



The Moroccan Health Data Bank: A Proposal for a National Electronic Health System Based on Big Data

Houssam Benbrahim*

*Corresponding Author, Engineering Sciences Laboratory, National School of Applied Sciences, Ibn Tofail University, Kenitra, Morocco. E-mail: houssam.benbrahim@uit.ac.ma

Hanaa Hachimi

Advanced Systems Engineering Laboratory, National School of Applied Sciences, Ibn Tofail University, Kenitra, Morocco. Email: hanaa.hachimi@uit.ac.ma

Aouatif Amine

Engineering Sciences Laboratory, National School of Applied Sciences, Ibn Tofail University, Kenitra, Morocco. E-mail: aouatif.amine@uit.ac.ma

Abstract

This work serves to propose a national electronic health system based on the Big Data approach. First of all, we assessed the practice of health information systems (HIS) in Morocco and their obstacles. We performed a survey that was founded on 24 questions to specify the necessary details on this topic. This study shows that there is a primary need for the establishment of an HIS that facilitates the control, analysis, and management of health data in Morocco. For this reason, we have proposed the implementation of the Moroccan Health Data Bank (MHDB). This system will be based on powerful big data technologies that save, manage, and process health data with greater efficiency. The information present in this proposed system can provide the necessary resources for several actors to exploit this wealth, which is embodied in this massive data. We have developed a general description of the MHDB, its components, its conceptual architecture, and an example of a use case.

Keywords: National Electronic Health System, Big Data, Health Information System, Survey, Moroccan Health Data Bank, MHDB



Introduction

Health information system means any system that serves to capture, store, manage, analyze, control, secure, or transmit health information concerning persons or the actions of communities operating in the health sector. The HIS plays a vital role in diminishing costs and promoting individual and public health outcome measures, as well as scientific research. It also aims to evolve the quality of care services and impact policies and decision-making processes.

Today, the healthcare sector produces a significant amount of data. Usually, this information is large and stored in many different forms and types. Driven by the potential to evolve the level of health care, increase the quality of health services in general, reduce prices, and meet other mandatory necessities, big data technology keeps the promise of supporting a large variety of medical and health care functions, such as population health management, disease surveillance, and clinical decision support. Big data in healthcare refers to data that is so huge that it is difficult to administer with traditional support. It is additionally recognized by the types of information and the speed at which it should be managed. The totality of data in healthcare is related to the data generated by patients as well as clinical data, sensor data, emergency care data, news feeds, social media posts, articles in medical journals, etc.

However, in the current situation, the quality of health services in Morocco suffers from several problems, and the use of HIS is linked to a certain number of constraints. Although multiple projects have been launched to promote and enhance the quality of services in this area, they have not yet reached the aspirations of Moroccan citizens. While the Ministry of Health in Morocco does not have a national HIS that manages patient data and all types of information collected in hospitals, clinics, pharmacies, medical analysis laboratories, medical insurance, etc.

For all these reasons, we first created a survey to evaluate the use of HIS in Morocco. Second, we proposed a national electronic health system in Morocco. The latter will be based on advanced big data technologies and will therefore become the solution to store, manage, govern, and exploit health data in Morocco. This system will be named the Moroccan Health Data Bank (MHDB). We will make the data accessible to several entities interested in this field, such as hospitals, insurance companies, research laboratories, pharmaceutical companies, and others, in order to produce new ideas, new methods, and new approaches in the interest of patients while preserving their anonymity. In this paper, we first define the context and profile of new technologies in the health area. Afterwards, we present the study of the survey carried out on the internet and the results obtained in order to accept or reject certain hypothesis. Finally, we will establish the architectural framework of the proposed system.

Literature Review

Big data definition

The term "big data" refers to a significant amount of structured, semi-structured, and unstructured data that is too big and challenging to process using relational databases and traditional software techniques (Gantz et al., 2011). In most enterprises, the volume of data is too large, it moves too fast, or it exceeds current processing capacity. Big data technologies have the potential to help companies improve operations and make faster, more intelligent decisions. This data, when captured, formatted, manipulated, stored, and analyzed, can help a company gain useful insight to raise revenues, attract or retain customers, and enhance operations (Muni, 2014). In the early 2000s, and more specifically in 2001, Meta Group (now Gartner) analyst Doug Laney noted in a research report that there are three challenges and opportunities for data growth: V³: Volume, Velocity and Variety (Laney, 2001; Sagiroglu & Sinanc, 2013). Some authors have highlighted the characteristics of Big Data in 5 V (Anuradha, 2015) and others have proposed 7 dimensions (Hassan et al., 2019).

Supported Technologies for big data

The big data technological environment is based on very advanced tools, among them the two frameworks Apache Hadoop and Apache Spark. The most important big data platform is Apache Hadoop. It is Java-based open-source software that supports storage with the HDFS model and massive data processing with the MapReduce model for version 1 of Hadoop (Hassan et al., 2019). All work is performed on a cluster that contains one or more machines. Since version 2, the Hadoop framework has adopted the third module, YARN, a distributed resource management tool. Hadoop has the capacity to process extraordinarily massive amounts of information with different structures. It provides an ideal environment for exploiting data that is difficult to manage and analyze. Apache Spark is an open-source data processing infrastructure characterized by its speed, and sophisticated analytics, and ease of use. Spark runs in memory on clusters; it is not tied to Hadoop's two-step MapReduce paradigm; and it has lightning-fast performance. Spark can run on top of Hadoop YARN, Mesos, in the cloud, or standalone, where it can read data directly from HDFS, Hbase, Cassandra, Hive, and Tachyon. In addition to its in-memory processing, graphics processing, and machine learning, Spark can also handle streaming. Spark uses resilient distributed data sets (RDD) to store data in memory (RAM). Users can read data from disk and write to RDDs to create a job with the iterative operation, as well as run multiple queries on the same subset of data while continuously keeping it in memory with the interactive mode. Regarding batch processing, Spark offers incredible speed advantages, which can reduce memory usage. Spark streaming is a good stream processing solution for workloads that value throughput over latency. To solve all the problems encountered in the big data era, a set of tools has been grouped together under the name of the Hadoop ecosystem. The Hadoop ecosystem offers a

collection of technologies particularly dedicated to simplifying the creation, deployment, and support of big data solutions.

