_____

Spatial Decision Support System for Animal Diseases

Mahmoud Bakr / Research assistant
Dept. Information Systems,
Faculty of Computers & Info. ,
Helwan University, Egypt
mbakr_2005@yahoo.com

Mona Nasr / Associate prof.
Dept. Information Systems,
Faculty of Computers & Info. ,
Helwan University, Egypt
m.nasr@helwan.edu.eg

Maryem Hazman / Assistant prof.
Dept. Testing and Training,
CLAES,
ARC, Egypt
m.hazman@claes.sci.eg

Abstract— in this paper, a Spatial Decision Support System (SDSS) was presented to help decision makers in the decision making process. The proposed spatial decision support system utilizes the capabilities of Geographic Information System (GIS), Data warehouse and Online Analytical Processing (OLAP) to provide decision makers with their needed information about the infected animals, infected places and diseases outbreaks. This information is displayed as reports or charts or allocated in a map which illustrates the most and the least affected places in an easy and fast way.so that, decision makers can take the right decision to control the spread of diseases outbreaks. The proposed SDSS consists of three databases namely: TADinfo, BOVIS and Climate databases. Data warehouse generated from integrating those three databases and diagnosis data mart is subset of that data warehouse. OLAP capabilities integrated with data warehouse to enable decision makers browse diagnosis data from different views and generate needed reports and charts. The proposed SDSS enhanced with GIS capabilities to make various spatial analysis on diagnosis data and visualize the results as maps. The experimental results show that the proposed system can provide the decision makers with their needed information in a fast and easy way.

Keywords- *Animal Disease; spatial decision support system (SDSS); Geographic Information System (GIS); Data warehouse; Online Analytical Processing (OLAP).*

_____*****_____

I. INTRODUCTION

Many countries face the food gap and insufficient food due to many reasons, like high population growth and production/resources lack. Governments try to maximize its resources usages. So, animal husbandry is one of the most important activities in developing countries. Where, their governments pay special caring of plant and animals production as they have impact on their economies and the gross national product. They entered in all industries, which aim to achieve food security for the community and its members.

In Egypt, one of the government sustainable strategic objectives is to improve agricultural productivity for both land and animal [15]. Improving animal production with the lack of resources need to preserve animal. So, Egypt established the General Organization for Veterinary Services (GOVS) to be responsible for animal preserve. Also, there are large number of endemic diseases appeared in Egypt in the recent period, such as Foot and Mouth Disease (FMD) and avian flu. These diseases caused losses in a large number of animals, medications and immunizations. And hence, these losses lead to lack of production of meat and milk [1, 2]. That was happened although; GOVS has many transaction information systems which record animals and its diseases. Since, these systems are lack of clear reports that help the decision makers, e.g., reports about places of disease outbreak and geographical representation and visualization of the spread of specific disease.

So, the goal of this work is to utilize the existing animals' information and databases systems, weather database and geographical information database to provide decision makers with their needed information in the suitable time. It investigates on how to use these animals' diseases data and integrates them into a data warehouse and uses the capabilities of OLAP and GIS Spatial Analysis and Visualization to build an Animal Diseases Spatial Decision Support System (ADSDSS). ADSDSS must has the capability to provide decision makers with numerical reports about animal diseases in different places with specific criteria, a representation and visualization of these affected animals numbers on the map. So, it assists decision makers in taking appropriate decisions to control the spread of disease in the right time.

ADSDSS consists of three components namely: data warehouse, OLAP, and GIS. In the data warehouse component, data from different resources (TADinfo, BOVIS and Climate database) are collected and integrated into one repository data mart called diagnosis data mart. In the OLAP component, diagnosis cube was generated to include different needed dimensions and C1 OLAP tool was used to browse the cube data and generate charts and reports for the selected dimensions data. In GIS spatial analysis and visualization component, specific analysis methods were applied to selected diagnosis data to generate and visualize output map.

The rest of this paper is organized as follows. A brief overview of related work is presented in section 2. Section 3 presents definitions for the used technologies for building ADSDSS. Section 4 describes the problem analysis. In section 5 we explain the proposed (ADSDSS) architecture. A case study is presented in section 6. Finally, section 7 presents the conclusion and future work.

