Spatial Data Infrastructure Components for Early Warning Systems

Kiehle Ch.
lat/Ion GmbH, Aennchenstr. 19, 53177 Bonn, Germany,
eMail: kiehle@lat-ion.de

Abstract
The Earthquake Disaster Information System for the Marmara Region, Turkey (EDIM) is a current interdisciplinary research and development project in the context of Early Warning and Decision Support Systems. Within this project several components based on the concept of Spatial Data Infrastructures have been developed to support the relevant stakeholders in gathering concise information for early warning and disaster management purposes.

The components considered are based on the Open Source Spatial Data Infrastructure framework deegree. It is outlined which components comprise the system, how spatial information can be accessed by users and which technologies are relevant. Based on the concept of grid-computing, a proposal for ensuring fail-safe operation is laid out.

Keywords: Spatial Data Infrastructure, Early Warning System, Turkey, Sensor Observation Service, Web Processing Service

Introduction
Early warning systems should be based on a great variety of data and information sources, especially with a direct or indirect relation to space. Spatial data may include topological data sets, elevation models, satellite imagery and data collected through sensors. To provide accurate decision-supporting information, early warning systems should be capable of hiding the complexity of the underlying data sources and information models through well-known interfaces (component interfaces as well as user interfaces). Within the Earthquake Disaster Information System for the Marmara Region, Turkey (EDIM), the Open Source framework deegree (Fitzke et al. 2004; Müller & Fitzke 2008) provides an implementation of interfaces for data, services, and user interaction. Deegree provides an Open Geospatial Consortium (OGC) compliant set of server and client components to support the design and implementation of sustainable early warning systems. Besides classic service interfaces for raster-, vector-, and mapping data (i.e. Web Coverage Service, Web Feature Service, and Web Map Service), deegree also provides access to sensor data collections through a Sensor Observation Service (SOS, Na & Priest 2007) and accompanying user interfaces.

The user interfaces are based on portal technology to provide access to distributed information sources through a consistent interface. Maps can be overlaid with aerial imagery as well as just-in-time information from sensor data. A chart module is capable of visualisation of numerical datasets as charts (e.g. pie charts, line charts or scatter plots) within a map to support decision makers. For event-driven workflow (e.g. issuing early warnings after seismic events) a Web Processing Service (WPS, OGC 2007) is available to propagate alerts.

Being primarily developed for early warning systems related to seismic events, the components demand to be prepared for generic purposes. Focussing on the information technology aspects rather than the functional aspects of the EDIM application, the components pre-
sented are of vital interest to adjacent projects\(^1\) with a focus on spatial data processing and visualisation.

**The EDIM Project**

Earthquake Disaster Information System for the Marmara Region\(^2\) (EDIM) is a joint research and development project funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF) and the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) within the special programme Geotechnologien\(^3\) (Stroink&Mossbrugger 2008). It is carried out by a project consortium consisting of geo- and computer scientists and practitioners in spatial information systems:

- University of Karlsruhe (TH), Geophysical Institute (project lead)
- GeoForschungsZentrum (GFZ) Potsdam
- Humboldt University Berlin, Computer Science Department
- lat/lon GmbH
- DelphiIMM GmbH
- Kandilli Observatory and Earthquake Research Institute of Bogazici University, Istanbul, Turkey

The project partners work jointly on three subprojects. The overall goal is to investigate the expansion of an already existing Istanbul Earthquake Rapid Response and Early Warning System (IERREWS) to the regional scale of the Marmara region as well as to enhance the quality of the provided information.

The Marmara region is Turkey's most densely inhabited area, inhabited by nearly 18.000.000 people. Since a large geologic fault crosses the Marmara Sea, Istanbul has been regularly shocked by earthquakes (up to magnitudes exceeding 7.0). This raises the need for improving early warning and rapid response systems in this area. The associated

Kandilli Observatory ensures that the research and development undertaken in EDIM suits the needs of the stakeholders in Istanbul.

