

Enterprise GIS Solution for census operation: A case study from Kingdom of Saudi Arabia

<http://www.planning.gov.sa/statistic/cdsonnet/index.asp>

Aymen A.Solyman
GIS Project Manager (GTZ)
admin@cairolocator.com

Introduction

Census operation is a very important event in the history of a nation. This operation covers every bit of land and property in the rural and urban areas in the country. In Kingdom of Saudi Arabia, Central Department of Statistics (CDS) conducts census operation every 10 years. The last census operation was conducted in (1413H-1992) .The preparation of the next census operation started one and half year ago (1423H-2003) and expected to be finished by the end of (1425H-2005).The strategic plans of CDS, that census information must be shared by all divisions within the organization to support mandated functions and also to supply other ministries and organizations in KSA with this information.

GIS and census data

When we look at major forces of the 21st century like population growth, economic development and consumption of natural resources, we see that all of them have a spatial component. Data produced by census is a primary source of information needed for effective development, planning and monitoring of population, services distribution and socio economic.

The census database contains a lot of attribute information which can be linked to spatial data by spatial referencing. Relating the spatial component along with non-spatial attributes of the existing organization data enhances user's understanding and gives new insights into the patterns and relationships in the data that would otherwise not be found and to enhance the decision making process. Taking into consideration that many potential users of spatial information don't know about GIS technology, the development of an Enterprise GIS solution would serve two crucial purposes. Firstly, it would allow the users to operate the system without having to go the underlying intricacies of GIS and RDBMS technology. Secondly, it would allow sharing of information and technical experience between a wide range of users.

Enterprise GIS

Over the past few years, the terms Enterprise GIS and Enterprise-wide GIS are becoming more common in the GIS community, in organizational strategic plans in requests for services. Enterprise GIS is not a method to provide automated map plotting capabilities to the entire organization. Enterprise GIS is the sum of a coordinated GIS effort and its federated systems working together to support and promote coordinated geo-spatial data development and access across jurisdictional and organizational boundaries.

The vision and promise of Enterprise GIS is its ability to be an essential or core technology that integrates information across the many "islands of computer information systems" that exist in any organization.

The critical capabilities associated with Enterprise GIS include:

- a GIS that integrates geographic data across multiple departments and serves entire organization.
- a GIS that provides desktop access via LAN and WAN connection to anyone who needs access, i.e. Enterprise GIS fully supports client-server operability.
- a GIS that provides access to other information systems in the organization by using a map or application as the integrator of the organization's information.

-a GIS that can be communicate with other entities outside of the enterprise itself.

System Architecture Design

While designing any Enterprise GIS solution with the help of client's requirements, the salient factors those are important to be included are: functional module, graphic user interface, downloading time, system performance and the cost of the solution.

The system must be portable and extendable to accommodate the future changes in hardware, software and networking. Keeping in view these factors ,the solution has been done based on two components :the server side and the web client side, which runs in the browser. The server is a set of applications that serve the requests made by a client. It is divided into two sections: the map server which responsible for creating maps from the spatial database based on client requests and sends them back to the client, and the other section is the data server which is responsible for managing tabular data at the server side and sending the information to the client. The web client is partitioned into two parts ,the map console(map area, map tools, dynamic legend and the query items) and the side information table.

In this solution, the client is a thin client user, i.e. the analysis procedures is processed on demand on the server side. Data transmitted to the web-client are in standard HTML format embedded with typical PNG (Portable Network Graphics)-8bit image format(35K-45K) that can be viewed in any standard web browser. The advantage of this approach that there is no need to download any plug-ins at the client side.

The performance of the system is measured by the ability to respond the requests faster and the reliability of the system. Typically the response to a certain query must not take longer than 25 second. The overall system performance of internet GIS is dependant on the combination of client, server and networking performance, not the individual components and is controlled by the weakest component within the total solution.

In this phase of the project ,one server machine is used for the functionality of the web server, map server and data server, which have installed together. In phase two of the project after the completeness of the census data project, it is planned that the web server, application server and data server will be on separate machines.

The diagram below illustrates a general view of the system design architecture.

