Role of Big Data Technologies in Water Information System

Hicham Jamil[1], Bouabid El Mansouri[1] and Aniss Moumen[2]
[1]*Natural Resources Geosciences Laboratory, Kenitra Faculty of Sciences, Ibn Tofail University, Morocco*
[2]*Systems Engineering Laboratory, Kenitra ENSA, Ibn Tofail University, Morocco*

Keywords: Information System, Big Data, Water Resources

Abstract: Water management is an essential vector that Morocco has adopted as an integral part of its government policy since the dawn of independence. It also initiated an important project for the collection and analysis of data relating to water resources.

Data and systems (Water Department, 2021). Necessary for water-related problems. The Hydraulic Basin Agencies, which constitute the body producing data relating to water resources, through measurement stations produce hydrological data, which will be consolidated at the Directorate General for Water level, thus making it possible to make decisions. The volume of these data and the diversity of the actors leads us to think of a water information system that integrates the concept of Big data to ensure good governance of these water resources by exploiting various data from several databases.
In this article, we will compare existing architectures of water information systems according to a review of systemic literature and propose an architecture by exploiting Big data technology.

1 INTRODUCTION

Water is considered an essential component of the Moroccan economy. The combination of growth in demand for water, climate, and hydrological gap pushed water resource managers to search for strategies for the management of water resources (Loi, 2016).

So the management of water resources is a strategy initiated by the Kingdom of Morocco since independence.

The Directorate General for Water, relying on the Hydraulic Basin Agencies, is required to measure and collect data relating to water resources, to process, consolidate it, to be able to make decisions, and " ensure perfect governance of water resources.

The Water Act of 1995 contributes to watershed agencies' (ABH) several functions, including collecting data and information on the situation of water resources in a watershed (Loi, 2016).

Since its creation, the hydraulic basin agency has not ceased to achieve strategic objectives by its attributions;

- Operation and maintenance of hydraulic structures, including dams
- water resources management at the hydraulic basin level
- Issue authorizations and concessions for the use of the public hydraulic do-main
- Taking measures relating to water resources
- Develop the Integrated Management Plan for Water Resources (PDAIRE)

The Hydraulic Basin Agency (ABH) produces and manages several Data, Has several applications and databases for data management. By its Mandate to take decisions based on the results of its Information system (Moumen et al., 2015).

Based on this, the Hydraulic Basin Agency plays a vital role in producing data relating to water resources before their processing, modelling, and analysis, allowing decision-making: the part of decision-making information systems in this process is essential.

Also, the volume of data produced at the HBA level makes their exploitation difficult, leading us to think of using the latest technologies relating to big data.

The following points, therefore, represent the problem:

Jamil, H., El Mansouri, B. and Moumen, A.
Role of Big Data Technologies in Water Information System.
DOI: 10.5220/0010735800003101
In *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning (BML 2021)*, pages 419-425
ISBN: 978-989-758-559-3

- Which decision-making information system is adequate for water management
- What are the components of this system?
- How to integrate the notion of big data into this system, and at what level

2 METHODOLOGY

The literature review is an essential step for rich documentation of the research. It is explained as follows:

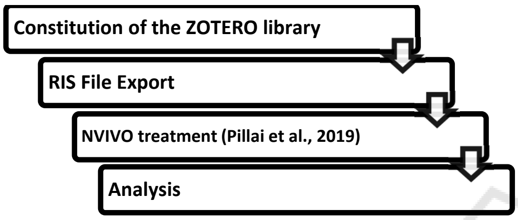

Figure 1: Working Methodology

In the end, we found that 93 references were identified in the various databases and were stored in the Zotero library. This number of references is created using the following method

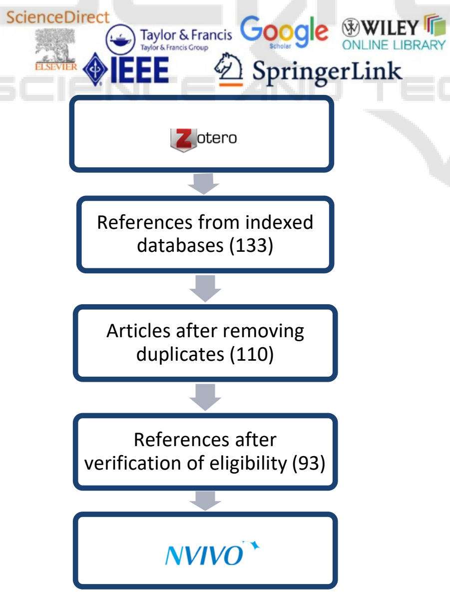

Figure 2: Constitution of the Zotero Library

Based on searches using the keywords used: **Information System, Water Resources, Big Data**.

