

Geotechnical Geospatial Semantics and Interoperable GIS : the Case of *GeoSeism* for Automated Seismic Microzonation Studies

Maria Apostolos Papadopoulou^{1,2}

¹Department of Civil Engineering, Alexander University of Applied Sciences
of Thessaloniki, Macedonia, Thessaloniki, Greece

²Thessaly Region Local Government
Larissa, Greece

Email: *m.papadopoulou* [AT] *thessaly.gov.gr*
info2.dm.papadop [AT] *gmail.com*

ABSTRACT— *GeoSeism* is a GIS-based application software which is intended to provide more automated seismic microzonation studies by utilizing the interoperable GIS technology. Proving the applicability of the LCIM (Levels of Conceptual Interoperability Model), it portrays an implemented combination of technical along with semantic and conceptual interoperability maturity, in which the modelling level of interoperability consists of more than one data end-users. Based on the experiences distilled from an extensive case study, this case faces problems of data inconsistency related to semantics. Inconsistency restricts the capability to derive geotechnical meanings from the employed geospatial data. The role of the semantic interoperability level is to appropriately process the meanings of the exchanged data so that the end-users can efficiently use these data. Otherwise, data are not fully applicable on geotechnical methodologies. The interoperation with the expert knowledge algorithm Semantic Interpreter Pythia (SI) proved especially helpful to undertake the role to improve the semantic interoperability in an automated manner. It modeled the semantics of data prior to data exchange. This innovative type of semantic interpretation is not related to the conventional search processes but to the spatial data process and the overcoming of a variety of many semantic-related data inconsistency. Distributed systems could evolve into geotechnical interoperable GIS combining typical application software with an SI. *GeoSeism* implemented a first concept of interoperable GIS which can model geotechnical semantics to ensure semantic interoperability, as well as allows for current or future interchange of geotechnical operations and applicable data for the benefit of engineers and GIS.

Keywords— Geotechnical Engineering, Geographical Information System (GIS), Levels of Conceptual Interoperability Model (LCIM), Seismic Microzonation, Semantic Interoperability, Semantic Interpreter (SI)

1. INTRODUCTION

The isolated information systems (IS) and the geographic information systems (GIS) of the past decades are gradually evolving into interoperable GIS. Especially the distributed web-based and the mobile GIS for domain applications tend to combine the multitier architecture with the interoperable technology in order to access and use spatial data and services [1]. Among the advantages which make the interoperable GIS a sought-after concept is that modelling can extend continuously to lead to current or future interchange of operations and that it can improve the quality of spatial data for the benefit of the open GIS data sources. The benefit is that these gain the elaborated GIS retrieving, visualization, and processing methods, as well as the flexibility of the distributed technology for sharing (data, programs, and services) and the abilities to interact and cooperate with various sources and system components. This challenging trend is enhanced by the rapidly growing Geospatial Web (GW), the active attempts for open geospatial technologies and data by the Open Source Geospatial Foundation (OSGeo), such as the implementation of open standards for geospatial content and services, sensor web and Internet of Things, GIS data processing and sharing by the Open Geospatial Consortium (OGC) and the related foundations and associations. In general, interoperable GIS seem to be a more advantageous and productive software for flexible, fast, less expensive, reliable, transparent, participatory, accountable, and multi-tasking services. The maturity of the research in this field is now concerned with the modeling of data semantics and the management of their meanings by interoperable GIS [2].

In this direction, the present research presents the interoperable GIS called *GeoSeism* (Geotechnical Seismic) which is an application software developed to make automated seismic microzonation studies. This model-driven software inherently focuses on the study of geotechnical subsurface conditions [3]. It initially developed in the context of a

doctoral thesis. The paper outlines the features of this interoperable software which implements a combination of technical along with semantic and conceptual interoperability maturity, in which the modelling level of interoperability consists of more than one methodology (sections 3). It proves the applicability of the LCIM (Levels of Conceptual Interoperability Model [4]) classification (section 2.2) to the interoperable GIS of this domain. An innovative feature is that it is designed to interoperate with the expert knowledge algorithm “Semantic Interpreter *Pythia* (thereafter SI)” [2] [1]. The extensive case study (sections 4) found that this interoperation is very productive. SI models geotechnical semantics prior to data exchange, solves problems related to semantics of data and feeds with ready multi-thematic geotechnical soil profiles (GSP) of applicable geospatial data the various geotechnical methodologies constitute *GeoSeism*. The term “multi-thematic” and the SI’s GSP are outlined in more detail (section 2.3). The term “applicability” [2] is used to denote the internal data consistency which is critically dependent upon meaningful and composable data. The innovative geotechnical semantic interpretation of SI is advantageous for many reasons (sections 4). It attempts to improve the content of geotechnical databases, to create new data in cases of data inadequacy, to minimize the wasted effort on the application (modelling) level, to automate the demanding preparatory data process, to share reusable and up-to-date data contributing to the creation of global geotechnical spatial data infrastructures. In addition, by combining typical application software with an SI, distributed systems could evolve into geotechnical interoperable GIS. For the concern of the engineering and technology community, *GeoSeism* implemented a first concept of interoperable GIS which can model geotechnical semantics to ensure semantic interoperability, as well as allows for interchange of geotechnical services and applicable data for the benefit of engineers and GIS.

2. BACKGROUND

2.1 GIS Interoperability, Interoperable GIS, Geotechnical Geospatial Semantics

The word “**interoperability**”, by etymology, means an ability for mutual and inverse operation (inter – operation - ability). The variety of definitions in the GIScience shows that the term rates depending on the means and instances. It is used to describe the ability of two or more systems or components to exchange information and to use the information that has been exchanged [5]. In the ICT (information and communication technology), it indicates the ability of different types of computers, networks, operating systems, and applications to work together effectively, without prior communication, in order to exchange information in a useful and meaningful manner. It is also used to describe the capability of different programs to exchange data via a common set of exchange formats, to read and write the same file formats, and to use the same protocols (the ability to execute the same binary code on different processor platforms is not contemplated) [6]. Especially after the Web [7] and the advances of the engineering industry, the term is increasingly expanding in cross-border eGovernment (electronic government) services. However, such services are so far rather inconsistent to provide adequate data in the civil engineering studies because management is paying attention to the straight knowledge management and most often lacks the non-technical levels of information maturity.

