A Cloud-Based Data Integration Framework for E-Government Service

Rashed Salem², Samah Abd Elrahman^{1*}, Mohammed Hassan¹

¹ Information Systems Dept. Faculty of Computers and Informatics, Zagazig University, Egypt.

² Information Systems Dept. Faculty of Computers and Information, Menoufia University, Egypt.

* Corresponding author. Email: samahelazazy@zu.edu.eg

Manuscript submitted October 10, 2017; accepted January 8, 2018.

doi: 10.17706/jcp.13.8.913-923

Abstract: Today, every government is seeking to offer the best services to its citizens to improve competence, translucence, responsibility, access, and other services at low cost. Most governments are using communication technology and Web-based applications in most sectors to provide and ease the services between the state organizations and citizens. However, the government faces a variety of challenges towards delivering its e-services including data heterogeneity, data duplications, infrastructure scalability, IT budgets and costs, etc. These challenges always lead to failure of e-government projects. Currently, there is a strong direction for most applications to move onto the Web, particularly in the cloud. In this paper, we propose a cloud-based framework for integrating data from different data sources of government sector's databases into a unified repository, which deployed into the Cloud. Therefore, the proposed framework delivers several e-services to the government, citizens, and business communities. The proposed framework tackles several challenges of delivering e-government services. Due to warehousing data, the framework enables business intelligence applications to provide several services both for supporting decision makers and provide daily routine services for citizens. Finally, the proposed framework is validated using syntactic operational data for different seven governmental sectors, such data are integrated and warehoused into galaxy schema. The data warehouse is deployed into Cloud and different OLAP queries are performed for checking the consistency and integrity of data.

Key words: Business intelligence (BI), data warehouse, data integration, DW architecture, e-government, cloud.

1. Introduction

The main aim of the e-governments is to deliver the governmental services to the customers online, *i.e.*, citizens, enterprises, public organizations [1]. E-government was built considering the Web and data warehouse technologies for maintaining the huge amounts of the data. Data warehousing (DW) and Business Intelligence (BI) are used for supporting decision-making. BI employs set of applications and technologies for aggregation, storing, analyzing, and delivering data [2]. BI applications exploit the stored data in the data warehouse or data marts which were extracted from the operational databases [3]. A data warehouse is a repository that stores a comprehensive information from several databases to fulfill decision makers' requests [4]. Furthermore, the online analytical processing (OLAP) uses the data warehouses to browse different kinds of data satisfying the needs of the users from different perspectives [5].

Data warehousing operations involve integrating, storing and analyzing business data [2]. Data integration is regarded as the critical and time-consuming task for DW/BI applications. It brings the data from diverse

resources, unifies their format and delivers them into a centralized repository for enabling analytic tasks. In the data integration platform, data are filtered and combined with performing extraction-transformation-loading (ETL) tasks [6].

Indeed, governmental sectors face many challenges in exploiting their e-governments' services and their infrastructure in order to improve their current services or develop new services. These challenges include the scalability of data and hardware resources, IT costs and budget, software licensing, systems integration, duplication of data, migration difficulties, performance, disaster recovery, security and privacy, which may lead to failure of e-governments' projects. Cloud Computing plays a vital role to cope with these challenges. Therefore, most of the today's organizations move into the Cloud and migrating their data and services.

In this paper, we propose a framework to integrate governmental sectors' operational databases with the aim of establishing a complete information environment, which able to intake the rules of national data to fulfill complementarity between different sectors of the country. The proposed framework has several influences such as:

- It allows providing modern and sophisticated services for the government sectors and individuals in an integrated manner,
- It contributes in decision support in all governmental sectors through the audited statements that are linked to the national identification number of the citizen,
- It helps in the formulation of policies on the basis of careful analysis of information, as well as the strengthening of the overall vision for planning and address duplication in databases,
- It contributes in the fight against various forms of corruption, for instance, the purification of the beneficiaries lists of the social support to ensure the delivery of the financial support for those who deserve it and saving the dues of the country.

Finally, the proposed framework is deployed into the Cloud providing several benefits such as reduced cost, increased storage, scalability, accessibility and availability, privacy and security [7].