Context and profile of health in Morocco

The use of information and communications technology (ICT) for health is rapidly transforming the delivery of health services and systems globally. In this context, the health sector in Morocco has multiple achievements to take into consideration and is also suffering from several problems. The implementation of big data techniques and technology in this area in Morocco will undoubtedly aid in solving the majority of its problems and improving its accomplishments (Benbrahim et al., 2019)

Achievements

The health field in Morocco undergoes major changes to meet the increased needs of patients and professionals. In this regard, it appears that IT constitutes a strategic component in the development of hospital services in Morocco. For this reason, several initiatives have been launched to reform the health sector in Morocco, for example:

- The project to set up a hospital information system at the new Fez University Hospital (Berraho et al., 2006). This project was adopted in 4 years and was a part of an integrated action between the University of Sidi Mohamed Ben in Fez and the University of Valenciennes in France.
- The information system of the Sheikh Zaid Hospital in Rabat (Makhloufi et al., 2009) It's a new integrated solution (Sivsa / Hosix) that has been applied and which replaces computerized maintenance management, the economic and financial management system.
- The computerization project of the kingdom's military hospitals in collaboration between the Royal Armed Forces health service and the leading French medical software publisher MedaSys. It is intended to support the deployment of local projects and support institutions in the country in their clinical computerization (Bouhriz et al., 2010)
- The IbnSina University Hospital Center (CHUIS) was part of a process of computerization and modernization of its system medical information (SIM). The expected goal was to have a single patient-centered medical record that is instantly shared among multiple stakeholders to unify the data collection system and reduce redundancy of information and administrative procedures. The duration of the project was 30 months (Xu et al., 2004)

Considering the demands and requirements necessary for the growth of Morocco's health sector, all of these projects are very modest.

Problems

The health sector in Morocco suffers from multiple issues, as we find:

- The Moroccan health system is ranked 91st in the world, 7th in Africa, and 10th in the Arab world by Numbeo, an independent Serbian survey center (Kumar et al., 2022).
- Morocco achieves a low percentage in the use of big data in the health sector, according to the World Health Organization (WHO) report, which is based on the results of the 2015 global survey, the global observatory for eHealth, and the use of eHealth to support universal health coverage (LK et al., 2021).
- Accreditation with a high rate of manual management of the various supports in medical organizations in Morocco. This problem causes losses of information, redundancy, as well as great complexity in terms of data analysis (WHO, 2016).
- So far in 2023, Morocco doesn't have a national EHR system. There is a primary and immediate need for this type of system by patients as well as all stakeholders in the health sector.
- Morocco's Ministry of Health does not devote a large budget to investing in new information technologies; it gives priority to both human resources and medical devices (Benbrahim et al., 2018)

For all these reasons declared at the top, we can say that Morocco has a primordial need for big data technologies in the health sector to make this sector more active and dynamic. The implementation of a national HIS based on the context of big data offers a real-world work environment. In the following part, we will propose our proposed system, its definition, its general structure, its components, its architecture, and its working methodology.

Methodology

To assess the use of EHR in Morocco, we published a survey via the Internet. Accordingly, we utilized 27 items in a survey that 199 people responded to. This research was conducted with several choices that must be ticked to avoid the loss of data. The survey was created using Google Forms and distributed on social networks like LinkedIn and Facebook, and we sent it by email. The target population consists of patients, doctors, nurses, employees of pharmaceutical companies, medical analysis laboratories, scientific research laboratories, health insurance, and the administrative staff of the health services in Morocco. We were able to get the essential data about the survey respondents, including their age, gender, and status, as well as their responses to the questions. The majority of questions had a yes or no response; for other questions, we selected the grades excellent, very good, good, fair, and poor. In another question, words like "no change," "evolving," and "getting worse" were used. The participants also had a choice for one question: lack of money, lack of funding, lack of

regulations and standards, or no idea. In this study, we asked 24 questions (Dyu et al., 2018) about the surveyed population. Figure 1 illustrates these questions.

