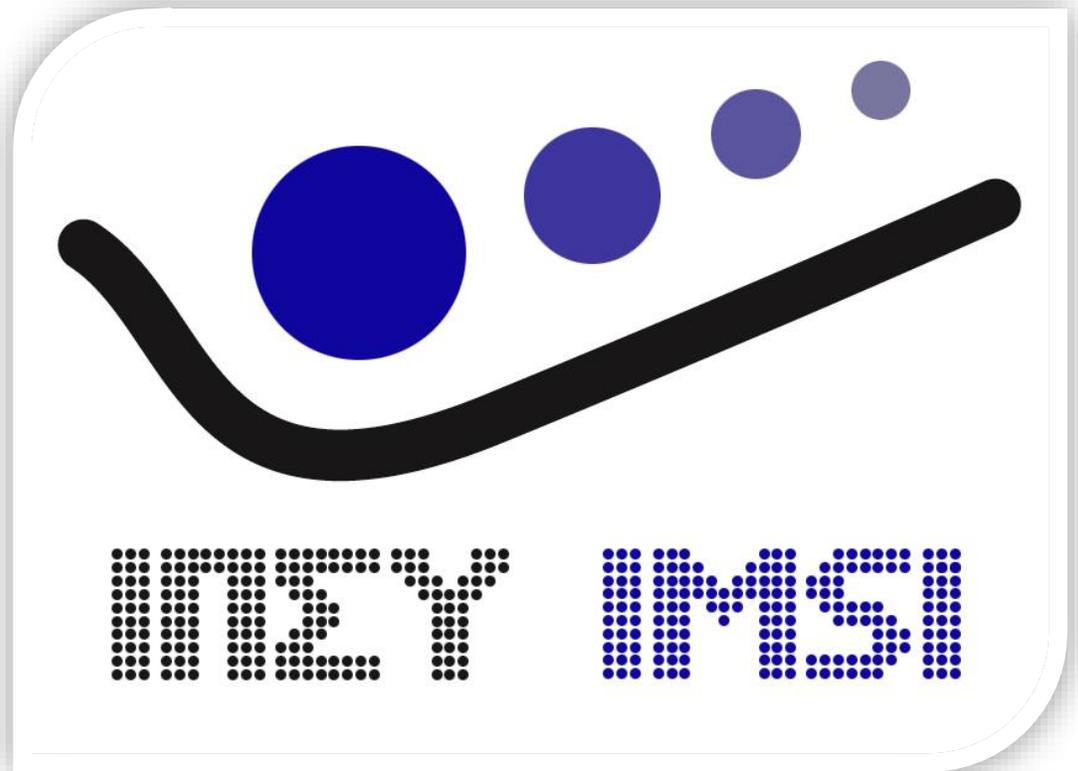


# IMSI annual report 2019



# Preface

---

Welcome to the Information Management Systems Institute (IMSI) of the ATHENA Research and Innovation Centre! Established in 2007, IMSI is today one of Greece's premier research centers in the areas of large-scale information systems and Big Data management. Over the past few years, IMSI researchers have been very successful in attracting and implementing numerous cutting-edge research & development projects, at both the national and international level; furthermore, IMSI has created strong collaborative ties with top European research institutions and has successfully promoted the development and use of state-of-the-art information systems in both local industry and various Greek government organizations.



As in previous years, in 2019, IMSI has significantly expanded the scope of its research efforts in a number of focus areas, including Big Data and Scalable Data Analytics, Web Data Management and Semantic Web Technologies, Geospatial Data Management, and Digital Curation and Research Infrastructures. In addition, IMSI researchers have led and/or participated in numerous activities promoting research and educational excellence in the areas of information systems and data management, as well as the development of novel software platforms and services, made available to the research community and employed by both local and international users.

Over the next few years, IMSI aims to continue to strengthen its collaborative ties with local and international industry and academia, promote the transfer of state-of-the-art information technology to national organizations and industry, and continue to strive for excellence further increasing the visibility of its research efforts and results.

***Prof. Minos Garofalakis***

*Director, Information Management Systems Institute (IMSI)*

*ATHENA Research and Innovation Centre*

*Athens, Greece, 1/10/2020*



# Contents

---

- Preface ..... 1
- Profile ..... 3
- Infrastructures..... 6
- Highlights ..... 8
- Research Directions ..... 10
- Projects ..... 23
- Publications ..... 60
- Dissemination Activities..... 65
- Systems ..... 67
- Education..... 76
- Facts and Figures ..... 78
- Contact ..... 84



# Profile

---

The Information Management Systems Institute (IMSI) is one of the research institutes of the Research and Innovation Center in Information, Communication and Knowledge Technologies "ATHENA", which is the first research center in Greece with a focus exclusively on Information Society. Established in Athens in 2003, it is a research and technology organization supervised by the General Secretariat for Research and Technology of the Ministry of Education and Religion. IMSI was founded in 2007 with the mission to conduct research in the area of data management and large-scale information systems. In 2009, the Digital Curation Unit, which had been established in 2007 in "ATHENA" R.C, became a unit of IMSI.

The research at IMSI has a strong collaborative aspect, and ranges from basic to applied research. The collaborative aspect is expressed in that research is conducted with national and international partners from industry as well as academia, often also in the context of novel and innovative projects.

The mission of IMSI is to conduct research, develop applications and products and to offer services in the areas of information management and large-scale information systems. IMSI is particularly interested in the areas of database technology, software engineering and development methodologies, managing and exploiting information resources, as well as their large-scale applications.

The objectives of IMSI are:

- to participate and carry out research and development projects in the area of information technology and to cooperate with the academic community in topics of research, education and knowledge transfer to the industry.
- to produce experimental and industrial prototypes, and to develop new innovative products in cooperation with industry.
- to transfer knowledge on research and development topics and to offer training and certifications on information technology and related processes for the management of information and information systems.

To attain these objectives, IMSI is structured as follows:

- The Department of Database and Business Intelligence is responsible for basic and applied research in (a) database management systems, and (b) business intelligence systems. The focus is on design, optimization, scalability, security, privacy and high availability of databases and data warehouses, data mining and ETL.



- The Department of Distributed and Web Information Systems conducts basic and applied research on data management problems rising in Web applications and distributed computing in general. Focus is on methods and technologies to support search and exploration on the Web during a creativity cycle: from the abstraction describing the search domain, the information harvest and retrieval tasks and the adaptation of results to user needs, to the classification of results and exchange with other users. One of the use-cases we adopt is based on R&D for Web applications for biosciences. Strong interest also exists for modeling and data management issues for the Semantic Web and ontologies, heterogeneous data source integration problems, web services, and semistructured data storage and querying. Finally, focus is also given on real-time information systems, like sensor networks and p2p/grid systems.
- The Department of Geoinformatics conducts basic research and development of prototypes in the areas of geospatial data management, i.e., Geographic Information Systems, Spatial Data Infrastructures, Location-based services, geospatial data management on the Web, and human-machine interaction with spatiotemporal information.
- The Department of Scientific Databases and Simulation conducts basic and applied research on the modeling and efficient implementation of data management systems that support scientific applications and simulation processes. The focus is on biological data, environmental data, the management of evolution in scientific databases, and the management of imprecise information. Moreover, the Department concentrates on issues of design, development, and management of Digital Libraries.
- The Department of Software Engineering and Project Management Software is responsible for developing innovative applications that integrate cutting-edge research results and responding to needs of the Public and the Private Sector, based on professional standards and methodologies. It also conducts studies and provides consulting services and technical support to entities involved with IMSI. It also offers certifications to individuals and institutions with respect to the proficiency in software development and management of software project lifecycle (technical project description, project contract, project coordination, project delivery). Furthermore, it has the responsibility of training programs in specialized software issues, software development environments and project management software.
- The Digital Curation Unit (DCU) has the mission to conduct research, develop technologies and applications, provide services and training, and act as a national focus point in the field of digital curation. Digital curation encompasses a set of activities aiming at the production of high quality, dependable digital assets; their organization, archiving and long-term preservation; and the generation of added value from digital assets by means of resource-based knowledge elicitation. To ensure the adequate capture of the context of digital resources and their subsequent creative and effective use, the DCU adopts a multidisciplinary approach that



considers the full lifecycle of digital assets, such as records, digital surrogates and scholarly/ scientific datasets.

The activities of IMSI departments are supported by the “ATHENA” R.C. Economic and Administration Office.



# Infrastructures

---

IMSI IT Infrastructure provides the necessary resources and services to support the organizational needs of the Institute and the development and support of quality IT services and solutions. It consists of the hardware, software and network components that are used in order to achieve the above goal.

In an effort to combine the best services IMSI Infrastructure is built upon both cloud computing and traditional locally implemented solutions. Local IT Infrastructure consists of servers hosted in IMSI owned facilities in a dedicated server room and is largely based on Virtualization technologies in order to achieve server consolidation and maximize the hardware's efficiency. It is the base upon which most of the Institute's IT services are built while at the same time it provides the required resources for a large number of projects the Institute participates in. It is also used by the affiliated researchers and students for larger scale experimentation and research.

Along with maintaining its own private local servers, IMSI also takes advantage of the benefits of cloud infrastructures. The resources made available to the Greek academic and research community through GRNET "Okeanos" IAAS Service are heavily used by the Institute's users for research and development purposes. Also, in collaboration with Microsoft and its Academic program Office 365 is used for providing mail and collaboration services.

Through the described infrastructures IMSI provides to its members and affiliates a variety of services, such as:

- Mail Services
- Directory Services used for centralized authentication and authorization
- Source Control
- Shared storage
- Virtual Private Network
- Web publishing
- Project Management and Collaboration
- Communication Services

The above services are provided and implemented using both commercial and open source operating systems and software, such as but not limited to:



- Operating Systems: Debian Linux, Ubuntu Linux, CentOS Linux. Microsoft Windows Server
- Virtualization Software: KVM
- Database Server: PostgreSQL, MySQL, MariaDB
- Distributed Processing: Apache Hadoop

IMSI network infrastructure provides high speed connectivity to its users and the provided services. IMSI network connects to the Internet using a 1Gbps fiber optics connection to GRNET. It consists of several Gigabit switches that offer wired connectivity and takes advantage of the Research Center's Wireless Infrastructure to provide high speed and reliable Wireless Connectivity. Also, through the Research Center's participation to the Eduroam Initiative, IMSI members can use their account to gain wireless internet access in research and academic institutions in more than 70 territories in the world.



# Highlights

---

## Events

- The **17th Hellenic Data Management Symposium** was organized by [Theodore Dalamagas](#) (IMSI, Athena Research Center) and [Georgia Koutrika](#) (IMSI, Athena Research Center), who were the general chairs, and Stratos Idreos (Harvard University) as the program chair. The 17th Hellenic Data Management Symposium ([HDMS](#)), took place in Athens, July 8-9, 2019. The HDMS is the annual forum for the Greek database community to present and discuss recent advances for data-intensive applications. This year's HDMS covered a wide range of topics with an interactive, fast-paced program. We invited 3 keynote speakers, and hosted 2 industry talks, 21 research talks and 20 flash talks. We also invited 7 demos, organized 2 panels, an entrepreneurship panel giving the floor to 6 startups as well as an AI/DB panel. Over 140 people attended the event from academia and industry, from Greece and abroad (including USA, Germany and Switzerland).
- The 1st Summit on Gender Equality in Computing (GEC 2019) was held in Athens and organized by the ACM-W Greek Chapter (Georgia Koutrika is a treasurer in the chapter and General co-chair of the summit).
- The INODE - Intelligent Open Data Exploration (Grant No: 863410) - EU H2020 - H2020-EU.1.4.1.3. ranked first and started in November 2019. Project Coordinator from IMSI: Georgia Koutrika.
- The **INFORE H2020** project, ranked 1<sup>st</sup> out of 77 submissions, started on 1/1/2019. Project Coordinator from IMSI: [Antonios Deligiannakis](#).
- The EU H2020 research and innovation project **SmartDataLake (Sustainable Data Lakes for Extreme-Scale Analytics)** started in January 2019. SmartDataLake provides an adaptive, scalable, and elastic data lake management system. The data lake's contents are modeled and organized as a heterogeneous information network, implementing various functionalities for exploration and analysis, while also offering interactive visual analytics. The project is coordinated by IMSI, Athena RC ([Dimitris Skoutas](#)).
- A new EU project titled **"Novel EOSC Services for Emerging Atmosphere, Underwater & Space Challenges – NEANIAS"** funded by H2020-EU.1.4.1.3 started in *November 2019*, which will develop innovative thematic services in three major sectors: underwater, atmospheric and space research. Project Coordinator from IMSI: [George Papastefanatos](#).
- A new collaboration between **IMSI, Intracom Telecom S.A. and Ericsson S.A.** started in *December 2019*. IMSI has been contracted to provide design and development services for Ericsson Device EcoSystem, a digital marketplace for collecting compliance specifications of IoT devices and manufacturers with international standards. Project Coordinator from IMSI: [George Papastefanatos](#).

- The EU (H2020-MSCA-ITN-2015) project ARCADES organized successfully the following three events during 2019:
  - a. The **2<sup>nd</sup> Software & Industrial Workshop**, that took place in Cambridge, UK on January March 28<sup>th</sup>.
  - b. The **3<sup>rd</sup> Learning Week on Growth and Harvest** that took place at the INRIA research center at Sophia Antipolis, France on March 27<sup>th</sup>.
  - c. The **Final Open Workshop** of the Project that took place at the Technical University of Vienna, Austria on November 27<sup>th</sup>.

Project scientific coordinator: Ioannis Z. Emiris.

- Organization of the [1st Summer School in Digital Humanities](#), Athens, 15-19 July 2019.
- The DSIT Program is organized by the ATHENA Research Center, the Biomedical Research Foundation of the Academy of Athens (BRFAA) and the Department of Informatics & Telecommunications of the National & Kapodistrian University of Athens (NKUA). IMSI Directors of Research, Theodore Dalamagas and Georgia Koutrika are responsible for two classes, and supervise several diploma theses.
- Cooperation with the MSc Programme in Digital Methods for the Humanities, Athens University of Economics and Business.
- European Commission Training school on Economics of Sustainable Water Management of Intermittent Rivers and Ephemeral Streams: SMIRES, ATHENA Research Centre.
- 19-18/02 Prof. Koundouri organizer and coordinator, [5th GLOBAQUA Training Course](#): 'Economics of Sustainable Water Management in accordance to the Water Framework Directive (WFD), the Millennium Ecosystems Assessment (MEA) and Sustainable Development Goals of the UN Agenda 2030' Athens, Greece.

## Awards

- Phoebe Koundouri: Chair of the Best Dissertation Award of European Association of Environmental and Resource Economists ([EAERE](#))



# Research Directions

---

IMSI research activities fall into the following areas.

## Big Data and Scalable Data Analytics

*Current (ongoing) research directions in this area include:*

- **Analysis-aware Entity Resolution over Big Data**  
Exploration and analysis of dirty data have gained great attention recently due to the emergence of data aggregators; i.e., organizations that harvest, aggregate and analyze data containing overlapping and usually contradicting information from multiple sources. Analysis-aware data processing refers to an exploratory analysis scenario, where users apply traditional data integration methods, such as cleaning, repairing and deduplicating, during query time. Analysis-aware Entity Resolution (ER) is a special case which aims at extending the results of the query by resolving duplicate entities (records that represent the same real-world entity) during query time. In this direction, we have proposed methods and a Framework, called *QueryER*, that integrates ER operations into the planning and execution of SPJ queries. To achieve that, we propose novel (ER-specific) query operators, which identify and resolve duplicates within a table by employing a schema-agnostic resolution approach with no configuration overhead; join duplicate entities between two or more tables and group/merge deduplicated entities into a single representation. We integrate these operators into normal query execution.
- **Physical optimization for large-scale, data science workloads**  
Modern Data Science processing workloads typically involve computations of extreme-scale analytics that can be encoded in various forms (e.g., queries, workflows, programs) and executed on more than one platform. Parts of the processing could be pushed to the edge level (e.g., input sensors), while other more computationally intensive parts (e.g., stock correlation in finance, gene simulations in life sciences) could be executed on one or more, potentially distributed, Big Data platforms or clusters (e.g., GPUs) of a supercomputer or in the cloud. The decision on what is the right platform and timing to execute a Data Science workload is based on a multitude of criteria and optimization objectives, including hardware and processing capabilities, scheduled and running workloads, available resources and pricing, and so on. In this project, we develop tools and techniques for optimizing (e.g., in terms of runtime, throughput, latency, scheduling, system resources, monetary resources) the execution of Data Science workloads across different computing platforms.



- **Transparent in-situ data processing**

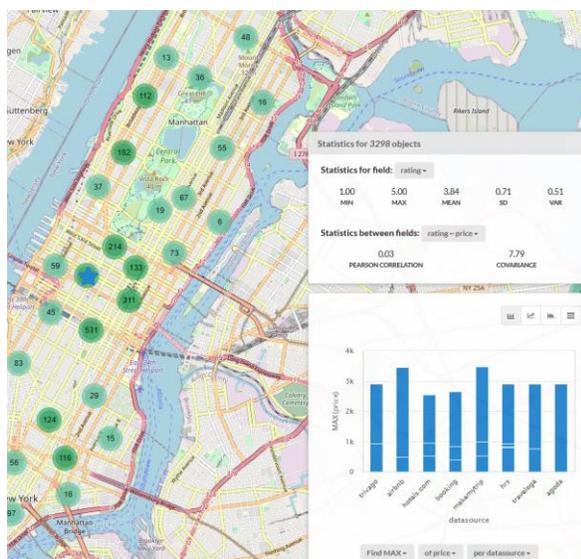
Bringing computation closer to data is critical when dealing with large datasets in highly distributed systems with specialized hardware components. Examples include supporting in-situ data transformations (e.g., decrypt on read) and tier crossings (e.g., select the form of compression based on physical medium to be moved to). The goal of this work is to improve the performance of data processing engines by leveraging hardware specificities, without affecting the interface to applications. Our approach improves scalability by decoupling the engine primitives from the underlying data store platform, avoiding data transfers through the various levels of the stack, and leveraging specific hardware.

- **Self-service scalable visual analytics**

One of the major challenges in the Big Data era is that it has realized the availability of a great amount and variety of big datasets for analysis by non-expert data analysts, such as research scientists,

data journalists, policy makers, SMEs and individuals. The level of difficulty in transforming a data-curious user into someone who can competently access, analyze and consume that data is even more burdensome now for a great number of users with little or no support and expertise on the data (pre)processing part. We are currently working in developing a scalable platform for self-service visual analytics. Self-service visual analytics is a new

paradigm, in which users are enabled and encouraged to directly manipulate (explore, blend, analyze) underlying data in rich visual ways, in order to derive insights from information as quickly and efficiently as possible. Our platform called *VisualFacts* (<https://visualfacts.imsi.athenarc.gr/>), enables the visualization of big geo-located data and helps data explorers perform ad hoc analysis of raw data files collected from different sources of varying quality (with duplicates or missing data) in rich visual ways. The backbone of the platform is a visual aware in-memory index, which is constructed on the fly and adjusted to user interaction, as well as a powerful deduplication engine which offers on-the-fly visual entity matching and clustering over dirty data. The platform can scale up the visualization, interactive exploration and analysis to million data points on a map, with the use of commodity hardware. Furthermore, we have developed *Socioscope* ([www.socioscope.gr](http://www.socioscope.gr)), a visual analysis tool, used by social scientists, for the visualization and exploration of social and political data. Another application developed within the project *CitySense* aims



to integrate city data from disparate sources and provide a visual way to combine them and filter city areas.

- **Mathematical Modeling and Analysis**

Mathematical modeling of the physical world is crucial in a number of applications. Despite the complexity of three-dimensional models, current algorithms and software are making enormous progress in efficiently representing, handling, exchanging, and operating on such models. The ARCADES Network contributes in this direction by exploiting cutting-edge research in mathematics and algorithm design so as to design and implement robust methods in Computer-Aided Design and manufacturing (CAD/CAM). However, geometric modeling is facing new challenges in modern engineering analysis, simulation, manufacturing, and construction. This is becoming evident in new sectors such as the movie and game industry, where CAD methods are not penetrating fast enough, or are facing new challenges arising from massive and fast point acquisition (e.g. by laser scanners), big data and mobile computing. This captures precisely the challenge taken up by the ARCADES Network, namely to build the next generation of CAD software based on strong mathematical foundations from computer algebra, geometric computing, numerical analysis, and algorithm design. The crux of our method relies on algebraic representations, understood in the widest sense of the term, namely relying on polynomial expressions, and including parametric, implicit, and semi-algebraic representations, which can be converted to/from further representations such as point clouds or subdivision surfaces. Compared to classical discrete representations, algebraic representations not only drastically reduce the size of the data, but also provide a small number of parameters (e.g. control points). The Network participants represent a multidisciplinary and multisectoral spectrum for implementing this vision, thus also offering an excellent opportunity for career development to the ARCADES fellows.

