DEVELOPMENT OF A SERVICE ORIENTED ARCHITECTURE BASED GIS FOR EARTH SCIENCES

A. Garagon Dogru[a], *, G. Toz[b]

[a]Bogazici University, Kandilli Observatory and Earthquake Research Institute, Geodesy Department, 34680 Cengelkoy Istanbul, Turkey - garagon@boun.edu.tr
[b]ITU, Civil Engineering Faculty, Photogrammetry Division, 34469 Maslak Istanbul, Turkey - tozg@itu.edu.tr

Commission IV, WG IV/5

KEY WORDS: Internet GIS, Integration, Distributed Computing, Geophysics, Distributed GIS

ABSTRACT:

The studies of recent crustal movements are based on analyses of repeated geodetic measurements, and their combination with results of geophysical and geological investigations. It is obvious that a single data producer can not produce useful datasets and information without integrating data from others because one scientist's results become another's data. So, the problem to be solved naturally has an interdisciplinary character. However, Earth scientists traditionally work on one aspect of the problem and they have a tradition of sharing of data but they are willing to share it if asked. Because of this, the resources are being wasted in duplicative efforts. However, the goal is for data to evolve it into information, and then into knowledge as quickly and effectively as possible. In order to do this, calculations and analysis need to bring to the desktops of researchers, decision-makers, and educators. The aim of this study is to develop a service oriented architecture based Geographic Information System that enables linking and sharing multidisciplinary Earth science data, tools, and software and to provide a wide range of users (scientist, educators, policymakers, etc.) access to the system and in this way to build an easy-to-use interactive access to data and analysis environment to study earthquakes in Turkey.In this study, a system was developed to access the Earth science data that is available now and data which will be coming online, and to provide users easy access to computation and visualization tools. This study is a motivation for developing the necessary information technology infrastructure applicable to Earth science areas. It speeds up the scientific discovery process. It is scalable to accommodate future growth and changes since the framework is based on the service oriented architecture. This architecture allows the modules to be written in separate programming languages, and to be run on different computers over the Internet. GIS is used as a user interface with zooming, panning, and query capabilities and is executed over the Web. The functional requirements of the system are the various information technologies. It includes many types of data and algorithms from the field of the Earth sciences, computers, GIS, networking and databases.

1. INTRODUCTION

Earth sciences (a.k.a. Geosciences), is a term for the sciences related to the Earth. Earth sciences are a discipline that is strongly data and compute-intensive. There are duplicative efforts on data and tools which cause waste of labour and time. Building mechanisms which are capable to share data and tools is the key for the next generation of Earth science research. These platforms must include databases, networks, visualization, analytical tools, computational resources, and so on. There are a lot of efforts in Earth sciences such as data collection from field observations and sensors, database creation, software development, data integration, and data management. And each of them has its own various problems. The need is to provide access to all of existing resources and support interoperability among them by using information technologies (IT).

Turkey is an earthquake country. 96 percent of the land containing 66 percent of the active faults is affected by earthquake hazards and 98 percent of the population lives in these regions.
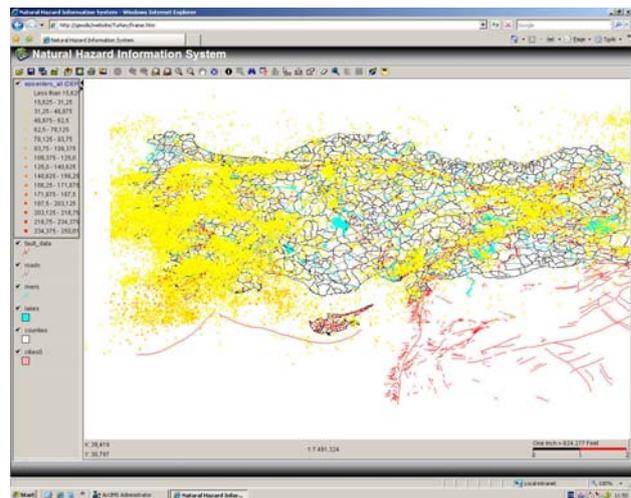
Figure 1. Seismicity of Turkey

Geoscientific data from earthquake studies in Turkey is stored and processed using various software. Furthermore a big amount of these data are available in different forms and projections. They come from various resources and they do not

---

* Corresponding author.

have a datum defined. And each scientist tries to create own tool to analyse them. However, the common platform can provide easy access to geospatial data and tools for internal and external users of them.

