ISSN:0975-9646

A Smart Analytical Tool for Digital Governance with a Global Perspective employing Business Intelligence

Aditya Vinay Jumde
Department of Computer Engineering
SKNCOE, University of Pune
Pune, India

Sayali Makarand ghaisas
Department of Computer Engineering
SKNCOE, University of Pune
Pune, India

***Abstract*—In the world of automation, advanced technology and ever expanding need for standards in multi-dimensional economies has led to a large amount of data being made available in the form of raw datasets. However efficient data analysis and potential decisions that can be evaluated from the data are not meeting the corresponding needs. Traditional database from various international surveys is in raw format that is neither useful for general public nor to the governing bodies unless analyzed properly. The paper employs Business Intelligence techniques and tools to analyze various Governance sectors such as Agriculture, Economy and Growth, Health Sector, Entrepreneurship, Environment, Urban development etc. The proposed system shall provide a comprehensive analysis and reflect the key growth and development indicators across various economies. The aim is to transform the traditional way of representing the available data into a smart analytical format that would prove helpful for the governing bodies to make faster and intelligent decisions thus resulting in smooth governance.**

***Keywords*—Business Intelligence; Analytical System; Digital Governance; Governance; Decision Support system**

I. INTRODUCTION

Business intelligence is a comprehensive term that encompasses multiple dimensions crucial to the evolution of strategic decisions in various sectors and domains. BI is an umbrella term that includes the applications, infrastructure, tools and efficient practices that enable access and analysis of information and optimize decisions and performance. As stated by Howard Dresner, the father of Business Intelligence, it involves concepts and methods to improve business decision making by using fact-based support system.

Digital Governance, an effective and time efficient governance using the various digital media and tools is a promising field and would be integral and inseparable part of all governments worldwide in near future. Today the Governments of various nations publish open datasets of crucial information based on surveys in different sectors. This is a part of the 'Open Data Initiative' that aims to promote awareness, transparency, in the form of actual facts and figures among citizens.

However the utilization of these datasets is limited as the archived format is difficult to comprehend, analyse and formulate conclusive decisions. This restricts its usage fro end users' perspective. The target audience include general public, governing bodies, entrepreneurs, business organisers and international ambassadors.

The intended system proposes to draw conclusions strategically across various sectors to ease the process of decision making. It aims to provide an interactive platform/interface so that even general public can understand and evaluate the government processes.

II. SYSTEM OVERVIEW

A. COMPARATIVE ANALYSIS MODEL

There are various digital governance models such as Broadcast Model, Critical flow model, Interactive Service model, Comparative Analysis model etc. We had a research on the Comparative Analysis Model (CAM). CAM is the least used but one of the most high potential digital governance model for developing countries. The model can be used to compare various sectors and sub-sectors of governance to give analysis on various issues. The analysis can be of a great help for governing bodies to formulate budgets, inculcate good governance techniques, verify and validate their governing decisions etc. The CAM can also be used by investors or Entrepreneurs for new investment opportunities in foreign nations, Check availability of resources, ease with which the business can be started etc. Furthermore, the Comparative Analysis model can be used by International ambassadors to just glance through the current state of their respective nation in various sectors of Governance. The CAM can be developed with the help of the authorized and standardised data that is published by various nations under the Open Data initiative. It explores the immense capacity of Information and Communication Technologies (ICT) to analyse information sets with comparable information available in public and private domain.

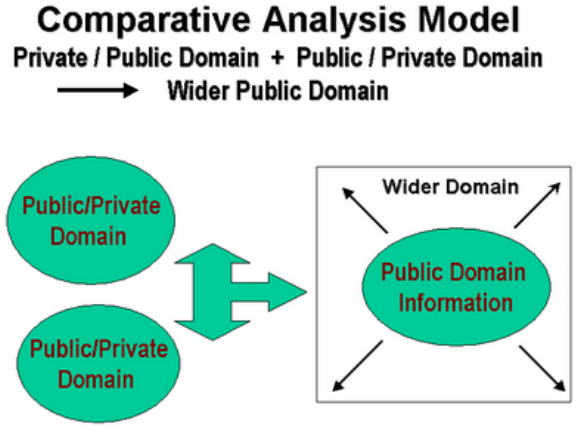

Fig. 1. Comparative Analysis Model

The Fig. 1 explains how the scope of information is increased if the public and the private domain are interlinked using the Comparative Analysis Model. Here the private domain can be governing bodies whereas the public domain can be general public.