- **Scalable Analytics for Social Data**

The increasing use of online social networks and microblogging platforms, such as Facebook and Twitter, and the content generated by millions of users, has led to the development of a multitude of innovative applications based on the analysis of big data from social networks. The selection of an appropriate sample for each different application data analysis is critical to the quality of the analysis results. In this direction, technologies supporting efficient collection and modeling of social network data are key components of an innovative information management platform for social networks. Such a platform fills the gap that exists between the restrictive and cumbersome interface offered by online social networks and the need of applications for easy collection and flexible sample selection. At IMSI, we are developing a novel platform, called *TwitHoard*, for managing information obtained from OSNs and we incorporating in it the results of relevant research activities. Moreover, the National (Synergasia) project *RealEstate2.0* develops an infrastructure



for collecting and integrating comments about specific geographical areas from various social media platforms.

- **Automatic Generation of Feature-Agnostic Datasets for Fake News Detection**

Fake news is growing into one of the most crucial issues for social media platforms, users, and news organizations. The development of efficient algorithmic solutions for detecting fake news in online social networks requires complete, up-to-date, and flexible training datasets. Fact-checking services can be very useful for providing fake news stories; however, existing datasets suffer from severe limitations and rely heavily on human annotators. In this work we developed PHONY, an infrastructure that leverages Twitter and fact-checking websites to automate as much as possible the generation of flexible, feature-agnostic datasets. This allows users to extract suitable feature-specific datasets according to the machine learning approaches used. Our feature-agnostic datasets support the wide range of features encountered in the literature, including semantic features and social network diffusion features, which have not received much attention.

- **Privacy Preservation**

The protection of user privacy in data analytics is one of the major challenges faces by modern information systems. IMSI has worked extensively in the field of data anonymization, where personal data are transformed to anonymous where the identities of individuals are hidden and sensitive properties can no longer be attributes to them. IMSI has developed a series of anonymization algorithms for complex data (tree structured, RDF, set-values) and a tool, AMNESIA, that allows non-expert users to use anonymization techniques on their data. AMNESIA is offered in the OpenAIRE infrastructure and it is also part of the European Open Science Cloud (EOSC). It is available at <https://amnesia.openaire.eu>. The work on privacy preservation has resulted in a series of publication and it is used in the *OpenAIRE*, *My Health My Data* and *MEDA* projects.

- **Scientific databases and bioinformatics**

We work on models, architectures and methods to store, preserve, process and query genomic and life science data. The vision is to provide high-performance computing methods and tools to perform real-time analysis and processing of big and complex life science datasets. In

the past, emphasis was given to genomic data related to miRNAs sequencing process and their analysis. IMSI, Univ. of Thessaly and GRNET have designed, implemented and

maintained a research infrastructure for genomic data management, oriented to processing, analysis and visualization of computationally predicted miRNA targets



(<http://diana.imis.athena-innovation.gr>). Part of the work is being integrated into the ELIXIR-GR research infrastructure that develops the Greek Node of the ESFRI European RI ELIXIR (<https://www.elixir-europe.org>), a distributed e-Infrastructure aiming to build a sustainable European infrastructure for biological information. In addition, we are currently leading the design and implementation of the ELIXIR-GR Cloud Infrastructure, which is based on containerization technologies to support all computational needs of the ELIXIR-GR community. The relevant research is supported by EXCELERATE, an H2020-INFRADEV project for fast-track ELIXIR implementation and ELIXIR-GR, a Greek project of the National Structural fund to support the Greek Roadmap for Research Infrastructures.

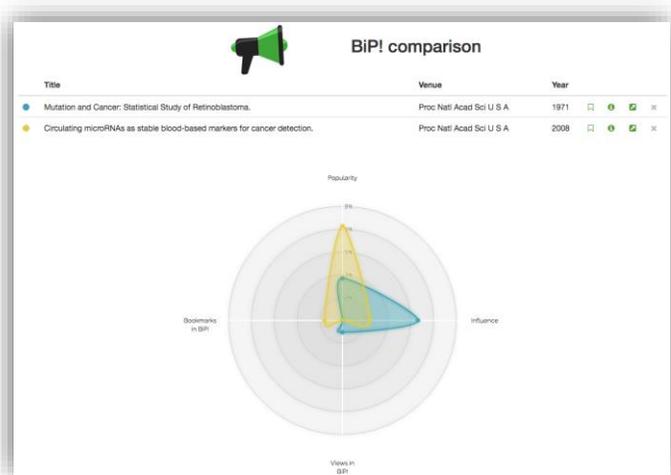
- **Text Mining and Information Retrieval for Scientific Publications**

Scientific publications retrieval and ranking based on their impact has been a long-established research topic,

especially due to the massive number of papers being published annually.

We are working on technologies to identify high-impact research output, focusing on the study of impact aspects like long-term and short-term impact. Our approach is based on PageRank adaptations for citation

networks and machine learning on paper metadata. We also apply text mining techniques on scientific texts to automatically extract useful knowledge. Finally, we develop powerful information retrieval systems that are based on the previously described technologies to provide useful services to the research & science community (e.g., BIP! Finder and SciTo visualiser).



- **Information and Communications Technologies for Big Data**

Specifically broadband networks and converged telecommunications and media services, cloud based large data processing, storage and network programming systems. The research is supported by EU and National projects related to scientific data infrastructures, currently ELIXIR-EXCELERATE and ELIXIR-GR. Note that IMSI, in collaboration with GRNET and Univ. of Thessaly, has designed, implemented and maintained an e-science platform for genomic data management, oriented to processing, analysis and visualization of computationally predicted miRNA targets (<http://diana.imis.athena-innovation.gr>). The platform is running on GRNET's cloud, providing reliable and high-performance data analysis and processing services to more than 2K active users.

*New (emerging) research directions in this area include:*

- **Big Data Analytics**

The attention Data Science receives recently is partly due to the availability of huge volumes of data and the opportunity to mine and extract useful knowledge. IMSI has already a significant activity and expertise around Big Data management, thus a natural next step is to focus on Data Science and Big Data Analysis. An interesting objective for IMSI is to examine how the machine learning techniques and statistical methods of Data Science can be combined with the Big Data need for scaling, particularly for non-conventional data types, like temporal, spatial, graph, stream, and scientific. In this direction we are working the on parallel execution of interval joins on modern CPU architectures. Our parallel implementation scales gracefully with the number of threads reducing the cost of the join to at most one second even for join inputs with tens of millions of rectangles We also plan to investigate the processing of deep learning operations (e.g., matrix factorization) in the same context.

- **Data-driven Circular Economy**

We have launched a multi-disciplinary research and innovation activity in data-driven approaches to unlock the circular economy potential. Contrary to the 'take, make, dispose' production model of a linear economy, in a circular economy, resource input, waste, emission, and energy leakage are minimised by narrowing material and energy loops through recycling, reuse, remanufacturing, repair, etc. Our aim is to build on state-of-the-art technologies and scientific solutions in data science and big data analytics, and provide novel methods, algorithms and tools to collect, process and analyze data to support key areas of circular economy, like smart waste management, automatic detection of resource/material flows in industrial networks, analysis and visualization of complex industrial symbiosis networks, etc.

- **Cloud Infrastructures**

Based on the experience from CONFES project that developed integrated wireless-wired transmission network infrastructure of ultra-high capacity optical technologies, project proposals are submitted to address challenges of the Future Internet especially in Network Function Virtualization - NFV / Software defined Networks –SDN for example by elaborating and constructing a dynamic Cloud infrastructure and in particular a converged Telco and IT node in access networks. One goal is to support advanced 5G/IoT services and applications with demanding QoS and edge analytics needs, with applications in many areas of great interest like in precision farming and 5G access and backhaul networks.

## Semantic Web Technologies

*Current (ongoing) research directions in this area include:*



- **OLAP Analytics in the Web of Data**

Another direction of research is related to the application of efficient analytics on multidimensional RDF data, i.e., data usually treated under the OLAP prism, where they are represented as observations that are instantiated over pre-defined dimensions and measures (similar to traditional DW modelling). The increasing volume and diversity of these data (statistical authorities, academic institutes, financial organizations and pharmaceutical companies publish such data) pose the challenge of finding hidden relations between them in a most efficient and accurate way, with the aim to detect inconsistencies or infer new facts. We have addressed this problem, by introducing new relationships (e.g., containment and complementarity of data) between multidimensional RDF data, and new algorithms for efficient and scalable computation of these relationships.

- **Scalable Query Processing in the Web of Data**

Another line of work addresses the need for efficient processing of SPARQL queries over voluminous RDF stores. Many indexing approaches have been proposed in this area; still few of them take into account the inherent structure of RDF graph data and how this structure can be exploited for efficient processing and optimization of SPARQL queries. For that, we have developed a scalable approach for query processing of RDF stores, based on a novel indexing technique, called Extended Characteristic Sets, which builds on top of the characteristics of the triples in an RDF dataset and enables the efficient processing of complex multi-join queries. We have also developed SRX, an extension of the popular RDF-3X system, which adds support for spatial data. SRX supports three types of spatial queries: range selections (e.g., find entities within a given polygon), spatial joins (e.g., find pairs of entities whose locations are close to each other), and spatial  $k$ -nearest neighbors (e.g., find the three closest entities from a given location).

*New (emerging) research directions in this area include:*

- **Distributed Storage and Parallel Query Execution**

Future interesting directions involve the distributed storage of the index and the parallel execution of such queries. Future directions include the extension of these techniques to cover more complex analytic techniques such as finding missing values, outliers and causalities on such data and make them scalable for very large volumes via parallelization. Some emerging directions, concern online entity resolution techniques, which aim at integrating the blocking\metablocking\entity matching tasks (in the form of query operators) in the query processing phase. The goal is to enable users to seamlessly analyze on query-time heterogeneous (e.g., different schemas, disparate data sources) datasets that involve highly noisy data of different quality.

- **Efficient storage management and query processing for Large Knowledge graphs.**



An emerging area of work is in the area of RDF Indexing and Query Processing for Big Knowledge Graphs. We are working in a [scalable approach](#) for storing in relational databases and querying RDF knowledge graphs based on the notion of Characteristic sets (CS). CSs organize graphs based on the set of properties associated with their subject nodes. This concept was recently used in indexing techniques, as it can capture the implicit schema of RDF data. While most CS-based approaches yield significant improvements in space and query performance, they fail to perform well when answering complex query workloads in the presence of schema heterogeneity, i.e., when the number of CSs becomes very large, resulting in a highly partitioned data organization. We work on a novel technique, for merging CSs based on their hierarchical structure. Our method employs a lattice to capture the hierarchical relationships between CSs, identifies dense CSs and merges dense CSs with their ancestors. We implement our algorithm on top of a relational backbone, where each merged CS is stored in a relational table, and therefore, CS merging results in a smaller number of required tables to host the source triples of a data set.

## User-Data Interaction Systems

*Current (ongoing) research directions in this area include:*

- **Intelligent Data Interfaces**

The volume of data and the need for data democratization call for shift from the classical «query-based information access paradigm» popularized by existing systems to a novel information access paradigm, where the system takes upon a more conversational and active role in helping users effectively explore data of varying quality, complexity, and relevance. We are looking into text-to-SQL systems, i.e., systems that allow users to ask queries using natural language. We have developed and experimentally compared several systems that employ a wide range of techniques. A new query benchmark is developed that is the first to enable fine-grained comparison of text-to-SQL systems.

A Spark-based architecture of an intelligent data assistant, a system that combines real-time data processing and analytics over structured and unstructured data with user interaction in natural language, is built and presented in the Data and AI Summit 2019 - Databricks. The system brings together several components, including natural language processing for understanding user questions and generating answers in natural language, automatic knowledge base construction techniques for learning about data sources and how to find the information requested, as well as machine learning methods to tackle query disambiguation and context modeling.

- **Intelligent Interactive Data Exploration**



We are working on algorithms that allow the system to actively guide the user through their information access endeavor by offering query recommendations, exploration options, and help as needed. We are looking into this area from three different perspectives:

- (a) The data perspective: where the objective is to explore non-conventional types of data and inputs to recommendations,
- (b) The problem perspective: where the objective is to study recommendation problems beyond typical ones focusing on user consumption problems (selecting movies to watch, products to buy, applications to download, and so forth). Such problems include data exploration, query optimization, visualization, data integration, and workflow design among others, where the purpose is to select tuples, queries, views, exploration actions, query plans, visualization graphs etc.
- (c) The methods perspective: where the objective is to build on the newest algorithmic developments in the field of recommender systems.

*New (emerging) research directions in this area include:*

- **Intelligent Data Interfaces to Databases**

Several aspects of such systems are emerging:

- (a) interaction in context: the system keeps track of the context (e.g., goals, background information, earlier queries) of a user quest, and it can respond and adapt accordingly
- (b) proactive systems: the system anticipates user needs proactively presenting information and suggestions without the user explicitly prompting the system, thus promoting continuous and easier learning.

- **Fair and Ethical Algorithmic Systems**

More and more aspects of our everyday lives are influenced by automated decisions made by systems that statistically analyze traces of our activities. It is thus natural to question whether such systems are trustworthy, particularly given the opaqueness and complexity of their internal workings. Although there exist several visualization methods and tools for a variety of datasets and algorithms, including AI systems, they focus on either presenting simplistic explanations to end-users or on trying to visualize complex models (deep networks) for experts. Further, the dimensions of fairness, automatic bias factors discovery and user-interactive fairness corrections and explanation are absent from such systems. Our ongoing work aims to increase trust in ML systems (e.g., machine-generated recommendations) by introducing new ways for interactive visualization of explanations behind ML models. The goal is to enable different stakeholders, with potentially varying levels of background and diverse needs, to query, understand, and fix sources of distrust.



## Geospatial Data Management

*Current (ongoing) research directions in this area include:*

- **Identifying Areas of Interest**

Large amounts of user-generated content are becoming available on the Web daily, with an increasing portion of it being associated with geospatial information. Typical examples include community-based mapping projects, such as OpenStreetMap, databases of Points of Interest (POIs), e.g., from Wikipedia or Foursquare, geotagged photos, e.g., from Flickr or Instagram, etc.



This information constitutes a valuable source for discovering and exploring locations and areas of interest, with numerous applications in location-based services, geomarketing, trip planning, and other domains. Given a set of geospatial objects, our work has focused on the Best Region Search (BRS) problem, which aims at finding the optimal placement of a fixed-size rectangle so that the value of a user-defined utility function over the enclosed objects is maximized. Recently, we had introduced the top-k BRS problem and we had presented a method for efficiently and progressively computing the next best result for any number  $k$  of results requested by the user. However, our previous algorithm was designed for a centralized setting, thus not scaling to very large datasets. To overcome this limitation, we have developed a parallel and distributed approach for computing the results. We have proposed a strategy that employs multiple rounds to progressively collect partial top-k results from each node in the cluster, while a coordinator handles the aggregation of the global top-k list, dealing with overlapping results. Moreover, we have devised a single-round strategy, where the algorithm executed by each node is enhanced with additional conditions that anticipate potential overlapping solutions from neighboring nodes. Additional optimizations have been proposed to further increase performance. Our experiments on real-world datasets indicate that our new algorithms are efficient and scale to millions of points.

- **Spatio-textual Data Integration and Annotation**

Spatio-textual data integration is an open and quite challenging research problem with high value in several commercial applications, including geomarketing, navigation and social networks. It mainly comprises interlinking (deduplicating) entity descriptions from different sources/datasets. A complementary problem, annotation of spatio-textual entities, e.g. POIs, with categories is a respectively important problem for the aforementioned fields.



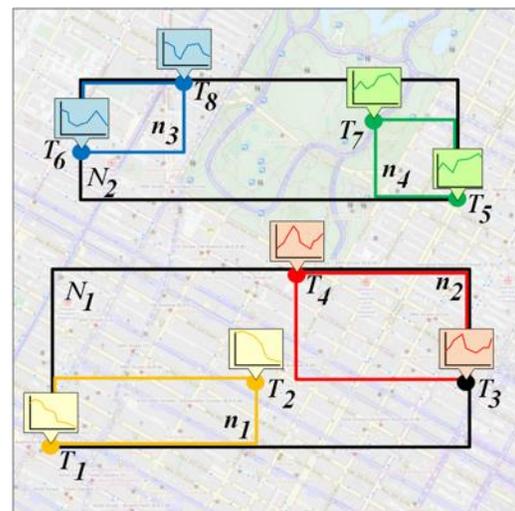
Spatio-textual data integration is significantly hindered by the large heterogeneity of entities: POIs, addresses, toponyms, land parcels are all types of spatio-textual entities, however, each type carries a different set of attributes, with varying importance in determining an entity's matching with another entity. The heterogeneity also expands to the namings of entities, resulting in quite diverse alternative descriptions regarding the same entity. Most methods in the literature do not delve in the specificities of the attributes (e.g. names, relations to other entities) of spatio-textual entities, rather than apply generic similarity measures and classification methods for interlinking and annotation. On commercial level, the shortcomings are even larger, since very simplistic similarity measures, rules and filters combined with largely manual processes are usually adopted to solve the problems.

Our work emphasizes on analyzing the spatio-textual data at hand (toponyms, addresses, etc.) and deriving novel, domain-specific similarity measures, training features and Machine Learning methods for solving several real-world spatio-textual data integration problems. Currently, our methods are materialized as open source python libraries for solving four distinct problems via Machine Learning techniques: toponym interlinking, POI classification, fusion of geocoding sources and land parcel integration (<https://github.com/LinkGeoML>). Our goal is to build a large pool of training features and Machine Learning processes/pipelines for handling an extended set of spatio-textual data integration and annotation problems, applicable in real-world processes and applications of GIS companies.

*New (emerging) research directions in this area include:*

- **Geolocated Time Series**

Time series associated with specific locations, such as visitor check-ins or sensor readings, have increased in size and popularity in various domains. Although several works have focused on efficient time series similarity search, there has been limited attention to the inherent challenge that geolocated time series introduce for hybrid queries on both spatial proximity and time series similarity. Our work addresses hybrid queries for retrieving geolocated time series based on criteria that combine both spatial distance and time series similarity. For the latter, unlike previous work, we allow filtering based on local similarity, which is computed based on subsequences rather than the entire length of each series, thus allowing the discovery of more fine-grained trends and



patterns. To efficiently support such queries, we first leverage our previously proposed BTSR-tree index, which utilizes bounds over both the locations and the shapes of time series to prune the search space. Then, we propose optimizations that check at specific timestamps to identify candidate time series that may exceed the required local similarity threshold. To further increase pruning power, we introduce the SBTSR-tree index, an extension to BTSR-tree, which additionally segments the time series temporally, allowing the construction of tighter bounds.

- Extracting value from (underutilized) Big Geospatial Data via Deep Learning**  
 Boosting the data economy and supporting the development of data value chains has been a major goal of the EC during the last years. In this spirit, our future work focuses on building Deep Learning and Transfer Learning methods for extracting and modeling knowledge from underutilized geospatial datasets and exploiting it to solve real-world problems. As an indicative example, cadastration companies have obtained, through years of manual efforts, big datasets of annotated aerial images on rural, semi-urban and urban areas. The volumes of these datasets allow the effective training of Deep Learning models for object detection, semantic segmentation and instance segmentation. Such models are directly applicable and can produce great value in several land management applications, including Cadastrals and Land Parcel Identification Systems. To this end, several research challenges need to be addressed, including the varying resolution and quality of aerial images, the low separability of certain object classes in rural areas and the appropriate/optimal application of Transfer Learning methodologies on the above processes. Another real-world application involves the existence of large open and proprietary databases of toponyms and addresses that can be utilized to build domain specific, distributed representations of spatio-textual entities, exploitable in various data integration and annotation tasks.