II. RELATED WORK

DSS is a rich discipline research area; many researchers try to enhance the capabilities of DSS by integrating it with other technologies like, GIS capabilities, Data warehouse, Data mining techniques and web technologies. Some researches focus on helping decision making process by monitoring a specific disease, like **Sanson** and their colleges introduced EpiMAN DSS to reduce the analysis time of FMD outbreak event details which leads to control and eradicate activities to be fine-tuned as the epidemic progressed, as individual

2405

_____

circumstances dictate. This system combines a database management system, Geographic Information System GIS, and a graphic user interface for helping its user to conduct descriptive analyses of infectious disease outbreak, manage resources (e.g. scheduling of patrol visits). Also, users can evaluate the effect of alternative control measures using a simulation model. EpiMAN was originally developed to manage data that is generated during the course of FMD outbreak [7].

While, Eisen and their colleges combine on Geographic Information Systems GIS and Decision Support Systems DSS to predict, prevent, and control of Vector-Borne Diseases. They focus on using the advanced of Geographic Information System, and Decision Support System technologies, with progress in spatial and space-time modeling to prevent and control animals diseases [5].

Elfangary and their colleges introduced integrate data mining with data warehouse techniques to develop a system that is used to analysis and measures the effect of climate on animal production. This system first combines the animal diseases database (BOVIS) with the weather database using the data warehouses techniques. Then, a set of interested variables are selected and prepared to be used in mining. After that, this variables set is mine to discover the relation between weather and bovine diseases. This work uses the predictive data mining modeling (classification and regression). By discovering the relation between weather and bovine diseases, the system can predict when the occurrence of disease [12].

III. PRELIMINARIES

A. DSS

Decision Support System (DSS) is a computer-based System which helps decision makers to confront ill-structured problems through direct interaction with data and analysis models [6]. DSS aims to support decision making process. DSSs avail the management, operations, and planning levels of an organization (usually mid and higher management) and help to make decisions, which may be rapidly changing and not easily specified in advance. DSS analyzes business data and presents it so that users can make business decisions more easily [4-8]. Typical information that a decision support application might gather and present would be:
- Comparative affected animal's numbers between one week and the next
- The consequences of different decision alternatives, given past experience in a context that is described.

DSS may present information graphically and include an expert system or artificial intelligence. DSS can Improve personal efficiency, Speed up the process of decision making and Increases organizational control [4-8].

SDSS differs from regular DSS in that SDSS deals with geographic or spatial data and aspects in decision making process.

B. Data warehouse

Data warehouse is a subject-oriented, time-variant, combined, and nonvolatile collection of data in assist of management's decision-making process. It differs from transactional database in that, it is a big store of nonvolatile data and can be un-normalized database structure. Date warehouse gathers data from different data sources likes, databases, spread sheets, text files … etc. Data warehouse structure composes of fact tables and dimensions with various

schemas, e.g. star schema and snowflake scheme. The preferred structure to improve the response and efficiency time of the Data Warehouse is the Star Schema. In star schema, data is maintained in a single fact table which is located at the center of the schema with additional dimension data tables, with all hierarchies collapsed [3].

C. OLAP

Online Analytical Processing (OLAP) is computer processing that provides its user with an easily way to extract and view data from different points of view. OLAP composes of three basic analytical operations namely: consolidation (roll-up), drill-down, and slicing and dicing. Consolidation includes the aggregation of data which can be accumulated and calculated in one or more dimensions. Where, the drill-down technique allows OLAP users for navigating through the details data. Slicing is a feature whereby OLAP users can take out a particular set of data of the OLAP cube. Dicing is a feature whereby OLAP users view the slices from different viewpoints. OLAP cube includes numeric facts (measures) which are categorized by dimensions [3].

D. GIS

Geographic Information System (GIS) is a computer based information system. It is used to represent the geographic features present on the Earth' surface in a digitalized form which allows its analysis. GIS links database (tabular) information and locational (spatial) and enables a person to visualize relationships, patterns, and trends. This procedure gives a new viewpoint to data analysis which cannot be realized in a table or list format. GIS include of five components: Data, Hardware, Software, Methods and People. GIS provides spatial data entry, management, analysis, retrieval, and visualization functions [8-11].