In particular, the concept “**GIS interoperability**” has been seen as a much important ability and still plays a significant role in the development of GIS. In previous years, data sharing between organizations with different GIS vendor systems was limited to use data converters, transfer standards, and open file formats. Sharing spatial data with external applications was rarely achieved. It is quoted [8] that, over the last 35 years, the concepts, standards, and technology for implementing for the Enterprise GIS have evolved through six stages: (1) Data converters; (2) Standard interchange formats; (3) Open file formats; (4) Direct read application programming interfaces (APIs); (5) Common features in a database management system (dbms); and (6) Integration of standardized GIS Web services. In recent years, the integration of standardized open GIS services on the Web allows for the sharing of geographic data, integration among different GIS technologies, and integration with other applications that are not GIS. **Interoperability** allows for data integration between organizations and across applications and industries by which it is able the creation and sharing of more useful information. Most GIS apps directly read and sometimes dynamically transform data with the minimum time delay. The basic principle is ever that a GIS is not an end in itself but it must produce useful information and share this information among multiple users. At the same time, it provides a consistent infrastructure to ensure data integrity. By its responsibility for gathering and managing geographic data, departments are able to share, coordinate, and communicate within an organization or between separate organizations often by using GIS as the core spatial data infrastructure. Research has been directed toward the development of relevant software and methods so that not only data but also programs and services are shared. It now also intends to make them all interact and cooperate. Especially to cooperate at the data level, so that an institution can profit from the data of another institution [9]. GIS interoperability aims to be able to operate on different platforms and databases and can scale to support a wide range of implementation scenarios by the individual consultant or mobile worker using GIS on a workstation or laptop in business implementations that support hundreds of users working across multiple regions and services. Thus, they focus on two major areas: (a) GIS data and technology; and (b) Interoperability of GIS technology with other technologies and systems.

The implementation of “**interoperable GIS**” today is extending to the semantics of spatial data and the management of their meanings. As aforementioned, the interoperable GIS concept allows for (the today or the future) interchange of

services (operations) and data for the benefit of the open GIS data sources [2]. Developments in GIS have long focused on the need for a technically unrestricted interchange of both spatial data and traditional GIS operations and analyses [10]. It is asserted [11] that while research in the fields of the interoperable GIS and the spatial data process is well advanced, the distinction between the two will be blurred. As these two fields of research advance, the development of the latter is becoming more reliant on a free exchange of both data and models. In addition, the development of relevant software and methods now puts forward semantic interoperability to make systems interact and cooperate. Based on the development of *GeoSeism* and the current advances on geotechnical geospatial semantics, the present research attempts a modern (descriptive) definition of the term “interoperable GIS” (section 4.4).

The term “**geospatial data**” (or GIS data or geodata or georeferenced data or geographic data) refers to data which are pertaining to space (spatial) and, at the same time, have explicit information about their geographic position within a GIS (on the **spatially enabled database** of the vector map or the geo-referenced satellite image). **Spatial data** refer to features or phenomena distributed in the three-dimensional space which have physical and measurable dimensions (e.g., the roof-depths and the space shape of a soil stratum, the space position and the spatial distribution of a variable, the earthquake-generated vibrations and the site effects). Accordingly, in the present research, the term “**geotechnical geospatial semantics**” [2] refers to the understanding of the meaning of geographic entities of the real world pertaining to the engineering semantics, both to the cognitive (human perception) and to the digital concepts of meanings (digital world). Note that, **geographic data and information** are defined in the ISO/TC 211 [12] series of standards as data and information having an implicit or explicit association with a location relative to the Earth.

2.2 Levels of Interoperability Maturity to Classify Technology Models

Technology models and measures for benchmarking the interoperability, are often classified into a hierarchical concept expresses their interoperability maturity. The word “maturity” here specifies that interoperability can be achieved at different levels so that a higher level includes the successful illustration of the lower ones.

Most of these classifications refer to digital business and organizational processes. Aggregate reports on models of interoperability maturity and standards for achieving interoperability in public service organizations (PSO) and electronic government organizations is discussed in the Proceedings of the European Conference on eGovernment (e.g., [13]). The maturity levels for interoperability in the above organizations are often described as a four-layered architecture includes: (1) syntactic; (2) functional; (3) semantic; and (4) user-task levels, and elsewhere, as a three-layered architecture includes: (1) organizational; (2) semantics; and (3) technical systems which later adds the: (4) political, and (5) legal levels. There are also models describe the five layers of: (0) performed; (1) modeled; (2) integrated; (3) interoperable; and (4) optimizing, and elsewhere, the layers of: (1) computer; (2) process; (3) knowledge; (4) value; and (5) goal interoperability, and elsewhere, identifies nine different constraints that influence interoperability including: (1) constitutional; (2) jurisdictional; (3) collaborative; (4) organizational; (5) informational; (6) managerial; (7) technological; (8) cost; and (9) performance. The literature quotes many examples of such implemented models. They often describe interoperability either as a mere “technological” issue or as a combination of organizational and technical issues (e.g. pragmatic, governance, organization, business, health, judicial, political process etc.). On the other hand, classifications which are quoted to the GIS or network communities, are often describing a six-layered schema which from lower to higher includes [14]: (1) network protocols; (2) hardware and OS; (3) spatial data files; (4) dbms; (5) data model; and (6) application semantics (information communities). In such a schema, applications interoperability and semantic interoperability are synonymous. In general agreement, the classification system ranges from the lowest level, devoted to network protocol interoperability, to the highest level of applications interoperability.

The research on composability has led to the classification known as “Levels of Conceptual Interoperability Model (LCIM)” [15]. It was recently reproduced by [4] as follows: (0) No – Interoperability; (1) Technical; (2) Syntactic; (3) Semantic; (4) Pragmatic; (5) Dynamic; and (6) Conceptual. This schema is a variant of the previous concept of LCIM. LCIM attempts to describe the stages through which systems should evolve to reach higher completeness in the realization of a modelling or abstraction, simulation or implementation, and network or connectivity [16]. [17] review the main maturity models. The discussion on this concept compares different aspects of its relevance to interoperability and with the use of maturity models (e.g., [18]). LCIM in this form or slight variants thereof has been applied not only in the simulation but also in many other domains (energy, health and human sciences, GIS, etc.) [19]. **Figure 2** illustrates a cross-classification schema for comparing the interoperability abstraction of *GeoSeism* with the LCIM concept (section 4.3). The maturity of the organizational levels is almost always far more vague than that of the technical levels. In order to use the internet for distributed simulation, procedures are needed for migration, alignment, and orchestration of the execution, which means higher levels of interoperation. Research introduces related concepts leading to the idea of the Dynamic Web, which will be a web of composable services [20].

