Information Technology (IT) Governance Process with Naïve Bayes algorithm to improve the success rate of the IT projects

M Maria Antony Ranesh[1], S Justin Samuel[2]

[1]Research Scholar, Sathyabama Institute of Science and Technology, India
[2] Professor, Dept of Computer Science and Engineering, PSN Engineering College, Trinelveli, India

[1]mantonyranesh@gmail.com, [2]dr.s.justin@gmail.com

ABSTRACT

Information Technology (IT) governance is a combination of processes and controls to increase the success rate of projects, programs, and portfolios. We have several research articles evidencing the stability of classification techniques into medical, security, sports, artificial intelligence, software project management, etc. The utilization of classification techniques such as Naïve Bayes algorithm in software governance has unique advantages in terms of accuracy. The approach in this article would help to provide key management insights to make precise decisions and improve the overall success rate of the software projects, programs and portfolios.

Naïve Bayes algorithm is used to logically validate the alignment of governance policies and project details through Weka (Waikato Environment of Knowledge Analysis). Weka is a collection of machine learning algorithms for data mining tasks. IT governance policies are used as a benchmark for validation and processed with Naïve Bayes algorithm in Weka

Governance team members would be an interface to communicate and set expectations with the project manager. Basic governance standards across Software Development Lifecycle (SDLC) are defined and used as a bench mark in the governance assessment form. These predefined governance assessment forms and project details are validated through Naïve Bayes algorithm. This proactive model and adherence to the governance standards would help to increase the success rate of IT projects.

Keywords: Classification Techniques, Naïve Bayes algorithm, IT Governance, IT Policies, Industry Standards, Project Success Criteria, Proactive reporting and Software Project Management, Weka

I. INTRODUCTION

IT governance can significantly influence project performance (Haq et al., 2019) hence governance process can determine the success of the project. A systematic literature review on project governance (Musawir et al., 2020) suggests the development of a mainstream model for software project governance.

Software organizations are committed (Bhatti et al., 2010) to multiple regulatory requirements and policies. The governance process would help to demonstrate the controls they have over the software projects and align with strategic objectives. Small scale to large scale organizations uses various industry-standard governance frameworks such as COBIT (Control Objectives for Information and Related Technologies), ITIL (Information Technology Infrastructure Library), COSO (Committee of Sponsoring Organizations), CMMI (Capability Maturity Model Integration method) and FAIR (Factor Analysis of Information Risk). The proposed model can be used within any of these standard governance frameworks.

Results in this model are determined based on the governance policies and project details shared to management throughout the project lifecycle. Any potential changes in the governance policies should be aligned and approved by all the key stakeholders in the organization. The recent "Pulse of the Profession 2021" (Pulse of the Profession, 2021) indicates the significant change in working and delivery. Organizations need to think beyond the standard project management process to align with strategic controls and increase the overall success of the projects.

II. Related works

The impact on the success and the implementation of industry approaches (Serra & Kunc, 2015) suggest that BRM (Benefits Realization Management) practices can help to support business strategies and successful execution. The findings also indicate that a benefits management strategy integrated into the governance processes helps organizations to increase their ability to define and manage their success criteria.

Projects are directed and controlled by project governance (Musawir et al., 2020) This paper also highlights that project governance's interest is a productive ground for future research to make meaningful contributions. The project/program manager needs to make sure the governance plan is implemented throughout the project to monitor and control the effectiveness of the governance plans. There are various governance components and Project Management Control processes. Effective project governance must play a central role in realizing organizational strategic objectives through projects.

"Predicting whether software activities will be performed successfully is crucial in software projects" (Ramírez-Mora et al., 2020) Senior management should tailor the governance framework to the organizational culture, types of projects, and the organization's needs.

(Hazil et al., 2020) conducted an analysis on "Software Project Management using machine learning techniques" and concluded that project risk assessment with machine learning techniques are more successful in reducing the loss of project and increasing the likelihood of project success.

