AI-Driven Test Automation for Healthcare Data Warehousing Projects

Arun Kumar Ramachandran Sumangala Devi

Architect II- Software Testing UST Global Inc, Glen Allen, Virginia, USA,
akumarrs@gmail.com

Abstract

Healthcare data warehousing test automation is becoming a success through the help of AI-based technology that results in accuracy, efficiency, and data integrity on the automated test. The one central to personalized patient data that forms the core of traditional data warehousing solutions frequently faces problems of complexity and dissatisfaction. Deep learning with test automation solutions makes data ingestion, processing, and testing conversant through machine learning algorithms [1]. These systems encompass separate acts such as testing for data integration, testing the quality of data as well as performance testing removing the factors of time and human error. In healthcare organizations, this approach delivers the message of enhancing quality, compliance and organizational efficiency in value-based care and precision medicine initiatives. As such, artificial intelligence tools can be used to address the essential issues regarding data warehousing, and this can in turn enhance decision-making, and thus, enhance the health of patient and the overall population.

Keywords: healthcare, artificial intelligence, machine learning, test automation, data warehousing.

1. Introduction

In the rapidly evolving era of the healthcare sector, the data warehousing process plays a crucial role in managing a large number of patients' personalized information. Traditional models such as virtual warehouses, data marts, and enterprise data offer complexity and often lack accuracy and efficiency. Data Warehousing refers to the method of business intelligence including storage and analysis of large volume information [2]. It is a central repository to aggregates data from various databases to provide a comprehensive overview for generating reports and informed decisions. A typical data warehouse architecture consists of three main tiers including the bottom tier, middle, and top tier. The bottom tier contains the warehouse servers and the process is called ETL (extract, transform, and load) to shift the data [3]. The middle tier consists of an OLAP (Online Analytical Processing) server that enables intelligent and fast query response time among the other services. This can use variety of OLAP models such as ROLAP (Relational online analytical processing), MOLAP (Multidimensional online analytical processing), HOLAP (Hybrid online analytical processing) [4]. Most of the time automation helps in predicting errors in routine tasks preventing errors before it is late. Artificial intelligence in test automation is emerging as the future of healthcare data warehousing. This work lays the foundation for the

analysis of AI based test automation application processes, advantages, difficulties, and further possibilities in the sphere of healthcare..

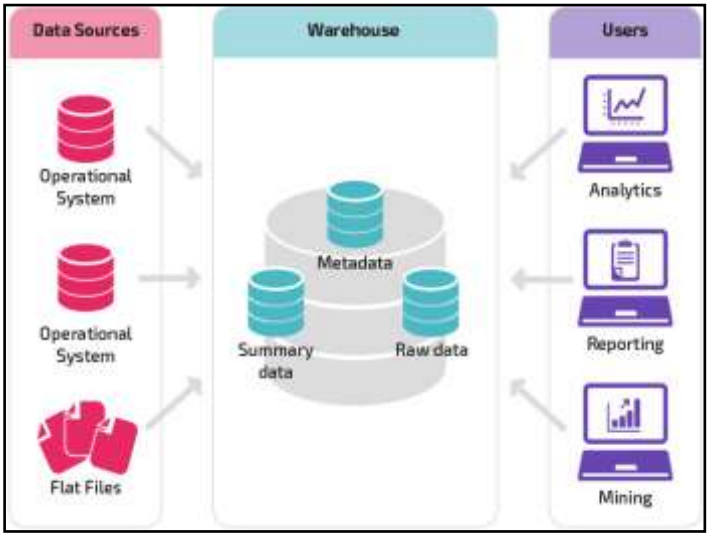

Figure 1: Steps of data warehousing projects [2]

2. SOLUTION

AI testing methodology for healthcare data warehousing project incorporates AI and ML in test automation to test the healthcare data. This solution entails the use of smart tools to perform data acquisition and analysis, alongside development of models for testing, with the primary goals of eliminating manual testing. The AI-dependent solution consists of the following major components:

Data ingestion: Subparts of an AI-based tool help with Patient Acquisition from other sources including the electronic health records (EHRs), and claims data, and other sources.