Developing measures and benchmarking on the interoperability of public organizations are nowadays more and more introduced (e.g., [13], [21]). The concern tends on how the maturity models could be used to measure the degree of interoperability between systems. Based on the present case study (section 4.4), a measure of semantic interoperability could be the degree of data elements that an interoperable GIS can productively use from the core database. The degree

of **semantic interoperability** is relevant to the extent to which authorized systems are able to exchange and interpret their shared data such as that these data can be understood by any recipient. In other words, it depends on the applicability of both the data model and the data content to represents information about the world. A measure of this semantic-related applicability could be the level to which the DIKW model can be effectively materialized by a GIS or by the help of a special mediator like Semantic Interpreter *Pythia* [2]. Further research should soon explore measures on the development of techniques for accessing and using remote distributed spatial databases through global GIS networks [22] [1]. The potential benefits of research in this area are the construction of platform-independent modelling methods. Researchers also propose collaborative spatial process modelling which is dependent on the free exchange of both data and models. It develops a system-independent tool capable of integrating the transfer of data and operations as well as other process modelling functions to complete desired outcomes [11]. However, the level of interoperability is relative and difficult to determine because it depends on the needs and objectives of services, processes, and procedures that a system has to offer to meet the objectives. On the other hand, each interoperability level corresponds to a broad field of technology. The degree of interoperability should be defined when referring to specific objectives. This is often achieved by adding data about the data (metadata), and by linking each data element to a controlled, shared vocabulary. In the case of *GeoSeism*, this is accomplished through the DIKW data fusion process called “(geotechnical) semantic interpretation”. The expert knowledge algorithm SI attempts to improve the applicability of the shared data in a reliable and automated manner (section 2.3).

In general confessions: **Conceptual interoperability** is achieved when the assumptions and constraints of the conceptual model (i.e. the meaningful abstraction of reality) are aligned and a fully specified, but implementation independent model is required [23]. Syntactic or syntactical Interoperability level refers to the ability of systems to use common structures to exchange information applying specified data formats, communication protocols, interfaces of descriptions and the like in order to ensure a clearly defined format of exchanged information and a common understanding of symbols [9]. **Semantic interoperability** is the ability of two or more systems or elements to exchange information and to use the information that has been exchanged taking advantage of both the structuring of the data exchange and the codification of the data (e.g., including vocabulary) so that the receiving ICT can interpret the data. It is therefore concerned not just with the packaging of data (syntax), but the simultaneous transmission of the meaning (semantics) with the data [24].

2.3 Multi-Thematic Geotechnical Soil Profiling (GSP) of SI *Pythia*

A **geotechnical soil profile** [2] is a fundamental cross-section concept in geotechnical engineering. It attempts to simulate a real-world cross-section of a soil deposit or rock mass under the ground surface, either in a schema or in a spatially enabled database for application use. There is no strict standard on how to create an appropriate soil profile and what to include in it. For most applications, it is a limited input of data which are manually typed. It is included in the input stage and there is no processing stage to improve these data. It is modified to suit a particular individual software, for its specific syntactic and semantic tasks, without being able to be used in another application. Depending on the application, it may be a one-dimensional (1D), or a two-dimensional (2D), or a three-dimensional (3D) soil profile. It may transmit much or less information. The anyway goal of GSP is to determine the soil geometry and properties as a function of depth that the section crosses. Geometry includes the thickness, the inclinations, the boundaries, and the discontinuities of the soil layers. Properties include the natural and engineering properties (and parameters) of the soil layers; determined on the basis of the available actual samplings investigated the stratigraphy at various depths. An integrated GSP also include the bedrock and the groundwater levels. A site-specific profile typically displays a vertical soil column beneath a required geographic point of the ground surface. In any case, the GSP is a fundamental input for most geotechnical applications (e.g., site characterization, seismic ground response analysis, liquefaction potential, spatial analysis). It is a fundamental because it provides detailed data, soil characterization and distribution of properties in relation to depth.

A **multi-thematic geotechnical soil profile and profiling** (thereafter GSP) is a modern concept and methodology, respectively, introduced by SI *Pythia* [2]. It is an automated special process which subdivides soil deposits into thin sub-layers and then relates the multi-thematic soil data to determine the soil properties and geometry as a function to depth. In the end, the process automatically organizes the output into a permanent data structure so that this output is anytime accessible for a variety of purposes. SI developed a specific data fusion model, called DIKW, to process data from geotechnical databases. Advantages of the GSP of SI *Pythia*, are: (1) Multi-Thematic Data; (2) Thin Sub-Layers; (3) Deep SP; (4) Maximum Information; (5) Geographical-Referenced Sites; (6) Detailed Automated Information; and (7) Improved Data Applicability (Data Consistency and Composability).

From the point of view of geographic information science, the subject of the expert SI is to relate multi-thematic sets of data from databases, to interpret these data with the specialized data fusion model and, ultimately, to lead to unified information in a core relational database. From the viewpoint of geotechnical engineering, the subject of this innovative type of semantic interpretation of databases is to automatically determine the geometry, the properties and the

stratigraphy of the site subsoil. The aim of SI concept is to fully process data and share GSP with the modern interoperable GIS so that GIS understand geotechnical semantics and disseminate useful and up-to-date data content for wide use. The present case study (section 4), confirmed the applicability of this concept (section 4.3).

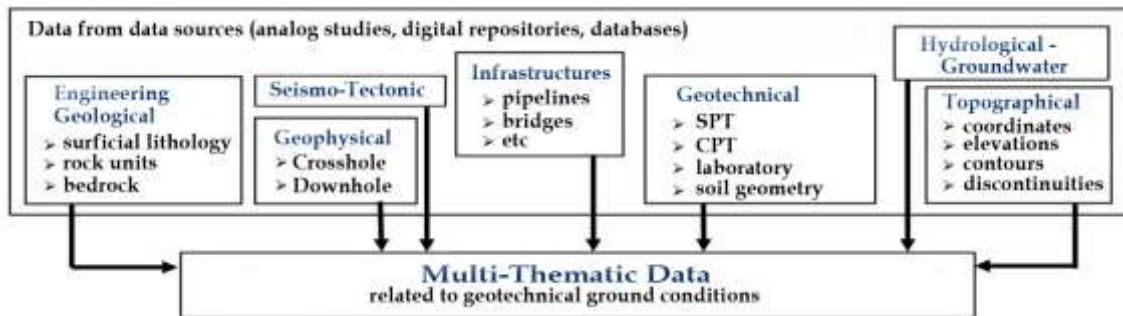


Figure 1: Multi-Thematic Geotechnical Data

The term “**multi-thematic**” geotechnical data (or information) is related to geotechnical ground conditions (subsurface soil) and includes: geotechnical data from geotechnical field tests (SPT, CPT) and from laboratory tests, geophysical data (Crosshole, Downhole) from geophysical field tests, geology data (surficial lithology, rock units, bedrock) from geology descriptions, hydrologic data from groundwater measurements, topographic data (coordinates, elevations, contours, discontinuities), and relevant data about the properties and geometry of soil strata, the geology and the aquifers. **Figure 1** depicts the themes of multi-thematic geotechnical data. Details about the geophysical and geotechnical tests can be found in many geotechnical engineering books (e.g., [25]). These data come from a number of related thematic databases (lithology, laboratory, geophysical tests, etc.) of actual investigated locations (IL). An available IL is the site which the actual investigations refer to.