(Shrivastava et al., 2013) assessed different types of classification techniques in the data mining process using WEKA (Waikato Environment of Knowledge Analysis). Concluded that data mining process has importance in finding patterns, forecasting and discovery of knowledge.

(Kawamura et al., 2018) analyzed the data of 332 projects and concluded that success/failure prediction procedure, including the prediction algorithm, help significantly to specify projects that an organization needs to participate in as priority.

III. PROBLEM STATEMENT

Guidelines of how IT governance can enable organizational strategy implementations through projects is under researched area in the literature (Musawir et al., 2020).

Systematic literature reviews on IT Governance indicates that need for developing a main stream model for project governance. At an organizational level, governance should extend beyond the project life cycle to implement best practices and key learnings in subsequent projects.

Majority of research of IT governance is conceptual. Lack of scientific evidence to prove the value of IT governance.

Standish Group, Chaos report 2020 indicated that only 31% projects are currently successful. 19% failed and 50% challenged. Hence, we have strong demand to improve the success rate of the projects through various Governance Strategies. We have significant gaps in converting organizational goals at a projects level.

Research Questions

Research Question 1: Can we build machine learning based governance models to increase the success probability of the software projects?

Research Question 2: What are the key parameters to ensure projects are aligned with organization goals?

Research Question 3: What is the governance process to create more visibility and accountabilities?

Research Question 4: What is the governance checks to enable transparency in project management?

Research Question 5: What is the effective system to meet the governance need of the organization?

IV. METHODOLOGY

Proposed process can be grouped into following steps

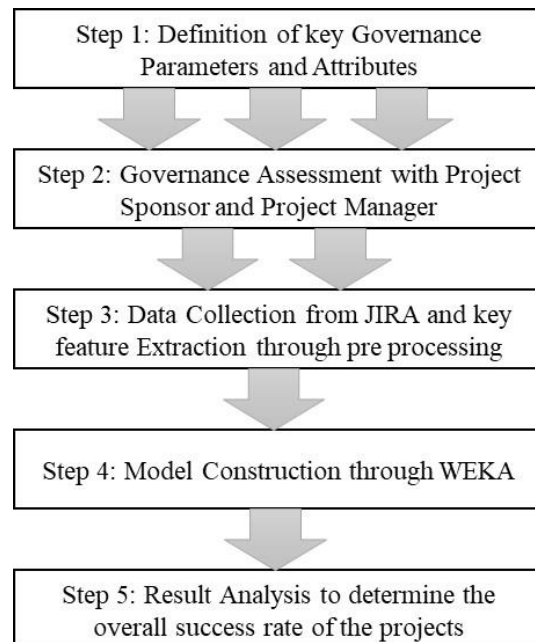**Step 1:** Definition of Governance Parameters and Attributes

Step 2: Governance assessment with project manager

Step 3: Data Collection from JIRA and key feature Extraction through pre processing

Step 4: Model Construction through WEKA

Step 5: Result Analysis to determine the accuracy

Figure 1: Representation of the overall methodology

Step 1: Definition of key governance parameters and Attributes

All the activities in a software project are expected to be aligned with IT Governance policies. Any changes in the governance policies should be agreed and signed off by the senior management team / key stakeholders. These signoffs would help to ensure the standards across multiple projects in an organization.

Following are the indicative list of governance parameter and attributes used to validate across the project from initiation till successful closure. "Table 1: Overall list of software project parameters" is the high-level list of project parameters in the governance process.

Table 1: Overall list of software project parameters		
S.No	Project Parameters	Data Type
1	Number of Risks, Issues & Dependencies	Number
2	Sponsor Approval	Binary

834

3	Reporting Schedule	Number
4	Number of key milestones	Number
5	Number of change request	Number
6	Number of rebaselining	Number
7	Milestone dates	Date

Risks, issues, and dependencies:

"Risks are the uncertain conditions that may appear unexpectedly during any phase of a project and may hinder the governance strategies to deal with risks" (Haq et al., 2019) This is a critical governance parameter to track the project-specific risk, issues, and dependencies. The governance team can determine the minimum and maximum numbers of risks across the organization based on the project's type.