The remainder of this paper is organized as follows. In Section 2, the literature of e-government services in different countries is reviewed. Section 3 presents the framework for data integration through validating a case study of providing e-government services. Section 4 develops the proposed framework through building data warehouse and deployment this services onto the Cloud. Section 5 discusses the benefits of the proposed framework. Finally, we conclude and highlight future trends in Section 6.

2. Related Work

Today, e-Governments are based on comprehensive resources of related data which are integrated from different sectors [8], [9].

Data integration approaches. In general, there are two principle groups of approaches to deal with integrating data, *i.e.*, virtual views (mediation) and materialized views (warehousing). In a virtual view approach, data integration is based on lazy fetching. Data are integrated by creating a virtual view on the different data sources. These sources are queried at runtime in order to reach and integrate their data [1]. In a materialized view approach, data is integrated in advance by fetching data from different sources, into one centralized database. Pertinent data are filtrated from data sources and pre-stored into a warehouse that can be later queried by users [2].

Furthermore, most of the e-governments are based on an open architecture, integrating heterogeneous resources with the united interface, logically centralizing and enabling unified management for these resources [8]. In literature, there are several case studies of e-governments which based on data warehousing and business intelligence technologies.

Case studies: In Turkey, Kurfal *et al.* utilize UTAUT (Unified Theory of Acceptance and Use of Technology)

model that enriched with trust of Internet and trust of government factors to explore the determinants of e-government adoption of Turkish citizens, and to determine factors that play key role in citizen's decision to use e-government services [10]. Therefore, the authors validate this model through a survey conducted with Turkish citizens who are from different regions of the country and collected answers through purposive sampling and the replies were evaluated with the SEM (Structural Equation Modeling) technique.

In India, Bhanti *et al.* propose e-governance execution for higher education system using data warehousing and data mining techniques [11]. The developed framework utilizes data mining techniques to tackle the problem of accumulated massive data which was far from decision and forecasting. In Oman, Al Salmi *et al.* explore the factors that drive the e-government implementation and adoption that would affect the e-government performance to decrease and remove as much as possible the failure of e-government projects [12]. Therefore, the authors evaluated the e-government sustainable development in the Oman through secondary data which were obtained from various sources like the United Nations data of e-government execution maintainable development and readiness. In Tanzania, Rukiza *et al.* present a case study for securing e-government services to compare the adoption process, both private and public organizations was judgmentally studied [13]. The study uses a mixed research method that combines the use of both qualitative and quantitative methods for questionnaire preparation, data collection, data processing, and data analysis. In Hong Kong to develop the information and communications technology infrastructure, Lee *et al.* utilize XML data interoperability framework to facilitate e-government data standardization through XML schema design methodology and the essential schema management infrastructure [14]. Business enterprises and other governments can readily adapt it to establish their own data interoperability. In Japan, Chatfield *et al.* introduce a case study of Japan's National Tax Agency (NTA) and its core e-government initiative, *i.e.*, "e-Tax" [15]. It is an online income tax system which enables filling and propulsion services for individuals and companies. It also allows studying the relationship among public service repair through e-government and actual government performance. In China, Yang *et al.* propose a service-grid-based framework for e-government interoperability, namely eGovGrid, which facilitates resource sharing among "horizontal" organizations and interoperability between "vertical" subsystems of e-government [8]. Therefore, the authors validate the framework through a case study of e-government with seven departments, two countries, and five companies to exchange economic information.

3. The Proposed Framework

In this paper, we propose a framework to build an integrated database with the aim of establishing information environment to achieve complementarity between disparate sectors of the government. The proposed framework consolidates data from heterogeneous and disseminated sources into a unified repository as in data warehousing and provides users with a consolidated view of these data. In this framework, data are integrated from several governmental sectors such as utilities, financial, healthcare, civil and legal sectors in order to deliver different e-government services. Moreover, the framework can be exploited for decision-making and solve some critical problems for state government such as determining the people who deserve financial aids and backing from the government.