## Digital Curation and Research Infrastructures

*Current (ongoing) research directions in this area include:*

- Digital Curation**  
 Digital curation encompasses a set of activities aiming at the production of high quality, dependable digital assets; their organization, archiving and long-term preservation; and the generation of added value from digital assets by means of resource-based knowledge elicitation. To ensure the adequate capture of the context of digital resources and their subsequent creative and effective use, the DCU adopts a multidisciplinary approach that considers the full lifecycle of digital assets, such as records, digital surrogates and scholarly/ scientific datasets.
- Digital research infrastructures**  
 A strategic action line is the development of digital research infrastructures for the humanities at the national and european levels. At the national level, IMSI leads the



APOLLONIS Greek Infrastructure for Digital Arts, Humanities and Language Research and Innovation (P. Constantopoulos, coordinator) resulting from the unification of CLARIN-EL and DARIAH-GR. At the European level, IMSI actively participates in the European Digital Research Infrastructure for the Arts and Humanities (DARIAH) since the preparatory phase, currently with leading roles in VCC2, DARIAH's Virtual Competence Centre for Research and Education (A. Benardou is VCC2 Co-Chair, C. Dallas chairs the Working Group "Digital Methods and Practices Observatory"). IMSI, through DCU, has been heavily involved in building the ARIADNE infrastructure for archaeology. It provides advanced aggregation services for the Europeana ecosystem. It has developed the ESF NeDiMAH methods ontology (NeMO), now used to drive the automatic extraction of research processes. Collaborations with other research infrastructures, also in areas besides the humanities, are actively pursued.

*New (emerging) research directions in this area include:*

- **Digital storytelling**

The initiation of project SHARE3D marked a new line of work on issues of information delivery and publication as an integral part of the digital content lifecycle and establishing an efficient continuum from data curation to creative uses and dissemination. The specific project focuses on enabling users to explore, create and experience 3D objects as stories. This project was the first in which we employed methods of digital storytelling. SHARE3D was one of the projects which offered hands-on digital storytelling exercises to students of the MSc in Digital Methods in the Humanities, Athens University of Economics and Business.



# Projects

## EU R&D Projects



### INSPIRED - The National Research Infrastructures on Integrated Structural Biology, Drug Screening Efforts and Drug target functional characterization

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	NHRF
<b>IMIS - funding</b>	140.000,00 Euros
<b>Programme</b>	2014-2020 (EPAnEK) - Operational Programme competitiveness, Entrepreneurship and Innovation
<b>Start date</b>	03/09/2018
<b>Duration</b>	4 years
<b>Website</b>	<a href="https://www.inspired-ris.gr/">https://www.inspired-ris.gr/</a>

INSPIRED is a national distributed research infrastructure unique in the field of Structural Biology that combines studies on bioactive (macro)molecules interactions and biomarkers identification. It offers services mainly in the field of biology, diagnostics and pharmacology, addressing the needs of the health sector with significant impact on agrofood and concerns a large number of organizations (potential users and collaborators). It bridges basic research with the Industry and SMEs supporting innovative actions by either providing services or in the frame of funded programmes. It comprises two complementary infrastructures: INSTRUCT-EL | UPAT-RISF. INSPIRED addresses the needs of the field of Health Sciences and Pharmaceuticals. The partners cover all the Greek Regions. It is a distributed Research Infrastructure across Greece.





## DESIRA- Digitisation: Economic and Social Impacts in Rural Areas

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	UNIFI
<b>IMIS - funding</b>	347.208,75 €
<b>Programme</b>	H2020-RUR-2018-2020 (Rural Renaissance)
<b>Start date</b>	01/06/2019
<b>Duration</b>	4 years
<b>Website</b>	<a href="https://desira2020.eu/">https://desira2020.eu/</a>

DESIRA will develop the concept of Socio-Cyber-Physical Systems to advance understanding of the impact of digitisation in rural areas, linking analysis directly to the United Nation's Sustainable Development Goals. Operationalising the Responsible Research and Innovation approach, DESIRA will enrol agriculture, forestry and rural stakeholders in co-developing scenarios and policies in Living Labs established in 20 European Regions, and a Rural Digitization Forum gathering 250 stakeholders from all Europe. A Virtual Research Environment tailored to the purposes of the project will connect all participants and allow to increase substantially the interaction within the network. DESIRA will provide a Taxonomy and Inventory of Digital Game changers which will be implemented into an online Visualization Tool, a Set of Socio-Economic Impact Indicators aligned to the Sustainability Development Goals implemented into an online Socio-Economic Impact Tool, a Pan-European Assessment of digitization in European rural regions, a Needs, Expectations and Impact appraisal report, a Comparative Scenario Report based on scenario development activities of Living Labs and the Rural Digitization Forum, a Policy analysis and Roadmap, an Ethical Code to be adopted by researchers and innovators and recommended by policy bodies, five Use Cases that will report a further analysis – co-created by Living Labs with the support of ICT specialists - of the most promising solutions identified by Living Labs, Showcase Technologies - including a Virtual Farm Platform - that will create a selection of proof of concepts suggested by Use Cases. A detailed, multi-media dissemination, engagement and communication strategy will accompany the project from the beginning, looking at research as a multifunctional (research, engagement and communication) process and at the same time involving communication specialists in the development of adequate messages and in the choice of the most effective media.



## NI4OS-Europe 2020 - National Initiatives for Open Science in Europe

<b>Project manager</b>	Natalia Manola
<b>Coordinator</b>	GRNET
<b>IMIS - funding</b>	807.000,00 €
<b>Programme</b>	H2020-INFRAEOSC-2018-2020 (Implementing the European Open Science Cloud)
<b>Start date</b>	01/09/2019
<b>Duration</b>	3,5 years
<b>Website</b>	<a href="https://ni4os.eu/">https://ni4os.eu/</a>

National Initiatives for Open Science in Europe – NI4OS Europe, aims to be a core contributor to the European Open Science Cloud (EOSC) service portfolio, commit to EOSC governance and ensure inclusiveness on the European level for enabling global Open Science. Support the development and inclusion of the national Open Science Cloud initiatives in 15 Member States and Associated Countries in the EOSC governance. Instill within the community the EOSC philosophy and FAIR principles for data Findability, Accessibility, Interoperability and Reusability. Provide technical and policy support for on-boarding of service providers into EOSC, including generic services (compute, data storage, data management), thematic services, repositories and data sets. NI4OS-Europe exploits and engages a strong human network covering a wide range of stakeholders, as well as the infrastructure and service providers in the region, to support the overall EOSC vision, architecture and governance. NI4OS-Europe will federate the existing EOSC-relevant services in the 15 target countries and include them in the EOSC service offering. NI4OS-Europe will collaborate with the other EOSC-related initiatives to contribute to the common EOSC platform including a set of policies, rules and principles for managing services and research data across the EOSC ecosystem. NI4OS-Europe will facilitate access to infrastructures, data, resources and services for users to benefit from know-how sharing and exploitation, thus creating opportunities for increasing innovation capacity of regional Science. In parallel, NI4OS-Europe collaboration with other EOSC related projects will leverage developments in the European Open Science landscape, contributing in a decisive manner to the EOSC vision of open and inclusive science and innovation. NI4OS-Europe will trigger EOSC-relevant technical advancements in the area by supporting a full stack of open and sustainable services. It will provide equal access to infrastructures, data and services for all European researchers, ensuring inclusiveness. NI4OS-Europe will enable collaborative and innovative research of

highest excellence by connecting the regional scientific and research community between them and to the wider EOSC landscape. A wide range of ICT and science professionals in the region will be involved in the provisioning and use of EOSC-relevant services and data. NI4OS-Europe will contribute to the retention of these professionals in the area, providing them with equal opportunities.



## **EOSC Secretariat- The European Open Science Cloud for Research Secretariat Project**

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	Technopolis Consulting Group Belgium
<b>IMIS - funding</b>	402.500,00 €
<b>Programme</b>	H2020-INFRAEOSC-2018-2020 (Implementing the European Open Science Cloud)
<b>Start date</b>	01/01/2019
<b>Duration</b>	2,5 years
<b>Website</b>	<a href="https://www.eoscsecretariat.eu/">https://www.eoscsecretariat.eu/</a>

EOSCsecretariat.eu addresses the call Support to the EOSC Governance subtopic (a), Setup of an EOSC coordination structure. It will deliver an EOSC Secretariat that is a proactive, dynamic and flexible organisational structure with all the necessary competences, resources and vision to match the ambition of the call. The 30-month project will maintain a practical approach addressing all the specific needs of the coordination structure required for the EOSC. The project will adopt a Co-creation approach working with the community to deliver many of the activities and has reserved a substantial portion of the budget for organisations not in the consortium. This approach will enable a high degree of flexibility in order to address any foreseen or unforeseen challenges that may arise during the project. EOSCsecretariat.eu is characterised by being neutral towards the community it is serving and by having a pragmatic approach that is fully dedicated to realising the outcomes of the EOSC design as stated in the Implementation Roadmap Staff Working Document and adopted Council Conclusions to deliver an operational open science cloud for all European stakeholders. The outputs of EOSCsecretariat.eu include: Secretariat organisational structure, processes & procedures, rules & legal framework; business models; press & media office; pan-European awareness increase; open consultation; knowledge base; coordination services to WGs; coordination with EOSC-related projects; organisation & support to Boards & events; two Stakeholders Forums; liaison with non-EU countries;

engaged community with all stakeholder groups represented. The Consortium is competent, lean and complementary: Coordinated by an independent research and consulting organisation with 30 years' experience and track record in services to the R&I community, and supported by 10 experienced partners, from academia & industry, with strategic and practical involvement in design and delivery of the EOSC.



## RISIS 2 - European Research Infrastructure for Science, technology and Innovation policy Studies 2

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	UPEM
<b>IMIS - funding</b>	141.250,00 €
<b>Programme</b>	H2020-INFRAIA-2018-1
<b>Start date</b>	01/01/2019
<b>Duration</b>	4 years
<b>Website</b>	<a href="https://www.risis2.eu/">https://www.risis2.eu/</a>

The European Research infrastructure for science, technology and innovation policy studies (RISIS2) aims at building a data and services infrastructure supporting the development of a new generation of analyses and indicators. To develop a deeper understanding of knowledge dynamics and policy relevant evidence, the project goes beyond established quantitative indicators, developing positioning indicators, which take into account critical features of knowledge dynamics i.e. the importance of asymmetries in producers, in places and in themes. RISIS datasets are built keeping information on these three dimensions. To exploit them, new services dealing with actor identification, geographical information and thematic foci are developed, as well as semantic analytical capabilities. This project builds on RISIS1 (2014-18), which has demonstrated the relevance of such an approach and opened access to a first set of databases and services. RISIS2 gathers 19 partners aiming to transform the field of STI studies into an advanced research community. This step change is achieved by: (i) developing an e-infrastructure that supports full virtual transnational access by researchers, (ii) providing a vastly enlarged set of services tailored to field-specific needs (for problem-based integration of datasets, for exploring open data, and for supporting analytical capabilities of researchers), (iii) maintaining datasets dealing with firm innovation capacities, public research developments, R&I outputs and projects, and policy learning, (iv) developing new datasets on 4 key issues for

research and policy (social innovation, non technological innovation, the role of PhDs in society, portfolios of public funding instruments). As reflected in the strong role of OpenAire in RISIS2, the infrastructure is fully inscribed into the open science movement. It is accompanied by a strong training, dissemination and communication effort to support the important widening of the community we aim at.



### **FRESQO - Freshness REcording System for fish Quality Observation**

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	IMIS
<b>IMIS - funding</b>	122.951,09 €
<b>Programme</b>	Fisheries and Maritime 2014-2020
<b>Start date</b>	04/05/2018
<b>Duration</b>	4,5 years

#### **Website**

The main objective of the FRESQO project is the construction of an innovative product (hardware and software) which will allow the automatic recognition of the freshness of the most important commercial catches of the Greek market with possibility of expanding the detection facilities to an unlimited set of marketable and non-marketable fish. This product will essentially consist of a small (portable) spectral camera which, with a simple photograph of the fish, can provide specific indications about the freshness of a fish. The camera will communicate wired or wireless with a small-sized controller that will communicate with the specialized repository to collect and manage the relevant data as a supporting independent infrastructure (multiple capture images, chemical analysis data and organoleptic measurements as well as auxiliary data for export safer conclusions).



**DARE- Delivering Agile Research Excellence**

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	NCSR
<b>IMIS - funding</b>	224.588,21 €
<b>Programme</b>	H2020-EINFRA-2017
<b>Start date</b>	01/01/2018
<b>Duration</b>	3 years
<b>Website</b>	<a href="http://project-dare.eu/">http://project-dare.eu/</a>

DARE (Delivering Agile Research Excellence on European e-Infrastructures) aims to provide scientific communities with a unifying hyper-platform and development context to allow for user-friendly and reproducible carrying out of huge data-driven experiments, and rapid prototyping. DARE specifically addresses the requirements of innovating teams of research developers and scientists, who work on the intersection of software engineering and scientific domains, and on data, complexity and computing extremes. The size and complexity of scientific data, as well as the difficulty in formulating domain-specific solutions in reproducible and reusable ways, may often lead to throw-away, unsustainable end-user products, or long release cycles. This complexity increases exponentially with the size and diversity of input and produced data. Furthermore, widely used big-data technologies and analytics, while they are known to lead to increased productivity in commercial settings, they are often not taken advantage of in scientific. The requirement to deal with diverse exascale data resources dictates the need to ensure and increase productivity through the controlled disruption of the current modus operandi of European RIs. DARE aims to be the technological pivot for this transition, while providing transparent, traceable and developer-friendly bridges over existing infrastructures and services. Building on extensive experience in research e-infrastructures, semantification and the handling of metadata, and on big-data technologies and domain applications, DARE will equip teams of innovators with meaningful abstractions and tools allowing for rapid prototyping of reproducible and efficient research solutions. DARE will improve further and integrate tried and tested programmatic dataflow specification APIs, big-data technologies and provenance/datalineage solutions to address the requirements of European RIs, initially of EPOS, on Earth science, and IS/ENES2, on climate.



**RDA 4.0- The European plug-in to the global Research Data Alliance**

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	Trust-IT
<b>IMIS - funding</b>	100.525,00 €
<b>Programme</b>	H2020-INFRA-SUPP-2017-1 (3rd Party)
<b>Start date</b>	01/03/2018
<b>Duration</b>	2.5 years
<b>Website</b>	<a href="https://www.rd-alliance.org/">https://www.rd-alliance.org/</a>

RDA Europe 4.0 addresses the INFRA-SUPP-02-2017 call targeting the area “European support to the Research Data Alliance, RDA” designing Europe’s contribution to implementation of an effective governance model and strategy in RDA global, while ensuring that RDA delivers on locally relevant issues. RDA Europe 4.0 focuses on the need for open and interoperable sharing of research data & on the need to build social, technical and cross-disciplinary links to enable such sharing on a global scale. It strives to do this with its community-driven and bottom-up approach launched since 2012. In fact, RDA Europe 4.0 directly builds on the current RDA Europe effort, by efficiently bringing in the organisations that implemented RDA Europe since 2012. The scope of RDA Europe 4.0 is to become the centrepiece for an EU Open Science Strategy through a consolidated European network of National Nodes, bringing forward an RDA legacy in Europe, providing skilled, voluntary resources from the EU investment to address DSM issues, by means also of an open cascading grant process. The ambitious, 27-month project is implemented by 5 beneficiaries (Trust-IT Services, Gottingen State University Library, the Digital Repository of Ireland at the Royal Irish Academy, the Digital Curation Centre and the RDA Foundation), skillfully supported by 9 National Nodes which carry out specific operational activities & act as national champions for their respective region. One of the specific goals of RDA Europe 4.0 is to complete a capillary European network by on-boarding additional 13 nodes by project end. Main Outputs: Expansion from 9 to 22 national nodes in EU; 7500+ individual members & 75 organisational members; 9 RDA recommendations as ICT technical specifications; consolidated programme for experts, early careers, ambassadors & adoption projects; 5 international RDA plenary meetings; integration with EOSC, ESFRI & other pan-EU initiatives; RDA Europe self-sustained after project completion.



**MH-MD - My Health - My Data**

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	LYNKEUS
<b>IMIS - funding</b>	419.375,00 €
<b>Programme</b>	H2020-ICT-2016-1
<b>Start date</b>	01/11/2016
<b>Duration</b>	3 years
<b>Website</b>	<a href="http://www.myhealthmydata.eu/">http://www.myhealthmydata.eu/</a>

MyHealthMyData (MHMD) is a Horizon 2020 Research and Innovation Action which aims at fundamentally changing the way sensitive data are shared. MHMD is poised to be the first open biomedical information network centred on the connection between organisations and individuals, encouraging hospitals to start making anonymised data available for open research, while prompting citizens to become the ultimate owners and controllers of their health data. MHMD is intended to become a true information marketplace, based on new mechanisms of trust and direct, value-based relationships between EU citizens, hospitals, research centres and businesses.



### OpenAIRE Connect - OpenAIRE - CONNECTing scientific results in support of Open Science

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	CNR
<b>IMIS - funding</b>	480.500,00 €
<b>Programme</b>	H2020-EINFRA-2016-1
<b>Start date</b>	01/01/2017
<b>Duration</b>	2,5 years
<b>Website</b>	<a href="https://www.openaire.eu/connect">https://www.openaire.eu/connect</a>

OpenAIRE-Connect fosters transparent evaluation of results and facilitates reproducibility of science for research communities by enabling a scientific communication ecosystem supporting exchange of artefacts, packages of artefacts, and links between them across communities and across content providers. OpenAIRE-Connect extends the technological services and networking bridges (human/social/support) today offered by the OpenAIRE infrastructure in order to foster the expansion of an Open Science publishing paradigm and facilitate the

emergence of shared solutions to it. Specifically, OpenAIRE-Connect introduce three classes of new services:

- Research community services: offering support for a uniform transition of research communities towards Open Science publishing;
- Content provider services: leveraging the transition of content providers towards Open Science publishing;
- Support services: building community capacities for European and global alignment on Open Science publishing

By introducing an Open Science as a Service (OSaaS) approach, OpenAIRE-Connect will deliver on-demand Open Science publishing services to research communities and content providers, aligning practices and mechanisms that address transparent evaluation and reproducibility. By complementing the technological efforts by networking activities that will strengthen the emerging Open Science social environment, OpenAIRE-Connect will facilitate a cultural and technological shift towards common Open Science publishing practices.



### ARIADNEplus - Advanced Research Infrastructure for Archaeological Data Networking in Europe - plus

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	PIN
<b>IMIS - funding</b>	136.875,00 €
<b>Programme</b>	H2020-INFRAIA-2018-1
<b>Start date</b>	01/01/2019
<b>Duration</b>	4 years
<b>Website</b>	<a href="https://ariadne-infrastructure.eu/">https://ariadne-infrastructure.eu/</a>

The ARIADNEplus project is the extension of the previous ARIADNE Integrating Activity, which successfully integrated archaeological data infrastructures in Europe, indexing in its registry about 2.000.000 datasets. ARIADNEplus will build on the ARIADNE results, extending and supporting the research community that the previous project created and further developing the relationships with key stakeholders such as the most important European archaeological associations, researchers, heritage professionals, national heritage agencies and so on. The new enlarged partnership of ARIADNEplus covers all of Europe. It now includes leaders in different archaeological domains like palaeoanthropology, bioarchaeology and environmental archaeology as well as other sectors of archaeological sciences, including all periods of human presence from the

appearance of hominids to present times. Transnational Activities together with the planned training will further reinforce the presence of ARIADNEplus as a key actor. The technology underlying the project is state-of-art. The ARIADNEplus data infrastructure will be embedded in a cloud that will offer the availability of Virtual Research Environments where data-based archaeological research may be carried out. The project will furthermore develop a Linked Data approach to data discovery. Innovative services will be made available to users, such as visualization, annotation, text mining and geo-temporal data management. Innovative pilots will be developed to test and demonstrate the innovation potential of the ARIADNEplus approach. Fostering innovation will be a key aspect of the project, with dedicated activities led by the project Innovation Manager.