The EDIM approach to early warning and rapid response is comprised of three subprojects, each targeting a specific area of earthquake detection, analysis and visualisation utilizing information and communication technology:

- **Part A** enhances methodologies for improvement of real-time information of earthquake source parameters and shakemaps.
- **Part B** enhances sensor technology and software for self-organising networks of sensor units.
- **Part C** uses the data and methodologies collected from subproject A and B and integrates them into a consistent Spatial Data Infrastructure. The remainder of this paper will focus the developments of subproject C with an special emphasis on components to set up a Spatial Data Infrastructure for the EDIM project.

**The EDIM Spatial Data Infrastructure**

The relevant data sets are comprised out of different kinds of data formats.

- Raster data displaying satellite imagery, topographic data, shakemaps, and damage classification data
- Vector data representing geological formations, faults, socio-economic data, etc.
- Sensor data representing measurements of relevant base data for earthquake detection

Data sets are in various source formats like ESRI shape files, georeferenced Tagged Image File Format (tiff) images and databases containing geometries. In order to provide a vendor neutral access to all available data, the interfaces have to be based on internationally accepted standards. For EDIM the information system should be based on standards as pub-

\(^1\) EDIM is carried out within the thread »Early Warning Systems Against Natural Hazards«. A complete project list can be found on [http://www.geotechnologien.de/portal/?$part=Geotechnologien&locale=en](http://www.geotechnologien.de/portal/?$part=Geotechnologien&locale=en)

\(^2\) The project homepage is available through [http://www.cedim.de/EDIM.php](http://www.cedim.de/EDIM.php)

\(^3\) [www.geotechnologien.de](http://www.geotechnologien.de)
lished by the Open Geospatial Consortium (OGC) to allow integration into a broad variety of information systems. The following subchapters describe the base services that make up the EDIM infrastructure.

**Data Services**

The infrastructure is comprised of several data services. Most importantly, a Sensor Observation Service is responsible for storage of and access to real-time data. Other important services are a Web Coverage Services, Web Feature Services and Web Map Services which are described in the following sections.

A Sensor Observation Service (SOS) is part of the Sensor Web Enablement (SWE) family of specifications (Botts et al. 2007). It aims at standardizing the access to data from observations and measurements. It defines three obligatory interfaces:

- **GetCapabilities** describes the capabilities of the service and provides metadata about the service instance.
- **Describe Sensor** returns a sensor description which is encoded using the Sensor Model Language (SensorML, Botts 2007).
- **GetObservation**: returns the actual values of a sensor. Those values are encoded using the Observations & Measurements (O&M, Cox 2007) encoding.

Sensor data in the EDIM project are taken from a test site in Istanbul, comprised of 20 sensors measuring earthquake-relevant data (e.g. STA/LTA, PGA, PGV, PGD, CAV). The data is recorded in the proprietary format of the sensors, which hinders the exchange of data. A SOS is used to access the data and provide the data in a well-known format (O&M). This allows generic clients to access the sensors through a standardized layer.

Figure 1 illustrates EDIM’s SOS architecture for storing and archiving data from sensor nodes. A shared database (prototypically based on ApacheDB, port to PostgreSQL/PostGIS is planned) is used to write raw measurements data to. This database is more for long-term storage of sensor data than to access real-time data. The purpose of this database is to allow scientists to access the data after seismic events to analyse the data recorded by sensors. A seismic event has not necessarily to be an earthquake causing damage to buildings or people.

The database structure can be tailored to suit the needs of other potential applications. On top of the database an adapter layer (db-connector) provides object-relational mapping. This layer is responsible for accessing the database, retrieving data, and restructuring the data into encodings provided by an SOS (Observations & Measurements). All the functionality defined by the SOS specification is provided by the interface deegree SOS, namely the operations GetCapabilities, DescribeSensor and GetObservation. Since GetObservation is the operation to retrieve data from the SOS, this operation is accessed by the information system (iGeoPortal).
A Web Coverage Service (WCS) is used to access coverage data (e.g. topographic maps) through the obligatory interfaces:

- GetCapabilities
- DescribeCoverage to access detailed metadata about the served coverages.
- GetCoverage returns a coverage data stream.