The CAM can be applied in the following ways:-

- To learn from the available data and derive analysis for future policy making or strategic planning.
- To evaluate the current trends in various nations and find out flaws or shortcomings.
- By enhancing the back-end knowledge more precise strategies can be implemented.
- To evaluate the performance of a decision maker or a decision making body.

Developing countries could effectively use this Comparative Analysis model as global and local knowledge is easily available through the Open Data Initiative. This model is very much based on the existing information but requires the ability to analyze and bring out strong arguments which could then be utilized for self-governance. Rules and facts can be stored in the database from which strategic decisions and policies can be formulated by cumulative and comprehensive analysis [1]. This model may become ineffective in absence of a strong civil society where the people themselves want to enforce the decisions and the policies on the government [2].

B. Graphical Analysis

The graphs shall include standard bar charts, pie charts, map graphs that could help in visualization and comparison of the complex datasets. Exhaustive graphical analysis can be generated using standard tools such as the Online Analytical Processing (OLAP) [3].

C. Multi-Dimensional Data Model

A data cube allows the data to be modelled and viewed in multiple dimensions. It is defined by dimensions and facts.

By providing multi-dimensional data views and pre-computation of summarized data, data warehouse system provides inherent support for OLAP operations such as Drill-down and Roll-up. These are used to present data at different levels of abstraction. It allows multiple combinations of dimensions at various levels of granularity in data mining to use intricate patterns to represent knowledge [4].

Broadly, dimensions are perspectives / entities with respect to the records that are kept. These are stored in the dimension table. Facts are numeric measures or quantities used to analyze relationship between dimensions. We will see the application of how the database can be designed to provide a comparative analysis in the following example.

III. IMPLEMENTATION – EXAMPLE

A. EXAMPLE 1

For our proposed system the multi-dimensional data model that we will be implementing is the Star Schema rather than the Snowflake model. The star schema can be further expanded to form the fact constellation schema.

Let us consider the Agricultural Sector for our analysis, one of the most important sectors of governance. We will see how the Star schema can be applied to develop a Comparative Analysis Model and formulate strategic decisions or decide policies. In the Agricultural sector we have considered four sub-sectors. Agricultural land (% land area), Arable land (% land area), Agricultural Machinery (quantity) and the Food Production Index (FPI).

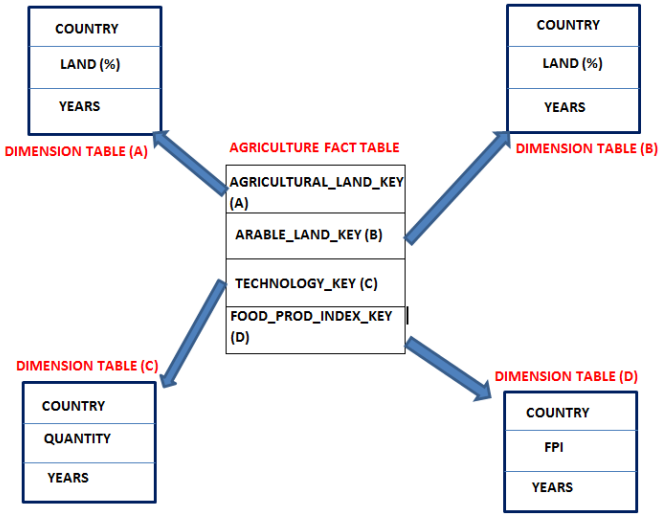

Fig. 2. Star Schema For the Agricultural sector.

Fig. 2 shows how the star schema can be used in the Agricultural sector. In this example the above four sub-sectors are used to develop a CAM under the main sector of Agriculture. Now let us consider the datasets in a single year (historical data) of various nations in the above the above sub-sectors.

