International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Tracing Data Flow Diagram for a Flood Early Warning System (FEWS) in Malaysia Using Prescriptive Big Data Analytics

Aziyati Yusoff[a*], Salman Yussof[b], Norashidah Md Din[c]

[a]College of Engineering, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia
[b]College of Computer Science and Information Technology, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia
[c]College of Graduate Studies, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia
[a]Email: aziyati@mohr.gov.my
[b]Email: salman@uniten.edu.my
[c]Email: norashidah@uniten.edu.my

Abstract

With the advent of big data era, it is commendable if this facility could also be a method of problem solving to the environmental issues, disaster management, and geographical sciences. In this research, the study of flood events particularly in Malaysia is using the approach of prescriptive big data analytics. The big data of flood events which is managed by more than one authorizing agencies in Malaysia is proposed to be tackled by designing a feasible smart engine that is able to integrate most data forms and sets that are available from the participating agencies. The critical part of this research is to conform the practicality of integrating those big data into a structured data management so that it is traceable and able to return the desired results. This article is deliberating on the possibilities of tracing the big data of flood events which has undergone the process of rigorous prescriptive data analytics and knowledge engineering to return the searched results.

Keywords: big data analytics; data flow diagram; knowledge management; prescriptive big data; smart engine.

----

* Corresponding author.

1. Introduction

The initiation of the needs on designing and developing an application for flood management emerges with the continuous studies on flood disaster and the ways to handle it especially for the authorities of Government of Malaysia. The study involves many aspects, disciplines and parameters that it requires an integrated and unified approach from the perspectives of civil engineering, hydrology, geography, and even in the world of Information System. This research is initiated due to the intensified study on the ongoing needs of improvements in the area of flood and disaster management for local authorities and authorizing agencies. This research is about the feasibility study of big data in flood disaster management by using the platform of smart retrieval engine which is equipped with expert systems including the semantic network, ontology and propositional logic designs. For performance purpose, a case study on flood incidence was selected particularly from selected flood reading stations in Malaysia [1,2,3,4].As a novel approach to this research, the big data involved is defining its modelling in prescriptive analytics and equipped users with not only what, when, and where did the incident happened, but also why and how did it happen. As a result, a smart retrieval engine was designed and developed to run an integrated system of managing the big data of flood events from various authorizing agencies data sources. Consequently, the need to verify the feasibility of the study has adopted the method of tracing it by data flow diagram and production rules.

2. Prescriptive Big Data Analytics of Flood Events in Malaysia

Big data analytics, until today, is undergoing four major phases: (i) descriptive, (ii) diagnostic, (iii) predictive, and (iv) prescriptive [5]. Descriptive big data carries the role of providing the information on what the event or the data is about. Diagnostic big data gives reasons and provides knowledge on why the event should occur. Predictive big data provides both information and knowledge on what will happen to the event in the future based on the available data. And prescriptive big data enables the event manager or people with roles and responsibilities to the event occurred to plan and implement on things that can make the event to improvise and to happen. A prescriptive big data contains wisdom within the data, information, and knowledge that are being processed. In flood events cases, the study on rainfall and weather reports are always the most important parameters to determine the prediction of flood. The data readings for rainfall and water level are not bound from one source of an agency. In fact, more than one agency can be referred to for the conformity of this phenomenon data feed. The agencies involved for the rainfall and water level data are the (i) Department of Irrigation and Drainage (JPS), Malaysia [6], (ii) National Aeronautics and Space Agency (NASA) Earth Observations (NEO) [7], (iii) Global Disaster Agency and Coordination System (GDACS) [8], and (iv) Google Earth [9] and ArcGIS. The data readings for tidal wave were taken from two main agencies. One is from the portal of local government office and the other is from the global site. The agencies involved are: (i) National Oceanic and Atmospheric Administration (NOAA) [10], and (ii) Department of Survey and Mapping (JUPEM), Malaysia [11].The satellite image data is tricky. Besides the use of reading the image data of clouds movement, it can also be used to read for other land telemetric weather measurement purposes including the ocean tidal wave and water level for lakes, rivers, and sea. The main portals to refer for this data are: (i) NASA Moderate Resolution Imaging Spectroradiometer (MODIS) [12], and (ii) Department of Meteorology (Met), Malaysia [13].There are secondary applications and portals that read the above-mentioned weather information services to carry out their explicit functions. These agencies are building their extension portals more specifically for the purpose of disaster management. In Malaysia, there are different government agencies that work hand-in-hand when a disaster occurred in a specific location. For flood study of this research, the following data from disaster management agencies were being integrated: (i) National Disaster Management Agency (NADMA), Malaysia [14], (ii) Public Works Department (JKR), Malaysia [15], (iii) Department of Social Welfare (JKM), Malaysia [16], and (iv) Kelantan State Development Office (*eBanjir*) [17].