IV. PROBLEM ANALYSIS

The goal of carrying out this analysis was to identify the needed solution of the existing problem. GOVS has many transactional information systems. These information systems concern with recording animals identifications and history for its transactions like, diseases and their vaccinations. They aim to enable decision makers to obtain statistical data about animals and help them to the tracing and management of Livestock. In this analysis, TADinfo [14] and BOVIS are the selected animal information systems. TADinfo stand for Transboundary Animal Disease Information System (TADinfo) was developed by Food and Agriculture Organization (FAO) to provide data management and decision support to national veterinary epidemiology units. It is a data management system fully dedicated to animal and zoonotic disease information management and analysis. TADinfo is used to store and analyze data related to observations, abattoir observations, active disease surveillance, livestock census and vaccination. TADinfo records the diseases for many species of animals like, bovine, buffalo, goat, and sheep. It provides numerical and detailed reports. Bovine Information System (BOVIS) is an information system which Developed by Central Laboratory for Agriculture Expert Systems (CLAES). It is used to record the identity bovine, buffalo animals and its historical transactions. It provides numerical and detailed crystal reports about the recorded animals. These two databases will be integrated with weather data and Geographical data. These databases systems are isolated, so there is no integration

between them and the generated reports are provided separately to decision makers.

In order to provide decision makers by their needed information in the suitable time, these systems need to be integrated and enhanced by using the SDSS integrated with the GIS. First, a data warehouse will be built for integrating the three databases and creating a data mart for animal diagnosis. Then OLAP analysis tool can use the animal diagnosis data mart for browsing and analyzing the diagnosis cube. Then a GIS tools can use the animal diagnosis data mart or OLAP data to represent and visualizes the results of the selected data on the map. The spatial analysis models will be used to analyze and visualize the results on the map.

V.   THE PROPOSED ADSDSS ARCHITECTURE

The aim of this work is to create a SDSS that provide decision makers with their needed information in the suitable time. In order to achieve this goal, the proposed system ADSDSS is produced. As shown in figure 1, the proposed ADSDSS architecture consists of three main components: data warehouse, OLAP, and GIS spatial analysis and visualization model. The following subsections introduce these components.
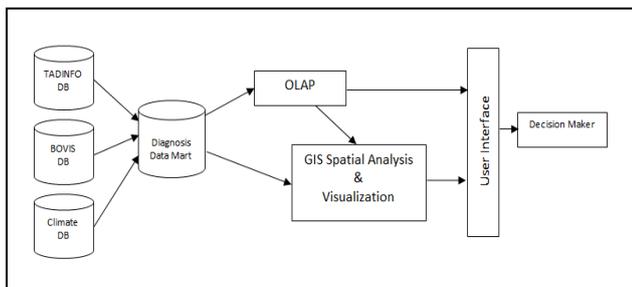

Figure 1. Proposed ADSDSS Components

A.   Data Warehouse Component

This component is responsible for integrating different data sources to enable generating integrated reports from these different sources. In order to integrate different data sources, the following steps are achieved:

- **Data Selection**: TADinfo, BOVIS and Weather databases contain many attributes which do not related to diseases. So, these databases were analyzed to select appropriate parts from both databases to be integrated in data warehouse. The diseases data mart was built using the most significant predictive attributes.
- **Data cleaning**: In this step, the data is analyzed to ensure its quality. It includes removal of null values, redundancy data; complete the missing values, identification and correction of data inconsistencies.
- **Data unified**: This step responsibility is to unify the used names for TADinfo and BOVIS databases to facilitate matching the data. For example, renaming places and disorders.
- **Data transformation**: in this step the selected data is transformed from the three databases into the data warehouse repository database.

The Diagnosis Data Mart is created with star schema structure which is diagnosis fact table and the related dimensions tables. Once the Data Mart is created, it can be used by both OLAP and GIS Spatial Analysis and Visualization components.