The use of GIS systems opens new avenues for comprehensive studies and solving complex problems related to integrated and dynamic earth systems. As we progress into the digital technology age, efficient ways of capturing, storing, organizing, manipulating, and updating data sets are needed so that we are not overwhelmed by the amount, diversity, and heterogeneity of the data. Clearly, GIS provides a convenient platform for data collection, organization, and research with multidisciplinary data sets. As more groups adopt GIS applications, the earth sciences community will be in a position to prepare a unified global database for more efficient, productive, and rewarding research (Seber, 1997). In the history of GIS, it has been evolved from Geographic Information System to Geographic Information Services. In the 1980's GIS was Geographic Information System; in the 1990's, Geographic Information Science was the preferred phrase; and now the trend is toward Geographic Information Services. GIS has benefited from the same improvements in the information technology (i.e. faster processing, increased bandwidth, greater storage capacity, mobile technologies, and real time networks) and has moved to workstations, then PCs, and now the Web.

Information technologies allow scientists to create appropriate solutions to meet the requirements of interdisciplinary Earth science projects which have multiple goals. One of them is called SOA (Service Oriented Architecture). SOA is basically a collection of services which communicate with each other. In another words, the implementation of SOA in the Web environments is called Web services. A Web service can be defined that a programmable application which is accessible using standard Internet protocols. Web services can be any piece of code that is available over the Internet and they can be written in any language. Reuse of existing tools, lower cost of maintenance and reduced impact of change are the most important benefits of Web services. In this study, the system was built using a service-oriented architecture for reusability and interoperability of each component.

2. APPLICATION

In this study, the developed system brings the complex strain analysis procedure developed by Haines and Holt, 1993; Haines et al., 1998; Beaven ve Haines, 2001; Holt and Haines, 1995; Shen-Tu et al., 1999; Kreemer et al., 2000 to a level that can be used by anyone efficiently and effectively by using Web services approach. The study creates an IT infrastructure to enable interdisciplinary Earth science research. To facilitate within this infrastructure, a portal based architecture has been adopted in the user interface and service oriented architecture in the middleware.
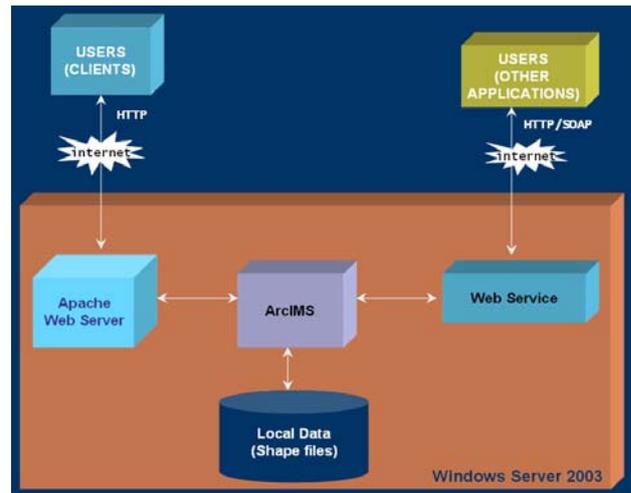
Figure 2. System architecture

The concept and implementation emerged from the needs of the scientific communities and decision makers. This system integrates data and tools from data- and labour-intensive, multidisciplinary Earth sciences and provides users from different levels (geoscientist, educator, decision-maker and public) to access it. It enables people, who want to work on earthquake hazard, to reach and use geoscientific data, as well as the computations and analysis environment by user-friendly interface and interaction. The scope of the study covers obtaining and processing Earth science data and tools, and integrating them in a GIS environment using information technologies, and then transmitting them to mass users via the Internet.