### SmartDataLake - Sustainable Data Lakes for Extreme-Scale Analytics

<b>Project manager</b>	Dimitris Skoutas
<b>Coordinator</b>	IMIS
<b>IMIS - funding</b>	€ 853,125.00
<b>Programme</b>	RIA, H2020-ICT-2018-2
<b>Start date</b>	1/1/2019
<b>Duration</b>	3 years
<b>Website</b>	<a href="https://smartdatalake.eu/">https://smartdatalake.eu/</a>

Data lakes are raw data ecosystems, where large amounts of diverse data are retained and coexist. They facilitate self-service analytics for flexible, fast, ad hoc decision making. SmartDataLake enables extreme-scale analytics over sustainable big data lakes. It provides an adaptive, scalable and elastic data lake management system that offers: (a) data virtualization for abstracting and optimizing access and queries over heterogeneous data, (b) data synopses for approximate query answering and analytics to enable interactive response times, and (c) automated placement of data in different storage tiers based on data characteristics and access patterns to reduce costs. The data lake's contents are modelled and organised as a heterogeneous information network, containing multiple types of entities and relations. Efficient and scalable algorithms are provided for: (a) similarity search and exploration for discovering relevant information, (b) entity resolution and ranking for identifying and selecting important and representative entities across sources, (c) link prediction and clustering for unveiling hidden associations and patterns among entities, and (d) change detection and incremental update of

analysis results to enable faster analysis of new data. Finally, interactive and scalable visual analytics are provided to include and empower the data scientist in the knowledge extraction loop. This includes functionalities for: (a) visually exploring and tuning the space of features, models and parameters, and (b) enabling large-scale visualizations of spatial, temporal and network data. The results of the project are evaluated in real-world use cases from the business intelligence domain, including scenarios for portfolio recommendation, production planning and pricing, and investment decision making. SmartDataLake will foster innovation and enable European SMEs to capitalize on the value of their own data lakes.



## INODE - Intelligent Open Data Exploration

<b>Project manager</b>	Georgia Koutrika
<b>Coordinator</b>	ZHAW
<b>IMIS - funding</b>	798.000 Euros
<b>Programme</b>	EU H2020 - H2020-EU.1.4.1.3. - Development, deployment and operation of ICT-based e-infrastructures, inode-project.eu
<b>Start date</b>	1/11/2019
<b>Duration</b>	3.5 years
<b>Website</b>	<a href="https://www.inode-project.eu/">https://www.inode-project.eu/</a>

The core principle of INODE is that users should interact with data in a more dialectic and intuitive way similar to a dialog with a human. To achieve this principle, INODE builds innovative services for exploration of open data sets that help users (a) link and leverage multiple datasets, (b) access and search data using natural language, using examples and using analytics (c) get guidance from the system in understanding the data and formulating the right queries, and (d) explore data and discover new insights through visualizations.





## INFORE - Interactive Extreme-Scale Analytics and Forecasting

<b>Project manager</b>	Antonios Deligiannakis
<b>Coordinator</b>	IMIS
<b>IMIS - funding</b>	834 687,50 Euros
<b>Programme</b>	H2020, ICT-12-2018-2020 - Big Data technologies and extreme-scale analytics
<b>Start date</b>	1/1/2019
<b>Duration</b>	3.25 years
<b>Website</b>	<a href="https://www.infore-project.eu/">https://www.infore-project.eu/</a>

At an increasing rate, industrial and scientific institutions need to deal with massive data flows streaming in from a multitude of sources. For instance, maritime surveillance applications combine high-velocity data streams, including vessel position signals emitted from hundreds of thousands of vessels across the world and acoustic signals of autonomous, unmanned vessels; in the financial domain, stock price forecasting and portfolio management rely on stock tick data combined with real-time information sources on various pricing indicators; at the fight against cancer, complex simulations of multi-cellular systems are used, producing extreme-scale data streams in an effort to predict the effects of drug synergies on cancer cells. In these applications, the data volumes are expected to dramatically grow in the future. Processing this data often requires not only using an HPC infrastructure, but also having data scientists, who are typically not expert programmers, program complex workflows, with a vast number of parameters to tune through time-consuming repeated programming and testing. INFORE will address these challenges and pave the way for real-time, interactive extreme-scale analytics and forecasting. The ability to forecast, as early as possible, a good approximation to the outcome of a time-consuming and resource-demanding computational task allows to quickly identify undesired outcomes and save valuable amount of time, effort and computational resources, which would otherwise be spent in vain. Consider, for example, the ability to forecast the outcome of a complex multi-cellular system simulation for tumor evolution, without the need to wait for the simulation to be completed. INFORE will also design and develop a flexible, pluggable, distributed software architecture that is programmable and set up by graphical data processing workflows. The INFORE prototype will be tested on massive real-world data from the life sciences, financial and maritime domains.





## Lawful evidence collecting and continuity platform development (LOCARD)

<b>Project manager</b>	Constantinos Patsakis
<b>Coordinator</b>	Athena Research Center
<b>IMIS - funding</b>	700.000
<b>Programme</b>	SU-FCT02
<b>Start date</b>	1/5/2019
<b>Duration</b>	36 months
<b>Website</b>	<a href="https://locard.eu/">https://locard.eu/</a>

Digital evidence is currently an integral part of criminal investigations, and not confined to pure cybercrime cases. Criminal behaviours like financial frauds, intellectual property theft, industrial espionage, and terrorist networks leverage the Internet and cyberspace. The very ubiquity of digital devices, e.g. smartphones, in modern society makes digital evidence extremely relevant for investigations about all kinds of criminal behaviour like murder, contraband activities, and people smuggling, to name a few. Due to its nature, the use of digital evidence in a court of law has always been challenging. It is critical that it should be accompanied by a proper chain of custody, guaranteeing its source and integrity. LOCARD aims to provide a holistic platform for chain of custody assurance along the forensic workflow, a trusted distributed platform allowing the storage of digital evidence metadata in a blockchain. Each node of LOCARD will be able to independently set its own permission policies and to selectively share access to digital evidence with other nodes when deemed necessary and upon proper authorization through fine-grained policies. LOCARD's modularity will also allow diverse actors to tailor the platform to their specific needs and role in the digital forensic workflow, from preparation and readiness, to collection, to analysis and reporting. LOCARD will have a crowdsourcing module to collect citizen reports of selected violations, a crawler to detect and correlate online deviant behaviour, and a toolkit for investigators that will assist them in collecting online and offline evidence. This will be powered by an immutable storage and an identity management system that will protect privacy and handle access to evidence data using a Trusted Execution Environment. Blockchain technology will not only guarantee that information about the evidence cannot be

tampered with, but allow interoperability without the need for a trusted third party.



**GRACIOUS** - Grouping, Read-Across, Characterisation and classification framework for regulatory risk assessment of manufactured nanomaterials and Safer design of nano-enabled products

<b>Project manager</b>	Georgia Tsiliki
<b>Coordinator</b>	Heriot-Watt University
<b>IMIS - funding</b>	124 037.5 Euros
<b>Programme</b>	H2020-NMBP-2017-two-stage
<b>Start date</b>	1/1/2018
<b>Duration</b>	3.5 years
<b>Website</b>	<a href="https://www.h2020gracious.eu">https://www.h2020gracious.eu</a>

The GRACIOUS project will develop a highly innovative science-based framework that supports the assessment of risk posed by the ever increasing array of nanomaterials on the market and under development. The framework will streamline the process for assessing their risk by logically grouping nanomaterials thereby allowing extrapolation between (read-across) nanomaterials and reducing the need to assess exposure to and toxicity on a case by case basis. The project will work continuously with stakeholders in an iterative cycle of design, testing and refinement to ensure that the Framework effectively meets the needs of both regulators and industry. Application of the Framework will allow movement away from the case-by-case risk assessment paradigm, thereby improving the efficiency of risk analysis and decision making for safer design of quality nano-enabled products.



**NEP4DISSENT – New Exploratory Phase in Research on East European Cultures of Dissent. European Cooperation in Science and Technology**

<b>Project coordinator</b>	Costis Dallas
<b>Coordinator</b>	Polish Academy of Sciences (PL)
<b>IMSI - funding</b>	About 50,000 euros (The funding of COST actions is adjusted yearly)
<b>Programme</b>	Open Call Collection OC-2016-2
<b>Start date</b>	16/10/2017
<b>Duration</b>	4 years
<b>Website</b>	<a href="http://www.cost.eu/COST_Actions/ca/CA16213">http://www.cost.eu/COST_Actions/ca/CA16213</a>

Resistance and dissent in former socialist Europe 1945-1989 constitutes a remarkable chapter of Europe's recent past, which not only informs in a decisive way the identities of post-socialist societies, but has also reshaped the continent as a whole and still provides an important reference for contemporary social movements worldwide. The proposers of this Action believe that, after a period of growth and consolidation, this field of study and the respective domain of cultural heritage have stalled and fell short of its true significance. This state of affairs results from (1) the inheritance of Cold War-era conceptual distinctions, (2) confinement of research within national silos and (3) neglecting the problem of access to original archival sources for digitally enabled research due to both their heterogeneity and uneven investment in research infrastructures.



### Novel EOSC Services for Emerging Atmosphere, Underwater & Space Challenges (NEANIAS)

<b>Project manager</b>	George Papastefanatos
<b>Coordinator</b>	National Kapodistrian University of Athens
<b>IMIS - funding</b>	470.625 Euros
<b>Programme</b>	H2020-EU.1.4.1.3
<b>Start date</b>	01/11/2019
<b>Duration</b>	3 years
<b>Website</b>	<a href="https://www.neanias.eu/">https://www.neanias.eu/</a>

NEANIAS is an ambitious project that comprehensively addresses the 'Prototyping New Innovative Services' challenge set out in the recent 'Roadmap for EOSC' foreseen actions. NEANIAS will drive the co-design, delivery, and

integration into EOSC of innovative thematic services, derived from state-of-the-art research assets and practices in three major sectors: underwater research, atmospheric research and space research. Each of these sectors engages a diverse set of research and business groups, practices, and technologies. Each thematic service will not only address its community-specific needs but will also enable the transition of the respective community to the EOSC concept and Open Science principles. In doing so, NEANIAS provides its communities with plentiful resource access, collaboration instruments, and interdisciplinary research mechanisms, which will amplify and broaden each community's research and knowledge generation activities. From a technological perspective, NEANIAS will deliver a rich set of services that are designed to be flexible and extensible; they will be able to accommodate the needs of communities beyond their original definition and to adapt to neighboring cases, fostering reproducibility and re-usability. From a sustainability perspective, NEANIAS identifies promising, cutting-edge business cases across several user communities and lays out several concrete exploitation opportunities.



### Wider Impacts and Scenario Evaluation of Autonomous and Connected Transport

<b>Project coordinator</b>	George Papastefanatos
<b>Coordinator</b>	University of Greenwich
<b>IMSI - funding</b>	-
<b>Programme</b>	COST ACTION
<b>Start date</b>	23/7/2017
<b>Duration</b>	4 years
<b>Website</b>	<a href="https://wise-act.eu/">https://wise-act.eu/</a>

Autonomous vehicle (AV) trials are currently taking place worldwide and Europe has a key role in the development of relevant technology. Yet, very limited research exists regarding the wider implications of the deployment of such vehicles on existing road infrastructure, since it is unclear if and when the transition period will start and conclude. It is anticipated that improved accessibility and road safety will constitute the primary benefits of the widespread use of AVs, whilst co-benefits may also include reduced energy consumption, improved air quality or better use of urban space. Therefore, the focus of this COST Action is on observed and anticipated future mobility trends and implications on travel behaviour, namely car sharing, travel time use or residential

location choice to name a few. Other important issues to be explored under different deployment scenarios are social, ethical, institutional and business impacts. To achieve this, it is essential to culminate co-operation between a wide range of stakeholders at a local, national and international level, including academics and practitioners. Consequently, this COST Action will facilitate collaboration within Europe and beyond about this emerging topic of global interest.



### **GLOBAQUA - Managing the effects of multiple stressors on aquatic ecosystems under water scarcity**

<b>Project coordinator</b>	Prof. Phoebe Koundouri
<b>Coordinator</b>	AGENCIA ESTATAL CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS Spain
<b>IMSI - funding</b>	368,360 euros
<b>Programme</b>	EU's 7 <sup>th</sup> Programme for research, technological development and demonstration
<b>Start date</b>	1 February 2014 (End date 31 January 2019)
<b>Duration</b>	4 years
<b>Website</b>	<a href="http://www.globaqua-project.eu">www.globaqua-project.eu</a>

The main aim of GLOBAQUA is to achieve a better understanding of how current water management practices and policies could be improved by identifying their main drawbacks and alternatives. Freshwater systems are under threat by a variety of stressors (organic and inorganic pollution, cover change, water abstraction, land use, etc...). The joint occurrence of many stressors (chemical, geomorphologic, biological) under water scarcity may produce novel and unfamiliar synergies and effects of unknown consequences. Therefore, it is crucial to understand deeply how water scarcity interacts with other stressors in freshwaters and to convey this information to managers, stakeholders and policymakers in order to minimize impacts, to adapt to oncoming changes, and to improve our management and policies. GLOBAQUA (Managing the effects of multiple stressors on aquatic ecosystems under water scarcity) has assembled a multidisciplinary consortium in order to study the interaction of multiple stressors within the frame of strong pressure on water resources. GLOBAQUA assesses the effects of water scarcity on aquatic ecosystems by focusing on six river basins

(Adige, Anglian, Ebro, Evrotas, Sava and Souss Massa). These basins encompass a rich set of socio-ecological conditions and a wide geographic coverage, and focus on a specific set of stressors to illustrate different management scenarios.



## SLIPO - Scalable Linking and Integration of Big POI Data

<b>Project Coordinator</b>	Spiros Athanasiou
<b>Coordinator</b>	IMSI
<b>IMSI - funding</b>	748 KEuros
<b>Programme</b>	IA, H2020-ICT-14-2016
<b>Start date</b>	1/1/2017
<b>Duration</b>	3 years
<b>Website</b>	<a href="http://www.slipo.eu">http://www.slipo.eu</a>

SLIPO ([www.slipo.eu](http://www.slipo.eu)) is an Horizon 2020 Innovation Action (IA) developing technologies for the scalable and quality assured integration of Points of Interest (POI) Big Data assets. POI data are the cornerstone of any application, service, and product even remotely related to our physical surroundings. From navigation applications, to social networks, to tourism, and logistics, we use POIs to search, communicate, decide, and plan our actions. The creation, update, and provision of POIs consists a multi-billion cross-domain and cross-border industry, with a value chain natively incorporating most domains of our economy. The evolved POI value chain introduces opportunities for growth, but also complexity, intensifying the challenges relating to their quality-assured integration, enrichment, and data sharing. However, the integration of POIs using current approaches remains labor-intensive and scalable only for domain-specific or small-scale efforts. In SLIPO, we argue that linked data technologies can address the limitations, gaps and challenges of the current landscape in integrating, enriching, and sharing POI data. Our goal is to transfer the research output generated by our work in project GeoKnow, to the specific challenge of POI data, introducing validated and cost-effective innovations across their value chain. Specifically, SLIPO develops effective and scalable software and processes for: transforming conventional POI formats and schemas into RDF data; interlinking POI entities from different datasets; enriching POI entities with additional metadata, including temporal, thematic and semantic properties; fusing Linked POI data in order to produce more complete and accurate POI profiles; assessing the quality of the integrated

POI data; offering value added services based on spatial aggregation, association extraction and spatiotemporal prediction.



### ELIXIR-EXCELERATE - Fast-track ELIXIR implementation and drive early user exploitation across the life-sciences

<b>Project coordinator</b>	Theodore Dalamagas (Deputy: Stelios Sartzetakis)
<b>Coordinator</b>	European Molecular Biology Laboratory - EMBL
<b>IMSI - funding</b>	40,000 euros
<b>Programme</b>	EU, INFRADEV-3-2015
<b>Start date</b>	1/9/2015
<b>Duration</b>	4 years
<b>Website</b>	<a href="https://www.elixir-europe.org/excelerate">https://www.elixir-europe.org/excelerate</a>

The precipitous drop in costs for high-throughput biology has enabled European research laboratories to produce a huge amount of complex and heterogeneous data. However, data will only generate long-term value if it is Findable, Accessible, Interoperable and Re-usable (FAIR). This requires a scalable infrastructure that connects local, national and European efforts and provides standards, tools and training for data stewardship. Motivated by this, ELIXIR, a distributed organisation comprising national bioinformatics research infrastructures and the European Bioinformatics Institute, has developed ELIXIR-EXCELERATE. The project will fast-track ELIXIR's early implementation phase by i) coordinate and enhance existing resources into a world-leading data service for academia and industry, ii) grow bioinformatics capacity and competence across Europe, and iii) complete the management processes needed for a large distributed infrastructure. ELIXIR-EXCELERATE will enable cost-effective and sustainable management and re-use of data for millions of users across the globe and improve the competitiveness of European life science industries through accessible data and robust standards and tools.



## ELIXIR CONTAINERS - Deploying Reproducible Containers and Workflows Across Cloud Environments

<b>Project manager</b>	Thanasis Vergoulis
<b>Coordinator</b>	EMBL
<b>IMIS - funding</b>	€ 3,231.25
<b>Programme</b>	Strategic Implementation Study (ELIXIR Commissioned Services)
<b>Start date</b>	24/10/2019
<b>Duration</b>	17 months
<b>Website</b>	<a href="https://www.imsi.athenarc.gr/en/projects/project/66">https://www.imsi.athenarc.gr/en/projects/project/66</a>

The project will realize a study to convene and establish a consensus on high-level community-driven standards for deploying reproducible containers and workflows across cloud environments.



## ARCADES - Algebraic Representations in Computer-Aided Design for complex Shapes

<b>Project coordinator</b>	Ioannis Z. Emiris
<b>Coordinator</b>	IMSI
<b>IMSI - funding</b>	485,000 euros
<b>Programme</b>	Marie Skłodowska-Curie Innovative Training Networks
<b>Start date</b>	1/1/2016
<b>Duration</b>	4 years
<b>Website</b>	<a href="http://arcades-network.eu">http://arcades-network.eu</a>

ARCADES aims at disrupting the traditional paradigm in Computer-Aided Design (CAD) by exploiting cutting-edge research in mathematics and algorithm design. Today, several approaches of the CAD industry are outdated, and 3D geometry processing is becoming increasingly the weak link. This is alarming in sectors where CAD faces new challenges arising from fast point acquisition, big data, and

mobile computing, 3D printing, but also in robotics, simulation, animation, fabrication and manufacturing where CAD strives to address crucial societal and market needs. The challenge taken up by ARCADES is to build the next generation of CAD software based on strong foundations from algebraic geometry, differential geometry, scientific computing, and algorithm design. Our game-changing methods lead to real-time modelers for architectural geometry and visualisation, to isogeometric and design-through-analysis software for shape optimisation, and marine design & hydrodynamics, and to tools for motion design, robot kinematics, path planning, and control of machining tools.



### GRAPES - Learning, processing and optimising shapes

<b>Project manager</b>	Ioannis Z. Emiris
<b>Coordinator</b>	IMIS
<b>IMIS - funding</b>	486,035 euros
<b>Programme</b>	Marie Skłodowska-Curie Innovative Training Networks
<b>Start date</b>	1/11/2019
<b>Duration</b>	4,5 years
<b>Website</b>	<a href="http://grapes-network.eu/">http://grapes-network.eu/</a>

GRAPES aims at considerably advancing the state of the art in a variety of fields ranging from Computational and Numerical Mathematics, to Geometric Modelling and CAD, up to Data Science and Machine Learning, in order to promote game changing approaches for generating, optimising, and learning 3D shapes. Research is articulated around 3 scientific work packages (WPs):

1. High-order methods and representations,
2. Algebraic & numeric tools in shape optimisation and analysis, and
3. Machine Learning for shapes.

Concrete applications include simulation and fabrication, design and visualisation, retrieval and mining, reconstruction, and urban planning. Our 15 PhD candidates shall benefit from both top-notch research as well as a strong innovation component through a nexus of intersectoral secondments and Network-wide workshops. Innovation and technology transfer rely on the active participation of SMEs, either as beneficiary, or as partner organisations hosting secondments.



## SHARE-3D

<b>Project coordinator</b>	Dimitris Gavrilis
<b>Coordinator</b>	Athena R.C.
<b>IMSI - funding</b>	154,615 euros
<b>Programme</b>	CEF , Action No: 2017-EU-IA-0139
<b>Start date</b>	1/9/2018
<b>Duration</b>	18 months
<b>Website</b>	<a href="https://share3d.eu/">https://share3d.eu/</a> <a href="http://www.dcu.gr/en/share-3d-sharing-new-perspectives/">http://www.dcu.gr/en/share-3d-sharing-new-perspectives/</a>

The Share-3D project is developing two applications, a Dashboard and a Storytelling tool, which will integrate APIs available from Europeana and Sketchfab to provide services for users to publish and re-use 3D and other digital materials. The aim is to use interest in 3D to encourage end-users to engage with content and make it accessible through Europeana in a creative and highly interactive way. In particular, the Action will create user friendly tools to upload 3D media and storytelling services that offer end-users interesting, informative and memorable experiences about European Culture. End-users will be able to browse cultural content, make links between related items, add descriptions and knowledge, and share the results with friends, colleagues, classmates, students, visitors and others. Lastly, the Action will showcase the potential of allowing users to both produce and consume cultural heritage content, and of giving users the opportunity to share their stories and share their discoveries.