B.   OLAP Component

The ADSDSS includes OLAP component to provide Decision makers with a flexible way to browse diagnosis cube data and create integrated reports and charts. Figure 2 shows the structure of diagnosis cube.
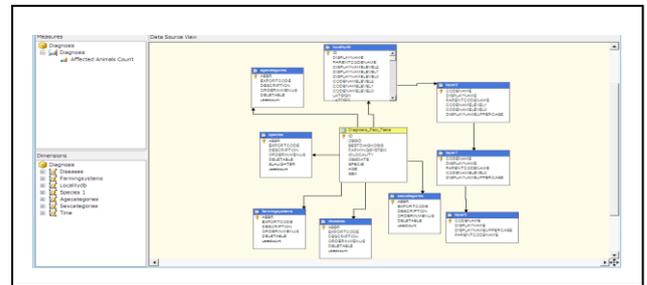

Figure 2. Diagnosis cube structure

In this component, the C1 OLAP tool is selected to be integrated in the proposed ADSDSS. It enables Decision makers to browse diagnosis cube data from different views. C1 OLAP tool helps Decision makers by an easy way to select the dimensions of diagnosis cube for generating the existing relation between them. Once Decision makers select the desired data dimensions, they can manipulate it by filtering these attributes, highlighting specific values that achieved certain condition and generating needed charts. Also, they can print the generated reports and export to other formats [16]. Figures: 3 and 4 show examples of C1 OLAP Grid and Chart. As shown in Figure 3, Decision maker can drag from the attributes that existing in the top left list to rows, columns and measures. The OLAP summarizes the relation between the selected fields as in the right OLAP grid.
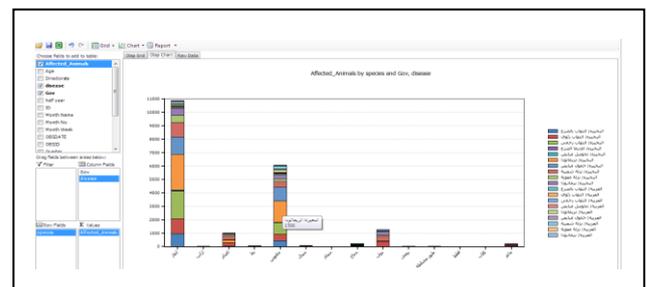

Figure 3. C1 OLAP grid integrated with ADSDSS

Figure 4. C1 OLAP chart integrated with ADSDSS

C.   GIS Spatial Analysis and Visualization Component

The main objective of this component is to analyze and visualize the diagnosis data according to geographical information on the map.

The spatial analysis phase uses a special map in gdb format. The geodatabase defines the generic model for geographic

2407

information. It is implemented as a group of files in a file system or a group of tables in a DBMS and consists of set of feature classes. These feature classes are homogeneous groups of common features, each has the same spatial representation, such as points, lines, or polygons, and a common set of attribute columns [11,17]. In Egypt, the feature classes are Governorate, Directorate and Village. The ArcObjects is used to represent these features, which is a set of platform-independent software components [18]. ArcObjects provides services to support GIS applications on the desktop – components to build our spatial analysis and visualization model used in AnimalSDSS.

ADSDSS enables decision maker to analyze and visualize the spread of specific diseases on the map and highlight the maximum affected places. The spatial analysis and visualization model, which represented in Figure 5 works as the following:
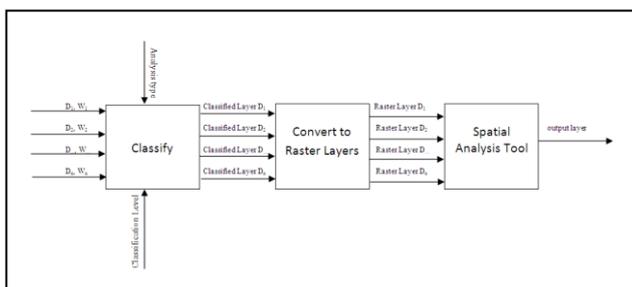

Figure 5. GIS spatial analysis and visualization model

First, from the data results from C1 OLAP grid or diagnosis data mart, decision maker selects one or more disease(s) and assigns weight for each one. The sum of all weights must equal to 100 in case of choosing weighted overlay analysis. In case of weighted sum analysis, decision maker can assign any weights. Then he/she selects analysis type from weighted sum, weighted overlay, unique values …etc. and classification level (Village, Unit, Directorate, Governorate). This process can be summarized as follow:

- Decision maker selects one or more diseases, assigns weight for each one and define spatial analysis method.
- The classifier in GIS component Generate classified layer for each disease.
- GIS component convert each classified layer to raster layer.
- GIS component Apply selected spatial analysis to the raster layers to generate one output layer.
- GIS component Visualize the output layer to decision maker.

