Business Intelligence Platform for COVID-19 Monitoring: A Case Study

Ricardo Duarte[1[0000-0001-6423-3442]], João Lopes[1[0000-0001-7854-8293]], Tiago Guimarães[1[0000-0001-6320-8878]], Susana Ferreira[2], Manuel Filipe Santos[1[0000-0002-5441-3316]]

[1] ALGORITMI/LASI, University of Minho, Braga, Portugal
b12723@algoritmi.uminho.pt, jlopes@dsi.uminho.pt,
tsg@dsi.uminho.pt, mfs@dsi.uminho.pt
[2] *Centro Hospitalar Universitário do Porto, Portugal*
sferreira@chporto.min-saude.pt

Abstract. With the emergence of the COVID-19 virus, the need to effectively and flexibly plan and manage the measures to be applied within a Healthcare institution has become imperative. New knowledge about the disease and legislative changes arose frequently and with high impacts on clinical practice and management.

The platform developed under this case study aggregates data from multiple sources within the Healthcare Institution and also from external entities, such as clinical analysis laboratories. Thus, it provides a set of dashboards that seeks to present in a simple and intuitive way, data analysis ranging from COVID tests and their results, bed occupancy, with an exceptional focus in occupancy in Intensive Care Units, to monitoring the positive movements of COVID patients in the *Centro Hospitalar e Universitário do Porto* (CHUP), in almost real time.

The development of this data collection, analysis and demonstration platform was created in response to a critical need for reliable information so that clinical and management decisions could be supported by solid foundations based on facts and current and reliable data.

Keywords: Business Intelligence; COVID-19; Healthcare; Monitoring; ICU; Intensive Care

1 Introduction

The unprecedented global spread of SARS-COV-2, a virus that has become better known as COVID-19, has placed severe strains on hospital systems worldwide [1].

The unknown clinical features, together with the speed of viral spread, create a difficult situation in which healthcare professionals lack important diagnostic tools, such as accurate predictive models and data-based information on disease progression factors.

Intensive Care Units (ICU) became bigger and additional were created within each hospital so to respond to all needs.

Business Intelligence (BI) and Business Analytics (BA) technologies are increasingly being used to extract knowledge and turn it into useful information and analytical tools for healthcare professionals. These data-based tools are not only essential for healthcare providers in terms of risk prediction, but also play an essential role for policy makers in terms of resource allocation strategy, as they would be an asset for population distribution and vaccination planning. Such tools could help in identifying high-risk patients for first-line vaccines, in addition to priority groups and cancellations or rescheduling as they arise [2].

This document describes a BI system developed for the *Centro Hospitalar e Universitário do Porto* (CHUP), a Hospital in the north of Portugal, to allow health professionals the extraction of knowledge through data analysis, serving as support in the decision-making process. To obtain the solution, it was necessary to perform all the tasks associated with BA projects, to create dashboards with efficiency and quality indicators for COVID patients' management.

This work is divided into several sections, starting with a brief introduction. Section 2 presents a review of the literature on COVID-19 combined with BI. Sections 3 and 4 present the methodology used as well as the developed proposal, respectively. Finally, the final conclusions are presented.

2 Background

The COVID-19 outbreak started in November 2019 in Wuhan, China. In March 2020, the World Health Organization (WHO) declared a global pandemic. The COVID-19 virus is extremely infectious and has spread rapidly to every country in the world [3, 4]. Infection rates reached record levels and led to a high death rate. Countries have begun imposing lockdowns and curfews for the first time in decades. WHO has started campaigns to educate people about the importance of social distancing and wearing masks in public places. In parallel, scientists began to study the virus and develop accelerated vaccines. To date, there are six vaccines authorized by the World Health Organization, namely Pfizer-BioNTech, Moderna, Johnson & Johnson, Astrazeneca, Sputnik and Sinopharm [3].

For health organizations not to collapse and to be able to cope with this outbreak, it is necessary to improve the quality of health care provision. With a good understanding of the generated data, it directly reflects the success of the organization, because a good understanding of these, transforms them into useful information capable of improving its entire environment. When dealing with a huge amount of data, the task of understanding it without the use of technology becomes difficult. This is where BA tools come into play, as they offer a process capable of transforming raw data into useful information for the organization. In addition to providing an interactive and intuitive visualization of the most relevant information, facilitating its analysis and understanding. Consequently, it leads to an improvement in decision-making processes, executing decisions more thoughtful and efficient [5].