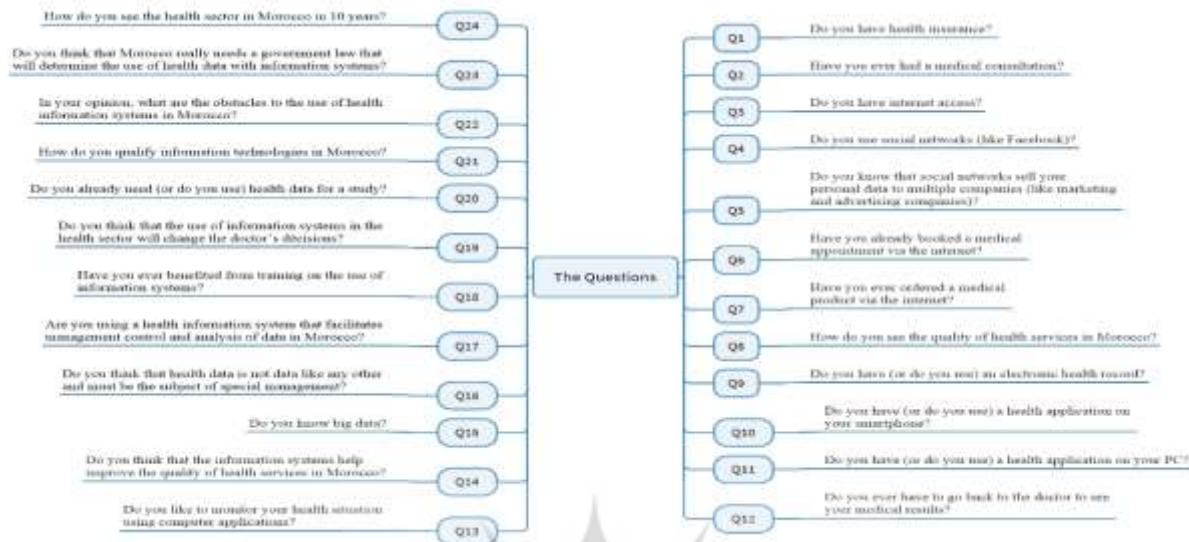


Figure 1. The 24 questions of the survey

To analyze the results of this survey, all collected data was saved in a CSV file, and we used SPSS software for Windows version 20. We employed this research to find out the current situation regarding the use of health services and, more particularly, the use of EHR in Morocco. Additionally, the purpose of this study is to identify the primary restrictions related to this topic. To solve the problems related to this subject, we provide several solutions to be implemented to enhance Morocco's health care system (Ahmed et al., 2022). In the next section of this paper, we will present the proposed system, its definition, its general structure, its components, its architecture, its working methodology, and a use case for the MHDB.

Results and Discussion

Questions and hypothesis

The deductive aspect of the study carried out allows us to ask a priori research questions and hypotheses to validate the results collected from the survey. The main research questions are:

- RQ1: How do study participants perceive the health situation in Morocco?
- RQ2: How do study participants see the need for an HIS that makes it easier to control and analyze health data in Morocco?
- QR3: What are the main obstacles to the use of HIS in Morocco?
- We proposed 12 hypotheses to analyze the previous research questions. Table 1 details the different hypotheses of the survey.

Table 1. Survey hypothesis

Hypothesis	Signification
H1	The current state of the quality of health services in Morocco is influenced by the monitoring of health status utilizing computer applications.
H2	The level of qualification for HIS in Morocco is correlated with the quality of health services.
H3	The decision of doctors to use HIS improves the current situation of the quality of health services in Morocco.
H4	The qualification of computer technologies is linked to the current situation of the quality of health services in Morocco.
H5	The current situation of the quality of health services in Morocco influences the future situation of this sector in 10 years.
H6	The need for an HIS that facilitates the control, management, and analysis of health data in Morocco depends on the status of the study participants.
H7	People who return to seeing doctors believe that the HIS facilitates the management of the health sector in Morocco.
H8	Obstacles to the use of HIS are linked to the quality of health services in Morocco.
H9	Obstacles to the use of HIS influence doctors' decisions.
H10	The use of health data in information systems is linked to a government law, which creates barriers to its usage.
H11	Obstacles to the use of HIS will influence the future situation of this sector in 10 years.
H12	The need for health data for a study is related to a government law that will determine the use of data in HIS.

To test the different hypotheses adopted in this subsection, we used a bivariate correlation analysis with SPSS, to describe and summarize the data collected during this study.

Hypothesis testing

In this subsection, we try to answer the initial questions, namely QR1, QR2, and RQ3, which are linked to the survey carried out in this chapter. To do this, we tested the bivariate correlation between the questions of the study (Q), which enters into the construction of each hypothesis. The different hypotheses were tested and recorded in Table 2.

Table 2. The result of testing the hypothesis

Hypothesis	Variables	Pearson correlation	sig. (bilateral)	N
H1	Q8 and Q13	,144*	,042	199
H2	Q14 and Q17	,206**	,004	199
H3	Q14 and Q19	,200**	,005	199
H4	Q8 and Q21	-,117	,100	199
H5	Q8 and Q23	-,011	,875	199
H6	Q17 and the status	-,015	,838	199
H7	Q17 and Q12	,156*	,028	199
H8	Q22 and Q8	-,174*	,014	199
H9	Q22 and Q19	-,145*	,041	199
H10	Q22 and Q23	,145*	,041	199
H11	Q22 and Q24	,209**	,003	199
H12	Q20 et Q23	,056	,433	199

* The correlation is significant at the 0.05 level.

** The correlation is significant at the 0.01 level (two-sided).

To accept or reject the hypotheses, we follow the following rules:

- Null hypothesis: The two variables are independent in the sample studied. In other words, there is no relationship between the two variables ($Sig. > 0.05$).
- Alternative hypotheses: The two variables are not independent in the sample studied, which means that there is a relationship between the two variables ($Sig. < 0.05$).

From the results obtained in Table 2, it can be deduced that:

- For H1: there is a relationship between Q8 & Q13. The nature of the relationship between the two variables is positive. Therefore, we accept hypothesis H1 (Pearson correlation = 0.144 and Sig. = 0.042).
- For H2: Pearson correlation (Q14 & Q17) = 0.206 and Sig. = 0.004, so it is bilaterally highly significant at the threshold of 0.001 and thereafter we can accept hypothesis H2.
- For H3: we have a positive correlation between Q14 & Q19 (Pearson correlation = 0.2 and Sig. = 0.005). We conclude that H3 is an alternative hypothesis.
- For H4: the Pearson correlation = -0.117 and Sig. = 0.100. There is no relationship between Q8 & Q21, so we reject hypothesis H4.
- Similarly, we reject the two hypotheses H5 (no relationship between Q8 & Q23) and H6 (the correlation between Q17 and the status is not significant at the threshold of 0.05).
- For H7, H8, H9, H10, and H11: We accept the five hypotheses because the nature of the relation that links Q17 & Q12, Q22 & Q23, as well as Q22 & Q24, is positively correlated; likewise, the relation that links Q22 & Q8, as well as Q22 & Q19, is negatively correlated.
- For H12: the Person correlation (Q20 & Q23) = 0.05 and Sig. = 0.433, so there is no relationship between Q20 & Q23, and subsequently, we cannot accept hypothesis H12.

