

Impact of Data Lake on User Satisfaction:

A prerequisite for adopting Artificial Intelligence in Bangladesh's Healthcare sector to achieve Sustainable Development Goal 3

Abstract

Background: Healthcare systems in developing countries, including Bangladesh, are grappling with a range of ongoing issues, including fragmented health data, unequal access to services, lack of skilled staff, and loss of public trust. While artificial intelligence (AI) provides a glimpse of potential solutions with predictive diagnostics and clinical decision support, the success of AI applications in healthcare relies on the quality and integrated value of healthcare data. The volume, variety, and velocity of healthcare data generated by electronic health records, medical imaging, and clinical information systems have led to the development of data lakes as a type of infrastructure that can handle such data. In low- and middle-income countries, however, data lakes are not adopted and underdeveloped because of poor digital infrastructure and governance issues.

Objective: This study aims to examine the effect of data lakes on user satisfaction, which is crucial for the successful implementation of AI in the healthcare sector in Bangladesh for attaining Sustainable Development Goal 3 (SDG 3). Specifically, the research investigates the role of information quality, system quality, service quality, trust and perceived privacy on intention to use, actual use, and finally user satisfaction in a data lake for clinical decision support systems.

Methods: The study is an extension of DeLone and McLean Information Systems Success Model (1992, 2003), adding the concepts of "trust" and "perceived privacy" as new mediators between the intention to use and actual use. A cross-sectional research design will be used, with semi-structured interviews being conducted with 15 healthcare IT professionals, administrators and providers who have been selected using judgmental sampling to gather qualitative data. This will involve the collection of quantitative data from around 330 patients and/or their attendants from primary, secondary and tertiary urban hospitals through structured questionnaires. A multistage probabilistic sampling method will be used where a combination of cluster and systematic random sampling will ensure representativeness. The data analysis method used is quantitative using partial least squares structural equation modeling (PLS-SEM) software of SmartPLS 4.0 and qualitative analysis for the data obtained from the interviews is used as thematic analysis.

Expected Outcomes: The anticipated outcomes of this study are that it will be one of the first empirical studies on data lakes as a prerequisite for the use of AI in Bangladesh's healthcare sector. The finding will inform the priorities for investments in data infrastructure by policy makers and healthcare administrators to enhance the quality of information, systems and services, build trust and privacy protections, and improve user satisfaction. This study provides empirical evidence to grasp the impact of a data lake infrastructure on user experience, which helps build AI readiness and contributes to SDG 3 in developing country context like Bangladesh.

Contribution: This research extends the existing information systems literature by taking the DeLone and McLean model in the context of data lake in Bangladesh using AI, alongside trust and perceived privacy as antecedents of intention to use, actual use, and user satisfaction, which was not explored in other related studies.

Keywords: Data Lake, User Satisfaction, Artificial Intelligence Adoption, Healthcare Sector, SDG 3, Bangladesh, DeLone and McLean IS Success Model, Trust, Privacy, PLS-SEM