For each disease, a classified weighted layer is created according to analysis type, and then these layers converted to raster layers. After that a spatial analysis tool is applied to them to result one raster layer which influences the spread of all selected diseases as the following:

In weighted sum analysis: for each affected animals count (C) with disease (D) at classification level or place (p) multiplied by the disease weight (W) a disease layer (L) created and classified according the new value. All diseases layers converted to raster then weighted sum spatial analysis tool

applied to diseases raster layers. Figure 6 shows the weighted sum analysis model, while Figure 7 shows an example for the generated layers.
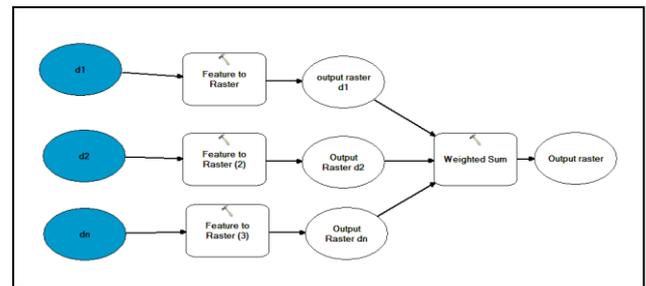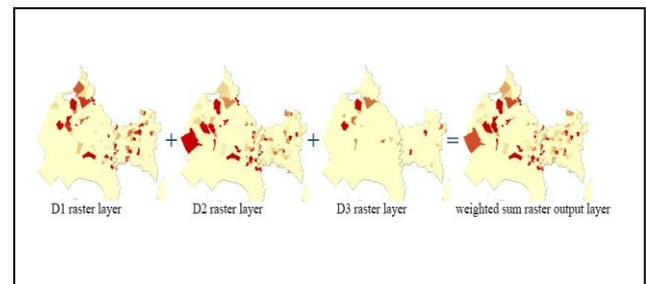

Figure 6. Weighted sum model

Figure 7. Maps of 3 diseases and the weighted sum map of them

Weighted Overlay analysis works the same way as weighted sum analysis. But in the classification step, the affected animals are classified into classes from 1 to 9 then weighted overlay spatial analysis tool applied. According to our meaning of classes places with large class numbers means large numbers of affected animals and vice versa. Figure 9 shows that how weighted overlay be calculated. First, the cell values are multiplied by their percentage influence, and the results are added together to create the output raster. For example, consider the top left cell. The values for the two inputs become $(2 * 0.75) = 1.5$ and $(3 * 0.25) = 0.75$. The sum of 1.5 and 0.75 is 2.25. Because the output raster from Weighted Overlay is integer, the final value is rounded to 2 [19].
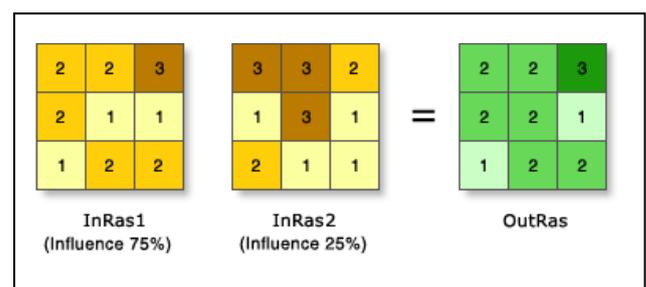

Figure 8. Raster weighted overlay calculation example

Also GIS Spatial Analysis and Visualization Component allows decision maker to visualize data that selected from diagnosis data mart or OLAP data directly on the map.

VI.   CASE STUDY

In order to evaluate the developed ADSDSS, it was applied using diagnosis data for two governorates namely: Algharbia and Elbehera. TADinfo, BOVIS are the selected animal diagnosis databases. The used TADinfo database contains more

2408

than 85000 diagnosis records for those two governorates. Where, BOVIS includes about 1 million diagnosis records.

