Journal of Advanced Multidisciplinary Research Studies and Development

ISSN : 2583-6404
Jan - Feb 2024

PERSPECTIVES ON SDLC MODELS USED TO DEVELOP AI ENABLED SOFTWARE

Nandkishor Mete[1], Akshay Gadhave[2]

[1] *Department of Computer Science, Padmashri Manibhai Desai Mahavidyalaya, Uruli Kanchan, Pune, 412202, India.*

[2] *Software Engineer, Freelancing Software Consultant, Keshavnagar, Pune, India.*

Abstract

This research paper aims to compare the different software development life cycle (SDLC) models used in building AI enabled systems and to learn what changes are needed in these models to handle the challenges thrown by AI technology. A systematic study was conducted to collect and analyze the technical knowledge about software engineering for AI enabled systems. Different SDLC Models were studied with regards to different properties of AI based applications. It was found that SDLC models need to evolve to suit the production of AI based applications. This study highlights the areas in which changes can be made in these models. The study also offers different perspectives to the decision makers to enable them to make better decisions regarding SDLC model to be followed.

Keywords: AI Enabled Systems, Software Development Life Cycle models, SE for AI, AI based Software.

► *Corresponding Author: Nandkishor Mete*

1. INTRODUCTION

In Software Engineering, the Software Development Life Cycle models (SDLC) provide framework and roadmap for building high quality software systems successfully [1]. Different prescriptive SDLC models like Waterfall, Agile, DevOps, Incremental Model, Prototype Model, V Model, Spiral Model, RAD were developed to engineer the software systems[2]. In recent years Software Systems are being built to implement Artificial Intelligence Technology. As the technology is ever changing, it is necessary to change the process of building the software. Software engineering for AI based systems poses challenges related to data, interpretability, scalability, ethics, robustness, security, compliance, continuous learning, and talent acquisition [3],[4],[5]. Addressing these challenges requires new approach to building the software. Some new models like CRISP-DM, TDSP were developed to cope with challenges thrown by AI technology. [6] This study aims to understand the different software development models (SDLC models) used in the context of building AI based applications, what are the challenges posed by new AI technology, and to learn how these models can be adapted to build AI based systems which are - accurate, robust, scalable, transparent, explainable, secure, adaptable, and developed with ethical considerations.

2. BACKGROUND

Artificial Intelligence is an innovative technology with transformative potential. Development of software systems with AI capabilities needs innovative Software Engineering approach and methods. Software development life cycle (SDLC) models are abstractions which specify the stages of software development process and sequence of activities. Different SDLC models serve different purposes like - simplicity, meeting the customer requirements, risk management, cost reduction, suitability for large and complex projects, faster releases, and adaptability [2]. Software

ISSN : 2583-6404
Jan - Feb 2024

Engineering for AI based systems must deal with unique challenges posed by the new technology [3]. Overcoming these challenges requires among other things, selecting appropriate software development life cycle model. This study tries to analyze various software project management models used in developing AI based applications, to present an overview of the existing practices, to understand the challenges thrown by the complex AI projects, to identify the changes needed and to offer a perspective in making decisions related to process of software creation. This was done through a study of existing practices. Also, a survey was conducted among practitioners of the software process models to understand challenges posed by AI.

The choice of SDLC model for building AI depends on various factors like Scope of the project, Changing requirements, Data Security, Explainability, Reliability of the product, Ethics, Resource and skill requirements, Quality, Innovation, Creativity. Different perspectives on the decision of selecting SDLC models and suggestions were provided.

3. LITERATURE REVIEW

AI based software systems have inherently different characteristics than traditional software systems. Software Engineering for AI enabled systems is different because of complexities involved in Data management, different skills needed in customizing AI models and difficulties in handling of AI components [3]. Various SDLC Models for AI are in use, these models combine features of traditional models like Waterfall, Spiral and Agile methodologies.

To provide framework for managing development of Machine Learning and Data Science projects CRISP-DM and TDSP models were introduced [6].

Cross Industry Standard Process for Data Mining (CRISP-DM) is a process model which combines features of Waterfall and Agile models. CRISP-DM breaks down a project into six phases- Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment [6].

Team Data Science Process TDSP is an agile, iterative methodology to deliver Machine Learning solutions efficiently. The original methodology includes four major stages, Business Understanding, Data Acquisition, Modelling and Deployment [6].