In this framework, integrating data into a unified repository faces several challenges including sources schema heterogeneity and interoperability. The heterogeneity of sources schema is tackled by schema mappings process. To solve this issue, sources schemas are mapped into a target schema through a source to target mapping. Another problem is interoperability, to solve this problem common data among public government sectors are effectively employed to match citizen records. Fortunately, several citizen records for different governmental sectors are identified using NIN (National Identifier Number). Therefore, NIN is employed as a unified identifier for different records integrated from several governmental sectors for the

same citizen. The overall architecture of the proposed framework is shown in Fig. 1. The architecture is composed of three essential phases which are data integration, warehousing, and data analysis. Data integration is the most crucial and effort consuming phase, it involves three main processes which are Extraction, Transformation, and Loading (ETL). ETL is a process in data warehousing responsible for extracting data from diverse sources, then transform them to be represented and loaded into a target repository. Several tasks are conducted in this step include summarizing, updating, sorting, filtering, merging, distributing, standardizing values, look up terms, data type conversion, aggregating, auditing, counting, and slowly changing dimensions. Data warehouse of the proposed architecture is the union of all data marts. A data mart is a simple shape of a data warehouse that is focused on a single subject such as finance, healthcare or other utilities. Further discussion of the architecture is introduced in following sections.

Fig. 1. Data warehousing framework for e-government services.

3.1. Conceptual Design of the Proposed Framework

A Data warehouse is generally modeled using dimensional fact models such as star schema, snowflake schema, a flat schema, constellation schema and galaxies schema. Such schemas contain one fact table (e.g., star-schema) or many fact tables (e.g., constellation and galaxy schemas) representing business measures and surrounded by dimension tables which represent business subject areas [16].

The galaxy schema which is used in this framework is a collection of snowflake schemas. In which there are many fact tables that may share some of their dimension tables as shown in Fig. 2. The galaxy schema is employed since there are many businesses are interested including financial (e.g., banking, insurance, and tax), utilities (e.g., electricity, water, and gas), healthcare, and vehicle data. Such businesses are dealt as facts which share the same dimensions, e.g., citizen information, geographical location information, date and time information. In galaxy schema, at the level of entity sets, each fact entity is associated with more than one dimension entity, and the same dimension entity can be associated with multiple fact entities as shown in Fig. 2. However, the relationship among dimension instances to fact instances is the one-to-many relationship like star-schema. Thus, the galaxy schema is the most appropriate model which supports some advantages to get the data easily which sharing among different data marts and decreasing heterogeneity through assisting database developers to determine, reuse, customize, and advertising related data [17].

Fig. 2. Galaxy schema for E-government.

3.2. Structural Design Issues

A Data warehouse is the central repository that stores data from different sources applying multidimensional model where the main concepts related to decision-making processes are stored as facts and their descriptions for analysis are stored as dimensions. Therefore, the structural design process of data warehouse consists of four main stages including 1) choice of business processes, 2) declaration of granularity, 3) determining the associated dimensions, and 4) identification of the measures [18].

In this work, the implementation of the proposed data warehouse framework follows the basic procedures mentioned above. In the current practice of e-government services, there are many main business processes which are related to e-services delivered to citizens and/or services that support decision-making for the government. The second process is the declaration of granularity, which is a detailed data level for business processes, for example, determining consumption of water, gas, and electricity every week, month, quarter and/or year. The third process is choosing the dimensions; dimensions tables contain descriptive attributes and a primary key column that related with a fact table. The dimensions tables of e-government case study are coming from several sectors including for example date, loan, policy, car accident, citizen, location, insurance, place therapy, arrears, discounts, receipt, wealth, activity, diagnosis, disease hierarchy, operations, water company, electricity company, branch, license, violations, vehicle, credit card, certification, account and bank. Finally, the last process is the identification of the fact measures, fact table indicates a business process and contains measures that are the additive attribute and foreign key from dimension table. The fact's table of e-government case study includes finance transaction amount, policy transaction cost, vehicle licensing cost, electricity consumption, water consumption, healthcare records and tax receipts.

4. Developing the Proposed Framework

Developing the proposed framework is based on building a data warehouse to collect data from sources of the different Egyptian governmental sectors. Then, the data warehouse is deployed on Cloud using IBM Bluemix cloud.