## ARKWORK - Archaeological practices and knowledge work in the digital environment

<b>Project coordinator</b>	Costis Dallas
<b>Coordinator</b>	Uppsala University (SE)

<b>IMSI - funding</b>	About 50,000 euros (The funding of COST actions is adjusted yearly)
<b>Programme</b>	COST European Cooperation in Science & Technology Programme, Action CA15201
<b>Start date</b>	6/10/2016
<b>Duration</b>	3 years
<b>Website</b>	<a href="https://www.arkwork.eu/">https://www.arkwork.eu/</a>

The COST action ARKWORK is aimed at forming an interdisciplinary network that will collect the knowledge and experiences of researchers and projects, national as well as European, in producing and exploiting archaeological knowledge. The network thus intends to contribute to the coordination of current efforts to study archaeological practices, and the potential for social and economic exploitation of archaeological knowledge in the digital environment. The network seeks to establish the state of the art in promoting archaeological knowledge and its application for the benefit of society, and to provide expert guidance in the production, validation, reproduction, management and use of archaeological knowledge. Apart from networking and information activities, the project collects and charts basic information and approaches concerning archaeological practices in field research, collections, and the production and validation of knowledge, through a combination of approaches from archaeological theory and methodology, information science, computer science, and social studies of science and technology.



**OpenAire Advance- OpenAIRE Advancing Open Scholarship**

<b>Project manager</b>	prof Yannis Ioannidis
<b>Coordinator</b>	NKUA
<b>IMIS - funding</b>	1.413.261,75 €
<b>Programme</b>	H2020-EINFRA-2017
<b>Start date</b>	01/01/2018
<b>Duration</b>	3 years
<b>Website</b>	<a href="https://www.openaire.eu/advance/">https://www.openaire.eu/advance/</a>

OpenAIRE-Advance continues the mission of OpenAIRE to support the Open Access/Open Data mandates in Europe. By sustaining the current successful

infrastructure, comprised of a human network and robust technical services, it consolidates its achievements while working to shift the momentum among its communities to Open Science, aiming to be a trusted e-Infrastructure within the realms of the European Open Science Cloud. In this next phase, OpenAIRE Advance strives to empower its National Open Access Desks (NOADs) so they become a pivotal part within their own national data infrastructures, positioning OA and open science onto national agendas. The capacity building activities bring together experts on topical task groups in thematic areas (open policies, RDM, legal issues, TDM), promoting a train the trainer approach, strengthening and expanding the pan-European Helpdesk with support and training toolkits, training resources and workshops. It examines key elements of scholarly communication, i.e., co-operative OA publishing and next generation repositories, to develop essential building blocks of the scholarly commons. On the technical level OpenAIRE-Advance focuses on the operation and maintenance of the OpenAIRE technical TRL8/9 services, and radically improve the OpenAIRE services on offer by: a) optimizing their performance and scalability, b) refining their functionality based on end-user feedback, c) repackaging them into products, taking a professional marketing approach with well-defined KPIs, d) consolidating the range of services/products into a common e-Infra catalogue to enable a wider uptake. OpenAIRE-Advance steps up its outreach activities with concrete pilots with three major RIs, citizen science initiatives, and innovators via a rigorous Open Innovation programme. Finally, via its partnership with COAR, OpenAIRE-Advance consolidates OpenAIRE's global role extending its collaborations with Latin America, US, Japan, Canada, and Africa.



## EOSC pilot

<b>Project manager</b>	Yannis Ioannidis
<b>Coordinator</b>	STFC-UKRI
<b>IMIS - funding</b>	500.650,00
<b>Programme</b>	H2020-INFRADEV-2016-2 (Grant Agreement no 739563)
<b>Start date</b>	01/2017
<b>Duration</b>	24 months (was given a 3 month extension)
<b>Website</b>	<a href="https://eoscpilot.eu/">https://eoscpilot.eu/</a>

The European Open Science Cloud will offer 1.7 million European researchers and 70 million professionals in science and technology a virtual environment with open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines by federating existing

scientific data infrastructures, today scattered across disciplines and Member States.

The **EOSCpilot project** has been funded to support the first phase in the development of the European Open Science Cloud (EOSC). It will:

- Propose and trial governance frameworks for the EOSC and contribute to the development of European open science policy and best practice;
- Develop a number of demonstrators functioning as high-profile pilots that integrate services and infrastructures to show interoperability and its benefits in a number of scientific domains; and
- Engage with a broad range of stakeholders, crossing borders and communities, to build the trust and skills required for the adoption of an open approach to scientific research.

## National R&D Projects



### VR-Park: Augmented reality platform for urban parks

<b>Project manager</b>	Yannis Stavrakas
<b>Coordinator</b>	H.A.O. "Demeter"
<b>IMSI - funding</b>	204,000 euros
<b>Programme</b>	RESEARCH - CREATE - INNOVATE, Operational Programme Competitiveness, Entrepreneurship and Innovation 2014-2020
<b>Start date</b>	9/2018
<b>Duration</b>	30 months
<b>Website</b>	<a href="https://vr-park.org/">https://vr-park.org/</a>

Urban parks and open green areas are important attractions of environmental interest to city residents and visitors. Careful and well-targeted promotion of these areas, not only enhances the importance of their existence in the urban space, but at the same time can assist in the development of alternative forms of "green tourism", and towards the direction of environmental awareness among citizens, which is particularly important nowadays and crucial for the future of the planet. New technologies are a key tool in enhancing the experience of touring urban parks, as they can make the tour much more attractive, highlighting interesting information about the flora and fauna of the park, as well as various other points of interest. At the same time, they can help guide the visitor inside relatively large

parks, and easily identify his/her paths, thus highlighting areas of the park that would otherwise be neglected. They can also assist park managers in organizing events, thus solving one of the key operating problems mainly of large-scale urban parks, which is the failure to exploit their entire site due to reduced or problematic accessibility. This project will build an integrated system that comprises an augmented reality mobile application for visitors of urban parks, and a corresponding park management web application for the managers of such park. Through the mobile app, an attractive, interactive touring environment will be created which will highlight the environmental and historical interest of those sites. At the same time, the web application will receive multimedia data from the users and will automatically collect anonymous data that may be useful to park managers to improve the visitors' touring experience and to better highlight the advantages of visiting such parks.



### CLIMPACT – Flagship Initiative for Climate Change

<b>Project coordinator</b>	Alkis Simitsis
<b>Coordinator</b>	National Observatory of Athens
<b>IMSI - funding</b>	-
<b>Programme</b>	GSRT project
<b>Start date</b>	2019
<b>Duration</b>	3 years
<b>Website</b>	-

The initiative brings together a National Network of 11 scientific organizations to study and analyze at scale scientific indicators and measures related to Climate Change, aiming at generating innovative scientific tools, methods, and technology to mitigate the results of the Climate Change. The Athena RC leads the design and development of the first National Repository for the Climate Change.



## Design and Development of the YouWeP platform

<b>Project manager</b>	George Papastefanatos
<b>Coordinator</b>	IMSI - Athena RC
<b>IMIS - funding</b>	24,800 euros
<b>Programme</b>	Contract
<b>Start date</b>	29/06/2018
<b>Duration</b>	2 years
<b>Website</b>	<a href="http://www.socioscope.gr/">http://www.socioscope.gr/</a> , <a href="http://www.youwho.gr">http://www.youwho.gr</a>

Following a long-standing collaboration between IMSI-ATHENA R.C and the National Centre for Social Research (EKKE) the project aims at extending and providing new features to the Socioscope platform, a data platform developed by IMSI and used for hosting and visualizing data collected from social surveys by the National Centre for Social Research (EKKE).

Moreover, the project will implement the YouWeP platform, which will be a web platform for social and political scientists to conduct online surveys and visualize the results in interactive ways. The platform will offer new novel ways of conducting surveys based on chat-based questionnaires for interacting with users and getting their feedback. The platform will be evaluated during the implementation of a large-scale online survey for the identification of trends and behaviors in young people aged 18-29 years old.



## Visual Facts - Democratizing Visual Analytics, A Self-Service Platform for Big Data Exploration

<b>Project manager</b>	George Papastefanatos
<b>Coordinator</b>	IMSI - Athena RC
<b>IMIS - funding</b>	172,000 euros
<b>Programme</b>	1st Call for Postdoc Research Projects - Hellenic Foundation for Research and Innovation (ELIDEK)
<b>Start date</b>	16/10/2018

**Duration** 3 years

**Website** <https://visualfacts.imsi.athenarc.gr/>

Self-service visual analytics is a new paradigm, widely promoted in modern corporate environments, in which business users are enabled and encouraged to directly manipulate (explore, blend, analyze) underlying data in rich visual ways, in order to derive insights from business information as quickly and efficiently as possible. Allowing less tech-savvy end users to make decisions based on their own queries and analyses, frees up the organization's business intelligence and information technology (IT) teams from the tedious work of data preparation.

The aim of VisualFacts is to develop a scalable platform for providing self-service visual analytic capabilities to a wide range of corporate and non-corporate users to access, explore, analyze open and privately-held data and collaborate on the analytic results of their work by sharing, annotating and reusing them in the form of visual facts.



### **LinkGeoML: Automatic and accurate interlinking of geospatial data using machine learning**

**Project manager** Giorgos Giannopoulos

**Coordinator** IMSI

**IMIS - funding** 199.300,00 Euros

**Programme** Competitiveness, Entrepreneurship, Innovation - ESPA 2014-2020 – Research, Create, Innovate – II. Partnerships between Enterprises and Research Organizations

**Start date** 9/7/2018

**Duration** 2 years

**Website** <https://linkgeoml.gr/>

The automatic and accurate interlinking of geospatial data poses an important scientific challenge, with direct application in several business fields. The major requirement is achieving high accuracy in identifying similar entities within datasets. For example, in a cadastral database, it is crucial that the land parcels, that were gathered from several different databases, are uniquely and clearly identified. In another example, for a geo-marketing company, it is of high importance to be able to accurately cross-reference the location/addresses of customers and companies, so that they are properly targeted. LinkGeoML aims at researching, developing and extending machine learning methods, utilizing the vast amount of available, open geospatial data, in order to implement automated

and highly accurate algorithms for interlinking geospatial entities. The proposed methods will implement novel training features, based on domain knowledge and on the analysis of open and proprietary geospatial datasets. Further, they will extend and specialize machine learning models on classification and similarity learning. The implemented technologies will be published as open source software and, also, will be integrated into existing, commercial applications for cadastration, geocoding and geomarketing, aiming at improving their functionality and increasing their commercial value and application domains. LinkGeoML comprizes a partnership between enterprises and research organizations, aiming to perform high quality, industrial research with a twofold purpose: Provide SMEs with useful geospatial data integration tools to solve real-world problems, and advance the state of the art on machine learning methods for geospatial data integration. To achieve these purposes, LinkGeoML identifies use cases based on real-world integration problems, elicited by its two industrial partners, and researches how machine learning-based, interlinking methods can be applied to these use cases and facilitate their handling, in real-world data. Check out our initially prescribed use cases, as well as our first results.



### Big Data in Monitoring and Analyzing Sea Area Traffic: innovative ICT and analysis models

<b>Project manager</b>	Theodore Dalamagas
<b>Coordinator</b>	IMSI
<b>IMIS - funding</b>	150,000 euros
<b>Programme</b>	RESEARCH - CREATE - INNOVATE, Operational Programme Competitiveness, Entrepreneurship and Innovation 2014-2020
<b>Start date</b>	7/2018
<b>Duration</b>	30 months
<b>Website</b>	<a href="https://www.imsi.athenarc.gr/el/projects/project/59">https://www.imsi.athenarc.gr/el/projects/project/59</a>

The project will design and develop an innovative ICT platform for the collection and analysis of big traffic data, spatial data, environmental data and meteorological data, to support sea area monitoring and observation. The vision is to provide effective and efficient data integration, processing and analysis technologies with the aim to deliver (a) a "Combined Real-time Operational Snapshot", and (b) a "Combined Historical Snapshot" of sea areas. The project will exploit state-of-the-art IT, and design and develop innovative IT based on in-memory database algorithms, models and methods for parallel computation, and

methods for big data analytics.



### ELIXIR-GR - Information Infrastructure for the Life Sciences

<b>Project manager</b>	Theodore Dalamagas (Deputy: Stelios Sartzetakis)
<b>Coordinator</b>	Biomedical Science Research Center Alexander Fleming
<b>IMSI - funding</b>	92,600 euros
<b>Programme</b>	National Roadmap for Research Infrastructures
<b>Start date</b>	9/2017
<b>Duration</b>	3 years
<b>Website</b>	<a href="http://www.elixir-greece.org">http://www.elixir-greece.org</a>

ELIXIR (<http://www.elixir-europe.org/>) is one of the ESFRI's pan-European research infrastructures (RIs) with key importance for the research and innovation in the life sciences. The aim of ELIXIR is to provide the facilities necessary for Europe's life scientists to manage and safeguard the massive amounts of data being generated every day by publicly funded research. Greece has recently joined ELIXIR, which has already entered its construction phase. ELIXIR operates a distributed research infrastructure based on ELIXIR nodes (at national level) which are connected to the ELIXIR Hub in order to provide infrastructure for data, tools, standards and training, as well as support for other ESFRI biological and medical science infrastructures. BSRC A. Fleming is joining forces with Athena RC and GRNET to lead the construction of the Greek component, ELIXIR-EL. Setting up ELIXIR- EL will be a dynamic process, taking into account user needs, available tools and datasets and technological advancements. The RI will be designed and setup using GRNET's Okeanos cloud IaaS service (<http://okeanos.grnet.gr/>) and ViMa service (<http://vima.grnet.gr>). The proposed RI aims to develop ELIXIR-EL RIs at the following levels: computing resources, data-intensive programming components, data resources, and tools.



### MEDA 2: Moving from Big Data Management to Data Science

<b>Project manager</b>	Theodore Dalamagas
<b>Coordinator</b>	IMSI
<b>IMSI - funding</b>	800,000 euros
<b>Programme</b>	National Roadmap for Research Infrastructures
<b>Start date</b>	9/2017
<b>Duration</b>	3 years
<b>Website</b>	<a href="http://web.imsi.athenarc.gr/projects/meda.2">http://web.imsi.athenarc.gr/projects/meda.2</a>

The Information Management Systems Institute (IMSI) has already established a strong technical and scientific background on big data technologies. IMSI has recently completed MEDA, a national funded R&D project focusing on the efficient management of big data, in three important data categories: operational, scientific and social data. The project comprised a significant development action for IMSI, offering added value to existing research results and delivering novel algorithms, models and prototype tools for managing big data. MEDA.2 makes a step ahead towards data science, exploiting S&T results from MEDA. The objectives of MEDA.2 are: (a) study and experiment with data science tools and technologies, (b) setting up a data science infrastructure (tools and data collections), (c) delve into pilot cases relevant to national RIS3 smart specialization requirements, and identify related data science problems to deal with (d) design and develop innovative technologies and data services to deal with problems in (c).



## APOLLONIS

**Greek Infrastructure for Digital Arts, Humanities and  
Language Research and Innovation**

<b>Project coordinator</b>	Panos Constantopoulos
<b>Coordinator</b>	IMSI
<b>IMSI - funding</b>	615,500 euros
<b>Programme</b>	ESPA
<b>Start date</b>	1/11/2017
<b>Duration</b>	3 years
<b>Website</b>	<a href="https://apollonis-infrastructure.gr/">https://apollonis-infrastructure.gr/</a> <a href="https://services.apollonis-infrastructure.gr/">https://services.apollonis-infrastructure.gr/</a>

APOLLONIS is the Greek Infrastructure for Digital Arts, Humanities and Language Research and Innovation, recently formed by the union of two existing ESFRI-related national research infrastructures: clarin:el, the CLARIN-related Greek network for language resources, technologies and services; and DARIAH-GR/DYAS, the DARIAH-related Greek network for digital research in the Humanities. The development of the APOLLONIS infrastructure advances the existing clarin:el and DARIAH-GR/DYAS services within a common framework that will ensure interoperability and reach to broader user communities, as well as promoting open science principles. Clarin:el will provide a permanent, stable infrastructure for accessing language resources and language processing web services, will support all kinds of language-related activities (regardless of subject), and collaborative workspace for application development environment. DARIAH-GR/DYAS will provide access to curated digital resources and services for the development, analysis and visualization of data, best practice guidelines, and dissemination and training activities on the use of digital methods and tools in the Humanities. A Digital Humanities Observatory will monitor the penetration of digital practices in the Humanities.



### eLib GGDE - eLib of Independent Authority for Public Revenue

<b>Project coordinator</b>	George Papastefanatos
<b>Coordinator</b>	Remaco S.A.
<b>IMSI - funding</b>	61,200 euros
<b>Programme</b>	NSRF - Public Tender
<b>Start date</b>	30/7/2015
<b>Duration</b>	7 years (ongoing)
<b>Website</b>	<a href="http://www.publicrevenue.gr/elib/">http://www.publicrevenue.gr/elib/</a>

The project eLib aims at developing a digital library for the Independent Authority of Public Revenue (I. A.P.R). eLib analyzes and provides information to public servants and citizens on the legislation concerning the areas of IAPR competence (taxation, public revenues, etc). Legal documents are automatically harvested from institutional sources (National Printing Office, Di@ygeia), their content and metadata are extracted, semantically analyzed and structured according to the

AKOMA NTOSO, and ELI) specifications. Moreover, content is indexed for full-text search, browsing and faceted filtering. Users can search and navigate the legal content and create personal collections with references to whole documents or parts of them.

## Other Projects



### TwitHoard - retrieving, modeling and analyzing Social Data

<b>Project coordinator</b>	Yannis Stavrakas
<b>Coordinator</b>	IMSI
<b>Programme</b>	Internal project
<b>Start date</b>	12/2015
<b>Duration</b>	4 years and 1 month

An increasing number of innovative applications use data from online social networks. In many cases data analysis tasks, like opinion mining processes, are applied on platforms such as Twitter, in order to discover what people think about various issues. In our view, selecting the proper data set is paramount for the analysis tasks to produce credible results. This direction, however, has not yet received a lot of attention. TwitHoard is a platform for supporting processes such as opinion mining on Twitter data, with emphasis on the selection of the proper data set. The key point of our approach is the representation of term associations, user associations, and related attributes in a single model that also takes into account their evolution through time. This model enables flexible queries that combine complex conditions on time, terms, users, and their associations.



### Europeana DSI-4

<b>Project coordinator</b>	Dimitris Gavrilis
----------------------------	-------------------

<b>Coordinator</b>	STICHTING EUROPEANA(EF)
<b>IMSI - funding</b>	97,372 euros
<b>Programme</b>	Service Contract-“Deployment and Maintenance of Europeana DSI Core Services-SMART-2017/1136”
<b>Start date</b>	1/9/2018
<b>Duration</b>	24 months
<b>Website</b>	<a href="https://pro.europeana.eu/project/europeana-dsi-4">https://pro.europeana.eu/project/europeana-dsi-4</a>

The Europeana DSI-4 project builds on and continues operation of the existing Europeana Digital Service Infrastructure’s (DSI) Core Service Platform (CSP) following on from the Europeana DSI-1, DSI-2 and DSI-3 actions under CEF. It is the fourth project for the completion of Europeana’s strategy 2015-2020. It is in line with the Connecting Europe Facility (CEF) Trans-European Telecommunications Network’s work programme for 2015, delivering interoperability, connectivity and coordination for digital cultural heritage at a European level and providing efficient solutions for access and distribution of multilingual and multi-domain resources in digital European cultural heritage. IMSI participates in DSI-4 both as a technical partner with the MORE platform (supporting the CARARE network as part of DSI) and also through Europeana Research.



## Digital Curation Services

<b>Project coordinator</b>	Panos Constantopoulos
<b>Coordinator</b>	IMSI
<b>IMSI - funding</b>	48,786 euros
<b>Programme</b>	Internal project
<b>Start date</b>	1/2/2016
<b>Duration</b>	4 years
<b>Website</b>	<a href="http://www.dcu.gr/en/υπηροεσίες-ψηφιακής-επιμέλειας/">http://www.dcu.gr/en/υπηροεσίες-ψηφιακής-επιμέλειας/</a>

The internal project Digital Curation Services supports the advancement of research and development efforts in the research directions and work programme of IMSI by using chiefly funds remaining from the execution of service provision projects. In view of parallel externally funded projects, this project on one hand

ensures the continuity of our work programme, while on the other it enables actions of topical interest for which external funds may be insufficient at the time. It also allows better integrating and employing results of completed projects in sustained and new digital curation services, thus supporting the capitalization of those results.