3. APPROACH

3.1 Specific-purpose Methodology and Architecture of the Program GeoSeism

The methodology and initial implementation of the software *GeoSeism* developed (and tested) in the context of a doctoral thesis [26]. The purpose is to elaborate automated seismic microzonation studies in the geospatial area of cities. Microzonation is a detailed and multidisciplinary study that aims at dividing the area of study into zones with respect to geological, geotechnical, geophysical and geometrical characteristics, and to provide reliable maps of seismic ground shaking parameters considering local site effects, as well as induced hazards like liquefaction and landslides [27].

GeoSeism mainly includes a series of analyses of the earthquake-generated ground response and lateral phenomena. The implementation interoperates with a range of standalone software applications and relational database management systems (dbms), such as uses an external GIS mapping environment for geospatial visualizations. The currently developed apps, are (**Figure 2**):

- (1) *Artemis*: The algorithm estimates equivalent linear analysis of 1D seismic ground response (propagation and frequency content of earthquake-generated ground motions).
- (2) *Polynoi*: The algorithm estimates liquefaction potential (a kind of earthquake-generated soil hazards).
- (3) *Themistoklea*: The algorithm calculates the Fourier spectra and Inverse Fourier Transform of motions.
- (4) *Penthesilea*: The algorithm estimates the elastic Response Spectra of Structures (earthquake-generated structural response).
- (5) Semantic Interpreter *Pythia*: This is an innovative expert knowledge algorithm, attempts to improve the semantic interoperability of the system [2]. A recent version is the mobile-SI built for mobile platforms [1].
- (6) *Kallipateira*: This algorithm includes a core dbms and a second management code, which undertake the permanent data store of the system and optionally manage the interoperations, respectively.
- (7) *HelGeoRDaS_uTH* [28]: This database system (rdbms) is used as an external standalone data source. This data donor can either permanently or temporary interoperate with the system in order to supply raw multi-thematic data come from one or more investigations. *HelGeoRDaS_uTH* and *Kallipateira* are compatible.

HelGeoRDaS_uTH currently filled with a remarkable relational content of raw multi-thematic data came from actual ILs. It currently includes 342 ILs. It is the maximum reliable information collected from all the available sources investigated the (multi-thematic geotechnical) ground conditions of an entire typical Greek city. The goal is to extend data for more cities, and to always registering any new IL is available to enrich the stored geospatial information. Note

that, any compatible data source can be such a data donor to the interoperable GIS. However, the need for related standardization arises.

The modular structure of *GeoSeism*, integrated the above apps into a single software implementation. The source code includes programmable nodes so that being flexible to allow code extensions and system integrations. The future aim is to collect data by interoperating with external dbms and the public-purpose network (i.e., geospatial data distribution within web-mapping and the large GIS of the official services). Interoperability is therefore the key by which *GeoSeism* integrates the above systems, procedures and data to provide complete information. This feature represents an implementation of exchanging models, apps and data.

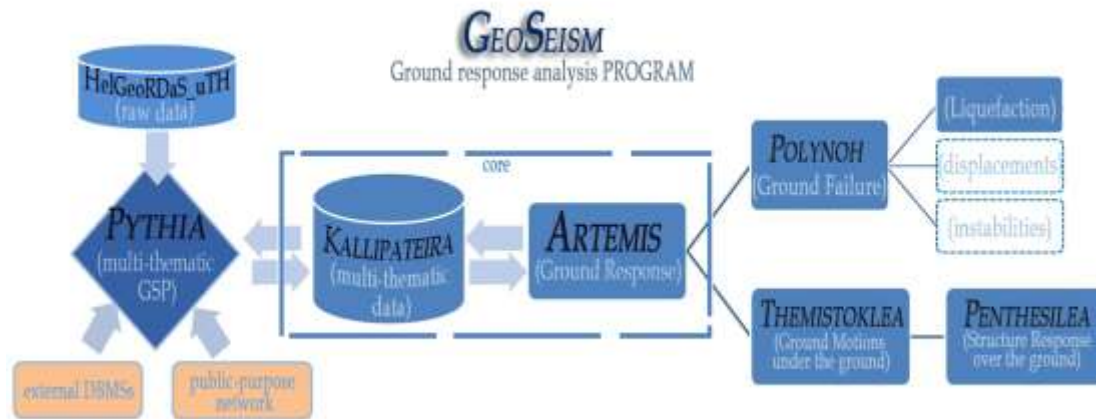


Figure 2: Program *GeoSeism* for Automated Microzonation Studies: Interoperates Specific-Purpose Apps and Databases Ensuring Semantic Interoperability at Conceptual Maturity Level

3.2 Technical Interoperability and Operation of *GeoSeism*

GeoSeism is an application software designed as a framework in which various nodes (here: apps, procedures) can optionally cooperate. The apps developed in environments: *MS Access*, *MS Excel*, *SQL*, *VB6*, *C/C++*. This technology is compatible with the widely used technical and syntactic standards. However, this system is not a bundle software product, which includes only a few compatible nodes. Instead, this program exploits techniques of a regular distributed system, with the difference that each node is a manager by itself communicates with other nodes in the network. Therefore, the technical (and syntactic) interoperability of *GeoSeism* is achieved by both the use of standards and the appropriate coordination between all the interoperable members. Table 2 (section 4.4) details some differences between the GIS data transfer, the distributed and the interoperable GIS considering a deployed *GeoSeism*.

Specifically, each node has been tuned so that it can manage a database and can call for another node to interact on the database. Each code here includes stored procedures through which it controls the interoperation. Queries and updates to a distributed node are transparently translated into operations on this node giving the user the impression of using a single program. Unlike a distributed dbms, which consists of loosely coupled sites that share no physical components, the collaborative apps here “manage” each other in the form of a communication. The interoperability technique is not a simple data transferring from one node to another. The “transfer” here aims to increase data productivity. Each node can offer open services specific to its field of expertise for the benefit of the larger system. First, the nodes share operations between them, in the sense that, these can cooperate during operations. Second, these can input intermediate and final outputs from others. This concept allows for both data independence and use of processed data (which SI has made more applicable).