4.1. Building the Data Warehouse

Building a data warehouse has several steps for integrating the operational data and storing them in the data warehouse. The major steps of building data warehouse are discussed as follows:

1. Data collection: Because data in the data warehouse are extracted from different sources, therefore the first step in data collection is to decide which sources are interested and determine the relevant data. There are exist many problems need to be resolved for collecting data from different data sources. The major problem is the heterogeneity, which classified into four categories: 1) structural heterogeneity due to diverse data models; 2) syntactical heterogeneity encompassing different languages and data representations; 3) systemic heterogeneity including hardware and operating systems, and 4) semantics heterogeneity encompassing distinct concepts and their interpretations. As mentioned earlier, the heterogeneity issues are tackled by schemas mapping between source schemas and target schema. Data collection requires several connections to different data sources such as connections for Oracle, SQL Server, dBase, Access, Excel, etc.

2. Transformation and Cleansing: This step is to clean data which extracted from data sources in the staging store. It resolves problems such as schema matching, schema integration, and schema mapping. Generally, achieving this phase is so hard and time-consuming. Therefore, it can be done with the assistance of ETL tools. The ETL operations are employed to populate the data warehouse and clean records from one or more data source, such operations are discussed as follows:

A. Extraction: This process extracts relevant data from the governmental data sources such as financial data sources (e.g., banking, insurance, and tax), utilities data sources (e.g., electricity, gas, and water) healthcare data sources, vehicle data sources, etc. Extraction is usually performed using SQL commands. Other than the initial extraction, the only delta of changes from the last load is extracted. Deltas of changes can be computed using different methods including updated timestamps and/or database triggers.

B. Transformation: This process involves several sub-processes including data cleaning, data quality, building keys, and transforming data into the multi-dimensional format. These processes are performed in a temporary area, i.e., Data Staging Area (DSA). Data cleaning and data quality ensure data deduplication, handling missing values, filtering attributes, handling misspellings, string manipulations, and there is no contradictory data. Moreover, transformation process includes data type conversion (e.g., date/time format conversion), normalization/de-normalization upon sources and target, text coding, matching and building keys.

C. Loading: This process populates the extracted and transformed data into the data warehouse. Both dimensions and facts of data warehouse model are loaded concurrently. Referential integrity is ensured during the loading process. Index on tables slows the loading process, thus indices are dropped and rebuilt after loading. Once loading processing is finished, end-users can perform efficient querying and analysis such as OLAP query, reporting, and data mining.

3. Multidimensional data modeling and data cubes: multidimensional data model is one of the most characteristics of the data warehouse. However, the multidimensional model is implemented using relational tables. Star, snowflake, constellation and galaxy schema are examples of multidimensional models. The proposed framework applies galaxy schema to fulfill requirements of e-government services.

Unlike relational database systems, the data are visualized as multidimensional cubes rather than tables. Each dimension of the cube represents the business concept or concept hierarchy which may contain multiple levels of granularity. Using OLAP tools, data cubes can be interactively manipulated for data retrieval and analysis in a user-friendly way. The multidimensional model is implemented using Relational OLAP (ROLAP) whereas facts are stored in fact table as one column for each measure and dimension, each dimension is stored in dimension table, and SQL is used for querying. The major advantage of using ROLAP is leveraging investments in relational technology, its scalability for several billions of facts, its flexibility and its easier change of the design.

4. Data analysis and aggregation: OLAP operators are used to formulating and executing user queries navigating data from many perspectives. OLAP operators are implemented as SQL Group By extensions and SQL window functions. Examples of SQL Group By extensions include operatives such as roll-up, cube and grouping set. The roll-up operation aggregates detailed measures from the most detailed level up to Grand (all) level for a specific hierarchy. whereas the cube operation aggregates all possible combinations as shown in Fig. 3, whereas this cube shows the average consumption of electricity (kWh) for different classes of consumers in each governorate monthly. Therefore, this cube is defined by three dimensions *i.e.*, consumers (citizens), date and location, whereas citizens are classified into classes according to their income. Unlike cubes, the grouping set aggregates just the needed groupings. Other operations include drill-down as the opposite of roll-up to aggregate data from grand level to most detailed one, slice for performing selection on one dimension, dice for performing selection on two or more dimensions, and pivot to rotate the dimensions of cubes to offer an alternate presentation of the data. Moreover, SQL window functions or analytic OLAP functions let us express many important query kinds containing time series analysis, ranking, percentiles, averages of movement, lagging, leadership, data rate, linear regression and progressive sum.