Table 1: Details of Data Type and Formats of Big Data for Flood Incident.

Data Type	**Data Format**	**Provided by**
Numerical	Comma Separated Values (CSV), Text (TXT)	• Department of Irrigation and Drainage (JPS) • National Aeronautics and Space Administration (NASA) Earth Observatory • Global Disaster Alert Coordination System (GDACS) • National Oceanic and Atmospheric Administration (NOAA) • Department of Survey and Mapping (JUPEM), Malaysia • Kelantan State Development Office, Malaysia
Image	Joint Photographic Experts Group (JPG), Graphics Interchange Format (GIF), Tagged Image File Format (TIFF), Portable Network Graphics (PNG)	• Department of Meteorology (Met), Malaysia • Twitter • Instagram
Text Retrieval	Form based or Content based	• Twitter • National Disaster Management Agency (NADMA), Malaysia • Public Works Department Malaysia (JKR) • Department of Social Welfare (JKM)
Mapping and geospatial	Keyhole Markup Language (KML)	• Global Disaster Alert Coordination System (GDACS) • Department of Survey and Mapping (JUPEM), Malaysia.

As a summary, Table 1 is illustrating on the big data of the flood events with details of data types involved in the computation, the formats and the agencies that are responsible to manage them.

3. Smart Retrieval Engine of Flood Early Warning System (FEWS)

To show that the integration of this smart retrieval engine is able to manage the big data demand, the inputs of numerous data types are also embedded within the Protégé environment. This Protégé software is used to design the ontology and semantic network of the smart retrieval engine for flood early warning system for this research. Documents and data files of CSV, PNG, and KML are embedded by the facility of data import in the Protégé. As a result, it is expected that this application will return the dynamic web pages of eXtended Markup Language

(XML) that is able to provide feed on flood prediction to the platforms that runs it. This process is illustrated in Figure 1.

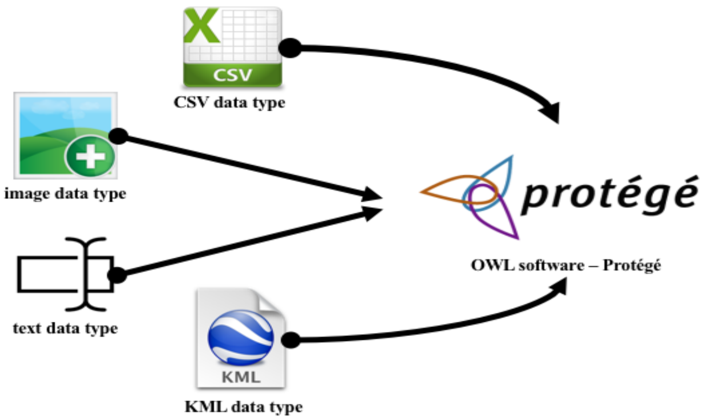

Figure 1: Embedding the csv, image, text, and KML data types into Protégé.

One of the works on the integration of Geographical Information System (GIS) with expert system was suggested by McCarthy. The integration concept was quite successful [18]. By this integration, the team has proven that geographical data was much accurate, the sensors return precise values, and images were much detailed. This is favorable when the longitude and latitude were also specific and exact location was able to be identified. Consequently, this research is utilizing the concept put forward by McCarthy with an additional method of semantic network and web ontology. The two methods were seen as relevant to support the advancement of GIS and expert system [19] for today's technology. Semantic network in this context is specifically for defining the classes, subclasses and inheritance of data, whilst the web ontology is to illustrate the properties and relationships of the classes which are dynamically built by XML to suit the news feed in the Internet platforms. Therefore, this smart retrieval engine for FEWS project is to suit the requirements of this research with a case study on the flood event in the Malaysia.