According to the tests carried out, we can conclude that the health situation in Morocco is really linked to the use of health information systems and their obstacles. Thus, the participants of the survey think that the need for a health information system in Morocco will modify the decisions of doctors, improve the quality of health services, and influence the management of this sector. Finally, the obstacles to the use of information systems in the health sector in Morocco are really related to the lack of a governmental law that will determine the use of health data, thus influencing the decisions of doctors and the future situation of this sector in Morocco.

Proposed solution

The Moroccan Health Data Bank (MHDB)

The MHDB is a suggestion for a national HIS in Morocco; it will be considered as an advanced module to store, manage, analyze, and enhance population health. What we want specifically is to design an architecture that will focus on health data. The MHDB will collect all types of data (structured, semi-structured and unstructured) that are related to patients, such as data from consultations, medical analysis data, medication, as well as the data that have a relation to the field of the health, for example, the scientific research, the statistical results, the causes of deaths and more. The MHDB will manage all forms of data like the reports (text), images, personal data, audios, videos... etc. The MHDB will link:

- The data of the Public and Private Health Insurance: The MHDB will consolidate data from banks and insurance companies. The aim is to exploit the data of their customers while guaranteeing the protection of individual data.
- Data from hospitals, clinics, pharmacies, medical practices: This part conceives the data of the patients who can be in the form of the medical reports (the text written by the doctors, the commentaries, the remarks and more), also the X-rays images, medical scanners, ultrasounds and all other materials of imageries, as well as the results of medical analyzes, prescription drugs, and many other things. It means detailed data about the institutions in question and the benefits they provide to patients.
- Data from research scientific, pharmaceutical laboratory, publications: The goal is to study and analyze the practice of the new scientific methods of treatments for example, or new pharmaceuticals launched in the market to inspire and get the best. It will be useful information to treat better or cheaper. But also, for pharmaceutical producers or medical equipment to better profile their marketing.
- Medical statistics: The MHDB will be responsible for annually compiling the national statistics of medical causes of death, medical errors in hospitals, disability data, and any other health-related statistical data.
- Web, social networking and connected objects data: MHDB is also interested in the data of people who publish or have published one or more elements concerning their health or that of a loved one on the internet (illness, symptoms, medication, and treatment ...), on social networks, online forums, commenting on an article in an online media, a blog, or patients who use an object connected as a mobile phone with a health application storing data (weight, blood pressure ...).

The second role of the MHDB will be dedicated to data processing and analysis. The BMDS will make health data manageable and usable to promote studies, research, or developments of a public interest nature and to contribute to one of the following objectives:

- Health information: Extract exploitable information that has a good quality value from data collection of different sources and that can change medical practice.
- The implementation of health policies: With the help of the MHDB, health sector leaders in Morocco can take major decisions based on the prevision methods provided by the MHDB system.
- Knowledge of health expenditures: Multiple actors finance health expenditures, including the state, local authorities, and complementary protection organizations (insurance companies, mutuels, and provident institutions). The MHDB will give a well-defined vision on health expenses in Morocco to organize and guide its charges in the right direction.
- Innovation in the fields of health and medico-social care: The MHDB aims to present cutting-edge concepts and technologies that cannot be ignored today in the health area in Morocco. its innovations will significantly improve the care of patients and will disrupt medicine in general.
- Surveillance, watchkeeping, and health security: MHDB will provide surveillance and continuous observation of the health status of the Moroccan population. The system will discover all health risks, including emerging risks, to permanently ensure the health security of the population, which includes reporting to the health authorities of any situation or event likely to threaten public health.

The patient will be at the core of the MHDB. Patient health data will become a gold mine to be used for analytical needs to better understand different aspects of health for better care and better medical treatments. With the MHDB the patient will control the access and the use of his data, and he will get read access to this information or use of the individual health record. This makes patients inside the MHDB ecosystem, to overcome the barriers that many patients currently experience and to encourage open information sharing in a secure environment. Access to data via MHDB should take place under conditions that guarantee data confidentiality and traceability of access. The BMDS specifies its conditions by adopting security rules guaranteeing control over all users. The figure 2 presents a general description of the proposed system.



Figure 2. A general description of the Moroccan Health Data Bank

This gold mine, which is going to be one of the largest medico-administrative databases in the world, comprises two categories of data: aggregated data, processed to obtain anonymous information about groups of individuals with common characteristics, and personal data.

Components of MHDB

Our proposed MHDB system will be composed of six major elements which are: storage, management, processing, governance, operations, and security. The expanded MHDB defines the different health data from several sources, and that is ingested by the authorized health entities in our system, and also the necessary analysis done with this data, Figure 3 defines the components of MHDB.