3 COMPARATIVE STUDY

Systematic Literature Review (SLR) has enabled us to find articles that offer water information systems incorporating the concept of big data. Indeed, several works have dealt with this problem, which shows the interest and importance of this subject, but no universal architecture is final apart from these works. We will try to present some actual results and try to compare them:

The choice of these architectures based on those related to the processing and management of data relating to water resources

3.1 Big Data Open Platform for Water Resources Management

By this architecture, the authors explain how we can exploit Big Data concepts for environmental sciences. (Driss et al., 2015)

The authors, in figure 3, presents the architecture of a big Data Open Platform used for supporting Water Resources Management; the architecture is based on nine blocks:

1. Decision Support Tools: It is a model that allows you to select the best decision module
2. Knowledge-Based System: This block concerns the collection and storage of data relating to water recooling and its exchange with stakeholders
3. Geographic Information System (GIS): Concerns the manipulation and analysis of geographic data
4. Big Data Analysis System: This block relates to big data processing and analysis for various data relating to water resources
5. Simulation Models: This module contains simulation models related to the GIS module
6. Computation and Processing: This part contains calculators and simulators for forecasting water resources
7. Communication System: The purpose of this block is to ensure communication between the different remote or local systems
8. Search Engine: This module ensures an indexed search at the big data level
9. Users Interfaces: Provide the user's capabilities of communicating with the platform

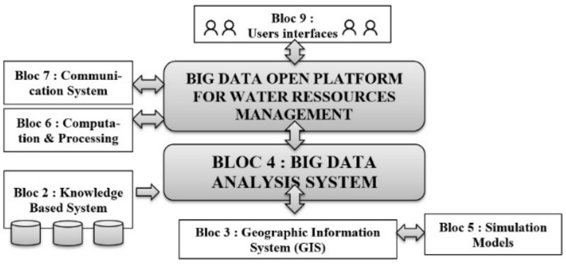

Figure 3: Conceptual architecture of Big Data Open Platform For Water Resources Management (Driss et al., 2015).

3.2 A Framework for Processing Water Resources Big Data and Application

The development of technologies and the addition of several aspects to water resources data makes the dynamic analysis of this data more complex. The authors' object in this paper is to present the application of big data in this process and offers a framework for processing water resources big data and application (Ping and Zhao, 2014).

This framework, presented in Figure 4, mainly consists of four layers:

1. Data acquisition layer:
   - ✓ Collecting sufficient quantity, density, and variety of real-time water resources data
   - ✓ These collected data are divided into structured data, semi-structured and unstructured data
2. Resource organization layer
   - ✓ Organization data
   - ✓ Using SQL and NoSQL tools for data extraction, integration, and transformation, to form the Master database finally
3. Data analysis layer
   - ✓ The core of big data processing
   - ✓ Support the big data analysis and application
4. Application service layer
   - ✓ Based on the data analysis layer
   - ✓ Provide comprehensive information services mainly for Water Re-sources System

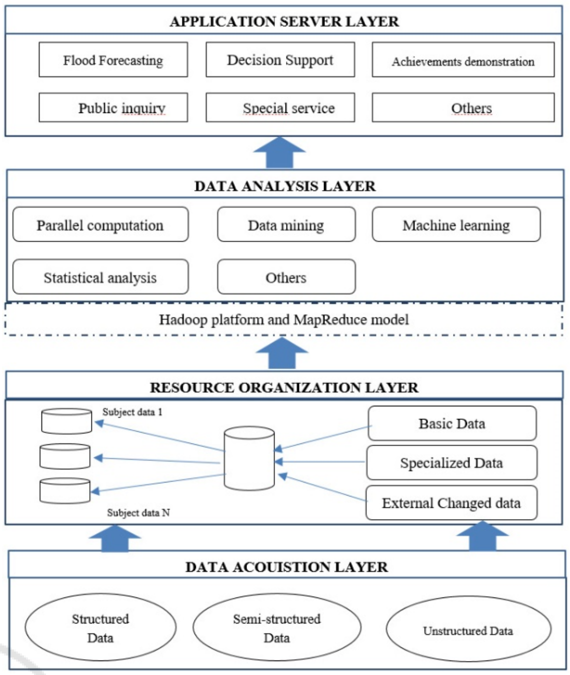

Figure 4: A framework for processing water resources big data (Ping and Zhao, 2014).