Fig. 3. An example of queried cube.

Several OLAP queries for different purposes can be conducted using IBM Cognos as follows.

Example 1. For each governorate determines average consumption of water and electricity monthly for citizens whose salaries between 1000 and 2500.

SELECT

L.governorate, D.month, trunc(avg(WF.current_reading-WF.previous_reading)) AS "average_usage_water",
trunc(avg(EF.current_reading-EF.previous_reading)) AS " average_usage_electricity"

FROM

Citizen C, Electricity_services_f EF, Water_services_f WF, d_date D, d_location L

where

C.SSN = EF.SSN and C.SSN = WF.SSN and D.date_id = EF.date_id and L.loc_id = C.loc_id
and (c.salary between 1000 and 2500)

GROUP BY

CUBE (L.governorate, D.month)

ORDER BY

governorate ASC;

Example 2. In each governorate determines citizens with chronic diseases where their salaries less than or equal to 1500 and compute the cost of their treatment which government offers to them every year.

SELECT

L.governorate, DA.year, DH.diagnosis_key , DI.diagnosis_type, SUM(F.cost_of_treatment) AS "cost_of_treatment"

FROM

Citizen C, Health_care_f F, d_date DA, d_diagnosis DI, d_location L, Disease_h DH

where

C.SSN = F.SSN and DA.date_id = F.date_id and DI.diagnosis_key = F.diagnosis_key and L.loc_id = C.loc_id and
DI.diagnosis_key = DH.diagnosis_key and DI.diagnosis_type = 'chronic diseases' and C.salary<=1500

GROUP BY

ROLLUP(L.governorate, DA.year, C.salary, DH.diagnosis_key, DI.diagnosis_type)

ORDER BY

governorate ASC NULLS LAST;

Example 3. In each governorate determines the type and the average value of the loan which taken by each class of citizens annually as well as determining their activity types.

SELECT

L.governorate, D. year, C.Income_class, DL. type_of_loan , A. activity_type, DL.rate_of_interest_loans,
AVG(FF.loan_amount) AS loan_amount

FROM

Citizen C, Finance_transcation_f FF, Tax_receipt_f TF, d_activity A, d_date D, d_loan DL, d_location L

where

C.SSN = FF.SSN and C.SSN = TF.SSN and A.activity_type_key= TF.activity_type_key and
D.date_id = FF.date_id and DL.loan_type_id = FF.loan_type_id and L.loc_id = C.loc_id

GROUP BY

GROUPING SETS ((L. governorate, D.year, C.Income_class, DL.type_of_loan,A.activity_type,
DL.rate_of_interest_loans),(C. Income_class, DL.type_of_loan ,A.activity_type, DL.rate_of_interest_loans))

ORDER BY

governorate ASC;

5. Predictive analysis, presentation, and visualization: predictive analysis which applies data mining techniques (e.g., clustering, classification, association, regression, etc.), interactive dashboards, as well as visualization are forms of business intelligence which based on data warehouses. Visualization technologies display results for end users, usually in the form of reports, graphs, 3D, etc. Visualization helps people to understand the importance the data by putting it in a visual context. Once data are loaded in the data warehouse format, several business intelligence tools (e.g., IBM Cognos Business Intelligence, Oracle Business Intelligence Suite, etc.) can be applied to discover the variety of knowledge from the data warehouse.