A WCS is currently not actively used inside EDIM. It is planned to integrate a WCS to serve radar images for estimation of land cover.

A Web Feature Service (WFS) provides standardized access to vectorial data through the interfaces:

- GetCapabilities
- DescribeFeatureType returns an description of the served features of the service
- GetFeature returns the actual geodata

A Web Map Service (WMS) gives standardized access to map images, allowing to overlay different map images from local as well as remote service instances. It defines three interfaces:

- GetCapabilities
- GetMap retrieves an map image
- GetFeatureInfo: retrieves additional information for a user-defined location on the map.

While WFS, WCS and SOS allow access to spatial data for the more technically experienced user, inside an information portal, the Web Map Service is likely the service accessed most. Inside EDIM, WMS is used inside a portal environment as the central source of information.

The EDIM Client Infrastructure

The EDIM client infrastructure is comprised of a set of loosely coupled client components to access the services explained above. The most relevant part of the application from a user's perspective is the web mapping portal iGeoDesktop. iGeoDesktop is an integrated environment to access web maps and provide functionalities for interaction with maps (e.g. pan, zoom, measure, add or remove layers, etc.) (see Fig. 2). The portal is configured with...
a security component to allow access to different sources of information only to authenticated users.

iGeoPortal serves as a single-point-of-entry to all information gathered inside the EDIM project. The workflow is based on an interactive map. A user may explore the relevant area by navigating inside the map and requesting further information on selected spatial features. Without authentication, the information provided is rather coarsely grained. Only satellite imagery, county boundaries and other non-sensitive information are available. After authentication through the iGeoSecurity component, a user may – depending on his rights – see more detailed information like sensor stations, status of sensors and other services. He is also capable of adding further map layers and to define a set of information layers (so called Web Map Contexts) based on his needs.

Figure 2 illustrates a layer containing sensor stations (gpik_station_coordinates). These stations, marked by red dots on top of the coastline can be queried for further information. Using iGeoPortal’s getFeatureInfo functionality triggers an opaque workflow to actually assess a single sensor station. The workflow is as follows:

– user selects a sensor by clicking on a sensor station in the map (utilizing the getFeatureInfo function of iGeoPortal)

Figure 3: prototype of SOS client

4 Further details on iGeoSecurity, degreee’s approach to security implementation, can be found on https://wiki.deegree.org/deegreeWiki/iGeoSecurity
– the system transmits an identifier (ID) of the sensor station to a prototypical Sensor Observation Client (see Fig. 3).
– The client allows the user to interactively request information from the sensor. Adjustable parameters contain the sensor information to be displayed and the requested time interval.
– The client displays the requested data in form of a table containing the relevant data. A charting module is responsible for integrating the measurements into a line chart diagram to better access the relevant information.

For more technical application, it is also possible to directly access the results as XML data.

The sensor station being operated by project partners from subproject B offer live data of seismic events. In fact, currently the access is only simulated, but the connection to the live sensor network will be realized within the final project phase.

System Stability through Grid Computing
Early Warning Systems require a high-grade of system stability. In case of an incident, the servers running the Early Warning System could be affected, too. It is indispensable that the system keeps running even under these circumstances. All components comprising the system must have a back-up, which contain all the most recent information. A simple back-up mechanism (e.g. by nightly data backup) would not be sufficient. The system should be based on a distributed computing environment, which allows a transparent replacement of single computing and storage instances in real time.

The concept of grid computing (Foster & Kesselmann 1998; Foster et al. 2001) seems to suit these needs. A grid provides a network of storage and computing power, usually based on a distributed architecture. The distributed parts are accessed through a grid middleware, which encapsulates the distributed nature of the underlying computing and storage facilities. Grid computing offers immense computing and storage power, since it is comprised of several grid-nodes, working together through a standardized protocol layer.

For EDIM, the Globus Toolkit middleware (Foster 2006) has been evaluated to provide access to a computing and storage grid. Globus is, like deegree for OGC webservices, an Open Source framework to set up a grid infrastructure. Within the project GDI-Grid (Padberg & Kiehle 2009), first experiences in coupling grid computing with deegree have been gathered.