3.3 Big Data Technology in Establishment and Amendment of Water Management Standard

Exploiting new technologies for water resources management, the authors build the integrated solution of intelligent water management standards based on big data in this paper.

In Figure 5, the authors present another architectural approach based on the following steps:

1. Primary data resource platform:
   - The data are stored in the backend database of all business systems
   - Support the regular stable operation of the business system
2. Unified data platform:
   - Data resources are managed in a unified way
   - Consolidate the data
   - Data warehousing of intelligent water management standards
3. Data resource utilization :
   - Provides multidimensional data analysis
   - Assisting business decisions and project approvals

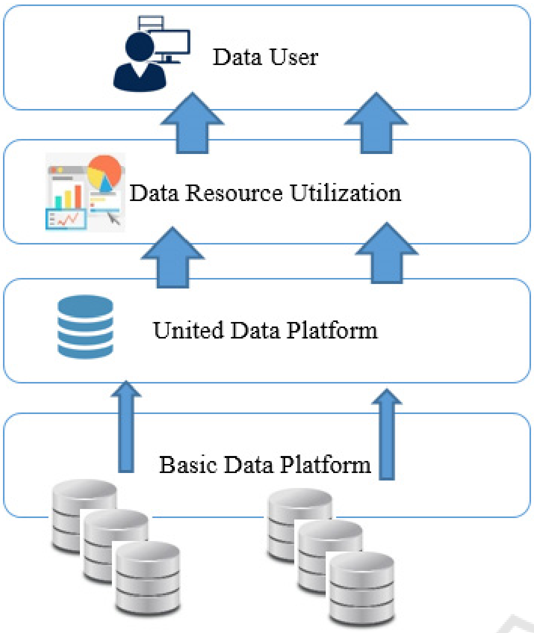

Figure 5: The technical framework of information resource (Bai et al., 2017).

3.4 Big Data Analytics for Water Leakage in Bangalore

In this article, the authors discuss water leaks that cause a significant water crisis in Bangalore (India). They propose using big data analytics for water management and the prediction of possible leaks (Pillai et al., 2019).

For this, they present this platform, in figure 6, based on:

Data acquisition - Scada model: This layer allows to acquire the water data from the sensors

1. Data transformation & storage - Database: The database that stores & transform the data
2. Analytics layer - Prediction & Forecast: Provides visualization of the data present in the database and allow to make possible predictions based on data processing
3. Business layer - Decision: The layer that allows reporting the data and monitoring the water leakage.
4. Presentation layer: this layer is at the top of the platform and allows the visualization of all the data present in the database
5. Presentation layer: this layer is at the top of the platform and allows the visualization of all the data present in the database (Pillai et al., 2019)

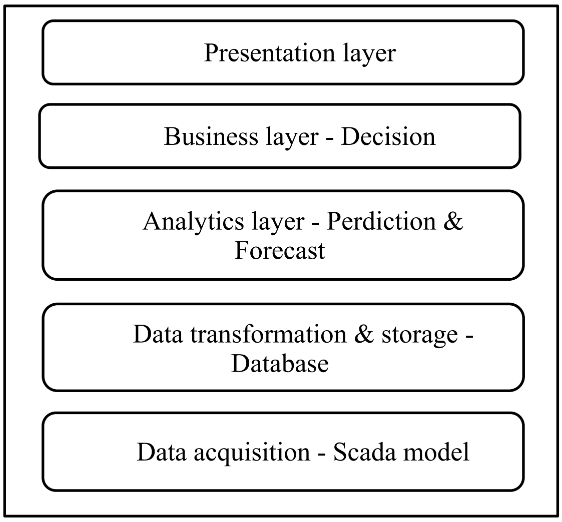

Figure 6: Architecture overview of water leakage data (Pillai et al., 2019).

In this article published in 2019, the authors limited themselves to designing an architecture by declaring that the implementation has not yet been carried out.

4 DISCUSSION

The examination of the different architectures in the previous phase allowed us to draw up the following table, highlighting the advantages and weaknesses of each architecture:

Table 1: Comparaison of big datas architectures.