Microsoft model studied by Amershi et al. (2019) had 9 stages - Model Requirements, Data Collection, Data Cleaning, Data Labelling, Feature Engineering, Model Training, Model Evaluation, Model Deployment, Model Monitoring. This model can be modified to include Feasibility Study, and Peer-reviewed Documentation. This model also emphasizes Model Scoring and Risk Assessment as a part of Model Evaluation [6].

Existing SDLC models used to build software with Machine Learning components are missing essential steps, such as feasibility study, documentation, model evaluation, and model monitoring. There is a need for proper tool for exploratory data analysis and data integration techniques [7].

An Agile, iterative approach called D7R4 was proposed for Intelligent Vision Systems. It is inherently test driven and consists of seven stages - discover, dig, describe, design, develop, demonstrate, and deploy. It also consists of four reviews from different perspective such as quality, user experience, ethics, and security [8].

AZ-Model, for software development was proposed by introducing new activities during software development life cycle. It was found to overcome the limitations of traditional models and significantly impact the production of a quality product in a time-box. AZ-Model was found to be effective in producing a quality product within a given time and budget [9].

AI needs technical expert teams from different domains, high sophistication of knowledge, costly frameworks, data and algorithms[3]. The Organization for Economic Co-operation and

ISSN : 2583-6404
Jan - Feb 2024

Development OECD has formulated the characteristics of the AI systems as follows- Robustness, Security, Safety, and Traceability. It also mandates Systematic risk management (risks related to AI systems, including privacy, digital security, safety and bias) [10]. There are different challenges in developing AI like - understandability and interpretability of AI algorithms, lack of specifications and defined requirements, need for validation data and test input generation, defining expected outcomes for testing, accuracy and correctness measures, non-functional properties of AI-based systems, self-adaptive and self learning characteristics, and dynamic and frequently changing environments [11]. Choosing the right methodology involves consideration of above-mentioned challenges along with other factors like - Technical expertise of developers, involvement of project owner, project complexity, budget and time constraints [12].

Existing studies have tried to offer suitable SDLC model for AI (mainly ML and Data Science). Some have tried to address uncertainty by proposing Waterfall Structure. Some have tried to focus on changing requirements by using Agile – Iterative process. But aspects like Innovation, Creativity, Expertise in different domains, Ethics, Security, and Government Regulations were not fully considered. Also matters relating to cost saving, project overruns, scope management, and team size are yet to be properly addressed.

4. DISCUSSION

4.1 Why SE for AI Enabled Systems Is Different

AI-based software, utilizing AI techniques like ML, NLP, Deep Learning, Reinforced Learning, Computer Vision, and Data Science, requires a different approach from traditional software. It must be flexible, scalable, easy to understand, and well-defined. The software development process must also address ethics and government regulations. Expertise in different domains and creative solutions are needed to build AI-powered systems that operate autonomously, make data-driven decisions, and interact with humans intuitively.

4.2 Methodology of Study

(a) Study of existing material: A study was carried out to understand different SDLC models like – CRISP-DM, TDSP, Microsoft Model, Agile Methodologies for ML, D7R4 model and AZ model for Intelligent Vision Systems. Various research papers were studied to find out the challenges arisen because of AI.

(b) Primary Sources: Based on the study of existing models and challenges faced while building AI; parameters for a survey were formulated. The focus group for survey was decided as follows: decision makers from software development companies which have adopted an SDLC model for managing its processes.

(c)

Parameters of Study were defined as follows:

(a) Dependent Variables: Quality of Software Produced, Meeting Clients expectations, Project Completion Time, Failure / success rate of project, Reliability of the product, Safety of the data.

(b) Independent Variables: Factors affecting software production, Cost, Availability of time, Domain Expertise of team, Creativity / Innovation, Government Regulations, Ethical Standards, Quality.

4.3 Characteristics of Existing Models

The Waterfall model is a sequential SDLC model, while the Incremental model is an iterative

ISSN : 2583-6404
Jan - Feb 2024

SDLC model. Agile methodologies prioritize flexibility, while the Spiral model focuses on risk management and cost reduction. RAD is an iterative SDLC model that focuses on rapid prototyping and quick development cycles. DevOps emphasizes collaboration, automation, and integration for faster releases. Special models for AI include CRISP-DM, TDSP, and Microsoft. D7R4 is an agile, iterative methodology for developing Intelligent Vision Systems, with phases such as Discover, Dig, describe, design, and review. The AZ-Model divides software development into design and communication, development, and deployment phases. From the survey conducted with software practitioners it was found that majority preferred Agile methodologies for developing software.