BI applications have been used to respond to existing needs in the industry. Governments and healthcare organizations should appreciate the use of these technologies to assist in tracking, monitoring, and controlling the pandemic, which includes the patient flow and vaccination process. The implementation of this system facilitates real-time communication, identifies any problems that may arise, and allows for better efficiency in the allocation of limited resources.

An intelligent system based on global knowledge must be incorporated to lead to a reliable vaccine delivery system. It follows the vaccine supply chain process from start to finish, that is, at all stages of the process. This includes the delivery process, and all the means used to follow up and monitor the entire process. In this way, assistance in scheduling vaccination appointments is improved, in addition to improved tracking and efficiency of vaccinations and vaccinated persons [3]. Control panels and charts are also essential and necessary for their implementation, as they are used to monitoring results and look for changes in processes. Applied to healthcare, control charts can be used to monitor outcomes such as patient status. Predicting disease transmission during an epidemic is a significant aspect of health management, as it helps to prevent and control disease transmission. Real-time climate data with the help of sensors, computing technologies and artificial intelligence were used to develop a big data-based surveillance system to monitor and control epidemics. The use of structured and unstructured hospital data, together with regional data, allows the construction of a strong forecasting model to help control the disease outbreak [6, 7].

3 Methodology

For the development of this solution, the Kimball methodology was used. The Ralph Kimball's methodological approach focuses on guiding the realization and development of projects with a Data Warehouse (DW), focusing on the needs and the data presented to users. Each implementation project must have a finite cycle with a specific beginning and end. A project management is carried out simultaneously, to follow the entire evolution of the solution, deadlines and duration [8]. Figure 1 shows all the phases of the Kimball's Methodology. The most important are described in the following chapter.

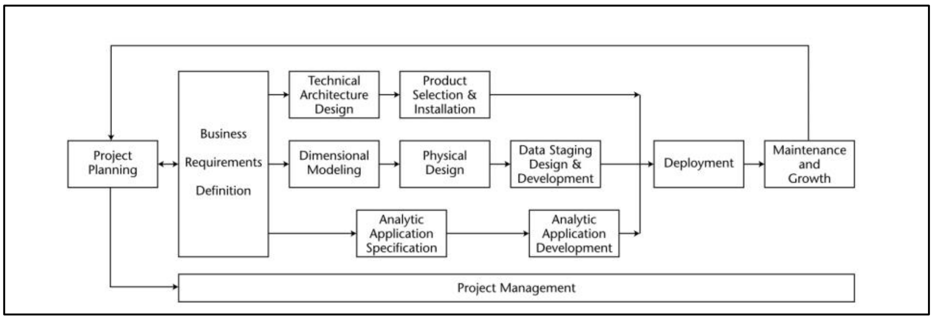

Fig. 1. Kimball's Methodology. Retrieved from [8].

4 Business Analytics Platform

Throughout this chapter, the phases of the Kimball methodology used to achieve the developed solution will be presented and described.

4.1 Project Planning

The project begins with the identification of development tasks to ensure its correct planning, meeting the defined deadlines. This product aims to develop a BI system to support health professionals at the CHUP, so that, through the visualisation of information and indicators, it allows better management of processes inherent to COVID-19 and all treatments to be provided to patients, accessing data with simplified and optimised visualisations to make decisions easier and more intuitive. Initially, the necessary tasks were identified and programmed, from the study of the available data to the design of solution, obtained by studying organisational requirements, understanding the data, data processing, ETL process and building the BI solution.

4.2 Business Requirements

The business requirements were defined, in order to identify the current needs that the CHUP has in the process of performance indicator management. Therefore, a study was carried out to determine what needs to be developed to meet the organisation's needs, being divided into functional and non-functional. These are:

1. Functional Requirements: Understand and implement performance indicators, such as the number of patients admitted, with clinical discharge, number of tests performed, flow of admissions over time, among others;
2. Non-functional Requirements: Intuitive solution, easy to interpret and access.

