Journal of Positive School Psychology
2022, Vol. 6, No. 4, 10331–10336

Data Analytics Model for Health Institutions

Fernando Molina-Granja[1], Silvana Vire[2], Patricio Inca[3], Wayner Bustamante[4]

[1] Universidad Nacional de Chimborazo, Ecuador fmolina@unach.edu.ec
[2] Universidad Técnica Particular de Loja, Ecuador, scvire@utpl.edu.ec
[3] Escuela Superior Politécnica de Chimborazo, Ecuador, g_inca@espoch.edu.ec
[4] Universidad Nacional de Chimborazo, Ecuador wayner.bustamante@unach.edu.ec

Abstract

In the environment of health management, it is very necessary to use data analytics for several aspects, from improving administrative management, the study of the behavior of diseases in the population to accurately diagnosing, even pathologies that can currently be identified early to make timely and accurate interventions on some ailments. The motivation for applying data analytics in health is to focus on a value-focused care model and not on episode care, increase the benefits of historical data and help solve possible complex pathologies. This article presents a proposal for a data analytics model for health institutions called ADHEALTH, which has been validated by the media of experts and has a statistical study that demonstrates a favorable trend for its implementation. The model presents seven phases in a specific order, the legal environment as the initial phase, the operational environment focused on health institutions, institutional strategic planning, thus defining the appropriate sources of information and generating the data warehouse, to generate the exploitation applications, and finally solve institutional requirements and establish institutional use.

Keywords: Data Analytics, health institutions, Adhealth model, health management

1 Contextualization

Data analytics refers to quantitative and/or qualitative perspectives to obtain useful and important information from data. Consider techniques such as data extraction, to categorize them and determine possible relationships, trends, or patterns of behavior. Currently, organizations of all types require a structuring of their data for their processing and obtaining information for effective decision making that allows optimizing the organization's business processes and improving performance.

Depending on the type of organization and its business line, various data analytic techniques and models can be applied to help meet organizational objectives. It is essential to determine requirements and data sources [1].

2 HEALTH DATA ANALYTICS REQUIREMENTS

In the environment of health management, it is very necessary to use data analytics for several aspects, from improving administrative management, the study of the behavior of diseases in the population to accurately diagnosing, even pathologies that can currently be identified early to make timely and accurate interventions on some ailments.

Today, there is talk of internal devices of the human body that monitor and record different organs and timely identify patterns with which a disease can be diagnosed and thus apply the timely treatment. Concomitantly, data analytics can be used to reduce costs of treatments, and medical services, anticipate epidemics, and improve efficiency in the management and administration of health facilities.

Electronic record systems and processes have been implemented in the health systems of several countries, and such a wealth of information requires an effective model for clinical and epidemiological data analysis to take advantage of the world's evident digital transformation. The motivation for applying data analytics in health is to focus on a value-focused care model and not on episode care, increase the benefits of historical data and help solve possible complex pathologies.

The correct response to the digital transformation in the health systems environment allows to determine insights, tends to accelerate discoveries, and increases the dominance of the subject in society. This transformation must be based on the flow of data and the constant interaction between the actors that make up the health system, be they patients, professionals, providers, etc. The handling of a large amount of structured and unstructured data, (annual storage is estimated in ranges of Petabytes of information[1]) in each phase of the ETL, thus presenting the data for analysis of relationships and patterns that allow solving the problem posed. A study shows that more than 70% of health actors are optimistic about the results that data analytics can provide in health care [2] [2].

From the management of health care, direct benefits are known in risk analysis in orthopedic surgical interventions, in the guarantee of patient safety[3], [4], and the management of scheduling of surgeries [5]. In the administrative efficiency of the infrastructure and the proper management of technology in health facilities, including optimization of maintenance of medical equipment [6], [7].

Among the most notorious challenges in health data analytics are:

- **Systems Integration:** Diverse IT solutions even within the same institution, with little or no compatibility, and a framework with low possibilities of integration between platforms and systems [8].

- **Regulatory framework:** In Latin America, there is a limited definition of policies and strategies for the management and management of big data and its analytics, including aspects **of data privacy** [8]**,** [9], **ethical issues** [9], and legal issues in the analysis of data in health[10].

- **Data management:** There is a tendency of criteria in ensuring that many health systems are not efficient due to their limited ability to collect, analyze and make timely and effective decisions with patient data. Despite an expenditure on information management of almost 30% of their budgets, the problem still exists due to the decentralization and diversity of data sources [11], [12], [13].