In the current version, the nodes shares (and process) all data directly in the core dbms (*Kallipateira*). It is a technique so that data are not managed by a central code but by their originators. Originators can thus anytime add, reuse and update data, while maintaining an optional independence. This core is coordinated to integrate both raw and processed soil information data. Thus, all input, intermediate, and output data are available to the users at any time. Note that, in the present research, “processed data” meant the intermediate and the final output, while “raw” data is the input from external data sources. Note also that, in a regular distributed dbms, it is up to the database administrator to manually determine how data collections (e.g. relational tables) are distributed over the nodes. The present system differs slightly in that the core dbms retains the alternative control over the network. This coordination allows for reliability control, ease of programming, diagnosis of process errors.

On the other hand, SI is the mediator between external databases and the core dbms. Interoperability with this knowledge expert enables data storage and dissemination both inside and outside the system. This offers many advantages. It also ensures the integrity of the external databases by the system itself. Unless the user seeks, for privacy reasons no processed data are allowed to adulterate the raw content of the external dbms.

However, the whole concept suggests a combination of technical features, which are: open source IDE, transmission protocols, compatible databases, files and data exchange formats, relational data model, record capacity, etc. Note that, syntactic (or syntactical) interoperability is achieved by the common syntax to exchange information (e.g., the same XML or SQL standards, the common core dbms, specified data files and data formats, communication protocols, interfaces of descriptions and the like) thus ensuring unambiguously defined formats of information exchange and common understanding of symbols.

3.3 Semantic and Conceptual Interoperability of GeoSeism

Modelling this interoperable GIS one aim was to be able to integrate and interoperate with more than one geotechnical methodologies. Future extensions are also included in this aim. It was therefore necessary to establish a common communication system at the level of semantic interoperability, so that the methodologies can understand (process) common geotechnical meanings by the employed geospatial data. SI found to be a much productive technique to undertaking this role of the mediator. It prepares the input which these methodologies use. **Figure 3** illustrates the logical structure of this architecture. Compared to LCIM [4], the present concept consists of more than one methodology at the conceptual interoperability level (modelling) and these all interoperate with each other.

To make the SI's role less complicated and able for generic application (so that it can be used for the future extensions), its methodology developed so as to prepare a sufficient input once. Therefore, SI designed to output a relational database of complete data content of a whole multi-thematic GSP. This output is ever sufficient to be input to the highest level of process (Level 4). On the other hand, the methodologies of this highest level are the end-users of data come from SI. However, their modeling has to be standardized so as to understand as input this database. The benefit is that these data-users will always avoid the demanding (long, know-how requiring, complex and uncertain) work of data processing which would be necessary in order to keep productivity. The notion of “**data applicability**” arises here meant “all input data (e.g. an input file) are able to be used and are able to be used the right way (understanding)”. This implies quality of data and differs from users' efficiency (data usability, ergonomics).

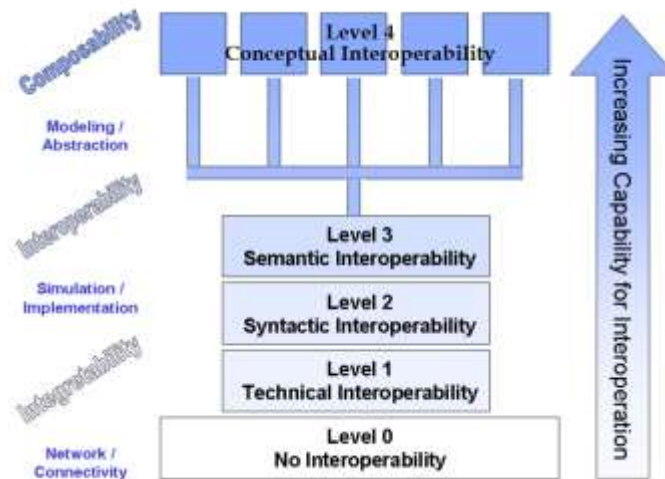


Figure 3: The Concept of *GeoSeism* Portrays a Variation to the Concept LCIM [4]:
More than One Methodology Interoperate in the Highest Level

The following principles outline how the semantic and conceptual interoperability are implemented in this concept.

1) Methodologies (apps or procedures) interoperate with each other.

The geotechnical methodologies (apps or procedures) which are the end-users of the data have a lot in common and are coordinated to interoperate with each other. Note that, these methodologies include GIS, logical and mathematical procedures. A representation of the flow-plan is illustrated in **Figure 4** in the form of an abstraction. The essential in this workflow is that the output database content from Level 3 is: (a) sufficient (meaningful) to serve a wide range of geotechnical methodologies, (b) independent from their modelling, and (c) always accessible by any (current or future) methodology and individual user.

Particularly, the procedures of *Artemis* elaborate a numerical analysis which estimates the seismic ground response due to earthquake-generated ground motions. Necessary inputs are the data about the expected bedrock motions (from the left app) and the SI's soil model (from SI below). *Artemis* is repeatedly interoperating with the right app exchanging data during the operation; sending frequency spectra and receiving corresponding time histories. The relational database output from *Artemis* can be shared to the apps lie above it so as they can make their own estimations or approximations (herein: structure response analysis, liquefaction potential analysis, pipeline vulnerability, soil-structure interaction, etc.). The apps of the interoperable GIS *GeoSeism* are mutually interoperating each other in order to elaborate automated microzonation studies (Level 4: conceptual interoperability level). SI (Level 3: semantic interoperability level) interoperates with *GeoSeism* to perform the demanding preparatory work of GSP. This way, SI establishes common geotechnical semantics to the apps which *GeoSeism* interoperates. However, the output from SI could be shared to a number of related geotechnical methodologies outside and inside the system. Note that, in geotechnical engineering, "ground response analyses are used to predict ground surface motions for development of design response spectra, to evaluate dynamic stresses and strains for evaluation of liquefaction hazards, and to determine the earthquake-induced forces that can lead to instability of earth and earth-retaining structures" [3]. *Artemis* is an 1D equivalent-linear analysis and could be replaced by any 1D or 2D non-linear analysis.

2) Data applicability depends to meaningful and composable data.

To improve the **applicability** of data the SI's process follows the DIKW model [1] [2]. It aims to create meaningful and composable GSP. The term "meaningful" meant that it relates multi-thematic data. The term "**composability**" meant that it creates complementary data.