The motivation is that the calculation process of strain programs and the other codes created during this study is pretty much complex. There are over 30 programs to be able to use mentioned method and to see the results obtained. Furthermore, each researcher who wants to use these algorithm has to develop own tools. There is so much redundancy of development efforts and little interoperability among these tools. Interaction among different tools requires extensive recoding for every effort. So the solution is to wrap these tools as Web services which are accessible to the scientific community worldwide. A wrapper is a design pattern where a piece of code allows classes to work together that normally could not because of incompatible interfaces. There were some modifications done to strain programs for study purposes. These modifications were primarily to increase portability and IO performance. There were a number of places where the codes require user interaction. These codes were removed, and such input parameters were arranged by user interface. The steps of this process are recoding the FORTRAN codes to remove user interaction in the codes, compiling them in Windows OS, scripting batch files to run all programs as one program, creating Web service, connecting the application to ArcIMS.
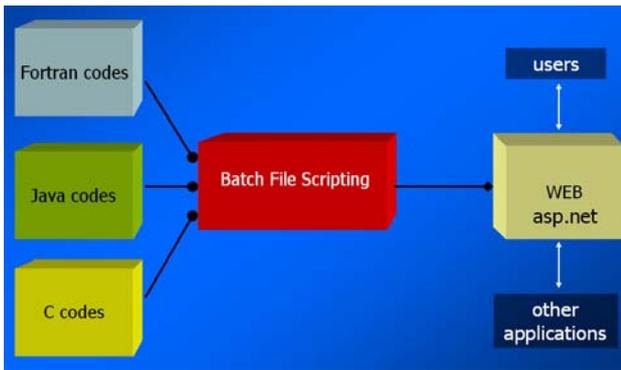
Figure 3. Running scientific tools as a Web service

Integration between GMT (Generic Mapping Tools) and ArcIMS (ESRI's Internet Map Server) programs for visualizing the results was performed. ArcIMS brings GIS to the Web by providing the ability to generate maps on the fly and also integrate data from different sources for display, query, and analysis. GMT program (Wessel and Smith, 1995) is widely used in the Earth science. Interactive maps served with ArcIMS include maps with layers that can be turned on and off, or with features containing attributes that can be queried. A web browser is just needed by the user, and the database is maintained on the server side. ArcIMS uses ArcXML to receive and respond to requests from the client. ArcIMS needs a Web server that can be extended to run Java code. Apache Web Server was used. ArcIMS installation also needs a servlet engine installation. Servlet engines extend Web servers with a common API and allow them to process Java code. Developing custom applications using standard Web development environments, sharing data, and implementing GIS portals are the important features of ArcIMS software. GMT is an open source collection of about 60 UNIX tools, which was developed by SOEST (School of Ocean and Earth Science and Technology) of University of Hawaii. It allows users to manipulate datasets and produce images.

In this integration process at the server side, after entering the parameters, each request is given a unique id number according to the time the request arrived at the web server. Then input files are created and stored in the workspace. The application runs the programs and output files are created. GMT is executed for each line in those files including strain and velocity information. GFW (world files for gif images) files are created for georeferencing strain and velocity images in ArcIMS environment. Batch files which run GMT creates PS files. Then ImageMagick program converts PS (postscript) files to GIF (Graphics Interchange Format) files. ImageMagick is software to create, edit, and compose bitmap images. It can read, convert and write images in a variety of formats. The functionality of ImageMagick is typically utilized from the command line. ImageMagick is free software delivered as a ready-to-run binary distribution or as source code. It runs on all major operating systems (ImageMagick, 2007). In the application for each request, ArcIMS's ArcXML file is created dynamically and GIF images are embedded into ArcXML file as a layer. ESRI.ArcIMS.Server.dll is to provide a connection between the client and ArcIMS Application Server.
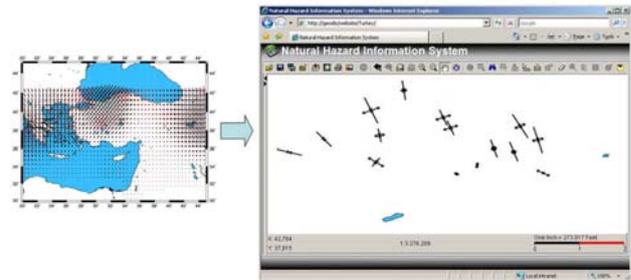
Figure 4. GMT images embedded into ArcIMS