	Modules	Status
Architecture 1 Chalh Ridouane (2015)	▪ Decision Support Tools ▪ Knowledge-Based System ▪ Geographic Information System (GIS) ▪ Big Data Analysis System ▪ Simulation Models ▪ Computation and Processing ▪ Communication System ▪ Search Engine ▪ Users Interfaces	▪ The proposed framework is in the design stage
Architecture 2 Ai Ping (2014)	▪ Data acquisition layer ▪ Resource organization layer ▪ Data analysis layer ▪ Application service layer	▪ The platform remained at the theoretical proposal stage
Architecture 3 Bai Y (2017)	▪ Basic data resource platform ▪ Unified data platform ▪ Data resource utilization	▪ The architecture was initially proposed for the management of water resources in China but posed problems in terms of implementation
Architecture 4 Deepthipriya R Pillai (2019)	▪ Data acquisition - Scada ▪ Data transformation & storage - Database ▪ Analytics layer - Prediction & Forecast ▪ Business layer - Decision ▪ Presentation layer	▪ The proposed framework is in the design stage

Table 2: Comparison of the strengths and weaknesses of big data architectures.

	+	-
Architecture 1 Chalh Ridouane (2015)	▪ Distributed on the web ▪ The presence of a module specialized in decision making ▪ Integration of the geographic module ▪ Presence of a particular research	▪ Architecture does not follow a process of data flow
Architecture 2 Ai Ping (2014)	▪ Classification of input data ▪ Architecture follows the pre-flow of information flow ▪ Presence of a layer for decision making	▪ Absence of a layer dedicated to GIS data ▪ data consolidation based on SQL only
Architecture 3 Bai Y (2017)	▪ Input data in several formats ▪ Architecture follows the pre-flow of information flow ▪ Standardization of data ▪ Possibility of multidimensional analysis	▪ Absence of a particular module for decision-making ▪ Lack of a layer dedicated to GIS data
Architecture 4 Deepthipriya R Pillai (2019)	▪ The presence of a lawyer specialized in decision making ▪ Presence of water prediction modules (leaks)	▪ Big data technology is not explained ▪ Platform designed for water leaks

Based on the comparison of the previous architectures, the first remark, which one can note, is that all the proposed designs remained in the state of design without implementation or real test.

Indeed, the first architecture (Chalh Ridouane, 2015), composed of blocks, contains a block of hydraulic models for data processing, a block for simulations, and a central block relating to the big data core.

In the second architecture (Ai Ping, 2014), the authors proposed a framework that aims to make decisions based on acquisition and processing data from several databases and accommodates the user from the acquisition phase to decision making, always passing through a relative layer of big data.

In the third architecture (Bai Y, 2017), the authors propose a platform that always follows the process from data acquisition for decision-making and a centre for the unification and standardization of data, which come from several business databases.

The authors, in the final design (Deepthipriya R Pillai (2019), present a model oriented towards water leaks, with a treatment model and prediction models but, like the last two architectures, always follow the process of Acquisition → Treatment → Decision.

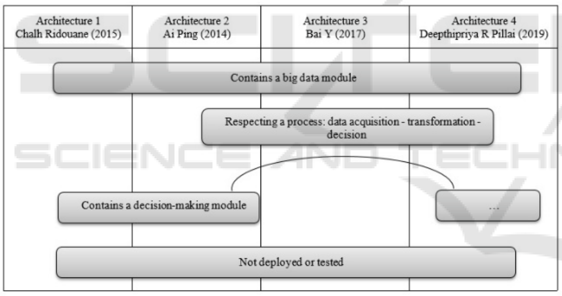

Figure 7: Comparison of big data water resources architecture.

After bibliographic research on the authors of these architectures, it seems to us that the authors, without any real implementation or experimentation in a real or quasi-real environment, proposed these architectures.

This observation raises questions about the possible obstacles and challenges to be presented to successfully introduce big data to the water information system of public actors in charge of water resources management

After comparing the different architectures of the authors, we came to the following conclusions for a water information decisional system:

- The architecture must be structured according to the flow of data processing flows
- It includes, as input, mechanisms for the acquisition of several data formats from several business databases
- It must have a block for the standardization of data based on the concept of big data
- It must contain a layer for multidimensional data analysis
- It must have a decision-making aid module

The treatment of articles through the nvivo tool allowed us to establish the following degree of similarity:

Figure 8: Degree of similarity of the articles.