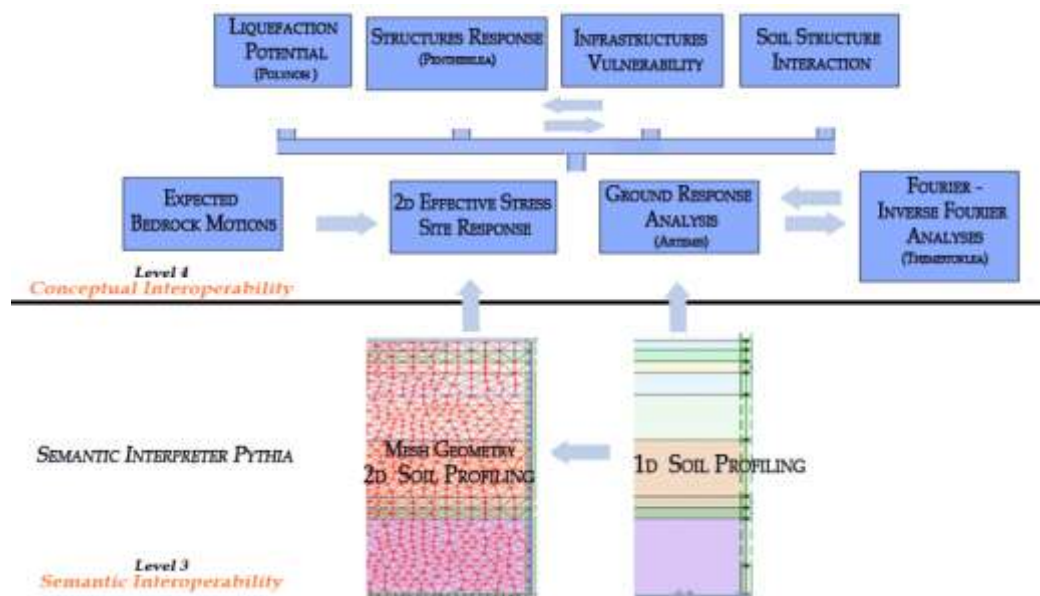


Figure 4: The interoperated Apps (Level 4) Interact and Cooperate to Share Data and Services, while SI (Level 3) Establishes Geotechnical Semantics for them

3) The GSP defines the common meanings in the larger system.

A GSP identifies the property and geometry of all soil layers and sub-layers. As explained above, the more complete and applicable this information is the better outputs all methodologies result. It represents the semantic interoperability level. It is the building block from which more complex operations may be constructed.

4) The GSP model is complete and defined to a spatial reference system.

It is an obvious benefit for the engineers to obtain ready information about a whole soil profile; rather than a flat map or schema. An additional advantage of SI is that, in the output database of SI, each soil column represents a standalone spatial reference system to which all multi-thematic data and soil formation ontologies can be referred.

5) The highest LCIM level can always be extended.

Because the highest LCIM level should be extendable, the aim of SI's GSP (Level 3) is to integrate as much information as possible (from available data sources). To this purpose, the system aims to share with a future public web-based community or a wide network of interconnected computers.

6) The outputs use a common database and relational key.

As mentioned above and **Figure 4** shows, the data exchange between the above-mentioned methodologies takes place through the common database. The methodologies read and write relational records in this core database. As a consequence, the aggregate content of data constitutes complete information, permanently stored, and available to be shared, reused and updated.

4. CASE STUDY, FINDINGS AND DISCUSSION

The case study includes the first wide performance of *GeoSeism*. It elaborated in the context of the same doctoral thesis. It was a microzonation study over an entire typical Greek city, which is a large-scale space. *GeoSeism* interoperated the currently developed apps (section 3.1). In this interoperation, SI's role was to provide ready GSP to these apps after processing the raw content of *HelGeoRDaS_uTH*. Apparently, it was an extensive case study. Consider that, there were analysed 342 IL and that the usual number of soil sub-layers that SI subdivides ranges from 30 to 150 in every IL. As expected, various research experiences distilled. A cumulative survey on the interoperability-related experiences is quoted below. Note that, the findings about the geotechnical methodology (seismic microzonation analysis and benchmark problems) of *GeoSeism* are subjects of subsequent related papers.

4.1 Applicability of LCIM

Figure 2 illustrates a cross-architectural schema of *GeoSeism*. As stated (section 3.3), the logical structure of *GeoSeism* is generally consistent with LCIM [4]. It is a variation of the concept LCIM. Note that, the dynamic and the pragmatic levels are not taken into consideration in the present research because these approaches are not related to the modeling of semantics. The purpose of the methodology is the conceptual level of interoperability maturity, assuming that problems at the lower levels are effectively resolved. The key to interoperate specific-purpose apps and databases at this highest level is the preparation of applicable data at the semantic interoperability level prior to data exchange. However, this requires demanding data processing as well as ongoing know-how and updating. SI is a very accurate and automated solution to performing this processing through permanent or temporary interoperability with the system.

4.2 Geospatial Data Inconsistency Problems Behind the Interoperable GIS

In the interoperable GIS, which consist of more than one data end-users (apps), the maturity of the system faces various problems of data inconsistency. This inconsistency leads to operational failure or to rarely detected errors in the output data. These in general include data inadequacy or ambiguity. The term “**data inconsistency related to geotechnical semantics**” is initially proposed by [2] to indicate these problems which restrict data applicability. Apart from the syntactic heterogeneity, which is always apparent, the most of these problems are related to the semantics and are rarely detectable. [2] outlines more this problem as well as presents a catalogue of inconsistency cases SI faces related to the semantics of data, along with actions and methods to deal with them.

Based on the experiences of the case study, this inconsistency is widely transmitted especially when more than one GSP are input together for analysis. For example, an automated study over the area of a city requires the input to include as many as possible soil profiles. It is thus obvious that, there is no “automated” seismic microzonation study and the so-called in the literature is rather an exaggeration. The role of the semantic interoperability level is to appropriately process the meanings of the exchanged data so that these are understood by the end-users. **Table 1** presents indicative data productivity measurements (applicability test) before and after improving by SI the raw data which are the input of *GeoSeism*. Note that, these results come from the currently materialized steps of SI (see [2], section 3.4, steps 1 - 5).

Table 1: Indicative data applicability test before and after improving data by SI

Problems against Applicability	Before Improvement (%)	After Improvement (%)
Empty Cells of Investigations (IL – Investigated Locations)	87	87
Cells of Sparse Investigations (< 6 IL/cell)	55	55
Depths < 30 % of Total Depth	83	83
Depths = 30 – 80 % of Total Depth	15	23
Depths ≥ 80 % of Total Depth	2	87
No Data about One or More Soil Layers	8	8
Certain Fields of the Records with No Data	88	78
Corrections over the In Situ Tests	0	100
No Soil Characterization	32	20
Insufficient Data for Soil Characterization	16	12
Approach a Soil Characterization	0	12
Alternative Classifications	0	12
Lack of Laboratory Tests	64	62
Lack of Groundwater Data	26	6
Up-to-Date (Active Communication)	0	3
Other Heterogeneity (e.g., Confuse Empty Fields with Zero)	3	0

4.3 Applicability of the Multi-Thematic GSP of SI Pythia

As aforementioned (section 3.3), the output from SI is a relational database of complete data content represents a detailed multi-thematic soil profile (here called GSP). The GSP which [2] proposes found to be sufficient for the present case study which elaborated a 1D ground response analysis.