Sponsor approvals:

"The project sponsor is a critical link between the executive and strategic levels of the organization and the effective delivery of the benefits the project/program was created to facilitate" (Too & Weaver, 2014) Approvals from sponsors are to ensure project alignment with the strategic structure with the required budget. Sponsors will also be key stakeholders in monitoring the overall progress and success rate of the project.

Budget

"Project performance was evaluated according to the planned budget, the schedule, the technical specifications (product/service requirements), and the ability to meet the customer service requirements" (Berssaneti & Carvalho, 2015) Used to determine the project's size and classify the type as large, medium, and small. If there are multiple projects in an organization, this grouping based on the budget will help us to predict the success rate based on the project size.

Reporting Schedule

Project visibility to the management comes only through consistent/periodic information provided by the project managers. The governance team can set directions on the reporting frequency (daily/ weekly/ fortnightly/ monthly) and monitor the project's progress. This is one of the most common measures in all the projects (Atkinson, 1999)

Number of change requests

"By adopting the formal change management process, the inconsistency in requirements and failures of the project can be avoided" (Bhatti et al., 2010) Based on the type of project, the organization can streamline the change request process. This process helps to ensure the minimum and the maximum number of change requests. The robust CR (Change Request) process will always help to clear all the dependencies and create more visibility to the organization's change.

Number of rebaselining

There are different types of baselines, such as scope baseline, schedule baseline, and cost baseline. This policy from the governance team will ensure that the project is baselined across the scope, schedule, and cost. "A baseline version can be changed only through a formal procedure" (Hanakawa et al., 1996)

Number of milestones and dates

"Progress control with milestones is one of the best-known approaches to managing software development" (Hanakawa et al., 1996). Progress of any project could be determined only through the successful completion of various key milestones. This parameter is a decision point for further investment in the project. Hence, this value in the tool will ensure that the minimum

number of milestones. "Table 2: Key project milestones in Governance policies" represents milestones for tracking the project progress.

S.No	Project Milestone Parameters	Data Type
Table 2: Key Project milestones in Governance policies		
1	Requirements Signoff	Date
2	Design Signoff	Date
3	Database Signoff	Date
4	Code Review and Signoff	Date
5	Testing / UAT Signoff	Date
6	Go Live	Date
7	Learning / Best Practices	Binary

Design standards:

This model can help to track the overall design quality of the project using these criterias. The software governance team can propose design standards across projects as per the organization's requirements and the type of the project. Design standards can include the latest technologies/versions of the tools to be used. E.g., W3C (World Wide Web Consortium) standards for web-based projects.

"Table 3: Software Design Standards" represents the indicative list of design standards. Project type and scope can change the design standards. This helps to address the design expectations from the stakeholders and fix the design concerns. (Engineering et al., 2009)

S.No	Design Parameters	Data Type
Table 3: Key design parameters in Governance policies		
1	HTML & CSS	Binary
2	JavaScript Web APIs	Binary
3	Graphics	Binary
4	Audio and Video	Binary
5	Accessibility (WCAG)	Binary
6	Internationalization	Binary
7	Mobile web	Binary
8	Privacy	Binary
9	Math on the Web	Binary
10	Design Signoff	Binary / Date

Database Standards:

The governance team can streamline the database standards used across multiple projects in an organization. Tracking the database structure will enable organizations to ensure standard practices such as naming conventions, consistency, transparency in the table structure, data normalization, documentation, data encryptions, security standards, etc. Control on the database can improve the overall data quality of the projects in an organization. The type and scope of the project can determine database standards. "Table 4: Key Database Standards" highlights the types of database standards that can be used with the governance validations.

Synod	Database Parameters	Data Type
Table 4: Key Database parameters in Governance Policies		
1	Standard Naming Conventions	Binary

2	Consistent Table Structure	Binary
3	Data Normalization	Binary
4	Data Encryption	Binary
5	Data Security	Binary
6	Database Signoff	Binary & Date

Coding Standards:

Coding Standard can reduce the hidden cost of developing software. Tracking of coding standards can increase technical reliability and understandability to reduce the complexity of the code. According to the project type, governance team can enforce and track the adherence to the coding standards. The governance team can also suggest standard naming conventions, indentation to increase the understanding of the code, exception handling, etc. "Table 5: Software Coding Standards" represents the various coding standards that can be used with the governance process.