In summary, the challenges to be solved in eHealth are framed in the lack of integration of systems, the generation of regulatory frameworks, the almost zero exchange of information, and with it the need for models and analytical methods adjusted to the reality of the sector [12] [13]**.**

Once the data has been reformatted and the results issued, you can decide between 4 models to apply Data analytics, they are differentiated by their approaches: descriptive, diagnostic, predictive, and prescriptive.

Descriptive: What's going on? - Provides a descriptive report of who, when, where, how much, and what has happened in a time, useful to provide basic information and a contextual view that allows understanding the data.

Diagnosis: Why is it happening? - Provides information focused on the reasons that led to a certain situation, which may require several data sources and even hypotheses to be resolved. By diagnostic analysis, patterns and related dependencies can be identified.

Predictive: What is most likely to happen? .- Focuses on trying to predict what will happen in the short term. It helps to determine trends and correlations between variables and possible

[1] Petabyte: Mbyte multiple equal to ten raised to the fifteen bytes

[2] ETL : Extract, Transform and Load. Prubbing to obtain data of multiple sources, transform them into a standard format, debug them and finally load them into a new database integral

causalities. It is supported by statistical models, machine learning mechanisms, and data mining.

Prescriptive: What do I need to do? .- It is based on a deep understanding of what has happened, why, and several possible future scenarios, options of what could happen, to provide information that would support timely and adequate decision making, and minimize the risk of future problems or taking advantage of new trends [13].

3 METHODOLOGY

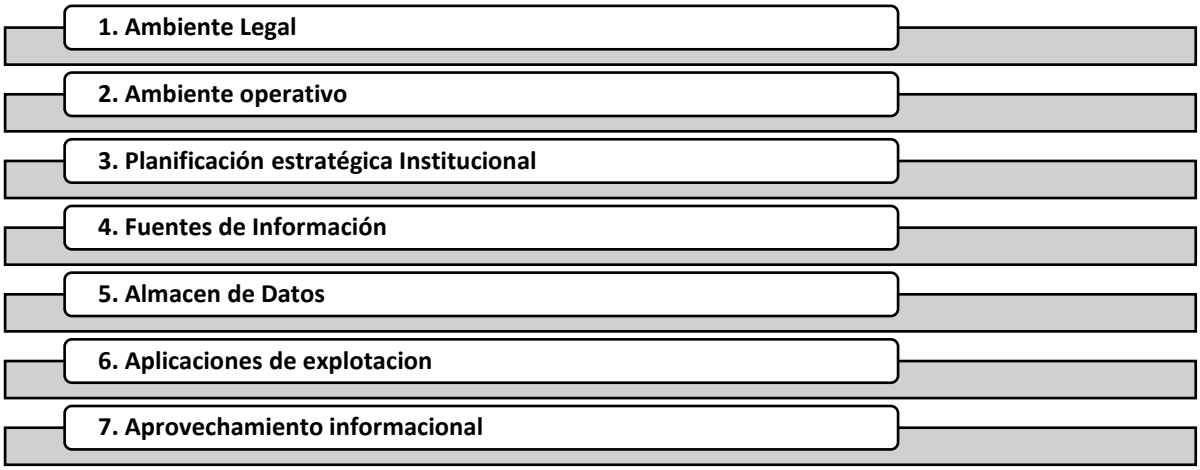
For the development of this research, a theoretical bibliographic analysis of the currently existing data analytics models was carried out. Then, an analysis of the data analytics requirements in the field of health was carried out. With this information, the present proposal for a data analytics model for ADHEALTH health institutions is proposed, which was validated by experts and with statistical analysis, a favorable result and a high trend for its implementation were obtained.

To validate the instrument, a reliability analysis of the survey was carried out, obtaining a Cronbach's alpha coefficient of 0.831, which indicates that the data collection instrument has an acceptable reliability, with which the aspects and dimensions that affect in data analytics in health institutions; he statistical treatment of the data was made using a spreadsheet, and Fisher's test was applied in the ANOVA and Tukey mean separation by analyzing the standard error and obtaining the percentage of good classification model; The corresponding exponential coefficients and confidence intervals were applied to the same 95% of reliability for ANOVA.

4 DATA ANALYTICS MODEL - ADHEALTH

According to the analysis carried out and the requirements in the health area, the following data analytics model is proposed, framed in 7 phases

4.1 Phases of the ADHEALTH model.



Source : The authors

4.2 Operationalization of the ADHEALTH Model

- **Legal Environment.-** Refers to the legislation and its characteristics that the institution has in terms of use and data analytics. Generation, preservation, and analysis of data, for this, at least a) Institutional regulatory framework, b) Current regulations in data analytics, c) Bioethics policies and codes, and an adequate guide of d) Analytical Governance must be available

- **Operating Environment.-** Refers to the functional and operational characteristics of the IT infrastructure of the institution, thus, the physical and technological infrastructure available and/or assigned; the available personnel, their training and skills; the

resources available to operate the infrastructure, comply with institutional regulations, institutional computing objectives, and institutional strategic objectives; analyzes additionally.