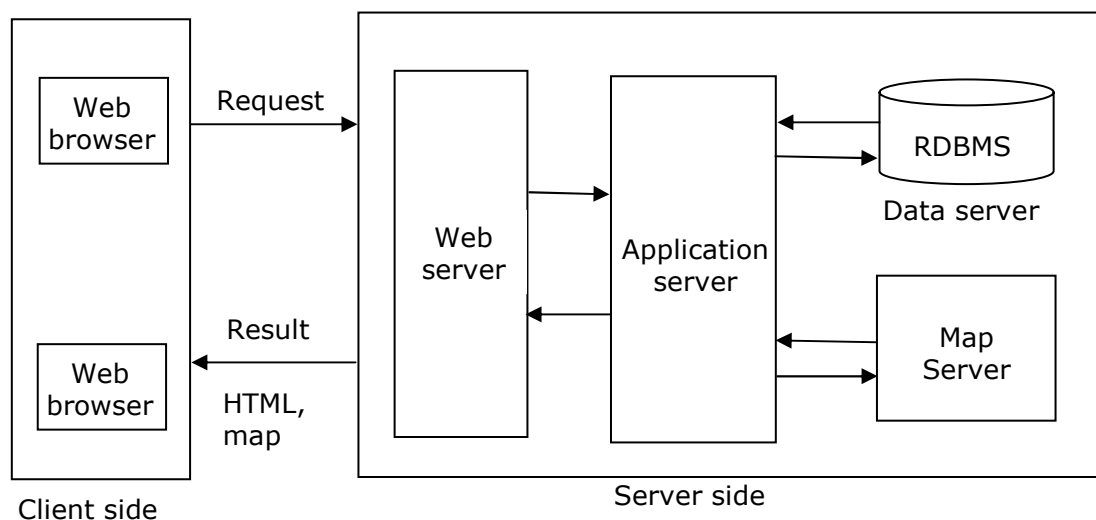


Figure (1) System Design Architecture

Database Design

The database design is the foundation for building any enterprise GIS solution. The enterprise GIS database design must be thorough, well documented, permit modifications and allow continual updates. Many enterprise GIS efforts suffer from either the complete lack of a database design or one that is so rigid that it can not grow. It must be noted that the database design is a factor affects the performance and download time of the solution. The database design defines the next components:

- 1-Spatial database content, structure and format.
- 2-SQL server attribute data content, structure ,format and constraints.
- 3-Relationships between spatial data and census data.
- 4-Relationships between attribute data tables in the existing RDBMS.
- 5-Updating process for both spatial and census data.
- 6-Data dictionary for both spatial and census data.

The design of the database includes three major elements:

- 1-Conceptual design, which is independent of hardware and software and could be a list of utilization goals.
- 2-Logical design, which is the specification of the database vis-à-vis a particular software.
- 3-Physical design, which pertains to the hardware and software characteristics and requires consideration of file structure, memory, disk space, access and speed.

The database in the solution can be classified into two categories (tabular data and geo-spatial data):

1-Tabular Data(attribute data)

Database contains the census information of KSA classified into four main categories, these categories are:

Population DB: Database stores data about the population categorized into saudi, non saudi and total number of population in each province.In each table in the population DB,the data are classified into categories due to the age range of both males and females.

Education DB: Database stores data about the number of educated people according to the level of education in each province.

Social DB: Database stores data about the marital status of Saudi, non Saudi and total in each province.

Services DB: Database stores data about the different services available in each province, these data include information about agriculture, education (primary-intermediate-secondary-high-others), health, administrative, public and social services.

2-Geo-spatial Data

Three initial data layers were determined necessary to build the foundation for the census data solution, these layers are:

International Boundaries Layer: Polygon layer contains the international boundaries of KSA. The source of this layer is the General Directorate of Military Survey.

Province Boundaries Layer: Polygon layer contains the boundaries of provinces in KSA,KSA consists of 13 province and 118 governorate. The source of this layer is the central department of statistics(CDS)-The mapping unit and was produced based on the settlement GPS points.

Grids Layer: Polygon layer contains the grids of 1:250,000 maps produced by the Ministry of Petroleum and Mineral Resources-Aerial Survey Department.

The design specifications are documented in a data dictionary that contains a description of each data layer, the data types used to model geographic features(point, line, polygon),tables structure, field definitions, coding schemes, and other information.

System Implementation

Based on the user needs assessment and database contents ,the census indicators are classified into 9 categories ,these categories are:

- 1-Population(Population distribution-Population density-Statistical operations) and each sub category is classified into(Saudi, non Saudi and total).
- 2-Education status (Saudi, non Saudi and total).
- 3-Social status (Saudi, non Saudi, total).
- 4-Education services (primary, intermediate, secondary, high and others).
- 5-Health services.
- 6-Agriculture services.
- 7-Social services.
- 8-Administrative services.
- 9-Public services.

