

OPEN GOVERNEMENT DATA MANAGEMENT AND TRANSPERANCY

Amelina Ahmeti ¹, Albina Basholli ², Brunilda Gjini ³

¹University of Agriculture of Tirana, Albania. E mail: amelinaahmeti@ubt.edu.al

²University of Agriculture of Tirana, Albania. E mail: albinabasholli@yahoo.com

³ University of Agriculture of Tirana, Albania. E mail: bgjini@ubt.edu.al

Abstract

Open Government is a worldwide movement which allows citizens to access its information and data. It increases interoperability, collaboration, participation and responsibility. The basic idea of Open Government is to establish a modern cooperation among politicians, public administration, industry and private citizens. One of the major factors on creating an open government is sharing its data. Related with this issue, I find this movement very important to be applied in Albania. The research study is focused especially: The reasons and the benefits of creating open government data in Albania. As there is a huge volume of government data, there is a need of management, administration, aggregation and process on it. The population find it difficult to gain and study all the required information. This movement of open government data increases transparency. The citizens can observe, check and control for example: how their government has spent the budget. According to their preference, they can check if the budget was spent properly and verify different details. As example, we can mention the election period in Albania. According to the above description, there can be studied these problems: There is a huge volume of open government data and the users are not able to discover directly their desired information. The process of this data should be presented to the user. There is needed computer science knowledge in order to create software that processes data quickly. After all government data is managed, administered and processed, the next important issue is presenting it to the population. This is very important as it strengthens the transparency of the government to its population. Not only that, but also the citizens can study how their government manage the profits from taxes and the expenses. Based on it, they can decide if next time the actual government deserves their vote or they need to change it. As result, the level of transparency, releasing and commercial value also the participatory governance, will be increased and will promote better governance which is really needed in Albania.

Keywords: *Open data, Governmental data, Cure, transparency*

Introduction:

Open Government is a worldwide movement which allows citizens to access its information and data. It increases interoperability, collaboration, participation and responsibility. The basic idea of Open Government is to establish a modern cooperation among politicians, public administration, industry and private citizens.

One of the major factors on creating an open government is sharing its data. As result, it can be available to everyone. Some people are interested, if their government has spent the budget properly. There are websites which shows the expenses of the government per year. They can study these websites and gain the right information. Related with this issue, I find this movement very important to be applied in Albania.

The term open data is referring to all data which is freely available to everyone without restrictions and it can be used, reused, shared and republished by them according to their needs and desire.

Open data is focused more on no textual information and knowledge such as maps, genomes, connectives, chemical compounds, mathematical and scientific formula, medical data and practice, bioscience and biodiversity etc.

There is a great importance on open data when it represents governmental information and in this case, it is called *Open Government Data*. Open government data is very important to show the transparency of the government to its population and as result, it will strengthen democracy. It also declares its work to the citizens that have elected it.

Based on Open Government Data movement, we find important to study and apply this movement in Albania. There is a need in Albania and in most of Balkan countries of representing government data to citizens. Government data should not just be available to anyone but also there should be studied a methodology that it can be processed and aggregated. As result, the population can gain the right information without spending much time according to their needs.

In this case, open data will increase transparency of the government to its population and strengthen democracy. Also, it will bring better governance and more responsibility from both sides between government and population.

Aim of study/research

Based on the topic stated above, we are inspired to research open government data and hopefully applying it in Albania in the future. The research study is focused especially:

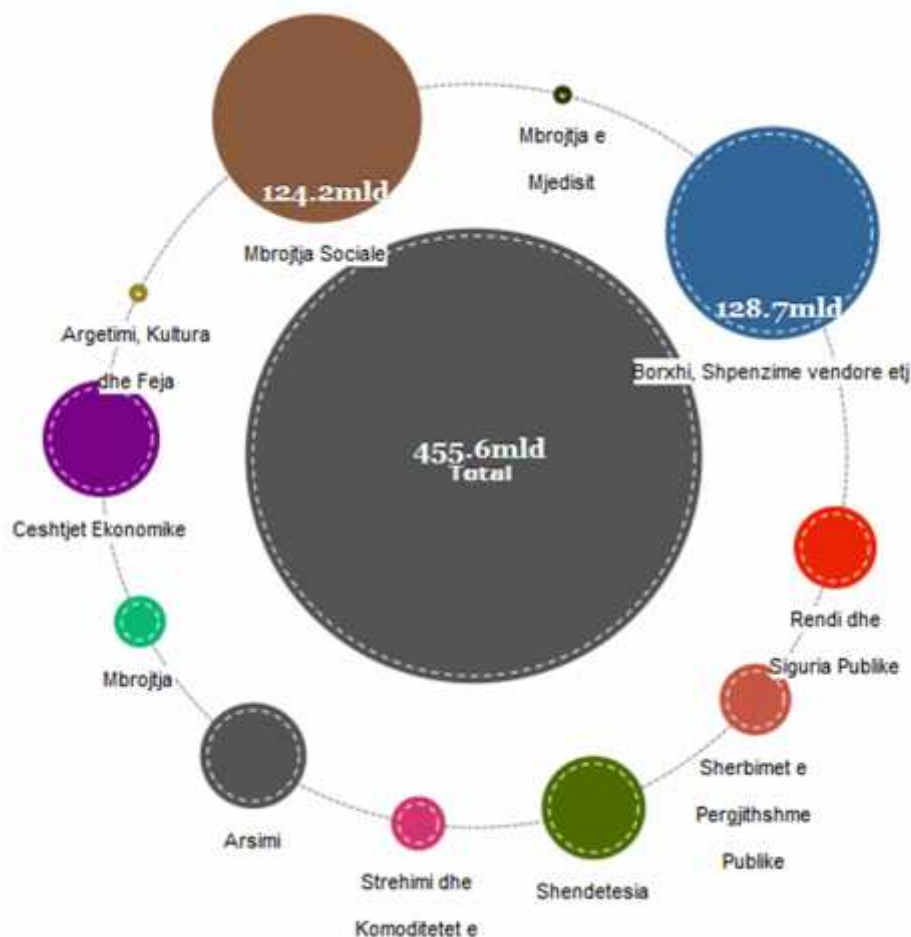
- The reasons and the benefits of creating open government data.
- As there is a huge volume of government data, there is a need of management, administration, aggregation and process on it. The population find difficult to gain and study all the required information.
- There should be found ways of aggregation and process which can be applied to Albanian government data. The aggregation must be based on algorithm, which can be applied by software.
- This movement of open government data increases transparency. The citizens can observe, check and control how their government has spent the budget. According to their preference, they can check if the budget was spent properly and verify different details. As example, we can mention the election period in Albania. Every party represents its program and if it wins the right percentage of people votes, it will govern the country. After one or two years, citizens can check referring to open government data, if this government has fulfilled its program or no. They can also criticize and cooperate with it. Thus, they can see how their government has worked to be prepared to the next election period.

Methods

Our research objectives are based on our European experience studies but also on observation of the situation in Albania. The population was asked in different cities related with which part of the governmental data is important to them and how they would like to be represented to them. There is a huge volume of governmental data and it will be not comfortable to observe all the governmental data without processing it. According to the above description, there can be studied these problems:

1. There is a huge volume of open government data and the users are not able to discover directly their desired information.
2. The process of this data should be presented to the user. There is needed computer science knowledge in order to create software that processes data quickly.

Based on the above problems, we found the inspiration to study the methods of processing open data and how they are applied also in Albania. The concept of open data is applied also in Albania. This kind of data is published at <http://open.data.al>.



The above picture shows the aggregation of data which represents the spending of Albanian government.

After all government data is managed, administered and processed, the next important

issue is presenting it to the population. This is very important as it strengthens the transparency of the government to its population. Not only that, but also the citizens can study how their government manage the profits from taxes and the expenses. Based on it, they can decide if next time the actual government deserves their vote or they need to change it.