- **Institutional Strategic Planning.-** Refers to the planning, execution, monitoring, control, and feedback of institutional objectives based on their strategic vision. The Mission, Vision, Values, Objectives, Goals, Impact Meters, and Risk Analysis should be analyzed.

- **Information Sources.-** Refers to the information available for an adequate data analysis procedure, you must know and have access to the integrated information systems available authorized in the institution, inventory, and access to own systems, national data of population censuses, and business among others.

- **Data Warehouse .-** Refers to the data storage system that allows supporting data intelligence, business intelligence, and analytics tasks as such. In this phase, the Data Cleansing Rules are defined and applied and the ETL (Extraction, Transformation, and Loading) processes are executed.

- **Exploitation applications.-** Refers to the definition, development, implementation, and implementation of Data Analysis Applications; application of computational statistics, artificial intelligence, and machine learning strategies; consolidation of information from all available and valid sources of information, Data Management processes and tools are determined and applied.

- **Informational use.-** It refers to the last step, once the data is organized, analyzed, and converted into strategic information, decision making is motivated, business opportunities are glimpsed, improvement plans are proposed, systems integration is determined, and an Analytical Strategy is defined, and data preservation is proposed.

5 Validation

To reduce the threat to the validity of the statistical conclusion, the following aspects are considered.

The questionnaires were answered with the same scale, in the same period, and by all stakeholders, without a change in the environment for the time the test was taken. The questionnaires were answered anonymously, however, there is an additional document signed by the actor as proof of application [18].

About the internal threat to the validity, this study not considered the effect of history, maturity, and testing because it does not yet have the model implemented. Moreover, there is no threat of selection bias; in the institution where the survey applies The participants have been selected at random, and not targeting a specific group, and all the players are interested in managing the data analytics and its application in health institutions.

About the external threat to validity, there is no risk of interaction, because the sample is the total population of the institution to which the study is applied.

To validate the instrument, a reliability analysis of the survey was executed, obtaining a Cronbach's alpha coefficient of 0.831, which indicates that the data collection instrument has acceptable reliability, thus determining the aspects and dimensions that affect data analysis in health institutions; he statistical treatment of the data was made using a spreadsheet, and Fisher's test was applied in the ANOVA and Tukey mean separation by analyzing the standard error and obtaining the percentage of good classification model; The corresponding exponential coefficients and confidence intervals were applied to the same 95% of reliability for ANOVA. Obtaining a level of significance less than 0.05, presents the conclusion that the ADHEALTH model is suitable for implementing data analytics processes in health institutions.

6 Conclusions

This paper presents a validated model of data analytics applied to health institutions, to meet the needs of the health area in terms of efficient information management, improvement of administrative management, the study of the behavior of diseases in the population, until accurately diagnosing pathologies to perform timely and accurate interventions on some ailments.

The ADHEALTH model presents seven phases in a specific order, the legal environment as the initial phase, the operational environment focused on health institutions, institutional strategic planning, thus defining the appropriate sources of information and generating the data warehouse, to generate the exploitation applications and finally solve institutional requirements and establish institutional use.

The ADHEALTH Model aims to support the overcoming of the challenges of system integration; data privacy, as well as its ethical and legal handling; of efficient data management.

It is essential to have data analytics models oriented to specific environments for adequate management and treatment of data.

7 Bibliography

[1] NewNet, "GOBIERNO, INTELIGENCIA Y ANALÍTICA DE DATOS," 2021. [Online]. Available: https://www.newnetsa.com/analitica-datos/.

[2] IBM, "IBM," 2020. [Online]. Available: https://www.ibm.com/mx-es/watson-health/learn/healthcare-data-analytics. [Accessed 1 12 2020].

[3] Moghimi, "Incorporating intelligent risk detection to enable superior decision support: The example of orthopaedic surgeries," *Health and Technology*, pp. 33-41, 2012.

[4] K. Giuliano, "Improving patient safety through the use of nursing surveillance," *Biomedical Instrumentation and Technology*, pp. 34-43, 2017.

[5] M. Dios, J. Molina-Pariente, V. Fernandez-Viagas, J. Andrade-Pined and J. Framinan, "A Decision Support System for Operating Room scheduling," *Computers and Industrial Engineering*, pp. 430-443, 2015.

[6] A. Luschi, L. Marzi and R. Miniati, "A custom decision-support information system for structural and technological analysis in healthcare," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013.*, 2013.