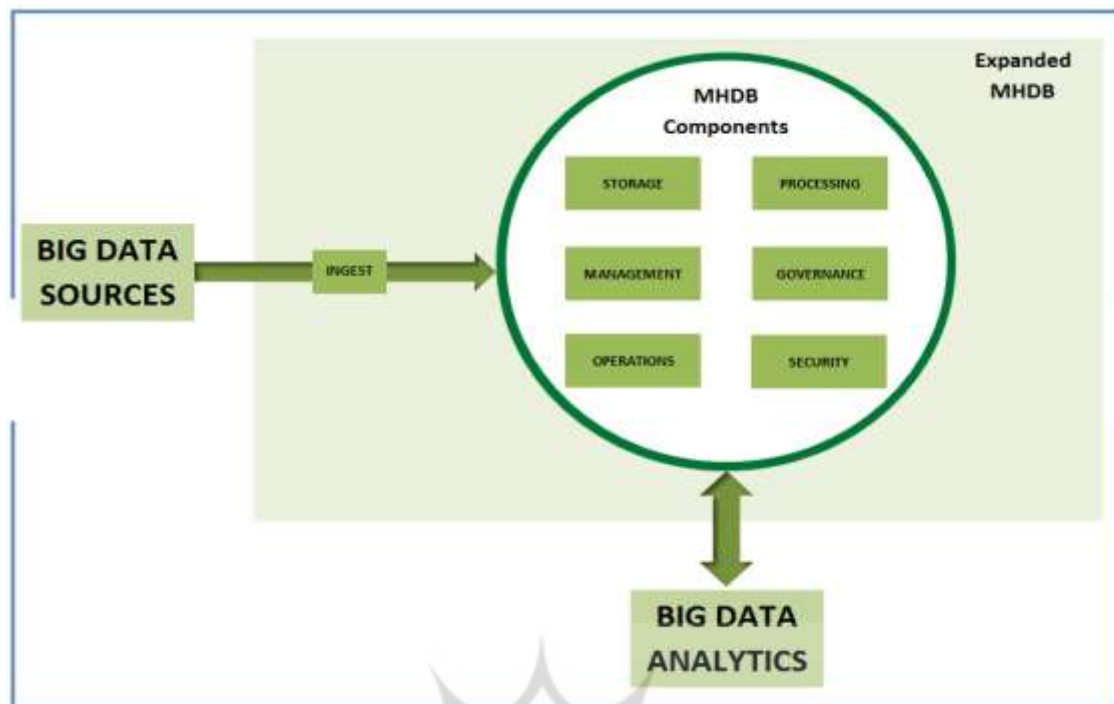


Figure 3. Components of MHDB

- **Storage:** The first component in our system is the health data storage module (different types and forms of data). It will be an effective module built on advanced big data technology. This module will provide scalable, cost-effective, and fault-tolerant storage.
- **Management:** This module will be responsible for the management of health data. It will provide a centralized architecture that can simultaneously handle a variety of workloads.
- **Processing:** MHDB Data Platform will have a large number of processing systems. Entities authorized to access the system can simultaneously use multiple methods of processing the same data.
- **Governance:** MHDB will expand data access and management with highly efficient data governance and integration tools based on Big Data technologies.
- **Operations:** Provisioning, administration, monitoring, and programming within the MHDB. Transparent adaptation to the environment of each entity.
- **Security:** The security of MHDB will be ensured in an integrated way at all levels. The vital features of authentication, authorization, accountability, and data protection will be within our system.

MHDB, ingest data from many sources and that have different types and forms. In the same way, actors of health can make analyzes adequate to their needs, all while keeping advanced security of the patients and recorded data.

Architectural framework

The general framework of a dedicated big data analytics system for healthcare is difficult because big data by definition, aims to capture, store, manage, and analyze huge numbers of data with different types and forms, and the processing is broken down and executed on a cluster that has one or more machines. The working environment of our architecture will be realized in the concept of distributed processing and will be used in the analysis of very large health care datasets, coming from several sources, to exploit them and gain insight into making more informed decisions about health. Also, several open-source platforms that are available such as Hadoop have encouraged the application of Big Data analysis in the health sector. Figure 4 defines a conceptual architecture of the MHDB.