As it is showed above, the aggregation of data is very important while it gives the conclusions and the facts of all the details of the data. It is needed a software which aggregates the data and process it. It will show the results to the people. The software should be based on CURE Algorithm which processes the data and gives the result.

Cure Algorithm

Cube is a multidimensional database with aggregated data. Let take as example a table which the columns of it are with A, B, C, and M.

A	B	C	M
1	1	1	10
1	1	2	20
2	2	3	40
3	2	1	45
3	3	3	35

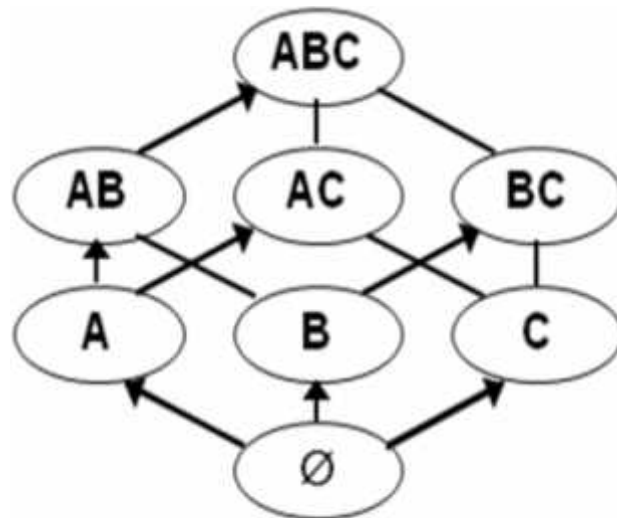
A, B, C are called dimensions while M is a measure if we decide to aggregate M. In a data cube, two basic relations among cells are roll-up and drill-down. Drill down means to move from summary information in detailed data by focusing on something while the roll up is the opposite of drill down which means from detailed information you aggregate and create summary information. In our case, we propose to use CURE for constructing complete data cubes over large datasets mostly for these reasons:

- It creates a novel lattice traversal scheme and it is optimized partitioning method.
- It supports relational storage model schemes for all forms of redundancy.
- It creates a cube with the highest density dataset.
- It is a pure ROLAP technique to construct the cube of the highest density datasets.
- It is fully compatible with relational model which makes it easy to implement over any existing search relational engine. Also, it is designed not only for at datasets but also for data whose dimensions are organized in hierarchies.

It gives the possibility to create cubes for large datasets whose size can exceed the size of memory.

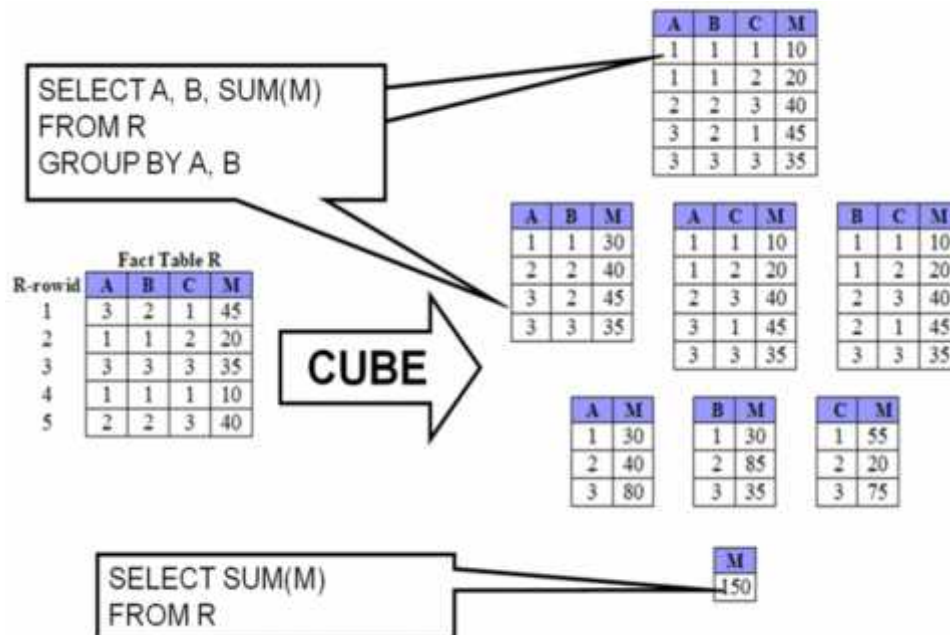
One important fact is that it stores cubes in a compact format from eliminating all kinds of the redundancy from the final result.