4.2. Deployment of the Data Warehouse on the Cloud

The National Institute of Standards and Technology (NIST) defines four cloud deployment models, *i.e.*, public cloud, private cloud, community cloud and hybrid cloud. In public cloud, services are being available to clients via a third-party service provider, which provides an access control mechanism for their clients. In private cloud, data and operations are managed within the institution without any restriction of network bandwidth, infrastructure controlling, security and legal requirements. In a community cloud, it is managed and controlled by a group of institutions that have shared interests, security requirements or common tasks. Finally, a hybrid cloud is composed of more clouds, *i.e.*, public, private, and/or community. The most appropriate model for the e-government is the private cloud or community cloud model due to their strength of managing data security and privacy. The private cloud is suitable when data of all e-governmental services need to be centralized, thus it fits with data warehouse when implemented as a union of all data marts. However, the community cloud is suitable when the data of e-governmental services need to be distributed to several governmental sectors, each sector provides its services and shares common dimension data among other sectors. In a community model, each sector manages and controls its own data marts.

Deploying the proposed framework into the Cloud has numerous advantages. The Cloud data warehouse can distribute workloads across different nodes to provide more improvements over traditional databases. Cloud data warehousing can bring structured data from legacy on-premises data warehouses together with these newer big data sources. It automates provisioning, administering, monitoring in less time and much cheaper than the traditional on-premises data warehouse. Moreover, it emerges advanced analytics workloads that combine traditional, transactional data with data of different types. Their elastic computation, storage, reduced IT costs and budgets, infrastructure scalability, availability and accessibility, disaster recovery, auditing and logging, data and application redundancy elimination, security, privacy and on-demand data analytics are another advantages of cloud data warehousing. Furthermore, the speed of deployment, enabling big data, and accelerated analytics are major reasons for provisioning the cloud data warehousing services.

In this study, the data warehouse of the proposed framework is deployed on IBM Bluemix cloud and IBM Cognos BI is used as predictive analytics tools providing business intelligence solutions.

5. Benefits of the Proposed Framework

The proposed framework provides several benefits not only for the government but also for citizen and business communities which some of them are listed as follows:

Benefits for government:

- Support decision making in all governmental sectors through the audited data linked to the national identifier of the citizen,
- Combat various forms of corruption, as well as to purifying the lists of beneficiaries of the support to ensure its arrival to those who deserve it,
- Provide accurate data for various sectors to ensure the gathering of country dues,
- Consolidate the obligation, punishment, and transparency in societies,
- Integration of formal, informal economy and determination of the wealth, including real estate and other wealth,
- Address the huge social problems, such as health insurance, education that loses the state billions of dollars,
- Facilitating public administration operations yields effort and time savings,
- Increasing the productivity and growing economy,

Benefits for the citizen:

- The provision of services to citizens electronically, facilitating administrative procedures and eliminating the bureaucracy,
- Facilitating the exchange of data among government sectors, which results in the simplification of procedures and enhancing the performance of services provided to citizens,
- Delivering the various types of financial support to those who deserve them, through the development of the integrated information system,
- Providing services to citizens, including health insurance, vehicle and driver licensing, etc.,
- Improving services delivery quality (e.g., healthcare, education, ration, etc.),
- The accomplishment of social justice among citizens,
- Saving time, money and effort.

Benefits for business communities:

- Providing an attractive investment environment and strengthening the commitment to transparency and integrity of the country administrative system standards,
- Expanding communication channels,
- Increasing the transparency of decision-making process and information accessibility,
- Enhancing the cooperation between business communities and the government,
- Creating new business and work opportunities,
- Availability of databases for investment and investors.

6. Conclusion and Future work

In this paper, we have proposed an integrated framework for warehousing data from different data sources of government sector's databases in Egypt into a unified repository, which is deployed into the Cloud. The proposed framework delivers several e-services to the government, citizens, and business communities. This framework tackles some problems such as schema mapping and interoperability. Deploying governmental data and providing e-government services via cloud yields several benefits to government itself, citizens, business communities and the whole society.

In future, we plan to address other data sources and integrating their data in real-time, as well as applying the framework using real-world data rather than empirical data.