## Industry – Funded R&D Projects



Design and Development of big data solution and methods for stream analytics on network data

<b>Project manager</b>	George Papastefanatos
<b>Coordinator</b>	Intracom Telecom S.A.
<b>IMIS - funding</b>	509,000 euros
<b>Programme</b>	Contract
<b>Start date</b>	16/10/2017
<b>Duration</b>	52 months (ongoing)
<b>Website</b>	=

A new collaboration between **IMSI, Intracom Telecom and Ericsson** started in *October 2017*. IMSI has been contracted to design and develop a big data solution and methods for stream analytics on network data coming from IoT devices. The solution is based on well-established big data and cloud platforms, i.e., Microsoft Azure, Cloudera Ecosystem and SPARK streaming and enables the collection of vast amount of network measurement streams from connected devices, their scalable processing, analysis and storage and the visualization of several KPIs.



Design and Development of Ericsson Device EcoSystem

<b>Project manager</b>	George Papastefanatos
<b>Coordinator</b>	Intracom Telecom S.A.

<b>IMIS - funding</b>	44,888 euros
<b>Programme</b>	Contract
<b>Start date</b>	20/12/2019
<b>Duration</b>	7 months
<b>Website</b>	=

The project concerns the provision of design and development services for Ericsson Device EcoSystem, a digital marketplace for collecting compliance specifications of IoT devices and manufacturers with international standards.



# Publications

## Journal Publications

- Nikos Giatrakos, Nikos Katzouris, Antonios Deligiannakis, Alexander Artikis, Minos N. Garofalakis, George Paliouras, Holger Arndt, Raffaele Grasso, Ralf Klinkenberg, Miguel Ponce de Leon, Gian Gaetano Tartaglia, Alfonso Valencia, Dimitrios Zissis: **Interactive Extreme: Scale Analytics Towards Battling Cancer**. IEEE Technol. Soc. Mag. 38(2): 54-61 (2019)
- Georgios Chatzigeorgakidis, Kostas Patroumpas, Dimitrios Skoutas, Spiros Athanasiou, Spiros Skiadopoulou: **Visual Exploration of Geolocated Time Series with Hybrid Indexing**. Big Data Res. 15: 12-28 (2019)
- Nikos Bikakis, George Papastefanatos, Olga Papaemmanouil: **Big Data Exploration, Visualization and Analytics**. In Journal on Big Data Research, Vol. 18 (2019). Editorial in Special Issue, <https://doi.org/10.1016/j.bdr.2019.100123>
- Agiatis Benardou and Leonidas Konstantelos: **A world of immersive experiences**. History Scotland, 19(2): 8-9 (2019)
- Apostolaki, Stella, Eburn Akinsete, Stella Tsani, Phoebe Koundouri, Nikittas Pittis, and Eleftherios Levantis. 2019. "Assessing the Effectiveness of the WFD as a Tool to Address Different Levels of Water Scarcity Based on Two Case Studies of the Mediterranean Region" Water 11, no. 4: 840. <https://doi.org/10.3390/w11040840>
- Eburn Akinsete, Stella Apostolaki, Nikos Chatzistamoulou, Phoebe Koundouri and Stella Tsani. **The Link between Ecosystem Services and Human Wellbeing in the Implementation of the European Water Framework Directive: Assessing Four River Basins in Europe**, Water, Special Issue Economics of Water Resources Management. Water 2019, 11, 508; <https://doi.org/10.3390/w11030508>
- Stella Apostolaki, Phoebe Koundouri, Nikitas Pittis. 2019. **Using a systemic approach to address the requirement for Integrated Water Resource Management within the Water Framework Directive**. Science of the Total Environment, Volume 679, 20 August 2019, Pages 70-79 <https://doi.org/10.1016/j.scitotenv.2019.05.077>
- Phoebe Koundouri, Stella Tsani, Nikittas Pittis, Eleftherios Levantis, 2019. **Socio-Economic Measures for Achieving Total Water Cost Recovery: A Brief Analysis with Illustration from the Evrotas River Basin, Greece**. Environmental Analysis and Ecology Studies, 2019, <http://dx.doi.org/10.31031/eaes.2019.03.000610>
- Tsiros P, Bois FY, Dokoumetzidis A, Tsiliki G, Sarimveis H (2019) Population Pharmacokinetic Reanalysis of a Diazepam PBPK Model. J Pharmacokinet Pharmacodyn, doi: 10.1007/s10928-019-09630-x.
- Basei G, Hristoov D, Lamon L, Zabeo A, Jeliakova N, Tsiliki G, Marcomini A, Torsello A (2019) Making use of available and emerging data to predict the hazards

of engineered nanomaterials by means of in silico tools: A critical review. *Nanoimpact*, doi: 10.1016/j.impact.2019.01.003

- I. Pietri, Y. Chronis, Y. Ioannidis: “**Fairness in dataflow scheduling in the cloud**”. *Inf. Syst.*, Vol. 83, pp. 118-125, July 2019.
- E. Kharlamov, et al: “An ontology-mediated analytics-aware approach to support monitoring and diagnostics of static and streaming data”. *J. Web Semant.*, Elsevier, Vol. 56, pp. 30-55, May 2019.
- Laurent Busé, Clément Laroche, Fatmanur Yildirim. Implicitizing rational curves by the method of moving quadrics. *Computer-Aided Design*, Volume 114, 2019, Pages 101-111. <https://doi.org/10.1016/j.cad.2019.05.019>.
- Eldon L Ulrich et al. : “NMR-STAR: comprehensive ontology for representing, archiving and exchanging data from nuclear magnetic resonance spectroscopic experiments”, *Journal of biomolecular NMR*, Springer, Vol. 73, No.1, pp.5-9, February 2019.

## International Conference / Workshop Publications

- Ioannis Z. Emiris, Vasilis Margonis, Ioannis Psarros. Near-Neighbor Preserving Dimension Reduction for Doubling Subsets of  $l_1$ . *APPROX-RANDOM 2019*: 47:1-47:13, <https://doi.org/10.4230/LIPIcs.APPROX-RANDOM.2019.47>
- Emmanouil Christoforou, Ioannis Z. Emiris, Apostolos Florakis. Neural Networks for Cryptocurrency Evaluation and Price Fluctuation Forecasting. *MARBLE 2019*: 133-149, [https://doi.org/10.1007/978-3-030-37110-4\\_10](https://doi.org/10.1007/978-3-030-37110-4_10)
- Ioannis Z. Emiris, Christina Katsamaki. Voronoi Diagram of Orthogonal Polyhedra in Two and Three Dimensions. *SEA<sup>2</sup> 2019*: 1-20, [https://doi.org/10.1007/978-3-030-34029-2\\_1](https://doi.org/10.1007/978-3-030-34029-2_1)
- Nikos Giatrakos, Alexander Artikis, Antonios Deligiannakis, Minos N. Garofalakis: **Uncertainty-Aware Event Analytics over Distributed Settings**. *DEBS 2019*: 175-186
- Vasilis Samoladas, Minos N. Garofalakis: **Functional Geometric Monitoring for Distributed Streams**. *EDBT 2019*: 85-96
- Xiaoying Wu, Dimitri Theodoratos, Dimitrios Skoutas, Michael Lan: **Evaluating Mixed Patterns on Large Data Graphs Using Bitmap Views**. *DASFAA (1) 2019*: 553-570
- Xiaoying Wu, Dimitri Theodoratos, Dimitrios Skoutas, Michael Lan: **Efficiently Computing Homomorphic Matches of Hybrid Pattern Queries on Large Graphs**. *DaWaK 2019*: 279-295
- Spiros Athanasiou, Giorgos Giannopoulos, Damien Graux, Nikos Karagiannakis, Jens Lehmann, Axel-Cyrille Ngonga Ngomo, Kostas Patroumpas, Mohamed Ahmed Sherif, Dimitrios Skoutas: **Big POI data integration with Linked Data technologies**. *EDBT 2019*: 477-488

- Spiros Athanasiou, Michail Alexakis, Giorgos Giannopoulos, Nikos Karagiannakis, Yannis Kouvaras, Pantelis Mitropoulos, Kostas Patroumpas, Dimitrios Skoutas: **SLIPO: Large-Scale Data Integration for Points of Interest**. EDBT 2019: 574-577
- Giorgos Giannopoulos, Konstantinos Alexis, Nikos Kostagiolas, Dimitrios Skoutas: **Classifying points of interest with minimum metadata**. LocalRec@SIGSPATIAL 2019: 1:1-1:4
- Georgios Chatzigeorgakidis, Dimitrios Skoutas, Kostas Patroumpas, Themis Palpanas, Spiros Athanasiou, Spiros Skiadopoulos: **Local Similarity Search on Geolocated Time Series Using Hybrid Indexing**. SIGSPATIAL 2019: 179-188
- Kostas Patroumpas, Dimitrios Skoutas, Georgios M. Mandilaras, Giorgos Giannopoulos, Spiros Athanasiou: **Exposing Points of Interest as Linked Geospatial Data**. SSTD 2019: 21-30
- Georgios Chatzigeorgakidis, Dimitrios Skoutas, Kostas Patroumpas, Themis Palpanas, Spiros Athanasiou, Spiros Skiadopoulos: **Local Pair and Bundle Discovery over Co-Evolving Time Series**. SSTD 2019: 160-169
- Giorgos Eftaxias, Nontas Tsakonas, Giorgos Giannopoulos, Nikos Kostagiolas, Andreas Syngros, Dimitrios Skoutas: **LGM-PC: A tool for POI classification on QGIS**. SSTD 2019: 182-185
- Giorgos Giannopoulos, Marios Meimaris: **Learning Domain Driven and Semantically Enriched Embeddings for POI Classification**. SSTD 2019: 214-217
- Vassilis Kaffes, Giorgos Giannopoulos, Nikos Karagiannakis, Nontas Tsakonas: **Learning Domain Specific Models for Toponym Interlinking**. SIGSPATIAL/GIS 2019: 504-507
- Giorgos Alexiou, George Papastefanatos: **Query Driven Entity Resolution in Data Lakes**. In 13th International Workshop on Information Search, Integration, and Personalization (ISIP 2019), Heraklion, Greece, May 9-10, 2019, 117-130.
- Giorgos Alexiou, George Papastefanatos. **Query driven Entity Resolution in Data Lakes** (Extended Abstract). In 13th International Workshop on Information Search, Integration, and Personalization, 9-10 May, 2019, Heraklion
- Marios Meimaris, George Papastefanatos. **Merging RDF Characteristic Sets to Optimize SPARQL Queries** (Extended Abstract). In 13th International Workshop on Information Search, Integration, and Personalization, 9-10 May, 2019, Heraklion
- Danae Pla Karidi, Harry Nakos, and Yannis Stavarakas. **Automatic Ground Truth Dataset Creation for Fake News Detection in Social Media**. 20th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL 2019), Manchester, 14-16 November 2019.
- Konstantinos Pontikakos, Myrto S. Barda, Konstantinos I. Kotsopoulos, Yannis Stavarakas, Konstantinos P. Ferentinos. **The Geospatial Database of the VR-Park System for Promoting and Touring Urban Parks: the "Pedion Areos" Case Study**. 3rd Conference on Geographical Information Systems and Spatial Analytics in Agriculture and Environment, Agricultural University of Athens, Athens, Greece, 11-13 December, 2019.

- Agiatis Benardou: **Immersive Experiences And Difficult Heritage: Digital Methods As Re-interpreters Of Historically Contested Sites**. In DH2019 Conference, Utrecht University 9-12 July, 2019.
- Lorna Hughes and Agiatis Benardou: **'The Ties That Bind': The Creation, Use, And Sustainability Of Community Generated Histories**. In DH2019 Conference, Utrecht University 9-12 July, 2019.
- Panos Constantopoulos: **APOLLONIS: The Greek Infrastructure for Digital Arts, Humanities and Language Research and Innovation**. In DH2019 Conference, Utrecht University 9-12 July, 2019.
- N. Klyuchnikov, D. Mottin, P. Karras, G. Koutrika, E. Müller. Figuring out the User in a Few Steps: Bayesian Multifidelity Active Search with Cokriging. ACM KDD 2019
- Maria Ilvanidou: **And The First One Now Will Later Be Last, For The Times They Are A-changin': Modeling Land Communication In Roman Crete**. DH2019 Conference, Utrecht University 9-12 July, 2019.
- Vayianos Pertsas and Panos Constantopoulos: **From Research Articles to Knowledge Graphs: Methods for ontology-driven knowledge base creation from text**. Tutorial, The Web Conference WWW2019, San Francisco, 13-17 May 2019.
- Panos Constantopoulos and Vayianos Pertsas: **From Publications to Knowledge Graphs**. In Flouris, G., Laurent, D., Plexousakis, D., Spyratos, N., Tanaka, Y. (Eds.), Information Search, Integration, and Personalization, Revised Selected Papers from 13th International Workshop, ISIP 2019, Heraklion, Greece, May 9–10, 2019, Springer, 2020, pp. 18-33. DOI: 10.1007/978-3-030-44900-1
- Thanasis Vergoulis, Serafeim Chatzopoulos, Ilias Kanellos, Panagiotis Deligiannis, Christos Tryfonopoulos, Theodore Dalamagas: **BIP! Finder: Facilitating Scientific Literature Search by Exploiting Impact-Based Ranking**. CIKM 2019: 2937-2940
- Thanasis Vergoulis, Ilias Kanellos, Anargiros Tzerefos, Serafeim Chatzopoulos, Theodore Dalamagas, Spiros Skiadopoulos: **A Study on the Readability of Scientific Publications**. TPD 2019: 136-144
- Serafeim Chatzopoulos, Panagiotis Deligiannis, Thanasis Vergoulis, Ilias Kanellos, Christos Tryfonopoulos, Theodore Dalamagas: **SciTo Trends: Visualising Scientific Topic Trends**. TPD 2019: 393-396
- Th. Mailis, Y. Kotidis, V. Nikolopoulos, Evg. Kharlamov, I. Horrocks, Y. Ioannidis: **"Mv-Index: An Efficient Index for Graph-Query Containment"**. Proc. ISWC Satellites, Auckland, New Zealand, pp. 193-196, October 2019.
- Th. Mailis, Y. Kotidis, V. Nikolopoulos, Evg. Kharlamov, I. Horrocks, Y. Ioannidis: **"An Efficient Index for RDF Query Containment"**. Proc. Int'l ACM SIGMOD Conference, Amsterdam, Netherlands, pp. 1499-1516, June 2019.

## National Conference / Workshop Publications

- Marios Meimaris, George Papastefanatos. **Merging RDF Characteristic Sets to Optimize SPARQL Queries**. In 17th Hellenic Data Management Symposium (HDMS'19), Athens, July, 2019
- Nikos Bikakis, Stavros Maroulis, George Papastefanatos, Panos Vassiliadis: **RawVis: Visual Exploration over Raw Data**. In 17th Hellenic Data Management Symposium (HDMS'119), Athens, July, 2019

## Other Publications / Technical Reports

- Vasilis Samoladas, Minos N. Garofalakis: **Functional Geometric Monitoring for Distributed Streams**. HDMS 2019.
- Georgia Angelaki, Karolina Badzmierowska, David Brown, Vera Chiquet, Joris Colla, Judith Finlay-McAlester, Klaudia Grabowska, Vanessa Hanneschläger, Natalie Harrower, Freja Howat-Maxted, Maria Ilvanidou, Wojciech Kordyzon, Magdalena Król, Antonio Gabriel Losada Gómez, Maciej Maryl, Sanita Reinsone, Natalia Suslova, Mark Sweetnam, Kamil Śliwowski, Marcin Werla. (2019). **How to Facilitate Cooperation between Humanities Researchers and Cultural Heritage Institutions. Guidelines**. DOI: 10.5281/zenodo.2587481
- Phoebe Koundouri (contributing author). Magand, C., Alves, M. H., Calleja, E., Datry, T., Dörflinger, G., England, J., Gallart, F., Gómez, R., Jorda-Capdevila, D., Marti, E., Munne, A., Pastor, V. A., Stubbington, R., Tziortzis, I. and Von Schiller, D. , 2020, Intermittent Rivers and Ephemeral Streams: What water managers need to know, Science and Management of Intermittent Rivers and Ephemeral Streams (SMIRES), Technical report – Cost ACTION CA 15113. 5281/zenodo.3888474. Available: <http://doi.org/10.5281/zenodo.3888474>
- T. Vergoulis, S. Chatzopoulos, I. Kanellos, A. Tzerefos, T. Dalamagas: "A comparative study on the gender representation in the Computer Science academic community". 1st Summit on Gender Equality in Computing. Athens, Greece, June 2019



# Dissemination Activities

---

## Invited / Keynote Talks

IMSI members participated in the following invited / keynote talks:

- Panel Coordination. **Modern Information Technologies: Artificial Intelligence, Big Data and Signals**. Athena RC: 30-year history, inventing the future. Athens, Greece, November 18, 2019. [A. Simitsis](#).
- **The Lifecycle of Big Data Processing Flows**. Infore, Barcelona, Spain, May 8, 2019. [A. Simitsis](#).
- **Challenges and Solutions for Applying ML, AI, and Big Data Technology on Medical Data**. ESTRO Think Tank Meeting: “Provocative Questions for Medical Physics (MP) in Radiation Oncology (RO)”, Budapest, Hungary, October 24, 2019. [A. Simitsis](#).
- **Infrastructure for Big Data Applications**. National Observatory of Athens, December 13, 2019. [A. Simitsis](#).
- Oct 2019. **“Building the catalogue of e-infrastructure services for the European Open Science Cloud”** [George Papastefanatos](#), Invited Speaker, In Symposium on “Open Access: From theory to practice”, Nicosia, Cyprus.
- June 2019. **“Enhancing the value proposition of the EOSC Portal for the Research Infrastructure community and their users”** [George Papastefanatos](#), Panel Speaker in International Workshop on “Building Open Science in Europe: The road ahead for the EOSC community and the EU Member States”, Tallinn, Estonia.
- [Constantopoulos, P.](#), **“From the catalogue to the virtual information space”**. Invited talk, IFLA World Library and Information Congress, Athens, 24-30 Aug. 2019.
- [Constantopoulos, P.](#), and V. Pertsas, **“From publications to knowledge graphs”**, World Health Organization Global Technical Meeting 2019, Seoul, 12-14 Nov. 2019.
- [Benardou, A.](#), **“The “de-siloization” of cultural heritage data: Musings on the European digital arts and humanities landscape”**, Invited Talk, Swedish Institute Athens, 9 April 2019, Invited Talk, Stegi Onassis Foundation, 28 May 2019
- [Benardou, A.](#), **“Βαρέθηκα να βλέπω την σκηνή, και σήκωσα τα μάτια μου στα θεωρεία»: Το Αρχείο Καβάφη και η ερευνητική κοινότητα των ψηφιακών ανθρωπιστικών επιστημών”**, Invited Talk, Stegi Onassis Foundation, 9 April 2019
- [Benardou, A.](#), **“Οι ψηφιακές ερευνητικές πρωτοβουλίες στην Ελλάδα και στην Ευρώπη και η εξέλιξη της επιστήμης της Ιστορίας”**, Invited Talk, National and Capodistrian University of Athens, 11 Oct. 2019
- [Benardou, A.](#), **“Immersive experiences and difficult heritage: Past lessons / Challenges ahead”**, Invited Talk, University of California Berkeley, 14 Nov. 2019

- 2019, “**Anonymization and GDPR**”, Manolis Terrovitis, invited speaker in 3rd Workshop on GDPR in conjunction with e-Democracy Conference

## Scientific Community Service

IMSI members have served in the **Program Committee** of more than 25 International Conferences and Workshops in 2019, including well-known Conferences like ACM KDD, EDBT, ICDE, VLBD, SSTD, SIGSPATIAL, CIKM, WISE, WWW and more.