Process of large volume data: The algorithm of AI and ML works to process the accumulated data by identifying patterns and predicting potential issues. In terms of artificial intelligence, it analyzes huge information in a short time to recognize the trends and anomalies affecting the data quality and integrity [5]. The algorithm of machine language helps in predictive analysis to assume any disparities in data processing and integration before occurring to initiate proactive measures.

Test automation framework: This framework concerns the processed data for the automation process including a few steps of data integration testing, data quality use, performance testing, and security testing. The technique of test automation involves continuous integration/continuous deployment (CI/CD) pipelines to ensure the happening of automation at

every stage of the development cycle [6]. In the case of scriptless testing, AI-dependent tools allow non-technical users to create and run tests without knowledge of extensive programming.

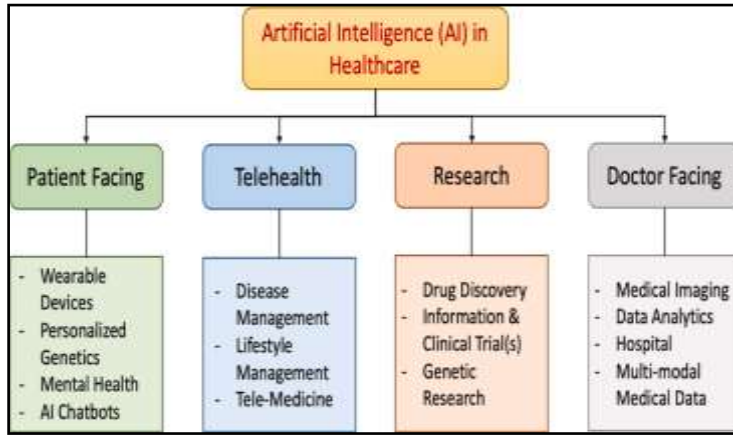

Figure 2: Impact of artificial intelligence in healthcare [7]

Evaluation and analysis of data quality: This component provides real-time reporting and analytics on testing results, enabling healthcare organizations to identify areas for improvement. AI-based tools used in data warehousing automation are able to automatically profile incoming data for quality assessment, completeness, and consistency. The algorithm derived from machine learning can detect outliers and anomalies in the dataset [8]. The evaluation of these actions is crucial in the healthcare industry to ensure the safety and compliance of patients.

Figure 3: Solution of using AI-based tools [9]

3. APPLICATION OF THE SOLUTION

AI-enhanced test automation improves the healthcare data warehousing effectiveness, dependability, and security to achieve more optimal patient results and operational performance in the following ways:

Data Integration Testing: The testing of data integration processes is made possible by using AI-powered tools. This guarantees that information will be converted well and coherently from one system to another, hence, avoiding the generation of a lot of unnecessary errors. For instance, if patient information from the EHRs and other sources such as lab results is pulled into a data warehouse, AI tools can check the correctness of such a procedure. Computer programs can mimic user behavior and produce facial data for system testing in various scenarios [10]. These can also access and interpret performance parameters in real time.

Data Quality Testing: The utilization of AI-driven test automation identifies and validates the issues in the data quality such as inaccuracies and inconsistencies. For example, it plays a vital role in accurately checking patient medical records such as missing values or repetitive entries, and fostering this requiring correction.

Performance Testing: Some AI-based tools mimic the users' interactions to assess the productivity of the data warehousing projects based on the offered load. This comes in practical use when evaluating how the system performs in terms of capacity and quantum of data as well as the number of user queries during the peak seasons. For example, it can simulate thousands of concurrent queries to be sure that the data warehouse is capable of processing them without any loss of performance quality. AI-driven testing can use techniques such as anomaly detection and data profiling, for instance, flagging outliers and or missing values of data sets [11]. The unstructured data entries can also be validated using Natural Language Processing (NLP).

Security Testing: In terms of intelligent testing helps to detect and verify the presence of risks associated with the provision of data warehousing systems. It reduces the possibility of leakage and access to sensitive information of health institutions. The application of AI tools entails the propagation of cyber threats to assess the exploited data warehouse for security loopholes and recommend how they may be addressed [12].