Table 5: Key coding Standards		
Synod	**Coding Parameters**	**Data Type**
1	Standard Naming Convention	Binary
2	Consistent Indentation	Binary
3	Code Grouping	Binary
4	Exception handlings	Binary
5	Supporting Documents	Binary
6	Code Review and Signoff	Binary & Date

Testing Standards:

Various industry-standard testing process are available across the software organizations. Testing standards would help an organization to improve the test documentation quality, optimize testing techniques and strategies. There are multiple types of industry standards for testing, such as ISO (International Organization for Standardization), IEEE (The Institute of Electrical and Electronics Engineers), CMM (The Capability Maturity Model), etc. This tool can ensure the project's adherence to the testing standards and increase the overall project success." Table 6: Key Testing parameters in Governance Policies" represents the list of key testing standards that can be expected by the Governance process to improve the overall testing standard of the projects.

Table 6: Key Testing parameters in Governance Policies		
Synod	**Testing Parameters**	**Data Type**
1	Test Plan	Binary
2	Test Cases	Binary
3	Unit Testing	Binary
4	Integration Testing	Binary
5	Black box testing	Binary
6	White-box testing	Binary
7	Interface Testing	Binary
8	Regression Testing	Binary
9	Testing / UAT Signoff	Binary & Date

Deployment Standards:

As part of this model set of activities planned for go-live can be grouped as deployment activities. The governance team can insist on project managers adhering to the deployment standards. These standards may change according to the type of project or organization. "Table 7: Key Deployment parameters in Governance Policies" represents few sample deployments standards.

Table 7: Key Deployment parameters in Governance Policies		
Synod	**Deployment Parameters**	**Data Type**
1	Milestone Signoffs	Date & Binary
2	Deployment Roadmap	Binary
3	Adequate Support	Binary
4	Access to Servers / Firewalls	Binary
5	User Testing	Binary
6	User Communication	Binary
7	Deployment Evidence	Binary

Above parameters can also be used for the support / maintenance projects based on the type of the requirements. Overall process can be customized and implemented to various activities such as resource management, number of planned projects, completed projects and InProgress projects

Step 2: Proposed Approach with Governance Assessment

Governance Assessment:

In most of the organization governance team / PMO manages the project creation request by the Project Manager. Hence formal meeting invite should be sent by the governance team to fetch the necessary project information. The governance team should have the authority to seek more information from the project manager to approve/reject the project creation request.

As part of the governance assessment, the project manager can also get the latest templates and understand the expected milestone standards in the project's execution. This model determines the success percentage based on the following industry-standard protocols.

Inputs: Governance Standards & Policies, Sample dataset from JIRA / Project details

Step 1: Governance assessment by Governance Team with Project Managers
▪ Governance Team to communicate the organizational governance policies
▪ Project Manager to confirm the project alignment with governance policies

Step 2: Validation of Governance Policies and Projects Specific information
▪ Data extraction from standard project / activity management tools (JIRA)

▪ Data Preprocessing to align with Project Governance Policies ▪ Data validation with Classification Techniques (Naïve Bayes) in WEKA tool SStep 3: Results Comparison and Analysis ▪ Result comparison of the Classification Techniques without Governance Policies and with Governance policies
Outputs: - Governance report to management based on the organization policies - Milestone / Project phase-specific recommendation to project managers to improve the success rate

Step 3: Data Collection and Feature Extraction: Required data set for experiments are collected from JIRA

JIRA is a software development tool widely used to plan, track, and release software. JIRA is commonly used as an Issue Tracking System (ITS) and many organizations have a public JIRA ITS to record and manage issues of their software projects / open source projects.