The EDIM infrastructure can benefit in several ways from grid computing:

– System stability: grid implements a distributed computing platform. In case a single grid node fails, another grid node is able to replace it. Since neither grid node is accessed directly but through an abstraction layer, the system will not even notice, that parts of the system are running on a different machine.
– System scalability: In case of huge computing and/or storage demands, additional grid nodes may be integrated without affecting the overall system during runtime. This offers a very flexible way to enhance computing and storage as required.
– System performance: grid computing systems usually offer computing and storage resources which exceed regular computers (servers, clusters, etc.) by far. The German D-grid for example currently offers about 20,000 Central Processing Units and more than 10 Petabyte of storage capacity. Similar grid infrastructures exist in many other countries (U.K., U.S.A., etc.)

In case of an incident, grid computing seems to be a good choice to avoid data loss, to ensure permanent uptime of the information system and to handle the increasing number of request to be expected.
Preliminary Results and Outlook

Within the EDIM subproject C, a couple of preliminary results have been achieved so far:

- **Spatial Data Infrastructure Components:** The deegree framework provides Open Source components for setting up custom tailored Spatial Data Infrastructures. Within EDIM, deegree has been enhanced in ways to provide access to sensor data through a standards-compliant Sensor Observation Service (SOS). A Web Processing Service Components to access vector- and raster data have been configured to serve data sets for the EDIM infrastructure in a standards-compliant way.

- **Client Components:** Available client components have been enhanced to support stakeholders from the field of Early Warning Systems in data and information retrieval. This includes the set-up of a specific Geoportal (based on deegree’s iGeoPortal software) as well as the integration of a Sensor Observation Client. In order to make the SOS data accessible more easily, the SOS client offers a variety of filter capabilities, basic statistics and graphical overview (i.e. chart functionalities). To support local authorities, the system can be configured to offer multi-language support.

- **User-Rights Management:** Since Early Warning Systems contain sensible information (e.g. building damage classification, land tenure, etc.) it has been necessary to restrict access to certain information layers for defined user-groups. Therefore a user-rights management system has been integrated, which only gives access to authorized users. Besides authorization based on roles, the authorization can be defined on spatial predicates. This can affect certain layers containing spatial data as well as certain features (i.e. vectors).

- **Concept for system stability:** In order to provide system stability even in case of an incident, a data and service backup system has been defined. The concept is based on grid computing paradigm in order to replicate data, services and computing tasks on a distributed computing platform.

- **Documentation:** To support the integration of the provided components into other Early Warning Systems as well as other information systems supporting decision makers, extensive documentation of the components on a technical level have been formulated.

During the final project phase until mid 2010, the focus will be on the finalization of the Spatial Data Infrastructure including the client-components. This involves fine-adjustment between the on-site sensor-network located at the test area and the Spatial Data Infrastructure components. Based on tests with real seismic data, thresholds for triggering alarm will have to be defined by the EDIM-project partners. These thresholds will be used to define an alarming event triggered by the integrated Web Processing Service.

The components developed in EDIM offer the main building blocks to set up early warning and decision support systems based on spatial data and information. The components are freely available under an Open Source Licence (i.e. GNU Lesser General Public License, Free Software Foundation 2007), which fosters reusability and sustainability. The underlying concept of an Service Oriented Architecture supports loose coupling of the components, thus enhancing the flexibility of the system.

Acknowledgement

The work presented in this paper would not have been possible without the funding within the EDIM project. Funding is provided by the German Federal Ministry of Education and Research within the special programme Geotechnologien, funding code 03G0650E.

References


Web Enablement: Overview and High Level Architecture.  
http://portal.opengeospatial.org/files/?artifact_id=25562

http://portal.opengeospatial.org/files/?artifact_id=22466


http://www.fsf.org/licensing/licenses/lgpl.html


http://portal.opengeospatial.org/files/?artifact_id=26667

http://www.opengeospatial.org/standards/wps


http://www.geotechnologien.de/portal/-;jsessionid=51B368E57793E568071D02610CC1C9F94?$part=binary&content&id=3367768&status=300&language=de