The following is a scenario for developing and using the ADSDSS:-

The diagnosis Data Mart: In this phase, the related attribute to diagnosis is selected. These attributes are diagnosis date, animal age, animal sex, species, governorate, district, unit, village, disease, temperature and humidity.

OLAP: Diagnosis cube was created from diagnosis data mart.it contains different dimensions which enable decision maker to browse the data from different views.

First, Decision Maker selects the specific criteria that meet his/her requirement. For example, Decision Maker can select TADinfo data source or BOVIS data source or both. From the selected data source, he/she selects the needed attributes like: diagnosis date, place, diseases. Also, the spatial analysis method is selected and visualization level.

Decision Maker can assign weight for each selected disease according to its influence in the analysis process. After that, the data with some statistics Meta data can be shown. These statistics Meta data is about maximum and minimum number of affected animals in each place, standard deviation and median for each disease. Figure 9 shows the selected criteria, data and the result statistics.
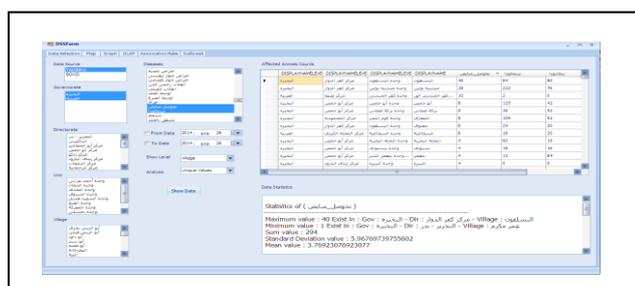

Figure 9. Data selection and statistical results

Decision maker can then view these results on the map according to the selected analysis type as in figure 10. Decision maker can save these maps as PDF format. ADSDSS generates classified map layer for each disease in which each place or the selected visualization level (governorate or district or village) assigned unique color. The most affected places with a disease to the least affected places assigned colors from red to green respectively.
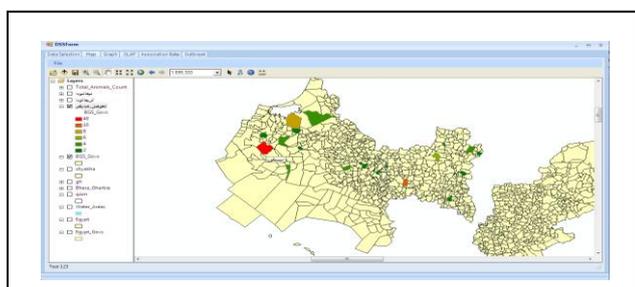

Figure 10. Visualization of affected animals' counts/village for each disease

ANSDSS provides decision maker with many types of charts that represent the selected data diseases. Decision maker can save, print and export these charts to many formats. Figure 11 shows example of these charts.
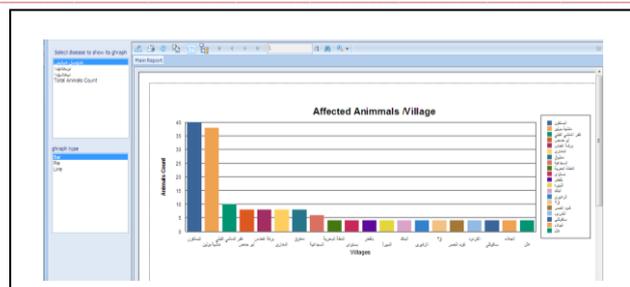

Figure 11. A chart of affected animals' count/village of a disease

ADSDSS has OLAP capabilities which enable decision maker to browse data from different views by dragging and dropping dimensions and measures, filtering specific data and highlighting specific values as in figures 3 and 4.

VII. CONCLUSION

In the last few years many animal diseases had appeared in Egypt in the form of outbreaks. There are huge losses of animal livestock, meat and milk due to the shortage of information in the right time like, the places of outbreak, and numbers of affected animals. Therefore, this paper has presented ADSDSS which utilizes the capability of GIS, Data warehouse and OLAP to help decision makers in the process of making decision. ADSDSS helps decision maker to control the spread of animal diseases outbreaks. It enables decision makers to browse and view diagnosis data from different viewpoints and view different dimensions of data. The proposed ADSDSS provides the decision maker with the capability to represent and visualize the affected animals' counts over different place levels (village, unit, directorate and governorate) and enables him to generate a map for disease(s) spread. Decision maker can use different types of spatial analysis to define the most and the least affected places with one or more diseases and the most and the least places appropriate for animal production and livestock activities or investment.