IMSI members have participated in the **organization** or co-organization of the following events:

- **ACM International Workshop on Data Warehousing and OLAP (ACM DOLAP)**. A. Simitsis, Steering Committee.
- **ACM SIGMOD Record**. A. Simitsis, Associate Editor, Research Center Column.
- **Data & Knowledge Engineering**, Elsevier. A. Simitsis, Editorial Board Member.
- Hellenic Data Management Symposium (HDMS) 2019, Athens, Greece. Georgia Koutrika. General co-chair.
- 1st Summit on Gender Equality in Computing (GEC 2019). Georgia Koutrika. General co-chair.
- Greek ACM-W Chapter: Georgia Koutrika is treasurer
- ACM SIGMOD: Georgia Koutrika is Associate Information Director
- **Elsevier Big Data Research Journal**. Special Issue on Interactive Big Data Visualization and Analytics. George Papastefanatos, Guest Editor.
- 2<sup>nd</sup> International Workshop on Big Data Visual Exploration and Analytics (**BigVis2019**) in conjunction with EDBT/ICDT, George Papastefanatos.
- **DARIAH-EU**, Agiatis Benardou as VCC2 (Research and Education Liaison) Co-Chair.



- **spaTScope**

spaTScope is a web application for visual exploration of geolocated time series developed in the context of the EU-funded projects SLIPO and SmartDataLake. spaTScope allows users to visually explore large collections of geolocated time series and obtain insights about trends and patterns in their area of interest. The provided functionalities leverage a hybrid index that allows to navigate and group the available time series based not only on their similarity but also on spatial proximity. The results are visualized using linked plots combining maps and timelines.

<https://github.com/smardatalake/spaTScope>



- **Mopseus**

MOPSEUS is a scalable, curation-aware repository system designed to be metadata schema agnostic. It can support any complex data model either at the digital resource level or at the collection level. This means that content can be organized using any structure ranging from simple hierarchies to complex graphs. The entire structure both at the digital resource level and at the collection level is stored and represented in RDF and can be accessed through a SPARQL endpoint. MOPSEUS employs an expressive data model that supports both intra- and inter- object relations thus allowing arbitrary organization of objects. In particular, all entities in MOPSEUS are digital objects each of which may include an unlimited number of metadata and/or data files and may be associated with multiple metadata schemas. A special class of digital objects, the containers, are used in order to organize information (digital evidence). Containers can be interconnected using semantic links thus giving rise to semantic graph structures of

arbitrary complexity. Metadata can be represented either as XML or RDF triples. MOPSEUS gives special focus on interoperability and digital preservation and is compliant with the PREMIS standard ensuring that the entire lifecycle of each digital resource is stored and semantically annotated. In order to access the MOPSEUS services a modular architecture is employed whereby the user interacts with the system through a set of Web-based interfaces that allow one to define metadata schemas and thesauri and to manage the entire content stored within the system. A SKOS editor supports maintaining term thesauri, while a linked data approach is adopted in associating terms with relevant data objects. Moreover, semantic relationships between objects are supported and can be defined through the GUI. Finally, all operations are organized in workflows which are also defined via the GUI. Supported site installations of the Mopseus repository system include: “Digital Academy” – Repository of the digital collections of the Academy of Athens, and “Pyxida” – Academic repository and digital library, Athens University of Economics and Business.

<http://mopseus.dcu.gr/>

- **MORe**

The Metadata and Object Repository (MORe) is a metadata aggregator designed to: harvest content (metadata records) from different sources and providers; enrich/curate; map to a target schema (e.g. EDM); and deliver the metadata using the OAI-PMH protocol to other systems, such as the Europeana library. MORe focuses on enriching / curating the aggregated content. This is accomplished through a set of micro-services that are streamlined in a workflow. These micro-services perform various curation actions like normalizing, associating records (e.g. those in close proximity to each other), transforming spatial coordinates to a given coordinate system (e.g. WGS84), creating elements like place labels, etc. MORe is OAI-compliant and preserves the whole lifecycle of each digital object. All ingestion and curation actions create new versions of metadata streams that are stored and semantically annotated, thus allowing to view the entire history of changes associated with each digital resource. MORe is tuned to support massive imports in the order of about 0.8 million records per hour. Online services with the MORe system are provided to Europeana and the related communities formed and supported through the projects CARARE, 3-D ICONS, ARIADNE, LoCloud and CEF Europeana. A total of approximately 70 content providers in over 20 countries, are being regularly served by the MORe aggregation service to aggregate, enrich and deliver content to Europeana, with approximately 10 million heritage asset records processed so far.

<http://more.dcu.gr/>

- **NeMO and SO**



The NeDiMAH Methods Ontology (NeMO) is a comprehensive ontological model of scholarly practice in the arts and humanities, offered and maintained by DCU/IMSI, the development of which was undertaken through the ESF Research Network [NeDiMAH](#). NeMO is a CIDOC CRM - compliant ontology which explicitly addresses the interplay of factors of agency (actors and goals), process (activities and methods) and resources (information resources, tools, concepts) manifest in the scholarly process. It builds on the results of extensive empirical studies and modeling of scholarly practices performed by DCU/IMSI in projects [DARIAH](#) and [EHRI](#). NeMO incorporates existing relevant taxonomies of scholarly methods and tools, such as TaDIRAH, the arts-humanities.net and Oxford taxonomies of ICT methods, DHCommons, CCC-IULA-UPF and DiRT, through appropriate mappings of the concepts defined therein onto a semantic backbone of NeMO concepts. It thus enables combining documentary elements on scholarly practices of different perspectives and using different vocabularies. NeMO was subsequently generalized to the domain-neutral Scholarly Ontology (SO).

<http://nemo.dcu.gr/>

- **BIP! Finder**

BIP! Finder is a tool that assists the discovery of high-impact publications in the field of life sciences.

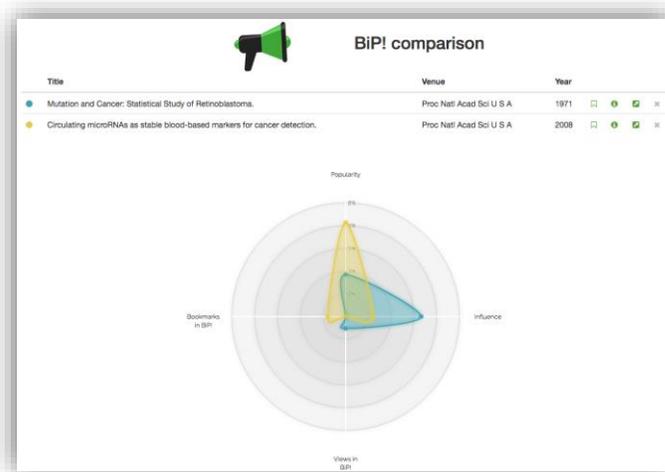
This tool supports ranking and comparing of scientific articles based on different aspects of their impact in their discipline, like their popularity (i.e., the current attention they receive) or

influence (i.e., their long-term impact in the discipline). Furthermore, the tool provides useful features like intuitive infographics for each article and a mechanism of bookmarks.

<http://bip.imis.athena-innovation.gr/>

- **DAIAD system**

DAIAD is the first *integrated residential demand management* system for water. It applies Big Data and Machine Learning technologies to leverage smart water meter data, engaging and informing consumers to induce sustainable changes in consumption behaviour, as well as providing novel large-scale analytics to improve short-, medium-, and long-term demand management for water utilities. DAIAD provides personalized pricing and non-pricing interventions to consumers through mobile and web



applications, adapted to their profile, individual determinant sensitivity, and consumption behaviour. Water utilities have access to several analysis services (segmentation, clustering, forecasting) enabling them to understand consumption behaviour at the household level, target specific consumer groups, and anticipate demand under various time scales. The average sustainable total water savings in residential water consumption achieved by the DAIAD system is -12%. DAIAD is available as an open source software under the Apache License.

<https://github.com/DAIAD>

<https://www.youtube.com/watch?v=YuLU9nitlss>

- **BUFET**

BUFET is an open-source software under the GPL v.3 licence, designed to speed up Bleazard's unbiased miRNA enrichment analysis algorithm. BUFET generates an empirical distribution of genes targeted by miRNA and calculates p-values for related biological processes. Benjamini-Hochberg FDR correction produces a '\*' or '\*\*' for significance at 0.05 FDR and 0.01 FDR respectively.

<https://github.com/diwis/BUFET>

- **PaperRanking**

PaperRanking is an open source library containing implementations of paper ranking methods that have been proposed in the literature. Our implementations utilise a suite of MapReduce scripts and can be used either on a single machine, or a Hadoop cluster. All codes were developed in the context of a paper ranking survey that aimed to evaluate each method's strengths and weaknesses.

<https://github.com/diwis/PaperRanking>

- **SciNeM**

SciNeM (Data Science tool for heterogeneous Network Mining), an open-source tool that offers a wide range of functionalities for exploring and analysing HINs and utilises Apache Spark for scaling out through parallel and distributed computation. SciNeM provides an intuitive, Web-based user interface to build and execute complex constrained metapath-based queries and to explore and visualise the corresponding results. Under the hood, all the supported state-of-the-art HIN analysis types have been implemented in a scalable manner supporting the distributed execution of analysis tasks on computational clusters. SciNeM has a modular architecture making it easy to extend it with additional algorithms and functionalities. Currently, it supports the following operations, given a user-specified metapath: ranking entities using a random walk mode, retrieving the top-*k* most similar pairs of entities, finding the most similar entities to a query entity, and discovering entity communities.

<http://scinem.imsi.athenarc.gr/>





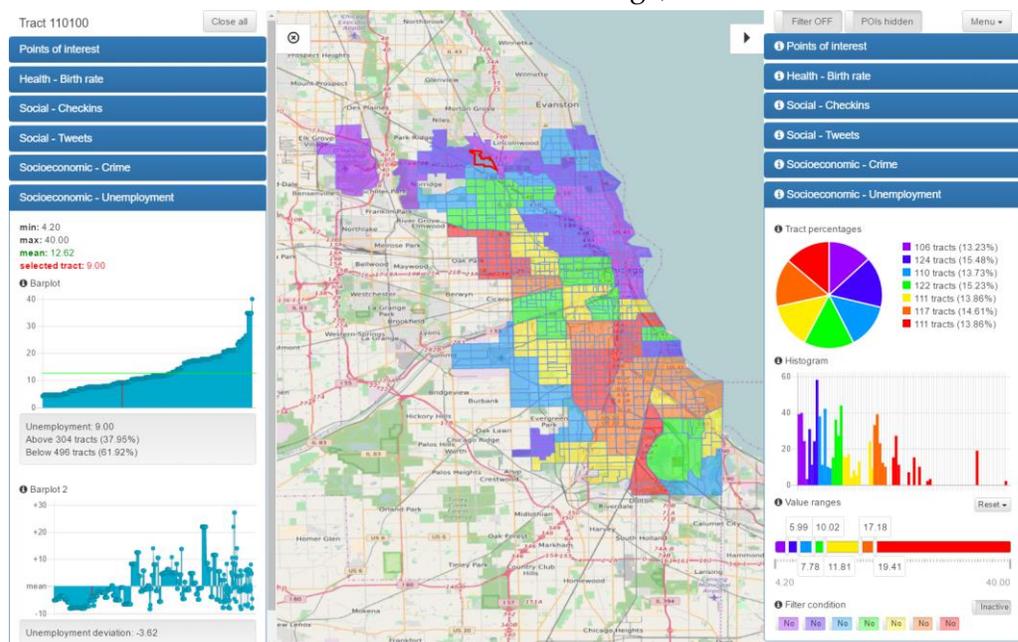
- **eLib**

The project eLib aims at developing a digital library for the Independent Authority for Public Revenue. eLib analyzes and provides information to public servants and citizens on the legislation concerning the areas of IAPR competence (taxation, public revenues, etc). Legal documents are automatically harvested from institutional sources (National Printing Office, Di@ygeia), their content and metadata are extracted, semantically analyzed and structured according to the AKOMA NTOSO, and ELI) specifications. Moreover, content is indexed for full-text search, browsing and faceted filtering. Users can search and navigate the legal content and create personal collections with references to whole documents or parts of them.

<http://www.publicrevenue.gr/elib>

- **GeoProfiler**

Social networks, available open data and massive online APIs provide huge amounts of data about our surrounding location, especially for cities and urban areas. Unfortunately, most applications and research usually focus on one kind of data over the other, thus presenting a biased and partial view of each location in question, hence limiting the benefits of such approaches. To remedy this, the GeoProfiler framework, developed within the CitySense project, combines data from administrative sources (e.g., public agencies), massive Point of Interest APIs (Google Places, Foursquare) and social microblogs (Twitter) to provide a unified view of all available information about an urban area, in an intuitive and easy to use web-application platform. Our use case shows how different sources of information can be combined to provide an accurate and diverse visualization for the urban area of Chicago, USA.



City of Chicago Demo: <http://geoprofiler.imis.athena-innovation.gr/>

- **Socioscope**

A visual analysis tool, used for visualization and exploration of social and political data ([www.socioscope.gr](http://www.socioscope.gr)) seeks to deliver a visual analytics platform for the social scientist

to explore and analyze social facts through a user-friendly visual interface.

The Socioscope platform offers a variety of interactive

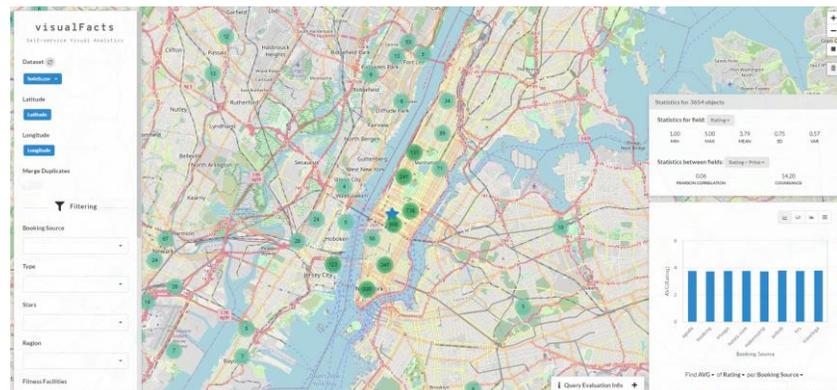


visualizations for each different type of data: charts and histograms, pies and tacked diagrams for numerical data; timelines for indices; and choropleth and point maps for geographical data. The platform is based on a multidimensional modeling approach and offers several visual operations for data exploration and analysis, such as filtering through faceted browsing, hierarchical representation of coded lists in charts, free keyword search of literal values, and capabilities for combining different datasets along common dimensions. Moreover, it makes knowledge reusable by making all data available for download in various formats including Linked Open Data.

[www.socioscope.gr](http://www.socioscope.gr)

- **VisualFacts**

VisualFacts is a self-service visual analytics platform for big geo-located data that helps data explorers perform ad hoc analysis of raw data files collected from different sources



of varying quality (with duplicates or missing data) in rich visual ways, even though they do not have a background in notebooks, data integration, or machine learning techniques. VisualFacts platform allows users to open their own data file(s) and via a map-centric Dashboard UI start visually interacting with the data without loading or indexing the data in a database. The backbone of the platform is a visual aware in-memory index (called rawViz), which is constructed on the fly and adjusted to user interaction, as well as a powerful deduplication engine (called QueryER) which offers on-the-fly visual entity matching and clustering over dirty data. The platform can scale

up the visualization, interactive exploration and analysis to million data points on a map, with the use of commodity hardware.

<https://visualfacts.imsi.athenarc.gr/software.html>

The two libraries are offered as standalone tools and as an integrated platform.

- **QueryER**

QueryER is the SQL engine behind VisualFacts which integrates entity resolution (ER) operations in the planning and execution of select-project-join queries. It offers three novel query operators, which (1) identify and resolve duplicates within a table employing a schema agnostic resolution approach; (2) enables joins between two or more tables containing duplicate entities; and (3) group/merge deduplicated entities into a single representation.

<https://github.com/VisualFacts/queryER>

- **rawVis: A System for In-situ Visual Exploration & Analytics**

RawVis enables efficient in-situ visual exploration and analytics directly over large raw data files without the need of an underlying DBMS or a query engine. RawVis exhibited low response time over large datasets (e.g., 50G & 100M objects) using commodity hardware.

<https://github.com/VisualFacts/RawVis>

- **TripleGeo**

TripleGeo is an ETL utility that can extract geospatial features from various sources (e.g. shapefiles, spatial DBMSs) and transform them into Basic Geo or GeoSPARQL compatible RDF triples, in several serialization formats. It copes with most common spatial data types, like points, linestrings and multi-linestrings, polygons and multi-polygons and supports on-the-fly transformations between different coordinate reference systems. Also, TripleGeo supports the transformation of INSPIRE-aligned spatial data and metadata into RDF, using XSL stylesheets, for selected INSPIRE data themes.

<https://github.com/SLIPO-EU/TripleGeo>

- **OSMRec**

OSMRec is a tool that trains on a set of spatial entities annotated with categories and provides category recommendations for new geospatial entities. OSMRec's goal is to exploit the richness of available geospatial datasets than contain entities already annotated with several categories (e.g., OpenStreetMap), to enrich new geospatial entities. OSMRec supports two modes of deployment: a generic command line, and a JOSM plugin, which allows the real-time recommendation of OSM categories for geospatial entities created within the JOSM user interface.

<https://github.com/SLIPO-EU/OSMRec>



<http://wiki.openstreetmap.org/wiki/JOSM/Plugins/OSMRec>

- **Amnesia anonymization tool**

Amnesia transforms a dataset with direct identifiers and quasi identifiers to an anonymized dataset, where formal privacy guaranties hold. Amnesia allows the use to customize the anonymization process, to choose the trade-off between data utility and privacy protection. Moreover, it allows users who are not IT experts to visually explore the data and the impact of different anonymization settings on them. It helps use to create supportive material to the anonymization process, like generalization hierarchies. Amnesia offers k-anonymity and km-anonymity and a parallel scalable anonymization algorithm, it is available through the OpenAIRE infrastructure.

- **FAGI**

FAGI is a tool that allows the fusion of geospatial Linked Data. It is designed to retrieve data through SPARQL endpoints, and implements a wide range of fusion actions both on spatial properties of the entities and on non-spatial metadata. These include moving, rotating, scaling and aligning the geometries of the entities, combining multiple, semantically related properties, maintaining both descriptions of a property of two linked geospatial entities, etc. It also supports batch fusion actions, automatic classification of fused entities using OSM categories, and provides a map-based UI.

<https://github.com/SLIPO-EU/FAGI>



# Education

## PhD / MSc / Diploma Thesis Co-supervision

IMSI members actively co-supervise several undergraduate and graduate university students, who often conduct their work in IMSI premises. As a result, in 2019 several MSc and Diploma thesis have been co-supervised by IMSI members, who often serve as members in the respective examination committees.

The following PhD students collaborated closely with IMSI members in their research during 2019:

- Vassilis Kaffes. Topic: **Keyword queries on road networks**. Joint supervision with the University of Peloponnese. Collaborating researcher: [Dimitris Skoutas](#).
- Giorgos Chatzigeorgakidis. Topic: **Scalable Indexing and Exploration of Big Time Series Data**. Joint supervision with the University of Peloponnese. Collaborating researcher: [Dimitris Skoutas](#).
- Theodora Galani. Topic: **Modeling and querying the data evolution and provenance**. Joint supervision with the National Technical University of Athens. Collaborating researchers: [Yannis Stavrakas](#), [George Papastefanatos](#).
- Danai Pla-Karidi. Topic: **Recommendation models using Social Networks**. Joint supervision with the National Technical University of Athens. Collaborating researcher: [Yannis Stavrakas](#).
- Giorgos Alexiou. Topic: **Entity disambiguation and data interlinking**. Joint supervision with the National Technical University of Athens. Collaborating researcher: [George Papastefanatos](#).
- Theofilos Belmpas and Oresti Gini (University of Athens) – In-depth comparison of NLIDBs - co-supervisor: Georgia Koutrika
- George Katsogiannis-Meimarakis (University of Athens) – Deep Learning for NL2SQL Translation co-supervisor: Georgia Koutrika
- Konstantinos Theocharidis, Topic: **Social Data Management**. Joint supervision with the University of Peloponnese. Collaborating Researcher [Manolis Terrovitis](#)
- Dimitris Tsitsigkos, Topic: **Join Operators for Complex Data**. Joint supervision with the University of Ioannina. Collaborating Researcher [Manolis Terrovitis](#)
- Serafeim Chatzopoulos. Topic: **Text mining and Information Retrieval for Scientific Texts**. Joint supervision with the University of Peloponnese. Collaborating researchers: [Theodore Dalamagas](#), [Thanasis Vergoulis](#).