JIRA data set - Apache's JIRA Issue Tracking System1 is an open Issue Tracking System that stores more than 600 software projects. Pre-processed to use only essential attributes to improve the success rate of the projects Analysis data extracted from Mendeley Data in the Issues, comments and projects from four popular Issue Tracking Systems - Mendeley Data

Spring and Atlassian data set are used for the analysis

JIRA Source	Bugs	Defect	Improvement	New Feature
Spring	11054	143	9786	2893
Atlassian	15533	0	744	9

Number of JIRA records for Bugs are consistent than defects, improvements and new Feature. Hence Bugs from Spring and Atlassian are used for the evaluation with WEKA.

Research Hypothesis:

Based the scope of the available dataset from JIRA following governance policies are validated as part of the data preprocessing.

Governance_Policy_1- All the issue resolution duration should be less than 365 calendar days. If the resolution duration is greater than 365 calendar days project / activity point of contact to provide required rustication to the governance team. Governance_Policy_2 – Issue watch count should be less than 4. If the issue watch count is greater than 4 project / activity POC to provide required rustication to the governance team

These IT governance policies are validated and results are loaded into WEKA. Governance_Policy_1 and Governance_Policy_2 would help the sponsors / management to address the following challenges.

Governance_Policy_1 – High ageing issue resolutions can be tracked by the management / sponsors. Implement the best practices from other resolutions to close the high ageing issues. Governance_Policy_2 - Track the watcher list to identify the active contribution in an issue. Keep track of the data security aspects and ensure only responsive watchers are included.

Both policies are validated in Spring and Atlassian dataset to ensure consistency in validation. Following are list of preprocessed datasets with corresponding attributes.

S.No	Field Name	Attributes
1	issue_id	Numeric
2	issue_resolution	Text
3	issue_status	Text
4	issue_type	Text
5	project_id	Numeric
6	issue_watchCount	Numeric
7	issue_success	Binary
8	issue_Resolution_Duration	Numeric
9	governance_Policy_1	Binary
10	governance_Policy_2	Binary

These governance policies are validated to determine the overall duration of the bug resolution time and number of watchers. Evaluate the success rate, accuracy and average through Naïve Bayes in Spring and Atlassian datasets from JIRA.

Naïve Bayes is one of the popular classification technique based on Bayes Theorem. It is used to determine the probability of Governance Policy 1 and Governance Policy 2 based on various attributes in the JIRA extract. This algorithim is mostly used in text classification and with problems having multiple classes.

Feature Extraction: Key features are chosen from the overall data set and inclusion on agreed governance policies

Step 4: Model Construction

Derived features are used as an input for WEKA to determine the data accuracy through Naïve Bayes classification technique.

Step 5: Result Analysis

Following results are analyzed to check the overall alignment with governance policy1 and governance policy 2

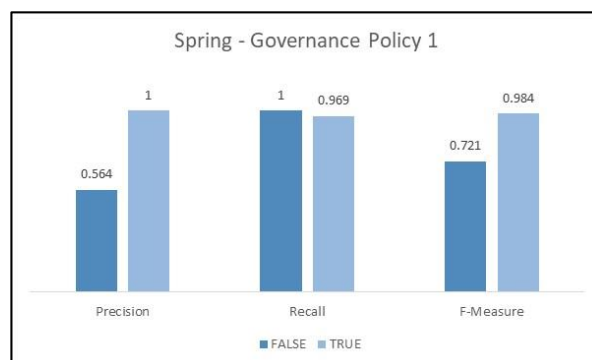Figure 2: Average performance values of governance policy 1 using Spring dataset

Figure 3: Average performance values of governance policy 2 using Spring dataset
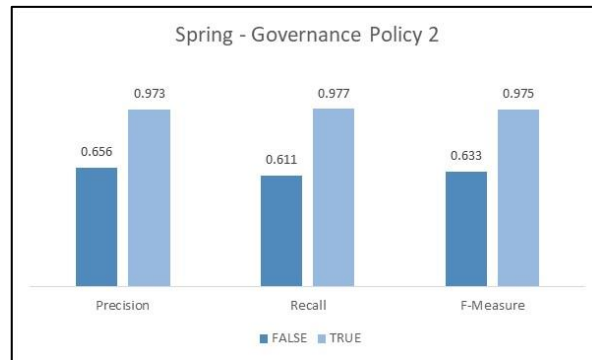