4.4 Challenges Introduced by AI Integration

Developers face challenges in building AI-based software, including data quality, model interpretability, scalability, ethical considerations, robustness, security, regulatory compliance, continuous learning, and skill gaps. Data quality is crucial for AI system performance, while model interpretability is crucial for understanding decision-making. Scalability is crucial due to computational demands and memory resources. Ethical considerations include fairness, accountability, and bias. Robustness is essential for AI systems, as models can be vulnerable to adversarial attacks. Security and regulatory compliance are also important. Continuous learning and model maintenance are essential for adapting to changing data distributions and environments. Recruiting, training, and retaining skilled software engineers is a challenge in the rapidly evolving AI field.

4.5 Parameters of Success of Software Development Project

The parameters for the success of software development project include accuracy, speed and efficiency, scalability, user satisfaction, cost effectiveness, and profitability. Accuracy is a critical parameter of success for AI applications, while speed and efficiency are important for AI systems that need to handle large volumes of data or requests. Scalability is an important parameter of success for AI systems that need to handle large volumes of data or requests, while user satisfaction is a key parameter of success for AI applications that interact with end-users. Cost effectiveness is a measure of how effectively the resources are used, which will ultimately influence the Profitability of the project.

4.6 Identification of AI Factors which Influence the Decision of Model Selection

After analysis of Primary Data following factors were found to have impact on the decision of selecting the SDLC model- Data Management, Requirement and Scope analysis, Domain Expertise, Data Safety, Robustness, Explainability / Transparency, Ethics, Govt. regulations, Innovation, Cost Effectiveness, Validation / Accuracy.

4.7 Mapping of AI Factors to SDLC Models

Existing SDLC models address some of the challenges posed by AI. Table 1 lists the existing models and AI factors or challenges these models address.

ISSN : 2583-6404
Jan - Feb 2024

Table 1 AO Challenges Addressed by the Existing SDLC Models

Model	***AI Factors / Challenges addressed***
Agile + Scrum	Adaptability, Flexibility
Incremental	Building ML model, Training the model, adding features
CrispDM	Business understanding, Model Building, focus on Data
TDSP	Business Understanding, Model Scoring and Risk Analysis
Microsoft Model	Focus on model training, documentation
D7R4	Discover + dig + demonstrate and review.
AZ	Quality

It can be seen that Agile models are adaptable and flexible. Incremental model allows the development of complex AI by adding features at each increment.
CRISP-DM a model focusing on Data allows better business understanding and model building. TDSP additionally incorporates Model Evaluation and Risk Analysis into the life cycle. Microsoft model focuses on model training and explainability by creating comprehensive documentation. D7R4 is test driven approach for enhancing Quality, User Experience, Ethics and Security. AZ model ensures quality by incorporating customer review of the prototype.
We have mapped the model characteristics required to address the various AI factors in Table 2

Table 2 Mapping of AI Factors to Model Characteristics Required

AI Factor	***Suitable Model / Model Characteristic required***
Uncertainty / complexity	Linearity, planning and design first approach, comprehensive documentation. Waterfall model.
Building reliable model (Robustness)	Prototyping and Incremental development
Domain Expertise	Adding more phases for algorithm / model building and also for Ethical Considerations. Spiral model.
Scope Management	Prototyping
Data Security	Testing at each level (Spiral), add a phase for security testing
Quality (of the model), ISO standards	Process Maturity Model (Microsoft)
Sophisticated frameworks and algorithms	Agile methods and Incremental approach

ISSN : 2583-6404
Jan - Feb 2024

Traceability, explainability for debugging	Documentation Heavy model
Lack of defined requirements	Agile and iterative approach
Self Learning + Creativity + Innovation	Spiral approach
Ethical Considerations	Spiral approach with Expert teams
Government Regulations	Add a new phase to check compliance
Budget + Cost Considerations	Agile + Scrum
Validation in absence of test specification	Develop new testing framework

5. RESULTS

This research paper explores various perspectives on the SDLC model for AI-based software development. It emphasizes the importance of adaptability, domain expertise, iterative nature, explainability, prototyping, validation, risk management, data security, ethics, regulatory compliance, quality, and innovation.