4.3 Technological Architecture

The technological architecture defined for the design of the final solution is shown in Fig. 1. The image illustrates all the technologies used and the process flow, beginning with the collection of information from different data sources, proceeding to the ETL process, where the data will be processed and loaded into a DW. When the data relates to the analysis tool, a visualization environment is created, which will be made available to the end user in the form of relevant information. Both areas developed are incorporated into the AIDABI application, a platform that, through the management of profiles and accesses, allows greater privacy and security of the information incorporated. Only users with appropriate access and permissions can access the information provided.

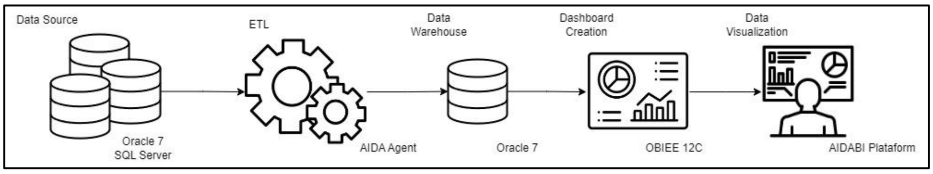

Fig. 2. Technological Architecture

4.4 Dimensional Modelling

Based on the existing data source, several relevant tables were selected to create a DW with the necessary dimensions and fact tables. The main objective of the fact tables is to manage the number of events regarding patients, collection results and hospitalisation flow. The dimensions contain all the important information to complement the existing information in the fact tables, highlighting patient demographics, tests performed, internal transfers, dates of occurrence and deaths. Before being loaded into the DW, all of them go through a data processing process, explained in the following sub-chapter.

4.5 Data ETL Design

When performing the analysis of the attributes that make up the different selected tables, some inconsistencies were identified, such as null values, excessive spacing between characters and spelling errors. The way to solve these problems is to replace the values by 0, eliminating the unnecessary spaces and replacing the errors by the correct nomenclature, respectively. All data processing operations are performed whenever new data is entered into the DW to maintain consistency in the type of data used in each of the dimensions and fact tables.

4.6 Analytical Application Development

The data presentation was carried out through dashboards in the Oracle BI tool. The solution is inserted and functioning at CHUP. It is a solution divided into two main components: Global data analysis and User interaction data analysis. In both, some additional columns and tables, metrics and filters were developed to control the visualizations and allow more robust and, consequently, more intuitive analysis. In the Global Data Analysis component, a series of dashboards were developed, each with its own specificities and contributions to the monitoring of COVID-19 within the institution. Each of the dashboards created are presented below, but only shown where data masking was reasonable:

1. **Synopsis** - Contains an overview of the evolution of the patient's status over time, by age group and condition. Additionally, it is possible to observe the status of hospitalizations and the evolution over time, stressing the number within ICU. The temporal evolution of accumulated cases per status is also shown, in addition to the number of collections made in the last 24 hours.

2. **Occupancy and Hospitalization** - Displays the number of vacancies available for occupancy, the number of occupied beds and the current occupancy in the different hospital units. A special focus was given to ICU so to easily identify the number of occupied beds and vacancies within these units.

Fig. 3. Dashboard for data masked Occupancy and Hospitalization

3. **Recovered** - Contains a detailed list of patients who have been flagged as recovered, according to the criteria defined by the *Direção Geral de Saúde* (DGS).
4. **Deaths** - Displays a detailed listing of patients who tested positive for COVID-19, indicating the date on which their death was declared and the Length of stay for either regular internment, intensive care units or domicile.
5. **Clinical Analytics** - Contains the number of tests that each patient has performed. For each test performed it is possible to determine the module in which it was requested, the request date, the collection date and the result (which varies between Pending and the date in which it became available, a color being assigned to each of the states obtained).