The results of any query can be visualized in a number of ways that enhance the user's understanding and interpretation of the data, some of which are:

- 1-compare multiple attributes of a feature by depicting the attributes as elements of bar or pie chart.
- 2-compare one feature to another by the relative size of each feature's chart.
- 3-Each category can be assigned a shade in a graded sequence of a user-defined color ramp.

As shown in figure 2,The user can choose a specific category such as the distribution of population for Saudi for a specific age range, the results are:

- 1-Thematic map classified into 5 equal categories according to the values of the query item(1).
- 2-Dynamic legend appears at the right part of the map console shows the color and the value of each category. The legend is classified into 5 equal categories(2).
- 3-Dynamic bar chart shows the percentage of values of values to the total in each province(3).
- 4-Table contains the values of the selected query item in each province(4).

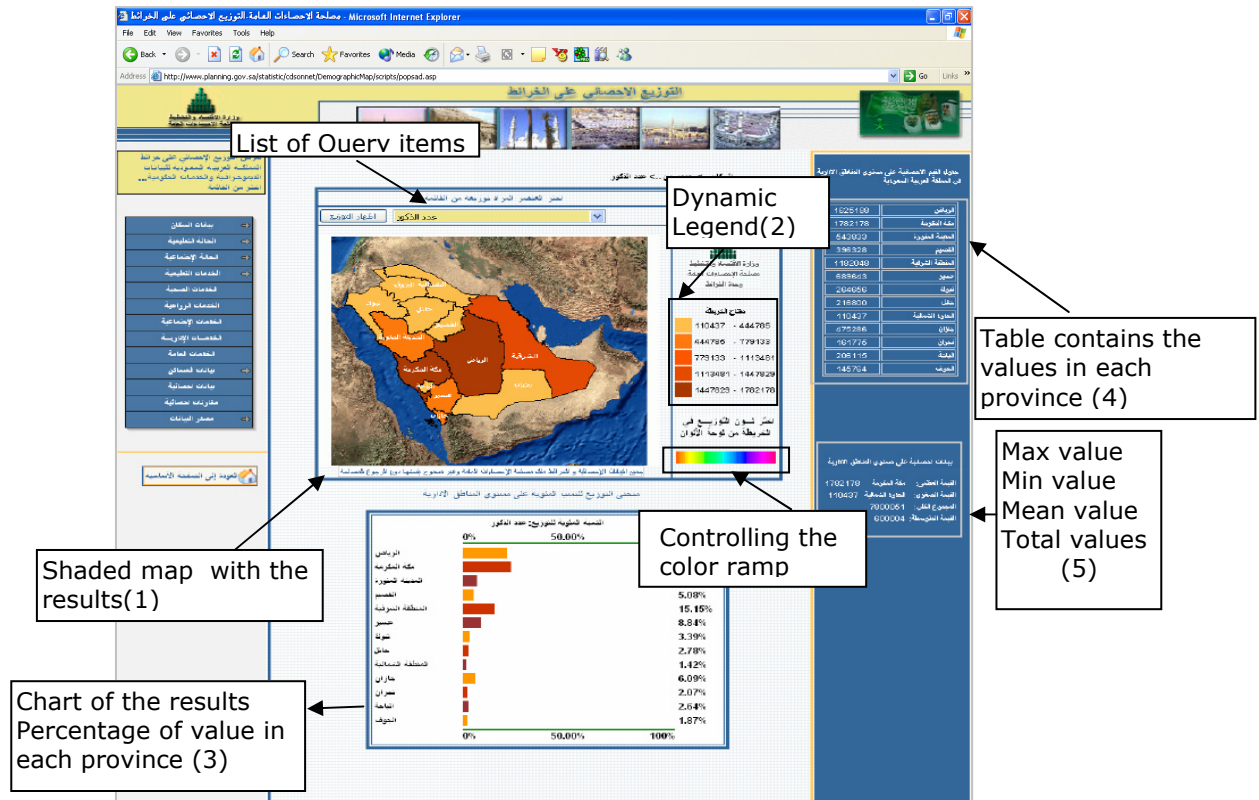


Figure (2) Census Indicators

5-Table contains the provinces with max value, min value, mean value and the total value(5).