Cure helps in creating data cube by aggregation in order to group together similar tuples. This was the inspiration of Gray who proposed a pre-computation of the data cube which is a data structure that is made of all the results of group by aggregating queries on all possible combinations of dimension attributes over a fact table. A simple representation of the data cube is cube lattice like below.



The above picture is an example of cube lattice of a fact table R with three dimensions (A, B and C). Every node represents a group by query and is labeled with its grouping attributes which is a subset of the combination of the dimensions. If we denote the number of dimensions of a fact table with D, then the number of all cube lattice nodes is 2^D . The implementation method computes each node and stores the result with exponential time and space complexity. This is a naive idea of cube and the first one on data cube creating. Just creating 2D tables is a simple data cube but also it contains redundant tuples. Also, just using this simple way of creating data cube there is no optimization and no efficiency. It needs much memory space. Consider again a fact table with D dimensions. If we denote the number of levels of the i-th dimension with L_i the number of cube nodes is given by the product for $i=1$ to d $\prod (L_i + 1) \gg 2^d$.

As example, how the naive implementation of Cube is created and its levels are shown below with a real fact table:



It is taken a fact table R with 3 dimensions and one measure as an example in this case. As a result, there are eight tables created in the end which form the data cube of the fact table R. This is a simple way of creating data cube but it is not practical and efficient as it needs a lot of data space when the fact table is a huge data. There is a need for an algorithm which gives a better solution. The algorithms which construct data cube are based on relational data. According to this algorithm, software can be created to process huge volume data. It will be very useful also in Albania.

Results:

Transparency strengthens the democracy and the responsibility of government to its population. Also data management will reduce time and make information to be gained properly from the population. As all data of the government will be available to everyone, it will increase its responsibility to use the budget properly. The people will not find difficult to gain the right information from this data. Thus, the software of processing and managing will help to understand and gain the right conclusion.

The open government data brings a new innovative movement and knowledge sharing, which will help to build a corporation between the citizens and government, among politicians, public administration, by enabling more transparency, democracy, participation and collaboration.

Conclusions:

The efficient management and administration of data is critical not only to the Albanian economy but for each country. In particular, the importance of sharing data will bring better decision-making and management from the government, ensure better performance and promote better finance mechanisms. There is a huge collection of government data that can be public to its population. This way the population can discover if their government has taken the right decisions. Otherwise, they can criticize not only personally but also by their vote in election. There is needed a software which will help on the aggregation of all data in order to help the population to gain a clear conclusion. As result, the level of transparency, releasing and commercial value also the participatory governance, will be increased and will promote better governance which is really needed in developed countries.

References:

- Florian Bauer, Martin Kaltenböck: Linked Open Data: The Essentials A Quick Start Guide for Decision Maker*
www.wikipedia.org
- Jrgen Kett Jan Hannemann. Linked data for libraries. 2010.*
<http://www.guardian.co.uk/news/datablog/2010/nov/19/government-spending-data>
<http://www.unescap.org/pdd/prs/ProjectActivities/Ongoing/gg/governance.asp>
<http://siteresources.worldbank.org/EXTPUBLICSECTORANDGOVERNANCE/Resources/PGPEbook121509.pdf>
<http://www.ids.ac.uk/files/dmfile/IETAAnnex4NatResGovMejiaAcostaFinal28Oct2010.pdf>
http://www.cs.ucsb.edu/~veronika/MAE/cure_efficientclusteringlargedatabases_guha2001.pdf