References

- [1] Breil, A., Hitzelberger, P., Da Silva Carvalho, P., & Feltz, F. (2012). Exploring data integration strategies for public sector cloud solutions. *EGOVIS/EDEM*, 271-278.
- [2] Salem, R., Boussaïd, O., & Darmont, J. (2013). Active XML-based web data integration. *Information Systems Frontiers*, 15(3), 371-398.
- [3] Ziegler, P., & Dittrich, K. R. (2007). Data integration — Problems, approaches, and perspectives. *Conceptual Modelling in Information Systems Engineering*, 39-58.
- [4] Marotta, A., & Ruggia, R. (2002). Data warehouse design: A schema-transformation approach. *Proceedings of 2002 IEEE International Conference of the Chilean on Computer Science Society, IEEE SCCC* (pp. 153-161).
- [5] Ghazali, D. P. S., Latip, R., Hussin, M., & Abd Wahab, M. H. (2015). A review data cube analysis method in big data environment. *ARPN Journal of Engineering and Applied Sciences*, 10(19), 8525-8532.
- [6] Geng, Y., & Kou, J. (2012). A design of data integration using cloud computing. *Advances in Future Computer and Control Systems*, 415-419.
- [7] Mirzaei, N. (2008). Cloud computing. *Pervasive Technology Institute Report, Community Grids Lab, Indiana University*, 1-12.

- [8] Yang, D., Han, Y., & Xiong, J. (2007). eGovGrid: A service-grid-based framework for e-government interoperability. *Integration and Innovation Orient to E-Society*, 2, 364-372.
- [9] Kurfalı, M., Arifoğlu, A., Tokdemir, G., & Paçin, Y. (2017). Adoption of e-government services in Turkey. *Computers in Human Behavior*, 66, 168-178.
- [10] Bhanti, P., Kaushal, U., & Pandey, A. (2011). E-governance in higher education: Concept and role of data warehousing techniques. *International Journal of Computer Applications*, 18(1), 0975-8887.
- [11] Sarrayrih, M. A., & Sriram, B. (2015). Major challenges in developing a successful e-government: A review on the Sultanate of Oman. *Journal of King Saud University-Computer and Information Sciences*, 27(2), 230-235.
- [12] Karokola, G. R., Kowalski, S., Mwakalinga, G. J., & Rukiza, V. (2011). Secure e-government adoption: A case study of Tanzania. *European Security Conference* (pp. 18-32).
- [13] Lee, T. Y., Yee, P. K., & Cheung, D. W. (2009). E-government data interoperability framework in Hong Kong. *Proceedings of 2009 IEEE International Conference on Interoperability for Enterprise Software and Applications China, IEEE IESA* (pp. 239-244).
- [14] Chatfield, A. T. (2009). Public service reform through e-government: A case study of 'e-Tax'in Japan. *Asymptotic and Computational Methods in Spatial Statistics*, 209.
- [15] Wang, J., & Kourik, J. L. (2015). Data warehouse snowflake design and performance considerations in business analytics. *ASBBS Proceedings*, 22(1), 472.
- [16] Mork, P., et al. (2009). Galaxy: Encouraging data sharing among sources with schema variants. *Proceedings of 2009 IEEE International Conference on Data Engineering, IEEE ICDE* (pp. 1551-1554).
- [17] Park, T., & Kim, H. (2013). A data warehouse-based decision support system for sewer infrastructure management. *Automation in Construction*, 30, 37-49.
- [18] Liu, X., & Luo, X. (2010). A data warehouse solution for e-government. *International Journal of Research and Reviews in Applied Sciences*, 4(1), 101-105.

Rashed K. Salem received his Ph.D degree in computer science from the University of Lyon 2, France in 2012. He has been a member of complex data warehousing and on-line analysis research group within the ERIC laboratory. He is an assistant professor at Faculty of Computers and Information, Menoufia University, Egypt. His current research interests mainly relate to database management systems, business intelligence (BI), data mining, data warehousing and big data analytics.

Samah Abd Elrahman is a teaching assistant in Information System Department in Faculty of Computers and Informatics in Zagazig University, Egypt.

Mohammed M. Hassan is an associate professor of information systems, head of the Department of Information Systems, Faculty of Computers and Informatics, Zagazig University, Zagazig, Egypt. He is specialized in computer aided system engineering. He received his Ph.D degree in computer aided system engineering from the Department of System Engineering, Faculty of Engineering, Zagazig University, Zagazig, Egypt; also The Pennsylvania State University, State College, USA. He has published more than 28 articles in international journals and conference proceedings.