Agile methods are essential for AI projects, as they emphasize flexibility, adaptability, and collaboration. Domain experts from different domains must interact with developers to ensure clear requirements are incorporated into the software.

The spiral model emphasizes continuous iteration and feedback, ensuring collaboration between stakeholders. Each phase must be well documented for transparency, especially in complex projects like finance and healthcare. Prototyping can help define project scope and requirements, while validation by domain experts ensures accuracy and reliability. The spiral model provides a structured approach to risk management, allowing developers to identify and address risks at each stage of the development cycle. Data security, ethics, and regulatory compliance are also considered, with a separate phase added to existing models to ensure compliance.

Quality is also a key consideration, so maturity models can be used to refine processes used in AI development.

6. CONCLUSION

Prototyping followed by a Spiral Model enhanced by extensive documentation can be considered for developing AI. New phases for Data Management, Data Modeling and Evaluation, Ethical Consideration, Innovation (creating new use cases) can be added. A new phase for evaluation of regulatory compliance can be added before deployment stage. Each cycle can follow Agile approach with Scrum for timely delivery of products at the end of each cycle.

ACKNOWLEDGMENTS

The authors express their gratitude towards the Principal, Padmashri Manibhai Desai Mahavidyalaya for providing all the support. The authors are indebted to the responders of the survey. They also express their gratitude towards the Research Karwaan Team for help in writing this paper.

REFERENCES

1. R. Pressman, Software Engineering a Practitioners Approach, Tata McGraw Hill, 2010

ISSN : 2583-6404
Jan - Feb 2024

2. Sabbir M Saleh, M Ashikur Rahman, K Ali Asgor, "Comparative Study on the Software Methodologies for Effective Software Development", International Journal of Scientific & Engineering Research, Volume 8, Issue 4, April-2017
3. Dr. Ipek Ozkaya, "What is Really Different in Engineering AI-Enabled Systems?" Engineering Intelligent Software Systems, Software Solutions Division, Carnegie Mellon University Software Engineering Institute, 2022
4. Lukas Fischer, Lisa Ehrlinger, Verena Geist, Rudolf Ramler, Florian Sobiezky, Werner Zellinger, David Brunner, Mohit Kumar and Bernhard Moser, "AI System Engineering—Key Challenges and Lessons Learned", mdpi, 2021
5. Hollen Barmer, Rachel Dzombak, Matt Gaston, Eric Heim, Jay Palat, Frank Redner, Tanisha Smith, Nathan VanHoudnos, "Robust and Secure AI", Carnegie Mellon University Software Engineering Institute, 2021
6. Mark Haakman, Lu´ıs Cruz, Hennie Huijgens, Arie van Deursen, "AI lifecycle models need to be revised, An exploratory study in Fintech", Empirical Software Engineering, 2021.
7. Silverio Martínez Fernández, Justus Bogner, Xavier Franch, Marc Oriol, Julien Siebert, Adam Trendowicz, Anna Maria Vollmer, Stefan Wagner, Software Engineering for AI-Based Systems: A Survey, ACM Transactions on Software Engineering and Methodology Volume 31 Issue 2 Article No.: 37 epp 1–59, 2021 https://doi.org/10.1145/3487043
8. J. I. Olszewska, Software Development Life-Cycle for Intelligent Vision Systems, School of Computing and Engineering, University of West Scotland, U.K., ScitePress
9. Muhammad Azeem Akbar, Jun Sang, Arif Ali Khan, Fazal-E-Amin, Nasrullah1, Muhammad Shafiq, Shahid Hussain, Haibo Hu1, Manzoor Elahi, Hong Xiang, Improving the Quality of Software Development, Process by Introducing a New Methodology AZ-Model, IEEEAccess, 2017
10. The OECD Artificial Intelligence (AI) Principles, [Online] available: https://oecd.ai/en/dashboards/ai-principles/P8, [Accessed: May, 16, 2023].
11. Lukas Fischer, Lisa Ehrlinger, Verena Geist, Rudolf Ramler, Florian Sobiezky, Werner Zellinger, David Brunner, Mohit Kumar and Bernhard Moser, "AI System Engineering—Key Challenges and Lessons Learned", mdpi, 2021.
12. John Mcgrath, Jana Kostalova, "Project Management Trends and New Challenges 2020", Project Management Trends and New Challenges 2020, Technological University Dublin, Dublin, Ireland