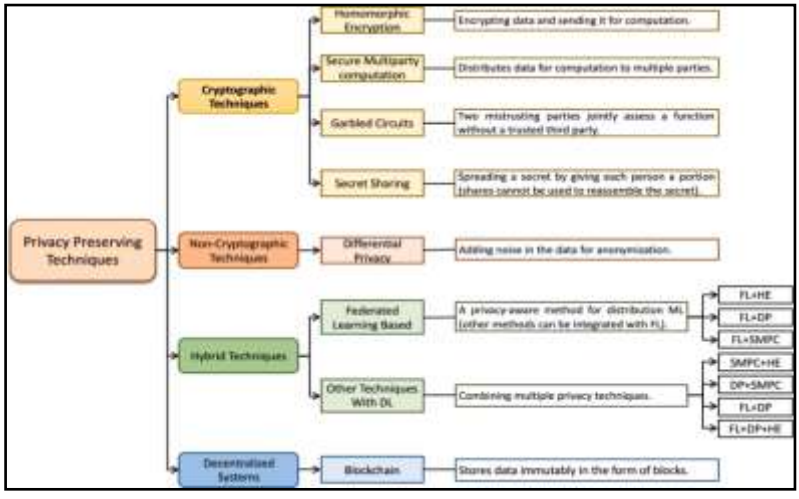

Figure 4: Lists of privacy-preservation techniques [7]

In the ML model, the protection of data privacy is the main concern by maintaining both the model parameters and training algorithms through the use of cloud providers. FL (federated learning) is an algorithmic framework for developing ML models in an online cloud setting for computation.

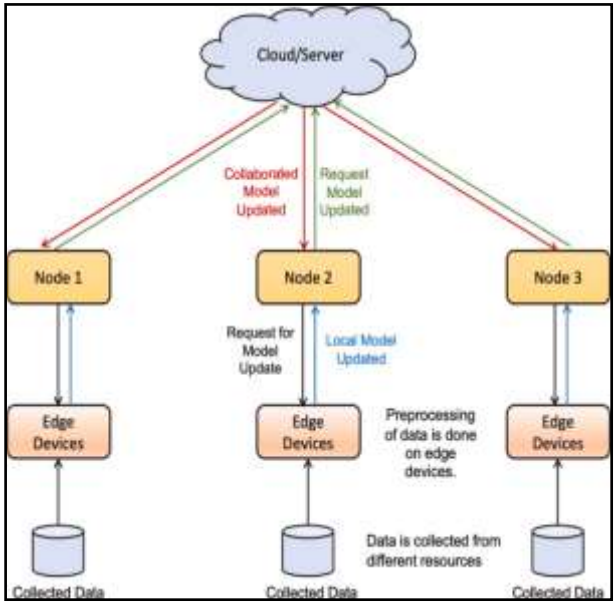

Figure 5: Steps of using cloud services [7]

4. BENEFITS OF THE SOLUTION

The implementation of AI-driven test automation (AITA) can significantly elevate the healthcare data warehousing projects in the following ways:

Population Health Management: AITA can help to easily test out any data integration and quality processes for population health management [13]. At the time of integration and automation, these processes tend to enable healthcare organizations to ensure that data from multiple sources including EHRs and social determinants of health are of high quality. This results in the improvement of population health management since huge and accurate data are available for processing and decision-making processes.

Value-Based Care: In value-based care initiatives, the use of AITA would enable testing of data integration and quality processes with regard to the quality of patient outcomes and total care quality. This ensures that the healthcare providers are in a position to meet the demands of the new payment reforms that have made it mandatory for measures of success to shift from a volume-based health system to a value-based system. Automation testing can make it easier for

an organization to find problem areas with data quality and prevent adverse effects on health care and expenditures [14].

Precision Medicine: For precision medicine initiatives, AITA can improve the validation of the data integration and quality processes so that a correct assembly of genetic, clinical, and environmental data can be performed [15]. This is relevant in carving out individual patient data that is used in the formulation of treatment procedures. These ways when automated systems need to assist the healthcare organization since they can make the numbers employed in precision medicine accurate resulting in improved performance amongst the patients.

Improved compliance: The current legal demands that healthcare firms have to obey comprise among others HIPAA (Health Insurance Portability and Accountability Act) and ICD-10 (International Classification of Diseases) [16]. These regulations can be met with the help of AI-Driven Test Automation. Such systems as AI can help check the compliance of the practices that are used in managing data with the laid down regulatory measures. For instance, they can verify that patient data is encrypted, and all access is being restricted to a certain group of people.