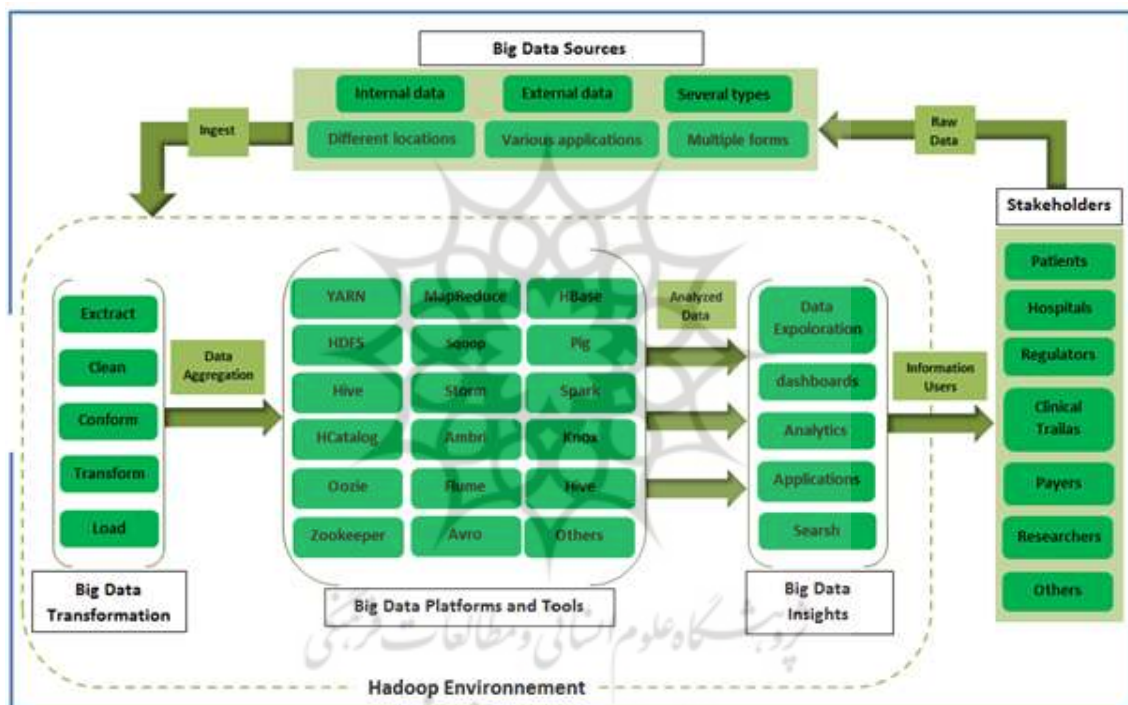


Figure 4. Conceptual architecture of the Moroccan Health Data Bank

Our Moroccan Health Data Bank system is composed of 5 demons that represent the bases necessary for the best functioning and the good progress of the data in a sustainable cycle and producer. The 5 demons are:

The first one is the big data sources, which are:

- Internal data like electronic health records, clinical decision support systems, etc.
- External data, for example, from government sources, laboratories, pharmacies, insurance companies, etc.

- Several types of data, as already defined at the beginning of our article, include Web and social media data, machine-to-machine data, big transaction data, biometric data, and human-generated data.
- Multiple forms: CSV, relational tables, text, etc.
- Different location.
- Various applications such as transaction processing applications, databases, etc.
- The second demon is the big data transformation, which we can define as:
 - Data extraction is the recovery of data from data sources.
 - Data cleaning is the operation of detecting and correcting (or removing) errors present in the data.
 - Conforming means resolving the labeling conflicts between potentially incompatible data sources so that they can be used together in our system.
 - Transforming is converting data from one format or structure into another format or structure.
 - Data loading is where the data is loaded into the final target database.

The third one is the big data platforms and tools: As already mentioned at the beginning, there are several platforms and tools to manage, store, and process massive data, and among these technologies, we find Yarn, MapReduce, HBase, DFS, Sqoop, Pig, Hive, Storm, Spark, HCatalog, Ambri, Knox, Oozie, Flume, Hive, Zookeeper, Avro, and others. Our HMDB architecture will be based on the Hadoop environment to benefit from its abilities and performance.

The fourth module is Big Data Insights. According to the use case, data in healthcare can be processed in batch, interactively, online, or streaming. Once data is stored and processed in Hadoop, it can either be analyzed in the cluster, explored, or exported to relational data stores for analysis there, for example, the enterprise data warehouse, surgical data mart, quality data mart, diagnosis data mart, clinical info data mart, and Neo4j graph database. Multiple data analysis and visualization applications can also work with the data directly in Hadoop, for example, EMR real-time analytics, Microsoft Excel, Metric Insights, RESTful Web Services, Patient Scorecards, Operational Dashboards, Research Portals, and Quality Dashboards.

The last demon is stakeholders. These are the entities that use our architecture for several reasons, either to consult, exploit, or benefit from the resources offered by our system. Among them are patients, hospitals, regulators, clinical, insurance, researchers, and others.

MHDB use cases

The MHDB will bring together vast amounts of medical data, merging the advanced cognitive capabilities of big data analytics with traditional analytics, to turn this wealth of data into knowledge. Take the example of Patient-X. It has a blood pressure condition that could worsen over time. He closely collaborates with his doctor to track the problem, and Patient-X's doctor starts using the MHDB ecosystem. It allows him to register the data generated from the consultation with Patient X and examine the personal and family medical histories provided by other doctors for Patient X. This offers Patient X and his doctor insightful information on his condition over time. The MHDB will operate in an ecosystem environment; it will combine huge amounts of data and knowledge and bring together patients, doctors, researchers, pharmaceutical companies, insurers, etc. on a secure and open platform. As Patient-X and his doctor continued to use the MHDB through web, mobile, or desktop applications, the application contributed data back into the MHDB ecosystem. This will make MHDB more dynamic, intelligent, and efficient for everyone. Using analytics and insights, the application makes the best recommendations for Patient-X's health care. The application alerts Patient-X's doctor and suggests a possible replacement medication. Patient-X will never have to worry about his medical privacy because the MHDB will go to great lengths to keep knowledge about patients private and anonymous. As the MHDB knowledge base continues to develop, this situation will enable health stakeholders to drive medical advancements and improve patient outcomes. Doctors can be more confident and accountable in their treatments, and patients can be more responsible for their health and wellness by taking preventative action. This section can be summarized in Figure 5.