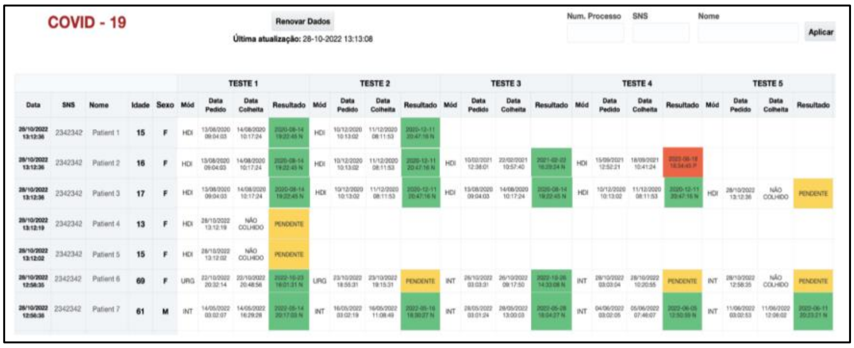

Fig. 4. Dashboard for data masked Clinical Analytics

6. **Patient Flow** - Shows the patient's evolution within the institution, i.e. from the moment they enter the CHUP until they leave. It is possible to see the specialties in which he/she has been, the place to which they belong, exams performed, admission and discharge date, as well as the episode registered.

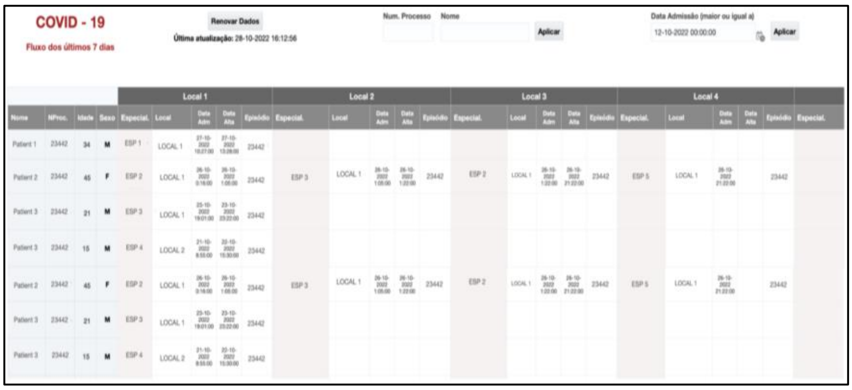

Fig. 5. Dashboard for data masked Patient Flow

7. **Geo** - Displays a geographical distribution by district of the number of positive cases, suspects and those awaiting the test result.

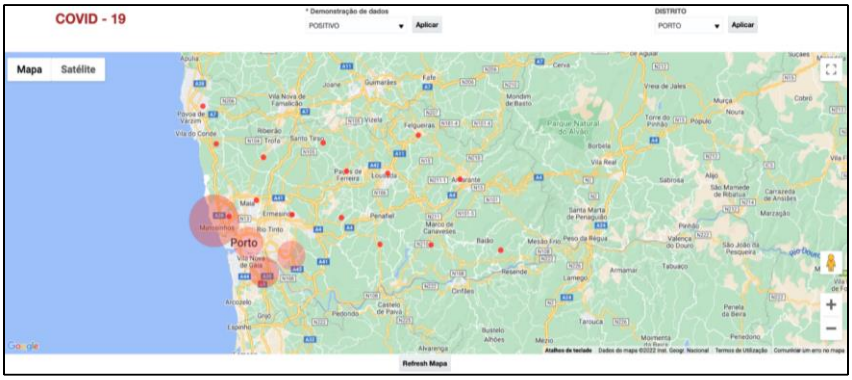

Fig. 6. Dashboard for data masked Geographical distribution

A series of dashboards were developed in the User interaction data analysis component, to maintain monitoring between the health institution and the patient about their status in relation to COVID-19. The dashboards developed are presented below:

1. **Webapp** - Some COVID positive patients whose symptoms were mild, were referred to confinement at home. In order to maintain a line of communication and constant monitoring of their symptoms, a web application was developed so that those patients could send daily updates of their health condition. Within

this scope of data analysis, a dashboard was developed to administer and construe the data sent by the patients.

Fig. 7. Dashboard for data masked Webapp

2. **SMS monitoring panel** - CHUP sends a message to the patient about the result of the test performed within the institution and what to do if the test is positive. In the panel developed for this monitoring, it is possible for the health institution to know the number of messages sent, the number of negative and positive tests (both for adults and Paediatrics) as well as the content of the message sent to each user.