It is applicable for a wide range of uses, because it relates quite detailed multi-thematic data and organizes in a core database the maximum possible information about the subsurface soil. At the same time, it is fully defined, implementation-independent and modelling-independent. However, the modelling of the end-users has to be standardized so as to read inputs by the SI's database. It remains to also develop a methodology to distribute this input from the site-specific concept into their potential 2D or 3D one. This is much effective today, because the technology trends tend to the interoperable sharing and impact and to the use of common standards on semantics [2].

4.4 Cross Specifications by the Present Research: Geotechnical Semantic Interpretation, Data Applicability, Composability, Applicability Measures, Interoperability, Interoperable GIS

The present case study confirms the following terms which the research [2] put under discussion:

Geotechnical Semantic Interpretation: It is an innovative type of spatial data processing designed to model the geotechnical semantics of multi-thematic data and provide ready applicable geotechnical information within the interoperable GIS. At the same time, it attempts to overcome a variety of interoperation failures related to semantics of data. The term emerged through the development and implementation of the expert knowledge algorithm SI *Pythia*.

The term “**applicable (geotechnical) data**” implies quality of data and differs from users' efficiency. It in brief reflects: (a) Data which are meaningful; (b) Data which prove composability regarding of fitness for use; and (c) Data which prove the organizational feasibility of the data structure.

Data composability meant the ability to evaluate (or approach) a target-element of data based on available data sets or elements. Composability is obviously increased in a meaningful database because the latter allows for new data evaluations. Problems associated with semantics (meanings) restrict composability. DIKW provides data values of all intermediate and final levels [2]. This advantage allows the independent use of each level's data to serve various purposes. It is also positive for the productivity of data.

Measures of data applicability: An applicable data model represents information about the world in a form that an interoperable GIS could productively use. A measure of applicability could be the level to which the DIKW model is effectively materialized [2]. Based on the case study, another measure of applicability could be the degree of data elements which an interoperable GIS can productively use from the core database.

Definitions which are most representative to the case of *GeoSeism* are: “**Interoperability** allows for the analysis of data in addition to the straight exchange” [11]. “Interoperability is the ability of systems to provide services to and accept services from other systems and to use the services so exchanged to enable them to operate effectively together” (ISO, [12]). “The development of interoperable GIS has long focused on the need for technically unrestricted interchange of both spatial data and traditional GIS operations and analysis” [10].

Semantic interoperability: It meant that, the meanings of the exchanged data are well established. This level is represented by the proposed semantic interpretation of the SI methodology. In essence, this level just prepares understanding data, does not estimate the adjective result of the system, albeit a GSP is so much meaningful an information that it can be used independently. It is the ability of two or more systems or elements to exchange information and to use the information that has been exchanged taking advantage of both the structuring of the data exchange and the codification of the data (e.g., including vocabulary) so that the receiving GIS can interpret the data.

Conceptual interoperability: It meant that, the modelling of the interoperated procedures efficiently impact and exchange data. It is concerned with using these meanings (as means) for the adjective purpose of the system. This definition is better understood by the descriptions in section 3.3.

Table 2 presents details some differences between the GIS data transfer, the distributed and the interoperable GIS considering a deployed *GeoSeism*. Based on experiences distilled from the development and applications of *GeoSeism*, a descriptive definition (see section 2.1) could be:

Research continues with the aim of evolving this technology. Extensions to the code could solve more semantic-related inconsistency problems for the benefit of the interoperable GIS. Based on the experiences distilled from the development and application of *GeoSeism*, a descriptive definition is: “**Interoperable GIS** is any GIS-based information system that comprises components which share data and impact over organized datasets, procedures or means in order to achieve commonly accepted goals. The advantages they offer are the diffusion of applicable for specific purposes geographic data, the more automated production of geographic information, the exploitation of the advantages of the internet, the saving of hardware, software or/and resource ware, such as the compliance (use and rights) to inter-operate

application algorithms with GIS data sources. They tend to an interoperable interpretation over available geographic data structures. They can impact and cooperate with GIS-based software applications and databases within one or more GIS operations (collection, storage, retrieval, management, visualization, and visual exploration in the Earth's space) and spatial data processes (modelling, process, analysis)".

Table 2: Differences between GIS data transfer, distributed GIS, and Interoperable GIS (compared to *GeoSeism*)

<i>Aspect</i>	<i>GIS Data Transfer</i>	<i>Distributed GIS</i>	<i>Interoperable GIS</i>
Scope	Data	Data and Process	
Semantic Interpretation (SI)	Manual	Depends on the Design	Interoperates and Improves Data Applicability
Architecture	Conventional	Multi-Tier	
Network Location	Same Physical Location	Nodes Can Far from Each Other	
Management	One Way	Unrestricted. The Nodes Are Managed by the Core Control	Optional. Each Node Is a Manager Itself. Alternative Core Control
Integration	In the Target System	In the Data Tier	In the Data Tier and Presentation Tier (Composable Data)
Database Access	Conventional	Through Remote or Distributed Way	
Data	Straight Exchange	Shares Output	Shares Input, Intermediate, and Output
Data Unit	Dataset	Objects (Datasets and Databases)	
Data Sources	User	Can Connect With External DBMS or Public-Purpose Network	
Applications	Blind	Interchange of Spatial Data and Processes. Traditional GIS and Specific-Purpose	Interchange of Data and Processes During Operations. Usually Additional Operations (Reasoning, Logical Decisions and Evaluations, Semantic Interpretation, etc.)

5. CONCLUSIONS

GeoSeism is a GIS-based application software which is intended to provide more automated seismic microzonation studies by utilizing the interoperable GIS technology. Proving the applicability of the LCIM (Levels of Conceptual Interoperability Model), it portrays an implemented combination of technical along with semantic and conceptual interoperability maturity, in which the modelling level of interoperability consists of more than one data end-users. Based on the experiences distilled from an extensive case study, this case faces problems of data inconsistency related to semantics. Inconsistency restricts the capability to derive geotechnical meanings from the employed geospatial data. The role of the semantic interoperability level is to appropriately process the meanings of the exchanged data so that the end-users can efficiently use these data. Otherwise, data are not fully applicable on geotechnical methodologies. The interoperation with the expert knowledge algorithm Semantic Interpreter *Pythia* (SI) proved especially helpful to undertake the role to improve the semantic interoperability in an automated manner. It modeled the semantics of data prior to data exchange. This innovative type of semantic interpretation is not related to the conventional search processes but to the spatial data process and the overcoming of a variety of many semantic-related data inconsistency. Distributed systems could evolve into geotechnical interoperable GIS combining typical application software with an SI. *GeoSeism* implemented a first concept of interoperable GIS which can model geotechnical semantics to ensure semantic interoperability, as well as allows for current or future interchange of geotechnical operations and applicable data for the benefit of engineers and GIS.