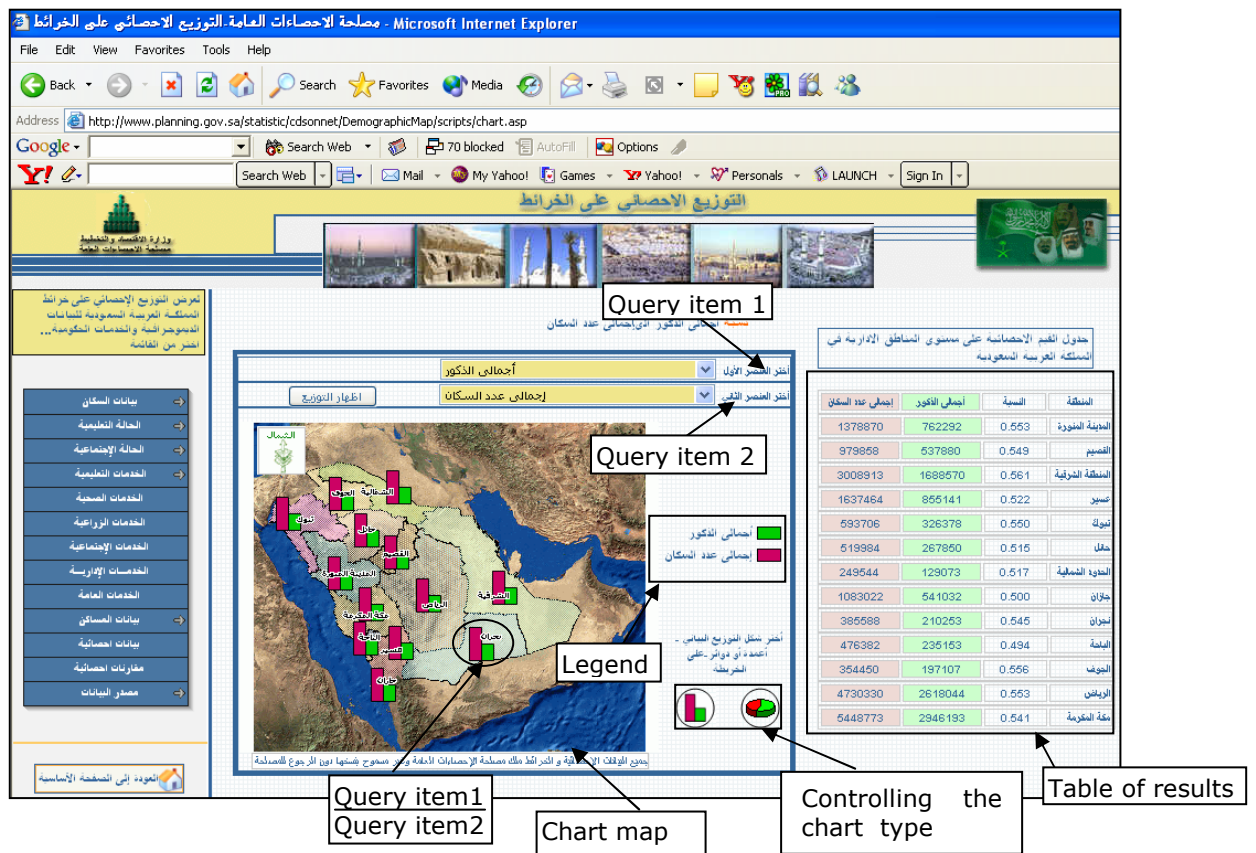


Figure (3) Statistical Operations

In the statistical operations, the user can choose any 2 items from the population data to compare together by relating the relative size of each feature's chart as shown in figure (3). For example, the user can choose to compare the number of male of Saudi (query item 1) to the total number of population (query item 2), the results are:

1-Chart map displays a relative bar or pie chart of query item 1 to query item 2 in each province. The user can choose to display the results as pie or bar chart.

2-Table contains the values of query item 1, query item 2 and percentage of query item 1 to query item 2 in each province.

The solution allows the users to print the map results in the standard template of the organization. Also the map can be exported as image to be used in other applications.

Conclusions

The key factors of implementing an Enterprise GIS solution are cooperation, management, knowledge, funding and experience. This paper discusses the methods for implementing an Enterprise GIS solution for the census operation in KSA. The objective of the solution as CDS strategic plans that census information must be shared by all divisions within an organization to support mandated functions and also to supply other ministries and organizations in KSA with this information.