Fig. 8. Dashboard for data masked SMS monitoring

4.7 Implementation and Growth

The implementation of the final solution was carried out in different phases, initially ensuring the correct functioning of the technological structure implemented, followed by each of the dashboards prepared, allowing the validation of all information. Thus, in general, this process was divided into three different phases:

1. Phase 1: Testing of the prototype, to ensure proper functioning and integrity between the different functionalities;
2. Phase 2: Implementation of the final solution, with validation of the data and information presented;
3. Phase 3: Training for health professionals, encompassing the interaction with the platform, allowing them to learn and understand all the available information. It is necessary to guarantee the continuous growth of the developed solution with the appearance of new data and processes of the CHUP.

5 Conclusion and Future Work

The main goal of this project was to develop a BI system that would allow health professionals to extract knowledge through data analysis. The visualization interface presents a set of reports that have metrics capable of responding to the objectives and needs of the institution, with insights being provided in real time.

Areas like Intensive Care were a major focus within this conceptualization and implementation given the urgent need to monitor and optimize its usage and efficiency.

The pandemic forced a complete restructuring of several clinical processes, none of which could be done without access to important information for decision making. Thus, it is considered that the main contribution of this work comes from the logical capabilities of this platform, which is able to monitor, in real time, all the events inherent to COVID-19, with a functional solution implemented in the Health institution itself. Additionally, the research team was concerned with developing a solution capable of being optimized and adapted to other healthcare units, to improve the quality of care and user satisfaction.

As future work, new areas of information are being prepared, with the integration of new data sources, capable of including predictive analyses, to anticipate certain scenarios, such as the limit of admissions, resources, among others.

Acknowledgments

The work has been supported by FCT – Fundação para a Ciência e Tecnologia within the Projects Scope: DSAIPA/DS/0084/2018

References

1. J. L. Izquierdo, J. Ancochea, and J. B. Soriano, “Clinical Characteristics and Prognostic Factors for Intensive Care Unit Admission of Patients with COVID-19: Retrospective Study Using Machine Learning and Natural Language Processing,” J. Med. Internet Res., vol. 22, no. 10, pp. 1–13, 2020, doi: 10.2196/21801.
2. E. Jimenez-Solem et al., “Developing and validating COVID-19 adverse outcome risk prediction models from a bi-national European cohort of 5594 patients,” Sci. Rep., vol. 11, no. 1, pp. 1–12, 2021, doi: 10.1038/s41598-021-81844-x.
3. S. Barakat and H. Al-Zagheer, “Business Intelligence application in COVID-19 Vaccine Distribution,” Ann. Rom. Soc. Cell Biol., vol. 25, no. 6, pp. 17973–17980, 2021, [Online]. Available: https://www.annalsofrscb.ro/index.php/journal/article/view/9161
4. S. Sen, S. Saha, S. Chatterjee, S. Mirjalili, and R. Sarkar, “A bi-stage feature selection approach for COVID-19 prediction using chest CT images,” Appl. Intell., vol. 51, no. 12, pp. 8985–9000, 2021, doi: 10.1007/s10489-021-02292-8.
5. Miguel Quintal, Tiago Guimarães, Antonio Abelha and and M. F. Santos, Business Analytics for Social Healthcare Institution, vol. 1161. 2020. doi: 10.1007/978-3-030-45697-9.
6. M. J. Ward, K. A. Marsolo, and C. M. Froehle, “Applications of business analytics in healthcare,” Bus. Horiz., vol. 57, no. 5, pp. 571–582, 2014, doi: 10.1016/j.bushor.2014.06.003.
7. S. S. Kamble, A. Gunasekaran, M. Goswami, and J. Manda, “A systematic perspective on the applications of big data analytics in healthcare management,” Int. J. Healthc. Manag., vol. 12, no. 3, pp. 226–240, 2019, doi: 10.1080/20479700.2018.1531606.
8. Kimball, R., & Ross, M. (2011). The data warehouse toolkit: the complete guide to dimensional modeling.