6. REFERENCES

- [1] **Papadopoulou M. A., Ioannidis G. S.**, "Mobilizing the Semantic Interpreter *Pythia* – Teaching Engineering Students to Integrate GIS and Soil Data During In Situ Measurements", In: Auer M., Tsiatsos T. (eds) Interactive Mobile Communication Technologies and Learning. IMCL 2017 (Advances in Intelligent Systems and Computing), vol. 725, Springer, Cham, 2018 a. DOI https://doi.org/10.1007/978-3-319-75175-7_19
- [2] **Papadopoulou Maria. A.** (2018b), "Semantic Interpreter *Pythia* : GIS-based Expert Knowledge Algorithm for Automated Geotechnical Soil Profiling of Applicable Data", Asian Journal of Engineering and Technology, Vol. 6, No. 6 (ISSN: 2321 – 2462).

- [3] **Kramer S. L.**, “Geotechnical earthquake engineering”, In: Prentice-Hall international series in civil engineering and engineering mechanics”, Prentice-Hall New Jersey, 1996.
- [4] **Turnitsa, C. D.**, “Extending the Levels of Conceptual Interoperability Model”, Proceedings IEEE Summer Computer Simulation Conference, IEEE CS Press, 2005.
- [5] **IEEE** (Institute of Electrical and Electronics Engineers), “Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries”, New York, NY: 1990.(iftikahr)
- [6] **Wikipedia**, “Interoperability”
- [7] **Berners-Lee Tim, Fischetti Mark., and Dertouzos T. Michael.**, “Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor”, Harper, San Francisco, 1999.
- [8] **ESRI**, “Spatial Data Standards and GIS Interoperability”, An ESRI White Paper, January 2003.
- [9] **Laurini R., Yétongnon K., Benslimane D.**, “GIS interoperability, from problems to solutions”, Advanced Geographic Information Systems, Vol. II, Encyclopedia of Life Support Systems (EOLSS).
- [10] **Goodchild Michael, Egenhofer Max, Fegeas Robin, Kottman Cliff**, “Interoperating Geographic Information Systems”, (Book), 1991. DOI <https://doi.org/10.1007/978-1-4615-5189-8>
- [11] **Marr J. Andrew, Pascoe T. Richard, and Benwell L. George.**, “Interoperable GIS and Spatial Process Modelling”, Proceedings of GeoComputation '97 & SIRC '97, 1997.
- [12] **ISO/TC 211**, “Geographic_information/Geomatics”, “Geographic_data_and_information”.
- [13] **Maheshwari D. et al.**, “Developing measures for benchmarking the interoperability of public organizations”, Proceedings of the 11th European Conference on eGovernment (ECEG), 2011.
- [14] **Bishr A.Y., Pundt H., Kuhn W., Radwan M.**, “Probing the Concept of Information Communities-A First Step Toward Semantic Interoperability”, In: Goodchild M., Egenhofer M., Fegeas R., Kottman C. (eds) (Interoperating Geographic Information Systems), The Springer International Series in Engineering and Computer Science, vol. 495. Springer, Boston, MA, 1999.
- [15] **Tolk, A. and Muguira, J.A.**, “The Levels of Conceptual Interoperability Model (LCIM)”, Proceedings IEEE Fall Simulation Interoperability Workshop, IEEE CS Press, 2003.
- [16] **Page, E.H., Briggs, R., and Tufarolo, J.A.**, “Toward a Family of Maturity Models for the Simulation Interconnection Problem”, Proceedings IEEE Spring Simulation Interoperability Workshop, IEEE CS Press, 2004.
- [17] **Guédria Wided, Naudet Yannick, and Chen David.**, “Interoperability Maturity Models – Survey and Comparison”, R. Meersman, Z. Tari, and P. Herrero (Eds.): OTM 2008 Workshops, LNCS 5333, pp. 273–282, Springer-Verlag Berlin Heidelberg, 2008.
- [18] **Tolk Andreas, Diallo Y. Saikou, Turnitsa Charles.**, “Applying the Levels of Conceptual Interoperability Model in Support of Integrability”, Interoperability, and Composability for System-of-Systems Engineering, 2013.
- [19] **Geographic Information Systems**, “GIS Design: A Review of Current Issues in Interoperability”, Geography Compass. **3**: 1105–1124. doi:10.1111/j.1749-8198.2009.00241.x.
- [20] **Tolk Andreas**, “What Comes After the Semantic Web - PADS Implications for the Dynamic Web”, PADS, pp. 55-62, 20th Workshop on Principles of Advanced and Distributed Simulation (PADS'06), 2006.
- [21] **Janssen M.**, “Measuring and Benchmarking the Back-end of E-Government: A Participative Self-assessment Approach”, In: Wimmer M.A., Chappelet J.L., Janssen M., Scholl H.J. (eds) Electronic Government. EGOV 2010. Lecture Notes in Computer Science, vol 6228. Springer, Berlin, Heidelberg, 2010.
- [22] **Thoen B.**, “Interactive Mapping and GIS Thrive on the Web”, GIS world, 8(10): 58-59, 1995.
- [23] **Davis P. K., and Anderson R. H.**, “Improving the Composability of Department of Defense Models and Simulations”, RAND Corporation, 2003.
- [24] **Wikipedia**, “Semantic Interoperability”
- [25] **Kulhawy F. H., and Mayne P. W.**, “Manual on estimating soil properties for foundation design”, Report EL6800. Electric Power Research Institute, Palo Alto, 306 p, 1990.
- [26] **Papadopoulou Maria. A.**, “Automated methodology for seismic hazard microzonation studies of interoperable geographic information systems – the case study of a Hellenic city” (Doctoral Thesis), Department of Civil Engineering, University of Thessaly, Greece, 2017. (in Greek at <https://www.didaktorika.gr/eadd/handle/10442/40605?locale=el>)
- [27] **SDGEE**. Research Unit of Soil Dynamics and Geotechnical Earthquake Engineering. “Microzonation”.
- [28] **Papadopoulou Maria. A.** (under review). “*HelGeoRDaS_uTH*: A Multi-Thematic Geotechnical Database System About The Subsurface Soil And Underground Infrastructures In Thessaly (Greece)”, International Journal of Spatial Data Infrastructures Research (<http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/487>).