. In other words, we can be about 99.999% and 99.999% confident that experts can complete Task 1 and Task 2 in less than 42 seconds and 28 seconds, respectively.

In the sample of final-year Computer Engineering students, with GM = 35,968 seconds for Task 1, and GM = 24,442 seconds for Task 2. Time to complete tasks was statistically significantly lower than the benchmark, $t(30) = 1.931$, $p = 0.031$ for Task 1, and $t(30) = 1.823$, $p = 0.039$ for Task 2. Therefore, the probability of observing an average completion time of 35,968 seconds for Task 1 and 24,442 seconds for Task 2, assuming the true population mean exceeds 42 seconds in Task 1, and 28 seconds in Task 2, is approximately 3.14% and 3.91%, respectively. In other words, we can be about 96.85% and 96.08% confident that students can complete Task 1 and Task 2 in less than 42 seconds and 28 seconds, respectively.

These results indicate that the system demonstrates a high level of efficiency, as experts and students can complete both tasks in significantly less time than the established benchmarks

3. Threats to validity

Construct and Internal Validity. To mitigate potential issues related to inadequate measurement of quality aspects in Human DevOps, well-established and reliable instruments, namely TAM, ISO 9241-11, and SUS, were adopted. Furthermore, the questionnaire underwent multiple iterative content reviews conducted by four authors, all of whom are researchers specializing in usability engineering. A portion of the survey respondents were industry professionals, which helps to counteract internal validity threats arising from confounding variables such as prior knowledge and experience in DevOps.

External Validity. The limited sample size used in the evaluation of Human DevOps (31 final-year Computer Engineering students and 19 experts) restricts the statistical generalizability of the findings. Future research should aim to replicate the study with a more representative and diverse sample to enhance external validity.

Conclusion Validity. A potential threat to conclusion validity stems from the design and administration of the expert survey. Although the experts and students provided informed and diverse perspectives, the absence of a comparative framework may have influenced their judgments by limiting contextual reference points. To mitigate this concern, the study employed a rigorous and transparent methodology, including predefined evaluation criteria, anonymized responses, and systematic analysis procedures, all aimed at enhancing reproducibility and minimizing bias.

Data Availability

The full dataset and statistical analysis (Excel file) are available at:

https://github.com/investigaciongiis/human-devops/raw/refs/heads/main/resources/usability-evaluation/data/HumanDevOps_ExpertStudy_Values.xlsx

https://github.com/investigaciongiis/human-devops/raw/refs/heads/main/resources/usability-evaluation/data/HumanDevOps_StudentStudy_Values.xlsx

https://github.com/investigaciongiis/human-devops/raw/refs/heads/main/resources/usability-evaluation/data/HumanDevOps_ExpertStudy_Calculations_ES.xlsx

https://github.com/investigaciongiis/human-devops/raw/refs/heads/main/resources/usability-evaluation/data/HumanDevOps_StudentStudy_Calculations_ES.xlsx

Please note that the sheets with the suffix '_Calculations_ES' use Excel functions in Spanish. For international use, refer to the sheets with the suffix '_Values.xlsx', in which all formulas have been replaced by their computed values.

References

[1] Moumane K, Idri A. Using ISO 9126 with QoS DiffServ model for evaluating software quality in mobile environments. 2014 Second World Conference on Complex Systems (WCCS), 2014, p. 134–9. https://doi.org/10.1109/ICoCS.2014.7060905.

[2] Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int J Man Mach Stud 1993;38:475–87. https://doi.org/10.1006/imms.1993.1022.

[3] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13:319–40. https://doi.org/10.2307/249008.

[4] Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward A Unified View1. Management Information Systems Quarterly 2003;27:425–78. https://doi.org/10.2307/30036540.

[5] Eccles JS, Wigfield A. Motivational Beliefs, Values, and Goals. Annu Rev Psychol 2002;53:109–32. https://doi.org/https://doi.org/10.1146/annurev.psych.53.100901.135153.

[6] Sauro J. A practical guide to the system usability scale: background, benchmarks & best practices. Denver: Measuring Usability LLC; 2011.

[7] Sauro J, Lewis JR. When designing usability questionnaires, does it hurt to be positive? Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA: Association for Computing Machinery; 2011, p. 2215–24. https://doi.org/10.1145/1978942.1979266.

[8] Sauro J, Lewis JR. Quantifying the User Experience, Second Edition: Practical Statistics for User Research. 2nd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2016.