Based on the above, we can propose the following architecture for a water decisional information system:

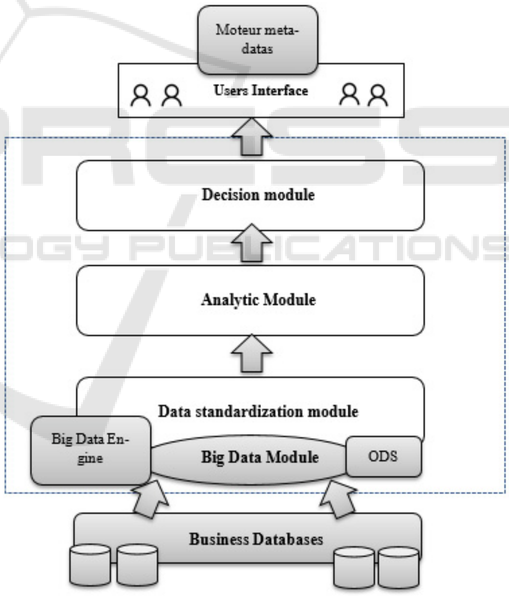

Figure 9: Proposed architecture.

The architecture, therefore proposed in Figure 7, results from the comparison of the different architectures studied. It is explained as follows:

- The data from business databases that contain data relating to water resources (measurement data, quality, GIS, etc.) are integrated into a standardization module.
- This module allows the processing and unification of the data format according to the big data concept.

- The analytics module allows us to do our various analyzes on this data to add a multidimensional visualization layer
- Next, there is the decision module, which allows you to define decision-making scenarios based on the result of the process.
- The user interface is the leading portal for using the platform, which contains an advanced search module (metadata search, etc.).

We can thus, based on this functional architecture, propose the following technical architecture

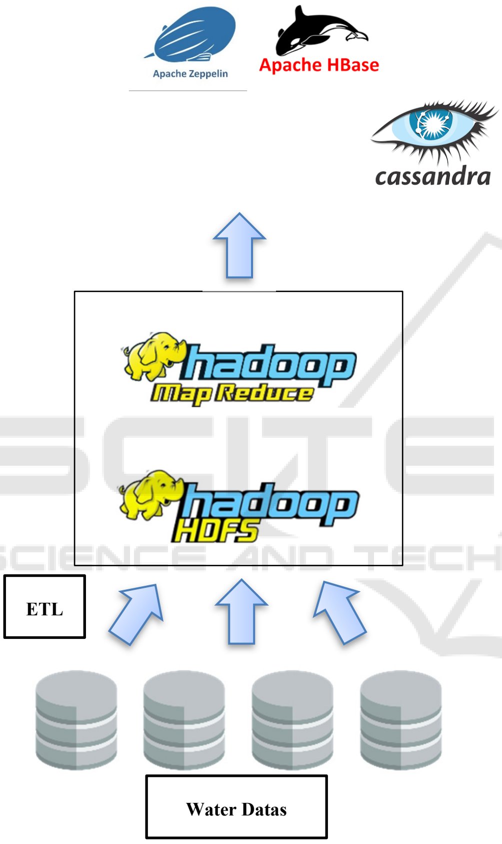

Figure 10: Technical architecture.

The proposed architecture contains the Hadoop tools for the organization and analysis of data collected via ETLs from the water basin agencies before being exploited by users by Cassandra and Hbase and interactively analyzed by Zeppelin.

5 CONCLUSIONS

The study showed us that several architectures are proposed for the realization of an information system for the management of water resources, but the implementation, as well as the test, remains a significant challenge to succeed to be able to present an exciting result the decisions-makers

The perspectives of our study would be to study in-depth the possible opportunities and challenges of the adoption of big data by the administrations in charge of water resource management in Morocco

REFERENCES

The website of the Water Department

Loi 36-15, Chapitre X (October 2016)

Aniss Moumen, Bouabid El Mansouri, Hassane Jarar Oulidi, Lamiaa Khazaz. 'Système d'Information sur l'Eau au Maroc: Etat d'art, Problématique, Approche et Prototype', (Conference Paper · November 2015)

Hasnaoui Moulay Driss, Ouazar Driss, Bakkoury Zohra, Chalh Ridouane 'Big data open platform for water resources management (2015 International Conference on Cloud Technologies and Applications (CloudTech), Pages 1-8, June 2015

Ai Ping, Yue Zhao' A Framework for Processing Water Resources Big Data and Application' Bai (Applied Mechanics and Materials 519-520, 2014)

Bai Y, Bai X, Lin L, Huang J, Fang H.W, Cai K 'Big data technology in establishment and amendment of water management standard' (Applied Ecology and Environmental Research 263-272 2017)

Deepthipriya R Pillai, P. Jeyalakshmi, Shwetha S P 'Big Data Analytics for water leakage in Bangalore' (IJ RAR Volume 6, Issue 2, May 2019)