In the future, we intend to integrate one of data mining association rules techniques with our ADSDSS. It will help decision maker to find if there are association rules between different attributes of data like place and disease or disease and weather and so on. Also ADSDSS can be integrated with knowledge based to provide decision maker with suggested alternative decisions and treatments for diseases outbreak. ADSDSS can be used with other governorate animal diagnosis data.

REFERENCES

[1] FAO. 2011. Good Emergency Management Practices: The Essentials. Edited by Nick Honhold, Ian Douglas, William Geering, Arnon Shimshoni and Juan Lubroth. FAO Animal Production and Health Manual No. 11. Rome (available at www.fao.org/docrep/014/ba0137e/ba0137e00.pdf)

[2] FAO. 2012. Foot-and-Mouth disease caused by serotype SAT2 in Egypt and Libya: a regional concern for animal health in North Africa and the Middle East. EMPRES WATCH, Vol. 25, March 2012.Rome(available at www.fao.org/docrep/015/an380e/an380e00.pdf).

[3]   A. Berson and S. J. Smith, " Data Warehousing, Data Mining & OLAP", Tata Mc Graw Hill Edition, Tenth Reprint 2007.

[4]   A. Adela, "A Cooperative Intelligent Decision Support System for Contingency Management", Journal of Computer Science vol.10, iss.2, pp.758-764, 2006.

[5]   L. Eisen and R. J. Eisen, "Using Geographic Information Systems and Decision Support Systems for the Prediction, Prevention, and Control of Vector-Borne Diseases", Annu. Rev. Entomol. 56:41–61, 2011.

[6]   E. Turban and J. E. Aronson, Decision Support Systems and Intelligent Systems, 6th edition.

[7]   R. L. Sanson, "The development of a decision support system for an animal disease emergency", Massey University, March 1993.

[8]   S. Gao, D. Mioc, F. Anton, X. Yi, D. Coleman, " Online GIS services for mapping and sharing disease information". Int. J. Health Geogr. 7:8., 2008.

[9]   S. Hay, R. Snow," The Malaria Atlas Project: developing global maps of malaria risk", PLoS Med.3:e473, 2006.

[10] J. Jiang , P. Wang, W. Lung, L. Guo, and M. Li," A GIS based generic real-time risk assessment framework and decision tools for chemical spills in the river basin ", Elsevier B.V, 2012.

[11] UK Essays. Survey About Geographic Information Systems Tools Computer Science Essay [Internet]. November 2013. Available from: http://www.ukessays.com/essays/computer-science/survey-about-geographic-information-systems-tools-computer-science-essay.php?cref=1.

[12] L. Elfangary, M. Hazman and A. Abdallah, "Mining the Impact of Climate Change on Animal Production", International Journal of Computer Applications (0975 – 8887) Volume 59– No.18, December 2012.

[13] X. Tan, Data warehousing and its potential using in weather forecast, Institute of Urban Meteorology, CMA, Beijing, China, 2006.

[14] FAO. "About TADinfo" [Internet]. June 2007. Available from: www.fao.org/ag/againfo/programmes/en/empres/TADinfo/about.html.

[15] MALR, Sustainable Agricultural Development Strategy towards 2030, Ministry of Agriculture and Land Reclamation summery [Internet]. 2009. Available from: http://idc.sci.eg/SADS_Book1.html.

[16] Component One. Component One OLAP for winform - summary [Internet]. 2011. Available from: http://www.componentsource.com/products/componentone-olap/summary.html.

[17] A. Abdoon. ESRI Geodatabase [Internet].2014. available from: https://www.academia.edu/6387099/ESRI_Geodatabase.

[18] ESRI. What is ArcObjects? [Internet]. Available from: http://edndoc.esri.com/arcobjects/9.2/CPP_VB6_VBA_VCPP_Doc/shared/ao_foundation/what_is_ao.htm.

[19] ESRI. Weighted overlay (spatial analyst) [Internet]. Available from: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Weighted_Overlay/009z000000rq000000/