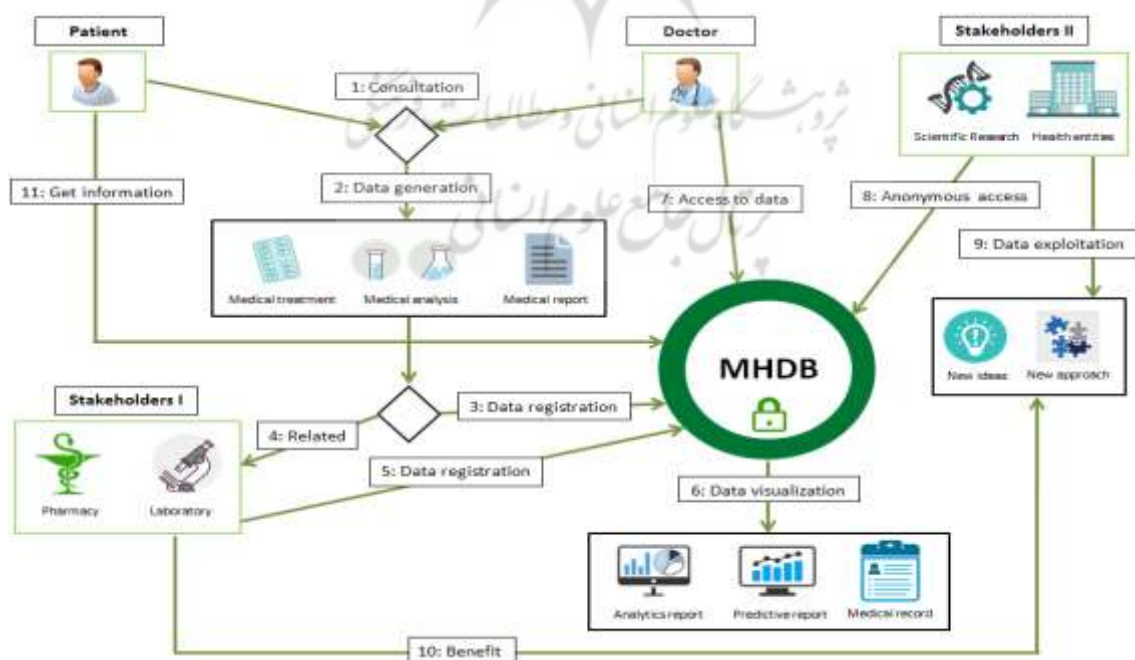


Figure 5. Use case of the Moroccan Health Data Bank

The MHDB will rapidly transform healthcare in Morocco; its advanced capabilities, interactive ecosystem, and secure environment will dramatically change the dynamic between patients, doctors, and the medical community.

In the world, several countries have adopted national EHRs as an innovative, effective, and modern solution to overcome constraints, improve the level of care, and guarantee a better management strategy for this area. They include Denmark, Canada, Australia, Finland, the United States of America, Sweden, the United Kingdom, and others (Louardi, 2012). Taking the following examples:

- Since December 2008, Estonia has been the first country in the world to achieve a national HIS exchange and reference environment, named the Estonian Electronic Health Record System, or "birth-to-death". The object of this project was to ameliorate the quality of health care, improve patient security, make the healthcare system most effective, enhance treatment processes, etc (Aaviksoo et al, 2010).
- In Singapore, since 2011, the National EHR has been adopted by public and private healthcare institutions. Under the leadership of the Ministry of Health and managed by the Integrated Health Information Systems, the system assembles all patient health records from multiple healthcare sources. It gives approved healthcare professionals a general view of their patient's healthcare history (Al-aswad et al., 2013).
- Launched in 2002 by the UK government and officially dismantled in 2011, the National Programme for Information Technology (NPfIT) known as the National Health Service (NHS) Care Records Service. It is a national EHR system that gathers patient records from across the UK (Muttitt et al., 2012).
- On July 1, 2012, the national Personally Controlled Electronic Health Record (PCEHR) system was released in Australia. It is a shared project created by the Australian government, with execution controlled by the National Electronic Health Transition Authority (NEHTA). It integrates information such as current medications, adverse drug reactions, and other data in a simple, readily available format.
- Since 2017, all Norwegian citizens have had a personalized Summary Care Record (SCR), known as the Kjernejournal in Norwegian. It is a new national system that permits all health professionals to select and share patients' health information across all establishments and levels of care in Norway.

Until now, in 2023, there was no national electronic health system in Morocco. All countries that have applied the EHR system have achieved exceptional, important, great, and satisfying performance. With the MHDB, a national system dedicated to the health sector in Morocco, access to data, improving their quality, and the production of new projects related to this sector, etc., will be done in a way that is reliable, easy, and intelligent. Indeed, the

MHDB relies on solid technologies for big data; we can analyze the data in three modes: batch, real-time, and streaming. Which will allow us to follow the health sector in our country with a lot of precision.

Conclusion

In this work, we analyzed the use of EHR in Morocco with a survey published online. Our objective is to collect as much information as possible to assess the current situation, know the different obstacles, and propose innovative applications and solutions for this sector. According to the tests carried out, it can be concluded that most of the participants in this survey are in favor of the urgent creation of an EHR that facilitates the management, control, and analysis of data in Morocco. This requires a very thorough intervention to enhance and strengthen this sector, either by providing the necessary human resources, by significant funding, or by the creation of laws and standards that govern this problem.

For all these reasons, we have proposed a national electronic health system founded on big data. The latter will provide a flexible work environment and be up to the task. The Moroccan Health Data Bank is a national system that will consolidate all health data. Its information can be massive and different. In the MHDB, we will store, manage, and analyze health data based on platforms like Hadoop, Spark, Yarn, MapReduce, HBase, HDFS, Sqoop, etc.

In our system, several active actors are both the sources of the data and the beneficiaries of the information. We will work with entities that are interested in the health sector, such as hospitals, research laboratories, insurance companies, pharmaceutical companies, and others. Its organizations will extract useful information, and using the MHDB, they will benefit from the advanced cognitive capabilities of big data analysis. The aim is to produce new ideas, new approaches, or new methods while preserving the anonymity of patients and improving the quality of health services in Morocco.