Figure 4: Average performance values of governance

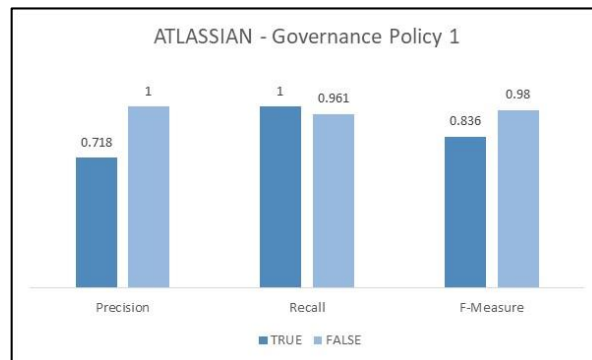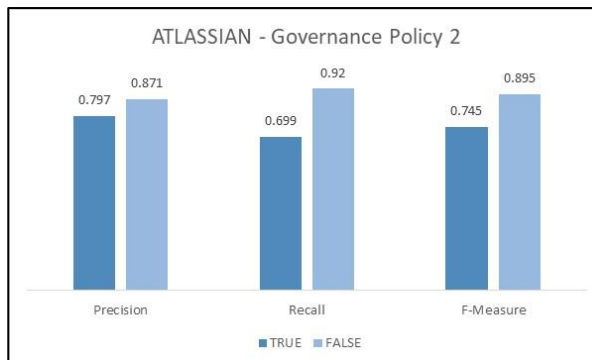policy 1 using Atlassian dataset

Figure 5: Average performance values of governance

policy 2 using Atlassian dataset

More statistics of accuracy measurements are detailed in appendix Tables

V. Conclusion and Research Answers

Precision, Recall and F-Measure score of predicting issue success varied from 0.0 to 1.0 for the governance policy 1 and governance policy 2. Both Spring and Atlassian dataset provided a consistent output for the dataset with governance validation. Based on the output consistency it determined that the overall issue status in Spring and Atlassian dataset with governance validation can improve the alignment with governance polices.

This approach could be used for validation in the across various key project activities at a project, program or portfolio level. Hence the overall alignment of governance policies and projects can significantly increased.

Organization governance policies and project details from a project manager are used as a critical input to improve the success rate of software projects. Mutual discussion between the Governance team and project manager based on the governance assessment form would trigger more insights and help the project manager to clearly understand the management expectations. Critical phases of projects such as requirement gathering, design, database, coding, testing, and deployment are tracked to validate the alignment with governance policies.

Research Question 1: Can we build machine learning based governance models to increase the success probability of the software projects?

Results of this IT Governance approach with Naïve Baye algorithm proves the feasibility of governance models with machine learning. Optimal use of project data and algorithms with governance policies ensures projects aligned with strategic objectives.

Research Question 2: What are the key parameters to ensure projects are aligned with organization goals?

Adherence with regulatory/compliance requirements, Budget, Timeliness and Risk Reduction are the key parameters to ensure projects are aligned with organization goals

Research Question 3: What is the governance process to create more visibility and accountabilities?

Consistent Management reporting and on time governance / management decisions would create more visibility   and accountabilities across the projects

Research Question 4: What are the governance checks to enable transparency in project management?

Project validations and reporting as per the governance policies would help to enable transparency in projects management

Research Question 5: What is the effective system to meet the governance need of the organization?

Project management and reporting should be aligned with governance polices. This would be an effective system to meet the governance need of the organization.

This approach will enable precise management information and critical decision at the right time to improve the success rate of the software projects

VI. Key Benefits and Recommendations

IT Project tracking and governance validation from industry-standard criteria such as budget, compliance, customer satisfaction, quality, and timeliness would help the organization's IT projects to succeed.