5. CONCLUSION

The healthcare data warehousing problem is well addressed with the help of an AI-driven test automation solution. Thus, the utilization of AI increases productivity and reliability concerning numerous large datasets containing sensitive information about patients through the automation. It potentially reduces the time-consuming and error-prone activities, including data extraction, data cleansing, and security testing. This strategy also ensures that the performance is better when the work is to maintain regulatory norms, for example under HIPAA and ICD-10 legal requirements. By assessing the probable problems that would occur in healthcare AI applications, the automation of decision-making processes helps in improving the management of population health across different payment models including value-based care, personalized medicine, and the welfare of patients. Finally, incorporating AI in test automation is expected to improve the healthcare data system and improve the general patient health, and organizational efficiency.

WORKS CITED

A. Bucher, E. S. Blazek, and C. T. Symons, "How Are Machine Learning and Artificial Intelligence Used in Digital Behavior Change Interventions? A Scoping Review," Mayo Clinic Proceedings: Digital Health, 2024.

"Data Warehouse Concepts: Traditional vs. Cloud," Panoply, 2024. [Online]. Available: https://panoply.io/data-warehouse-guide/data-warehouse-concepts-traditional-vs-cloud/. [Accessed: Oct. 21, 2024].

"Data Warehouse," IBM, 2024. [Online]. Available: https://www.ibm.com/topics/data-warehouse. [Accessed: Oct. 21, 2024].

T. M. Al Taleb, S. Hasan, and Y. Y. Mahd, "On-line analytical processing (OLAP) operation for outpatient healthcare," Iraqi Journal of Science, pp. 225-231, 2021.

Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, and A. Amira, "Artificial intelligence-based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives," Applied Energy, vol. 287, p. 116601, 2021.

A. Singh and V. Mansotra, "A Comparison on Continuous Integration and Continuous Deployment (CI/CD) on Cloud Based on Various Deployment and Testing Strategies," International Journal for Research in Applied Science and Engineering Technology, vol. 9, no. VI, pp. 4968-4977, 2021.

N. Khalid, A. Qayyum, M. Bilal, A. Al-Fuqaha, and J. Qadir, "Privacy-preserving artificial intelligence in healthcare: Techniques and applications," Computers in Biology and Medicine, vol. 158, p. 106848, 2023.

O. Koren, M. Koren, and O. Peretz, "A procedure for anomaly detection and analysis," Engineering Applications of Artificial Intelligence, vol. 117, p. 105503, 2023.

A. Aldoseri, K. N. Al-Khalifa, and A. M. Hamouda, "AI-Powered Innovation in Digital Transformation: Key Pillars and Industry Impact," Sustainability, vol. 16, no. 5, p. 1790, 2024.

D. S. AbdELminaam, A. M. Almansori, M. Taha, and E. Badr, "A deep facial recognition system using computational intelligent algorithms," Plos One, vol. 15, no. 12, p. e0242269, 2020.

D. Pentyala, N. Devarasetty, and V. C. Manduva, "AI-Powered Anomaly Detection in Distributed Data Engineering Systems," International Journal of Advanced Engineering Technologies and Innovations, vol. 1, no. 4, pp. 388-419, 2023.

R. Kaur, D. Gabrijelčič, and T. Klobučar, "Artificial intelligence for cybersecurity: Literature review and future research directions," Information Fusion, vol. 97, p. 101804, 2023.

A. Schneider-Kamp, "The potential of AI in care optimization: insights from the user-driven co-development of a care integration system," INQUIRY: The Journal of Health Care Organization, Provision, and Financing, vol. 58, p. 00469580211017992, 2021.

L. Budach, M. Feuerpfeil, N. Ihde, A. Nathansen, N. Noack, H. Patzlaff, F. Naumann, and H. Harmouch, "The effects of data quality on machine learning performance," arXiv preprint, arXiv:2207.14529, 2022.

T. Zayas-Cabán, K. J. Chaney, C. C. Rogers, J. C. Denny, and P. J. White, "Meeting the challenge: Health information technology's essential role in achieving precision medicine," Journal of the American Medical Informatics Association, vol. 28, no. 6, pp. 1345-1352, 2021.

J. I. Beik and J. Pepper, Health Insurance Today-E-Book: Health Insurance Today-E-Book, Elsevier Health Sciences, 2020.