- INDIA

1. Agricultural land (% land of total area) = 58%
2. Arable land (% land area) = 52%
3. Agricultural machinery, tractors etc. = 31,041
4. Food Production Index(out of 100) = 32.18

- GERMANY

1. Agricultural land (% land of total area) = 55.49
2. Arable land (% land area) = 35.01
3. Agricultural machinery, tractors etc. = 1027884
4. Food Production Index(out of 100) = 68.03

- UNITED STATES OF AMERICA

1. Agricultural land (% land of total area) = 48.86
2. Arable land (% land area) = 19.72
3. Agricultural machinery, tractors etc. = 4690000
4. Food Production Index(out of 100) = 45.45

The above data compares datasets of three different nations considering the above four sub-sectors [5]. Using comparative analysis the datasets clearly indicate that the better the agricultural technology the better the food production index (Food production index covers food crops that are considered edible and that contain nutrients. Coffee

and tea are excluded because, although edible, they have no nutritive value. [5]). The arable land can be used to its maximum capacity with the help of good agricultural technology. The data shows that the FPI of Germany is almost double that of India despite the fact that arable land in India is more than that of Germany.

Thus the system shall provide a conclusive decision that India can do much better in the agricultural sector with its valuable amount of resources if sufficient amount of finances are invested in better agricultural technology. The governing bodies can thus use the proposed system to validate their decision on the basis of true facts and figures. Thus the current scenario of a nation can be clearly visible in a particular sector with the above proposed CAM. Further the strategic decision that the system will infer will further strengthen the future decisions.

B. EXAMPLE 2

Another example that can be stated is with a perspective of self-governance. A lucrative & high potential field in which the system can be extensively employed is that of entrepreneurship. Large, medium & even small scale cottage industries exploring the overseas expansion opportunities can base their decision upon a number of different parameters. These include factors like economy of the host country, business density, and legal procedures linked to business permits including electricity, raw materials & permits, land & tax norms. Other favorable conditions like geographical location, political scenario, local labor & manpower should also be considered.

The system shall make a comprehensive analysis of all these factors independently and their interactions with one another to draw up conclusions on the overall favorability of starting the business venture or manufacturing plant in the said region. This favorability can also be given on a scale of 1-100 to compare the relative opportunities in various different regions of one country or even different countries & geographical locations.

IV. CONCLUSION

The system prototype can be implemented in several broad sectors as well as in niche domains as illustrated in the above examples. The Kimball life cycle model can be employed to execute the various stages of implementation [6]. It can be expanded further to include even parallel, seemingly unrelated fields. The interaction among dissimilar indicators & their analysis can also be employed to explore even the realms of predictive analytics. Predictions in the form of direct line regressions can be drawn up from the graphical representations [7]. It is represented in the form of following equation:

$$\text{Prediction} = a + b * \text{Predictor}$$

Based upon these values even certain future strategic decisions can also be taken successfully. This is within the scope of future enhancements.

REFERENCES

[1] Maria Antonina Mach and Abdel-Baheed M.Salem, “Intelligent Techniques for Business Intelligence in Health Care”, 2010 IEEE 10th Annual Conference on Intelligent System Design and Applications,DOI: 10.1109/ISDA.2010.5687209

[2] Website reffered : http://digitalgovernance.org/

[3] Krneta, D. ; Radosav, D.; Radulovic, B., “Realization Business Intelligence in Commerce using Microsoft Business Intelligence” Intelligent Systems and Informatics, 2008. SISY 2008. 6th International Symposium on DOI: 10.1109/SISY.2008.4664943

[4] Jiawei Han, Michelin Kamber, Jian Pei, “Data Mining”,2012 Morgan Kaufman,Elsevier publication

[5] Website reffered : http://data.worldbank.org/.

[6] Ralph Kimball,Margy Ross,Warren Thornwaite,Joy Mindy,Bob Becker, “The Data Warehouse Lifecycle toolkit(2nd edition) ",2009 Wiley Publication.

[7] Alex Berson, Stephen J. Smith, “Data Warehousing,Data Mining and OLAP ”,2012