The Senior management / Software Governance team can use this approach across multiple projects in an organization. On a large scale, an organization generally groups projects into portfolios and programs. This approach can be deployed at a project or program or portfolio level.

Grouping projects based on the business areas such as retail, banking, logistics, insurance, and healthcare would help to optimize the governance standards in the particular domain.

REFERENCES

1. Atkinson, R. (1999). Project management: Cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. *International Journal of Project Management*, *17*(6), 337–342. https://doi.org/10.1016/S0263-7863(98)00069-6

2. Berssaneti, F. T., & Carvalho, M. M. (2015). Identification of variables that impact project success in Brazilian companies. *International Journal of Project Management*, *33*(3), 638–649. https://doi.org/10.1016/j.ijproman.2014.07.002

3. Bhatti, M. W., Hayat, F., Ehsan, N., Ishaque, A., Ahmed, S., & Mirza, E. (2010). A methodology to manage the changing requirements of a software project. *2010 International Conference on Computer Information Systems and Industrial Management Applications, CISIM 2010*, 319–322. https://doi.org/10.1109/CISIM.2010.5643642

4. Engineering, S., Committee, S., & Computer, I. (2009). *IEEE Std 1016-2009 (Revision of IEEE Std 1016-1998), IEEE Standard for Information Technology—Systems Design—Software Design Descriptions* (Vol. 2009, Issue July).

5. Hanakawa, N., Iida, H., Matsumoto, K., & Torii, K. (1996). A framework of generating software process including milestones for object-oriented development method. *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*, 120–130. https://doi.org/10.1109/APSEC.1996.566747

6. Haq, S. U., Gu, D., Liang, C., & Abdullah, I. (2019). Project governance mechanisms and the performance of software development projects: Moderating role of requirements risk. *International Journal of Project Management*, *37*(4), 533–548. https://doi.org/10.1016/j.ijproman.2019.02.008

7. Hazil, M. Z. M., Mahdi, M. N., Mohd Azmi, M. S., Cheng, L. K., Yusof, A., & Ahmad, A. R. (2020). Software Project Management Using Machine Learning Technique - A Review. *2020 8th International Conference on Information Technology and Multimedia, ICIMU 2020*, 363–370. https://doi.org/10.1109/ICIMU49871.2020.9243543

8. Kawamura, T., Toma, T., & Takano, K. (2018). Outcome prediction of software projects for information technology vendors. *IEEE International Conference on Industrial Engineering and Engineering Management*, *2017-Decem*, 1733–1737. https://doi.org/10.1109/IEEM.2017.8290188

9. Musawir, A. ul, Abd-Karim, S. B., & Mohd-Danuri, M. S. (2020). Project governance and its role in enabling organizational strategy implementation: A systematic literature review. *International Journal of Project Management*, *38*(1), 1–16. https://doi.org/10.1016/j.ijproman.2019.09.007

10. Pulse of the Profession. (2021). *Beyond Agility*. 1–23. https://www.pmi.org/-/media/pmi/documents/public/pdf/learning/thought-leadership/pulse/pmi_pulse_2021.pdf?v=b5c9abc1-e9ff-4ac5-bb0d-010ea8f664da&sc_lang_temp=en

11. Ramírez-Mora, S. L., Oktaba, H., & Gómez-Adorno, H. (2020). Descriptions of issues and comments for predicting issue success in software projects. *Journal of Systems and Software*, *168*, 110663. https://doi.org/10.1016/j.jss.2020.110663

12. Serra, C. E. M., & Kunc, M. (2015). Benefits Realisation Management and its influence on project success and on the execution of business strategies. *International Journal of Project Management*, *33*(1), 53–66. https://doi.org/10.1016/j.ijproman.2014.03.011

13. Shrivastava, M. K., Chouksey, P., & Miri, R. (2013). *Exploring Data Mining Classification Techniques*. *2*(6), 2554–2559.

14. Too, E. G., & Weaver, P. (2014). The management of project management: A conceptual framework for project governance. *International Journal of Project Management*, *32*(8), 1382–1394. https://doi.org/10.1016/j.ijproman.2013.07.006