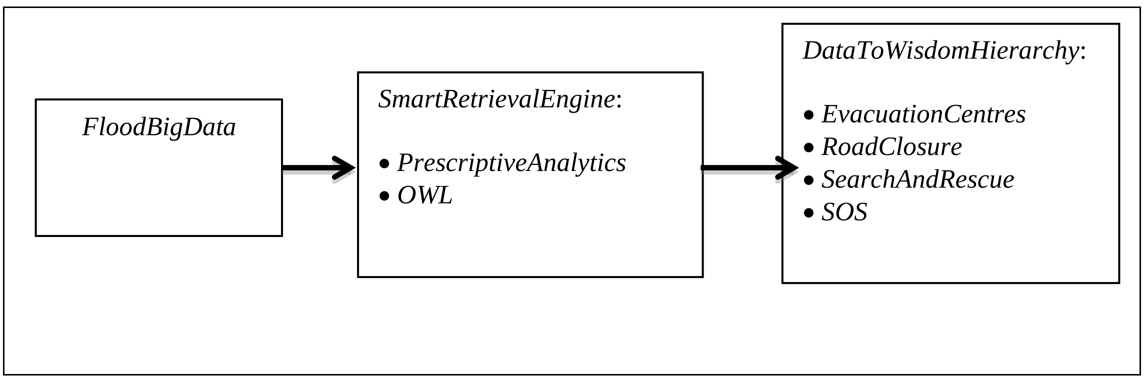

Figure 2: The input-process-output of Integrated GIS, and Expert System of smart retrieval engine of FEWS.

Furthermore, this engine is also meant to be in favor with the big data era that is influencing the field of data science and other research disciplines. As most of other engineering studies imply, Figure 2 is illustrating the input-process-output for this engine. The highlight of this process is that, the input of this project is the big data of flood reading from various sources and data type format that undergoes the computation of smart engine retrieval and yields the output of prescriptive analytics of the flood event.

4. Tracing Data Flow Diagram of Smart Retrieval Engine

To demonstrate the above result performance, given is the following example: During the monsoonal season of north eastern coast of Peninsula Malaysia, it is expected to rain heavily continuously for several days. This monsoonal season takes place between the months of November until January each year. When this happens, the authorizing agencies will be prepared with news, alerts, warnings and relevant information to the affected communities particularly in the states of Kelantan, Terengganu, and Pahang. The first agency to issue an alert is usually the Meteorology Department. From the readings of cloud movement, the wind speed and weather forecast, the Met agency will issue alerts and warnings to the affected communities on the prediction of flood. This alert would trigger other agencies to be prepared too including the JKR that will issue on the road safety and alternative routes, the JKM on opening active evacuation sites, the local authorities on the movement of the flood victims, and the JPS will keep collecting the data reading from all of the telemetric stations on the water level and the images that their cameras have captured on the event chronology. At this point, NADMA will launch its search and rescue teams and mitigation actions during flood emergency.