[7] Miniati, "Design of a decision support system for preventive maintenance planning in health structures," *Technology and health care: official journal of the European Society for Engineering and Medicine*, vol. 20, no. 3, pp. 205-214, 2012.

[8] F. Molina-Granja, R. Yánez and E. Cabezas, "IT Governance—Models and Application," *New Advances in Information Systems and Technologies*, pp. 467-480, 2016.

[9] Kaplan, "How Should Health Data Be Used? Privacy, Secondary Use, and Big Data Sales," *Cambridge Quarterly of Healthcare Ethics*, vol. 25, no. 2, pp. 312-329, 2016.

[10] F. Molina-Granja, G. Rodríguez, E. Cabezas and R. Lozada, "Implementation of the PREDECI model in the prosecution of Chimborazo in Ecuador: a case study evaluation," *International Journal of Electronic Security and Digital Forensics*, vol. 11, no. 1, pp. 28-45, 2018.

[11] Ishikawa, "Medical big data for research use: Current status and related issues," *Japan Medical Association*, vol. 59, pp. 110-124, 2016.

[12] A. Stylianou and M. Talias, "Big data in healthcare: a discussion on the big challenges," *Health and Technology*, vol. 7, no. 7, pp. 97-107, 2017.

[13] A. Urueña, M. Ballestero and E. Morais, Big Data en salud digital, España: Fundación Vodafone España y Red.es, 2017.

[14] OPS, Revisión de estándares de interoperabilidad para la eSalud en Latinoamerica y el Caribe, Organización Panamericana de la Salud, 2016.

[15] F. Molina-Granja and G. Rodriguez, "The preservation of digital evidence and its admissibility in the court," *International*

Journal of Electronic Security and Digital Forensics, vol. 9, no. 1, pp. 1-18, 2017.

[16] F. Molina-Granja and G. Rodriguez, "Model for digital evidence preservation in criminal research institutions-PREDECI," *International Journal of Electronic Security and Digital Forensics*, vol. 9, no. 2, pp. 150-166, 2017.

[17] Executrain, "Los 4 tipos de Analítica de Datos que Debes Conocer," 24 09 2021. [Online]. Available: https://executrain.com.mx/los-4-tipos-de-analitica-de-datos-que-debes-conocer/.

[18] F. Molina-Granja and G. Rodríguez, "Digital Preservation and Criminal Investigation: A Pending Subject," in *New Contributions in Information Systems and Technologies. Advances in Intelligent Systems and Computing*, vol. 353, Springer, 2015, pp. 299-309.

[19] G. Moreno, "¿Cuáles son los modelos de Data analytic y cómo utilizarlos?," 23 07 2019. [Online]. Available: http://datamall.com.ec/cuales-son-los-modelos-de-data-analytic-y-como-utilizarlos/.

[20] NewNet, "GOBIERNO, INTELIGENCIA Y ANALÍTICA DE DATOS," 2021. [Online]. Available: https://www.newnetsa.com/analitica-datos/.

[21] Immeria, "Web Analytics Maturity Model and Critical Success Factors," 20 11 2008. [Online]. Available: Http://blog.immeria.net/2008/11/web-analytics-maturity-model-and.html.

[22] G. Muñoz, "Modelo de Madurez en Analítica Digital," 2017. [Online]. Available: https://elartedemedir.com/blog/modelo-de-madurez-en-analitica-digital/.

[23] M. Spruit and K. Pietzka, "MD3M: The Master Data Management Maturity Model," *Computers in Human Behavior*, vol. 55, p. 1068–1076, 2015.

[24] C. Dremel, S. Overhage, S. Schlauderer and J. Wulf, "Towards a Capability Model for Big Data Analytics," in *13th International Conference on Wirtschaftsinformatik*, Switzerland, 2017.

[25] K. Gudfinnsson, M. Strand and M. Berndtsson, "Analyzing business intelligence maturity," *Journal of Decision Systems*, pp. 37-54, 2014.

[26] C. Khuen and M. Rehman, "A maturity model for implementation of enterprise business intelligence systems," in *In International Conference on Mobile and Wireless Technology*, 2017.

[27] T. de-Bruin, R. Freeze, U. Kulkarni and M. Rosemann, "Understanding the Main Phases of Developing a Maturity Assessment Model," in *16th Australasian Conference on Information Systems.*, 2005.

[28] J. Becker, R. Knackstedt and J. Pöppelbuß, "Developing maturity models for IT management.," *Business & Information Systems Engineering*, vol. 1, no. 3, pp. 213-222, 2009.

[29] Comuzzi, M;Patel, A., "How organisations leverage big data: A maturity model.," *Industrial Management & Data Systems*, vol. 116, no. 8, 2016.