Depending on the purposes, any Earth science data can be used in the study. Archived data such as geographic (cities, roads, rivers, etc.) and geological (faults), are currently integrated in the study. Focal mechanisms data is updated periodically. GPS velocity data are related to Marmara region of the country which involves strike slip faulting. The area has a high seismic hazard and risk because of the region's tectonics. GPS data from various sources can be currently uploaded into the system. Since the study is an open-end one, any kind of data, which contributes valuable information, can be added into the system. As new datasets are being added, new applications become possible which require new features to be added to the server. This loosely coupled development model can generate many innovative applications.

3. CONCLUSIONS

There is an overflow in all of the branches of science today, especially in Earth sciences. This situation causes problems on data storage and data processing, as well as accessing to these large size datasets and analysing them. Earth-related data are being collected every day using present-day technologies. By means of the satellite and computer technologies, it is now possible to study the Earth as a global system. And scientific instruments such as satellites generate terabytes and petabytes of data every day. Therefore, there is a rapidly widening gap between data collection capabilities and the ability to analyse the data. Once the solutions have been created for collecting, storing and accessing data, now it is a challenge to effectively share data, application and processing resources across many locations.

In Earth sciences, seismic hazard analysis can be defined as the integration of geophysical, geological, and geodetic data to estimate earthquake potential. If the seismic hazard can be analysed, the earthquake losses can be reduced. GIS is an important tool to analyse seismic hazards and risks by providing data from different disciplines of Earth sciences and IT enables successful solutions for delivering data and tools.

The ultimate goal of this study is to provide sustainable and user friendly on-line services that will enable risk management actors to better anticipate or mitigate earthquake hazard situations. Integration of IT resources into the Earth sciences allows not only researchers but also external users to access to state-of-the-art software and tools. This study is a motivation for developing the necessary IT infrastructure applicable to Earth science areas. It speeds up the scientific discovery process. It also holds a potential for similar projects. It is expected that this study will pioneer the development of such projects in other disciplines of science.

REFERENCES

Beavan, J., and Haines, J., 2001. Contemporary horizontal velocity and strain rate fields of the Pacific-Australian plate boundary zone through New Zealand, Journal of Geophysical Research, 106, 741-770.

Haines, A.J. and Holt, W.E., 1993. A Procedure to Obtain the Complete Horizontal Motions Within Zones of Distributed Deformation From the Inversion of Strain Rate Data, Journal of Geophysical Research, 98, 12,057- 12,082.

Haines, A.J., Jackson, J.A., Holt, W.E., Agnew, D.C., 1998. Representing Distributed Deformation by Continuous Velocity Fields, Sci. Rept. 98/5, Inst. of Geol. and Nucl. Sci., Wellington, New Zealand.

Holt, W. E., and Haines A. J., 1995. The Kinematics of Northern South Island New Zealand Determined From Geologic Strain Rates, Journal of Geophysical Research, 100, 17991-18010.

ImageMagick, 2007. ImageMagick Home Page, February 2007, <http://www.imagemagick.org>

Kreemer, C., Holt, W.E., Goes, S., Govers, R., 2000. Active Deformation in Eastern Indonesia and the Philippines From GPS and Seismicity Data, Journal of Geophysical Research, 105, 663-680.

Seber, D., Vallvé, M., Sandvol, E., Steer, D., Barazangi, M., 1997. Middle East Tectonics: Applications of Geographic Information Systems (GIS), GSA Today, 7, No. 2, February.

Shen-Tu, B., W. E. Holt, A. J. Haines, 1999. The kinematics of the western United States estimated from Quaternary rates of slip and space geodetic data, J. Geophys. Res, 104, 28927-28955.

Wessel, P. and Smith, W.H.F., 1995. New Version of the Generic Mapping Tools Released, EOS 76, 329.