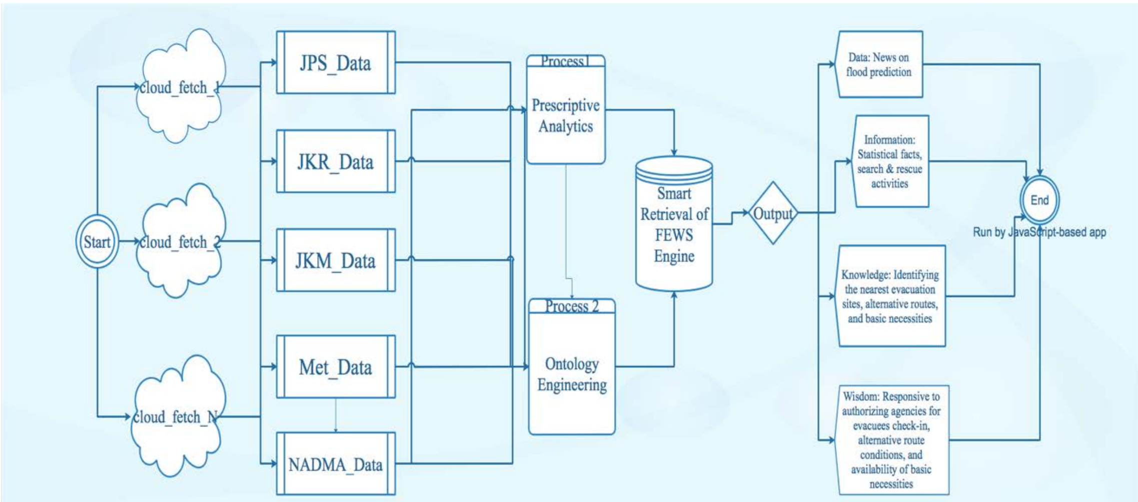

Figure 3: Data Flow Diagram (DFD) of prescriptive big data analytics for smart retrieval for FEWS Engine and ontology engineering for the output of DIKW of flood event.

Basically, with the implementation of the smart retrieval engine for FEWS, the above data and information are expected to be fetched and stored within this application to return the knowledge and wisdom of the event. In this way, the ontology engineering works in the sense that it computes the stored data of the flood from the Met, JKR, JKM, JPS, and NADMA to retrieve the output of complete package of DIKW to its user. Figure 3 is illustrating this Data Flow Diagram (DFD) of this performance result.

From the DFD it is learned that the clouds that contain and store the big data of flood events from authorizing agencies of JPS, JKR, JKM, Met, and NADMA are fetched and processed through two main process components of smart retrieval engine for FEWS i.e. prescriptive analytics and ontology engineering. This is summarizing the whole process of big data flow for the application that uses the smart retrieval engine for FEWS as its central processing unit and coherence to the research framework of this project.

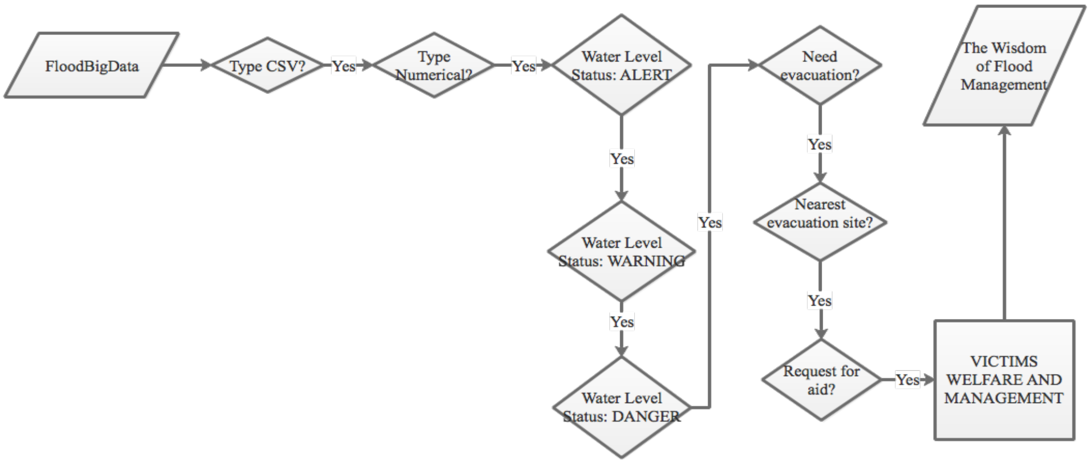

Figure 4: Example of reading the data type CSV from the *FloodBigData* input sources to return the data of Wisdom of Flood Management.