- Konstantinos Zagganas. Topic: **Efficient techniques for data intensive analysis and processing in life sciences**. Joint supervision with the University of Peloponnese. Collaborating researchers: [Theodore Dalamagas](#), [Thanasis Vergoulis](#).
- Ilias Kanellos. Topic: **Data extraction from scientific publications and research analytics**. Joint supervision with the National Technical University of Athens. Collaborating researchers: [Theodore Dalamagas](#), [Thanasis Vergoulis](#).

## Other Educational Activities

Other educational activities involving IMIS members include the following.

- **Organization of the 3rd ACM Europe Summer School in Data Science** 11–17 July 2019, Athens, Greece. The summer school invites researchers around the world (i.e., young faculty, postdocs, PhD candidates, young engineers from industry and in some special cases MSc students or even senior undergraduate students) to attend a series of interesting and constructive lectures by distinguished professors. Lectures focus on several forms of analytics, machine learning and data mining techniques, big data systems, and the ethics of data science.
- [Giorgos Giannopoulos](#) taught the course “**Big Data Management**”, for the Master of Science in “Space Science Technologies and Applications”, University of the Peloponnese.
- [Thanasis Vergoulis](#) co-taught the “**Data Management**” course of the post-graduate program “Data Science” of NCSR Demokritos.
- George Papastefanatos was an Invited Lecturer in the course “Analysis and Design of Information Systems” (Undergraduate course – Athens University of Economic and Business - Department of Informatics).
- Organization of the [1st Summer School in Digital Humanities](#), Athens, 15-19 July 2019.



# Facts and Figures

## Financial report

In 2019, IMSI continued its participation in EC and national funded research and development projects. The key economic indicators regarding the expenses and revenues in 2019 are shown in Table 1 and their distribution in categories is illustrated in Figure 1. We can see that the highest percentage of the revenue stream, about 66%, comes from EC funded projects. It is important to note that the revenues from the activities of IMSI (EC projects, national funded projects, Product and Service Sales) are more than 15 times the public expenditure received by IMSI.

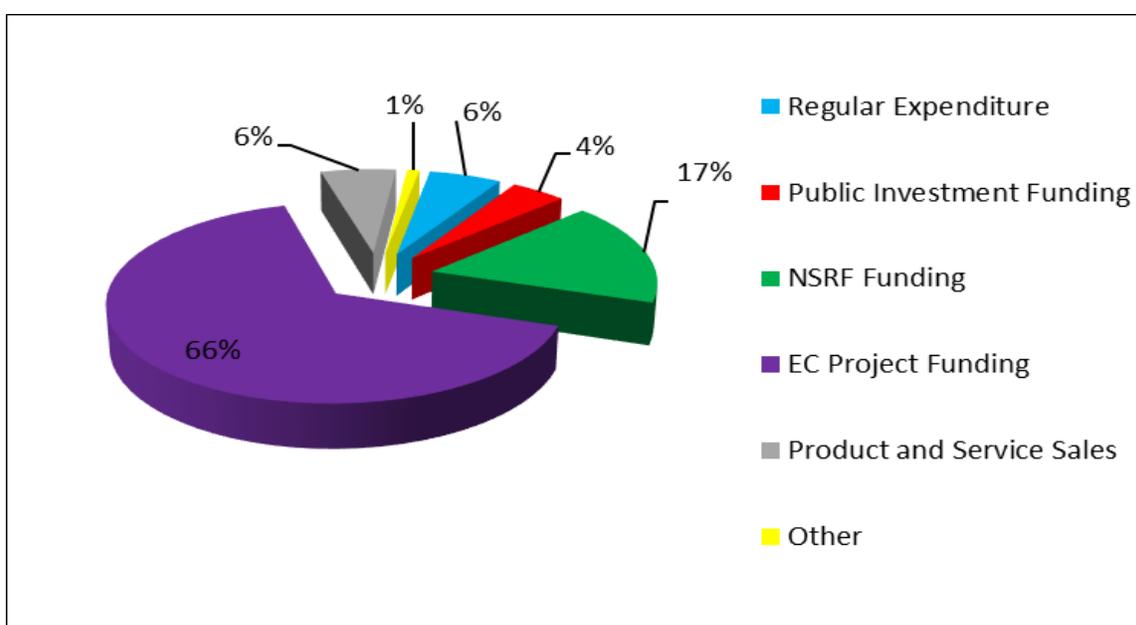


Figure 1. Distribution of revenues in 2019

Table 1. Expenses and Revenues for 2019

Expenses 2019 (in Euros)	
Travel Expenses	278.221,56
Operational Costs	50.506,08
Equipment	75.937,75
Other Expenses	200.774,39
Personnel fees and payments to third parties	3.297.793,84
<b>Total</b>	<b>3.903.233,62</b>
Revenues 2019 (in Euros)	

<b>Regular Expenditure</b>	346.777,65
<b>Public Investment Funding</b>	267.080,07
<b>NSRF Funding</b>	1.023.287,09
<b>EC Project Funding</b>	3.926.562,69
<b>Product and Service Sales</b>	359.999,93
<b>Other</b>	59.553,25
<b>Total</b>	<b>5.983.260,68</b>

## Staff

### Director & Researchers

- Garofalakis Minos, Professor, Technical University of Crete, Director
- Koutrika Georgia, Researcher A
- Sartzetakis Stylianos, Researcher A
- Simitsis Alkiviadis, Researcher A
- Dalamagas Theodore, Researcher A
- Stavrakas Yannis, Researcher A
- Skoutas Dimitris, Researcher C
- Terrovitis Manolis, Researcher C

### Adjunct or under contract Professors

- Alepis Efthimios, Assistant Professor, University of Piraeus
- Chatzopoulos Michael, Emeritus Professor, University of Athens
- Constantopoulos Panos, Professor, AUEB
- Dallas Costis, Assistant Professor, Toronto University
- Deligiannakis Antonis, Assistant Professor, Technical University of Crete
- Douligeris Christos, Professor, University of Piraeus
- Emiris Ioannis, Professor, University of Athens
- Ioannidis Yannis, Professor, University of Athens
- Kountouri Phoebe, Assistant Professor, AUEB
- Papandreou Andreas, Professor, University of Athens
- Papatheodorou Christos, Professor, Ionian University
- Patsakis Konstantinos, Assistant Professor, University of Piraeus
- Samoladas Vasilis, Associate Professor, Technical University of Crete
- Tsipouri Eleni, Professor, University of Athens
- Vassilopoulos Achilleas, Assistant Professor, University of Ioannina
- Vazirgiannis Michalis, Professor, AUEB

### Postdoctoral Researchers / Scientific Collaborators

- Arkoumanis Constantinos, Elec. & Comp. Eng. NTUA, PhD NTUA
- Benardou Agiatis, BA, PhD in Ancient History, King's College London
- Boile Maria, Department of History and Archaeology, UoA, MA Aristotle University of Thessaloniki, MA Hellenic Open University, PhD Aristotle University of Thessaloniki
- Casino Fran, BSc in Computer Science, URV, MSc in Computer Security and Intelligent Systems, URV, PhD in Computer Science, Rovira i Virgili University

- Dasaklis Thomas, Department of Industrial Management and Technology, University of Piraeus, MSc NTUA, University of Piraeus, PhD University of Piraeus
- Dimitropoulos Charalampos, Elec. & Elec. Eng. City University London, PhD University of London
- Dritsou Vicky, BSc in Informatics, AUEB, MSc in Information Systems, AUEB, PhD in Informatics, AUEB
- El Raheb Katerina, BSc in Chemical Engineering, NTUA, MSc & PhD, Dept. of Inform. and Telecom. UoA
- Gavrilis Dimitris, Elec. Comp. & Eng. University of Patras, PhD University of Patras
- Giannopoulos Giorgos, Elec. & Comp. Eng. NTUA, PhD NTUA
- Giatrakos Nikolaos, BSc in Informatics, University of Piraeus, MSc in Information Systems, AUEB, PhD, Department of Informatics, University of Piraeus
- Grypari Ioanna, BSc in Applied Mathematics & Economics, Brown University, MSc LSE, PhD in Economics, University of Minnesota
- Karagiannis Fotis, Elec. & Comp. Eng. NTUA, PhD NTUA
- Konaxis Christos, Dept. of Mathematics UoA, PhD . Dept Informatics & Telecoms, UoA
- Mailis Theofilos, Elec. & Comp. Eng. NTUA, PhD NTUA
- Meimaris Marios, Elec. & Comp. Eng. NTUA, Msc NTUA, University of Edinburgh, PhD NTUA
- Mihota Alexandra, BSc, MSc, PhD, Dept. of Digital Systems, University of Piraeus
- Ntantogian Christoforos, BSc, MSc, PhD, Dept. of Inform. and Telecom. UoA
- Papastefanatos George, Elec. & Comp. Eng. NTUA, PhD NTUA
- Pierrakos Dimitris, Dept. of Physics UoA, PhD UoA
- Samartzis Panagiotis, BSc in Mathematics, UoA, MSc in Mathematics & Finance, Imperial College London, PhD in Financial Econometrics, University of Piraeus
- Tsani Stella, Dept. Of Economics UoA, PhD University of Reading
- Tsiavos Prodromos, Dep. of Law UoA, Master of Laws (LL.M), Univ. of London, MSc Analysis Design & Management of Information Systems, LSE & Political Science, PhD LSE
- Tsiliki Georgia, BSc in Statistics, University of Piraeus, PhD in Statistics, Imperial College London
- Tzotsos Aggelos, BA in Rural and Surveying Engineering, NTUA, PhD NTUA
- Vergoulis Thanasis, Elec. Comp. & Eng. University of Patras, PhD NTUA
- Zacharia Eleni, Dept. of Inform. and Telecom. UoA, PhD UoA

### **Administrative and Technical Personnel**

- Apostolopoulou Nelly, Mech. Eng. NTUA, Msc NTUA, Head of Administration
- Bika Katerina, BA & MA, Department of International, European and Area Studies, Panteion University, Administrative Support
- Georgantas Panagiotis, Elec. & Comp. Eng. NTUA, Technician
- Paradissi Eva, Secretarial Support

### **Research Staff**

- Aglamisi Dimitra, BA in International and European studies, University of Piraeus
- Alexakis Michalis, Elec. & Comp. Eng. NTUA
- Alexandropoulou Vera, BA in Law, UoA, LL.M Heidelberg University
- Alexiou George, University of Aegean, Msc AUTH
- Athanasiou Spyros, Elec. & Comp. Eng. NTUA
- Balasis Ioannis, Dept. of Inform. and Telecom. UoA

- Bartzos Vangelis, Dep. of MBG, DUTH, Dep. of Math., Univ. of Patras, Msc Ecole normale superieure, Lyon, Msc Universite 1, Lyon
- Chatzigeorgakidis Giorgos, Elec. & Comp. Eng., Technical University of Crete, MSc in Computer Science, Denmark's Technical University, PhD Candidate, University of Peloponnese
- Chatzilenas Christos, Department of Informatics and Telematics, Harokopio University, Msc, Department of Informatics, AUEB
- Chatzopoulos Serafeim, BA, Msc, Dept. of Inform. and Telecom. UoA
- Chira Vassiliki, Department of Management science and technology, AUEB, MSc in computer science, AUEB
- Christoforidis Stamatios, BSc, MSc, Dept. of Inform. and Telecom. UoA
- Chronis Pantelis, Comp. & Inform. Eng. University of Patras
- Clement Laroche, BSc in Fundamental Mathematics, Toulouse III - Paul Sabatier University, MSc in Mathematics, Toulouse III - Paul Sabatier University
- Costagiolas Nikolaos – Stefanos, Dept. of Inform. and Telecom. UoA, Msc in Data Science, AUEB
- Dimakopoulos Nikolaos, Dept. of Inform. and Telecom. UoA, MSc in Data Science & Machine Learning, NTUA
- Diochnou Vasiliki, Department of Political Science and Public Administration, UoA, MSc in Southeast European Studies, UoA
- Diplas Ioannis, Dept. of Inform. and Telecom. UoA
- Emmanouilidi Sofia, Deree, Business Administration (Accounting and Finance)
- Foufoulas Ioannis Dept. of Inform. and Telecom. UoA, Msc UoA
- Galouni Constantina, Dept. of Inform. and Telecom. UoA
- Gentimis Thanassis, Elec. & Comp. Eng. NTUA, Msc UoA
- Giannakopoulos Anastasios, Dept. of Inform. and Telecom. UoA
- Giannoudis Asterios, Dept. of Inform. and Telecom. UoA
- Glenis Apostolos, BSc in Computer Science, University of Piraeus, MSc in Computer Science, University of Crete
- Iatropoulou Katerina, BA, Msc, Dept. of Inform. and Telecom. UoA
- Ilvanidou Maria, Faculty of History and Archaeology, UoA, Msc, University of Crete
- Kaffes Vassilis, Elec. & Comp. Eng. NTUA
- Kakalettris Georgios, BA, Msc, Dept. of Inform. and Telecom. UoA
- Kalampokis Panagiotis, BA, Computer Science, University of Ioannina, Msc in Mobile and Web Computing, IHU
- Kanakakis Panagiotis, Dept. of Inform. and Telecom. UoA
- Kanellos Ilias, Elec. & Comp. Eng. NTUA
- Karagiannakis Nikos, Comp. & Inform. Eng. University of Patras
- Karamanolis Georgios, Dept. of Inform, AUEB
- Karavaltsiou Maria, BA in Translation Studies, Ionian University, MA in Conference Interpreting, Aristotle University of Thessaloniki
- Kasomoulis Aristotelis, BSc, MSc, Dept. of Inform. and Telecom. UoA
- Katakazi Elissavet Olga, BA AUEB, Msc in Applied Environmental Economics, University of Kent
- Kavouras Lukas, School of Applied Mathematical and Physical Sciences, NTUA, MSc in Logic, Algorithms and Computation (NTUA, University of Athens, University of Patras), PhD Candidate NTUA
- Kokogiannaki Argiro, Dept. of Inform. and Telecom. UoA

- Konidaris Vissarion Berthold, Elec. & Comp. Eng., Technical University of Crete, MSc Technical University of Crete
- Kontaxakis Antonios, Elec. & Comp. Eng., Technical University of Crete, MSc Technical University of Crete
- Koukouli Mirto, Faculty of History and Archaeology, UoA
- Koulalis Antonios, BA Dept. of Inform. and Telecom. UoA
- Kouvaras Ioannis, Elec. & Comp. Eng. NTUA
- Kravvaritis Constantinos, Dept. of Inform. and Telecom. UoA, Msc, Dept. of Inform, AUEB
- Lebesis Antonis, BA, Msc, Dept. of Inform. and Telecom. UoA
- Legakis Antonis, BA in International & European Economic studies, AUEB, MSc in Electronic Business Management, University of Warwick
- Leonidakis Nikiforos, Dept. of Inform. and Telecom. UoA
- Levantis Eleftherios, Dep. of Law, UoA, Msc Environmental Management EAEME
- Liatas Christos, Electrical and Electronic Engineering, University of Nottingham, MSc University of Bristol
- Lougiakis Christos, BSc, MSc, Dept. of Inform. and Telecom. UoA
- Lykousas Nikolaos, Department of Informatics, University of Piraeus, MSc in Intelligent Interactive Systems, Universitat Pompeu Fabra
- Maili Evdokia, BA in Mathematics, University of Patras, Msc, Dept. of Inform. and Telecom. UoA
- Manola Natalia, Dept. of Physics, UOA
- Mantas Andreas, Department of Computer Engineering and Informatics, University of Patras
- Mantes Athanasios, BSc, MSc, Department of Computer Engineering and Informatics, University of Patras, MSc in Geoinformatics, NTUA, MSc in Bioinformatics, UoA
- Marinos Kouris Christos, Department Of Business Administration, AUEB, MSc Strathclyde University of Glasgow
- Maroulis Stavros, Elec. & Comp. Eng. NTUA
- Martziou Stefania-Varvara, BA, Msc, Dept. of Inform. and Telecom. UoA
- Mavrogiannakis Nikos, BA, Msc in Physics, University of Manchester
- Metaxas Omiros, BA, Msc, Dept. of Inform. and Telecom. UoA
- Mougiakos Ioannis, BA in Rural and Surveying Engineering, NTUA, Msc, University of Nottingham
- Mpikas Georgios, Dept. of Electrical Engineers, TEI Piraeus
- Mpiliris Grigorios, BA in Rural and Surveying Engineering, NTUA, Msc, University College of London
- Nakos Charalampos, Elec. & Comp. Eng. NTUA
- Nasikas Christos, BSc in Informatics, University of Piraeus
- Nomikos Vangelis, Dept. of Inform. and Telecom. UoA
- Oikonomou Leonidas, Architecture, School of Architecture NTUA, MSc NTUA
- Papadaki Aikaterini, Dept. of Geography, University of the Aegean, Msc, Rural and Surveying Engineering, NTUA
- Papadopoulou Elli, BA Librarian & Informatics Scientist, Alexander Technological Educational Institute of Thessaloniki
- Patroumpas Konstantinos, Comp. & Inform. Eng. University of Patras, Msc NTUA
- Pavlidou Androniki, BA Mathematics, AUTH, MSc on Science in Innovation Management, Univ. of Utrecht
- Petousi Dimitra, Department of Library Science & Information Systems, Technological Educational Institute of Athens, MA in Museum studies, UoA
- Petra Eleni, Comp. & Inform. Eng. University of Patras, BA University of Manchester

- Pikramenos George, BSc in Mathematics, Imperial College London, MSc in Computer Science, UoA
- Pittaras Nikiforos, BSc in Computer Science, University of Ioannina, MSc in Signal and Information Processing and Learning, UoA
- Pla-Karidi Danae, Elec. & Comp. Eng. NTUA
- Pristouris Costis, Elec. & Comp. Eng. NTUA, MSc in Geoinformatics, NTUA
- Psallidas Michail, Comp. & Inform. Eng. University of Patras
- Sifakaki Ilektra, BA in Mathematics, UoA, Msc, Department of Informatics and Telematics, Harokopio University
- Smyris Georgios, BSc in Physics, UoA, MSc UoA
- Soulis Athanasios, Dept. of Inform. and Telecom. UoA
- Spiliopoulou Panorea, School of Law, UoA, Master of Laws, University of Hamburg/Erasmus University/University of Vienna, Master of Laws, Bordeaux V-UoA
- Spiropoulos Dimitrios, Dept. of Inform. and Telecom. UoA
- Spirou Konstantinos, Dept. of Inform. and Telecom. UoA
- Stamatakis Giorgos, Elec. & Comp. Eng., Technical University of Crete
- Stamatopoulos Vassilis, Dept. of Inform. and Telecom. UoA
- Stergiou Marina, BSc in Chemical Engineering, NTUA, MSc in Design of Interactive and Industrial Products and Systems, University of the Aegean
- Svingos Christoforos, Dept. of Inform. and Telecom. UoA, MSc UoA
- Theocharidis Constantinos, Dep. of Electrical and Computer Engineering, University of Thessaly
- Toli Eleni, Department of History and Archaeology, University of Ioannina, MSc in Labour Economics Ruhr University Bochum
- Tsapatsari Mirto, BA in Law, UoA, LL.M in International Business Law, VU University Amsterdam
- Tsitsigkos Dimitris, Dept. of Inform. and Telecom. UoA
- Tziotzios Diamantis, Elec. & Comp. Eng. DUTH
- Vagias Christos, BA in Rural and Surveying Engineering, NTUA, Msc, University of Nottingham
- Vos Aggeliki, BA, Department of Library Science & Information Systems, TEI of Athens, Msc in Information Systems, Linnaeus University, Sweden
- Xiros Panagiotis, Elec. & Comp. Eng. NTUA
- Zagganas Constantinos, Elec. & Comp. Eng. NTUA
- Zeakis Alexandros, Dept. of Inform. And Telecom. UoA, MSc in Data Science, AUEB
- Zouros Michalis, Department of Digital Systems, University of Piraeus

# Contact

---

## Information Management Systems Institute (IMSI)

**Athena** Research and Innovation Center in Information, Communication and Knowledge Technologies

**Postal Address:**

Artemidos 6 & Epidavrou  
15125, Maroussi, Athens  
Greece

**Phone:** +30 210 6875403

**Fax:** +30 210 6856804

**e-mail:** [contact-imsi@athenarc.gr](mailto:contact-imsi@athenarc.gr)

Most current information about activities and contacts for the staff are available on the IMSI website:

<https://www.imsi.athenarc.gr/en>