Today, it has become indispensable to follow the global technological flow, in particular in the field of health, where the lives of people are at stake. The traditional tools currently used can't guarantee the expectations of Moroccan patients. The adoption of this system and its practice will overturn the health sector in Morocco. The Moroccan Health Data Bank will ensure and guarantee efficiency in terms of quality of care, relevance, and accuracy.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Aaviksoo, A. & Saluse, J. (2010). National Health Information System-Follow up. *Tallinn: PRAXIS Center for Policy Studies*.
- Ahmed, S. T.; Kumar, V. V.; Singh, K. K.; Singh, A.; Muthukumaran, V. & Gupta, D. (2022). 6G enabled federated learning for secure IoMT resource recommendation and propagation analysis. *Computers and Electrical Engineering*, 102, 108210.
- Al-Aswad, A. M.; Brownsell, S.; Palmer, R. & Nichol, J. P. (2013). A review paper of the current status of electronic health records adoption worldwide: the gap between developed and developing countries. *Journal of Health Informatics in Developing Countries*, 7(2).
- Anuradha, J. (2015). A brief introduction on Big Data 5Vs characteristics and Hadoop technology. *Procedia computer science*, 48, 319-324.
- Benbrahim, H.; Hachimi, H. & Amine, A. (2018). Moroccan Electronic Health Record System. *Proceedings of the International Conference on Industrial Engineering and Operations Management Paris, France*.
- Benbrahim, H.; Hachimi, H. & Amine, A. (2019). Survey on the Use of Health Information Systems in Morocco: Current Situation, Obstacles and Proposed Solutions. In *Advanced Intelligent Systems for Sustainable Development (AI2SD'2018) 4: Advanced Intelligent Systems Applied to Health* (pp. 197-204). Springer International Publishing.
- Berraho, M.; Begdouri, A.; Elrhazi, K.; Elfakir, S.; Boumane, A.; Talbi, A. & Nejjari, C. (2006, September). Projet de mise en place d'un Système d'Information Hospitalier au nouveau CHU de Fès (Maroc). In *Quatrième conférence francophone en gestion GISEH-Luxembourg, du* (13).
- Bouhriz, N.; Benabbou, F. & Benlahmer, H. (2015). Text concept extraction based on arabic wordnet and formal concept analysis. *International Journal of Computer Applications*, 111(16), 30-34.
- Centre Hospitalo-Universitaire IbnSina. (2016). *Lancement du projet de mise en place du Système d'Information Hospitalier (SIH) du Centre Hospitalo-Universitaire Ibn Sina*.
- Dyb, K. & Warth, L. L. (2018). The Norwegian National Summary Care Record: a qualitative analysis of doctors' uses of and trust in shared patient information. *BMC health services research*, 18(1), 1-10.
- Gantz, J. & Reinsel, D. (2011). Extracting value from chaos. *IDC iView*, 1142(2011), 1-12.
- Hassan, M. K.; El Desouky, A. I.; Elghamrawy, S. M. & Sarhan, A. M. (2019). Big data challenges and opportunities in healthcare informatics and smart hospitals. *Security in smart cities: Models, applications, and challenges*, 3-26.
- Kumar, S. S.; Ahmed, S. T.; Xin, Q.; Sandeep, S.; Madheswaran, M. & Basha, S. M. (2022). Unstructured Oncological Image Cluster Identification Using Improved Unsupervised Clustering Techniques. *Computers, Materials & Continua*, 72(1).
- Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META group research note*, 6(70), 1.
- LK, S. S.; Ahmed, S. T.; Anitha, K. & Pushpa, M. K. (2021, November). COVID-19 outbreak based coronary heart diseases (CHD) prediction using SVM and risk factor validation. In *2021*

- Innovations in Power and Advanced Computing Technologies (i-PACT)* (pp. 1-5). IEEE.
- Louardi, E. H. (2012). Stratégie sectorielle 2012-2016 Ministère de la santé. *Maroc Mars*, 71–75.
- Makhloufi, I.; Saadi, J.; Artiba, A.; El Hiki, L. & El Hassani, A. (2009). Système d'information hospitalier de l'Hôpital Cheikh Zaid: État des lieux et schéma directeur d'urbanisation. In *Conférence Internationale de conception et production intégrées CPI*.
- Muni Kumar, N. & Manjula, R. (2014). Role of Big data analytics in rural health care-A step towards svasth bharath. *International Journal of Computer Science and Information Technologies*, 5(6), 7172-7178.
- Muttitt, S.; McKinnon, S. & Rainey, S. (2012). Singapore's National Electronic Health Record (NEHR): The Journey to 2012 and Beyond. *Presentation, Singapore*.
- Sagiroglu, S. & Sinanc, D. (2013, May). Big data: A review. In *2013 international conference on collaboration technologies and systems (CTS)* (pp. 42-47). IEEE.
- Takian, A.; Sheikh, A. & Barber, N. (2012). We are bitter, but we are better off: case study of the implementation of an electronic health record system into a mental health hospital in England. *BMC health services research*, 12(1), 1-13.
- World Health Organization. (2016). *Atlas of EHealth Country Profiles: The Use of EHealth in Support of Universal Health Coverage: Based on the Findings of the Third Global Survey on EHealth 2015* (Vol. 3). World Health Organization.
- Xu, J.; Gao, X.; Sorwar, G. & Croll, P. (2014). Current status, challenges, and outlook of E-Health record systems in Australia. In *Knowledge Engineering and Management: Proceedings of the Seventh International Conference on Intelligent Systems and Knowledge Engineering, Beijing, China, Dec 2012 (ISKE 2012)* (pp. 683-692). Springer Berlin Heidelberg.

Bibliographic information of this paper for citing:

Benbrahim, Houssam; Hachimi, Hanaa & Amine, Aouatif (2024). The Moroccan Health Data Bank: A Proposal for a National Electronic Health System Based on Big Data. *Journal of Information Technology Management*, 16 (1), 79-97. <https://doi.org/10.22059/jitm.2024.96376>