As an example, if the situation of a flood event is about to occur in the targeted area, following the flowchart, production rules and the algorithms of the program will run such as illustrated Figure 4 on the flow of the data output. From the diagram, the production rule of the data flow is as the following listing:

Rule 1:

INPUT: FloodBigData is fetched

Go To ➔ Rule 2

Rule 2:

Input data is type CSV

Then data type numerical is checked

And Water Level status: ALERT is checked

Go To ➔ Rule 3

Rule 3:

Water Level status: ALERT is activated

Then Water Level status: WARNING is checked

When Water Level status: WARNING is activated

Then Water Level status: DANGER is checked

When Water Level status: DANGER is activated

Then evacuation need is checked

When evacuation need is activated

Then the nearest evacuation site is checked

When the nearest evacuation is located

Then request for aid is checked

When request for aid is activated

Update the Victims and Welfare Module Management run by NADMA and JKM

Go To ➔ Rule 4

Rule 4:

Update and Return on the Wisdom of Flood Management data to the smart retrieval engine for FEWS system

5. Conclusion

It is expected that the study of big data on multi-discipline is to be developed in many designs, platforms and applications. The apprehension of big data analytics from descriptive form to prescriptive form is also to be expanded towards something beyond the knowledge and wisdom of the data representations. From this research, it is found that the big data analytics though unstructured in nature, can still be defined and constructed for problem solving in numerous decision-making processes. Therefore, the data flow diagram which was normally being used to treat the structured data can also be utilized the big data analytics as long as the design and development of the data processing are able to be materialized in a structured way.

References

[1] Wing, Chia Chong. "Managing flood problems in Malaysia." Buletin Ingeniur (1971).

[2] Adnan, Nor Aizam, and Peter M. Atkinson. "Exploring the impact of climate and land use changes on

streamflow trends in a monsoon catchment." International journal of climatology 31, no. 6 (2011): 815-831.

[3] Global Disaster Alert and Coordination System, URL: http://www.gdacs.org/.

[4] Yusoff, Aziyati, MdDin, Norashidah, Yussof, Salman, and Khan, Samee Ullah, “The Semantic Network of Flood Hydrological Data for Kelantan, Malaysia”, In IOP Conference Series: Earth and Environmental Science (Vol. 32, No. 1, p. 012021), 2016, IOP Publishing.

[5] Gartner Hughlights Five Attributes of Cloud Computing, URL: “http://www.gartner.com/newsroom/id/1035013”, 2014.

[6] Public InfoBanjir, JPS Malaysia. URL: “http://publicinfobanjir.water.gov.my/”, 1 October 2014.

[7] NASA Earth Observatory, URL: “https://earthobservatory.nasa.gov/”, accessed on 1 September 2015.

[8] Global Disaster Alert and Coordination System, URL: http://www.gdacs.org/, accessed on 1 September 2015.

[9] Yusoff, Aziyati, Mustafa, Intan Shafinaz, Yussof, Salman and MdDin, Norashidah, “Green Cloud Platform for Flood Early Detection Warning System“, in The 5th IEEE National Symposium on Information Technology: Towards New Smart World, Riyadh, Kingdom of Saudi Arabia, February 2015.

[10] National Oceanic and Atmospheric Administration (NOAA), US Department of Commerce, URL: “http://www.noaa.gov/”, accessed on 1 September 2016.

[11] Department of Survey and Mapping Malaysia (JUPEM), URL: “http://www.jupem.gov.my/”, accessed on 1 September 2016.

[12] NASA EOSDIS Worldview, URL: https://worldview.earthdata.nasa.gov/, accessed on 1 September 2015.

[13] Malaysian Meteorological Department (Met), URL: “http://www.met.gov.my/”, accessed on 1 February 2016.

[14] Disaster Portal, National Disaster Management Agency, Malaysia (NADMA), URL: “http://portalbencana.ndcc.gov.my/”, accessed 1 February 2016.

[15] Public Works Department (JKR), URL: “http://www.jkr.gov.my/”, accessed 1 September 2016.

[16] Department of Social Welfare Malaysia (JKM), URL: “http://www.jkm.gov.my/”, accessed 1 September 2016.

[17] e-Banjir Kelantan Portal, URL: “http://ebanjir.kelantan.gov.my/”, accessed 1 February 2016.

[18] McCarthy, James D., Phil A. Graniero, and Steven M. Rozic. "An integrated GIS-expert system framework for live hazard monitoring and detection." Sensors 8, no. 2 (2008): 830-846.

[19] Krizhanovsky, Andrew, "Integration Of Ontology With Expert System: A Case Study",(2005).