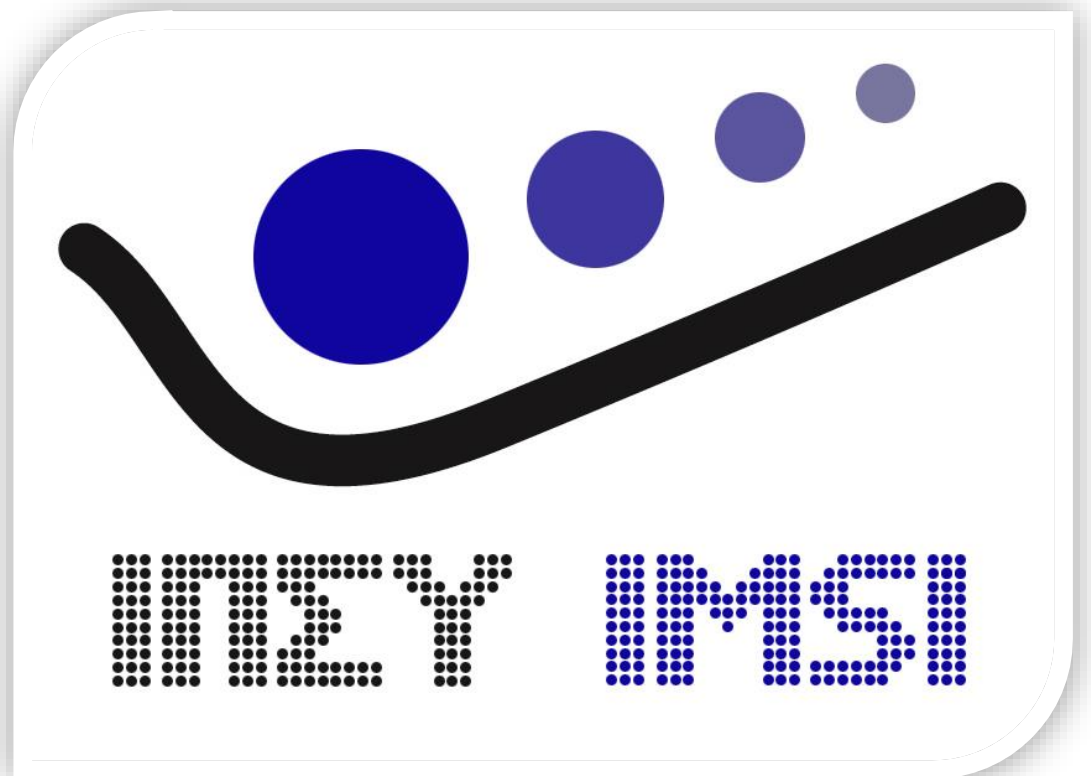# IMSI annual report 2021

# 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.
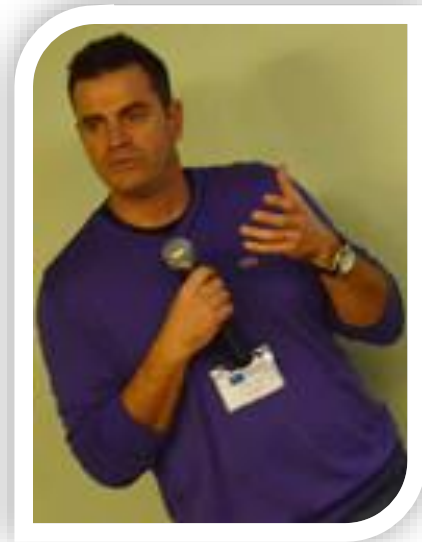
As in previous years, in 2021, IMSI has significantly expanded the scope of its research efforts in a number of focus areas, including Big Data and Scalable Data Analytics, Web Data Management and Semantic Web Technologies, Geospatial Data Management, and Digital Curation and Research Infrastructures. 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, 16/5/2022

# Contents

# 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 cutting-edge scientific research, and exploit research results in the development of novel core technologies, prototypes, applications, and products in its areas of expertise, including information and knowledge management, large-scale information systems and applications, databases and Big Data management systems, cloud-based platforms and services, Machine Learning and knowledge-extraction technologies, and Digital Curation. To this end, IMSI had brought together a team of internationally-known experts from a broad range of related disciplines (such as databases, systems, algorithms, and machine learning) under an environment promoting excellence in research, collaboration, and interdisciplinarity.

More specifically, IMSI's mission includes:

- Research and development in advanced computer and information systems with targeted applications in industry, society, and the real economy.
- Implement large research and development projects in the area of information systems technologies, in collaboration with academic, research, and industrial partners.
- Collaborate with the academic community in efforts involving research, education, and the transfer of knowledge and results to the information and software systems industry.
- Develop experimental and industrial information systems prototypes, as well as innovative products and services, in collaboration with industrial partners.

- Design, develop, operate, maintain, support, and evolve innovative infrastructures for data storage and analytics in various application domains (e.g., bioinformatics, precision medicine, natural sciences).
- Support international research and academic activities (at both the EU and global level) in computer science through competitive research programs and collaboration agreements/contracts with academic institutions and industrial partners.
- Transfer and exploit cutting-edge research results and technologies to industry, through the development of innovative products and services, as well as the founding of technology spinoff companies.
- Support the Greek public sector in developing novel technological solutions that emphasize the exploitation of Big Data and digital services for improving processes and services offered to society.

To attain these objectives, IMSI is structured in Departments as follows:

- **Department on Big Data Analytics and Machine Learning**
  The mission of the department is research, technological development and innovation in the fields of large-scale algorithms and systems for management, processing and analysis of large and heterogeneous volumes of static and dynamic data. In this context, key objectives include efficient, interactive analysis of Big Data for different application domains such as Complex Event Processing, the extraction and dynamic update of complex Machine Learning models and the development of Predictive Analytics models. Issues are also covered including analysis of continuous data streams, platforms and tools for scalable data analytics, algorithms and systems for large-scale supervised and unsupervised Machine Learning, as well as Privacy-Preserving Data Mining.
- **Department on Big Data Research Infrastructures**
  The mission of the department is to address research, technological and innovation challenges in the field of systems and infrastructures for the organization, storage, curation, and management of large data volumes to support a variety of important application domains. In this context, key activities involve efficiency and scalability issues for digital research infrastructures (RIs), including techniques and systems for complex information flow processing tailored to heterogeneous computation and data storage environments. Also, the group has a strong focus on producing reliable, high-quality digital assets, facilitating their archiving and long-term maintenance, and uncovering their added value via knowledge extraction tasks. In this context, our work demonstrates a strong interdisciplinary aspect, providing solutions for the effective exploitation of Big Data technologies in scientific areas and having a leading role in existing European and National RIs for several scientific domains. We carry out R&I activities to build scalable data infrastructures, either tailor-made for specific scientific domains (e.g., Health, Humanities) or generic enough for any RI (Generic Data Infrastructures, Open Science).
- **Department on Cloud Platforms and Data Services**

The mission of the department is to perform research, technological development, and innovation in the field of cloud computing and its exploitation for the development of innovative information systems and services. In this context, issues such as the development and support of various cloud service delivery models (e.g., Function as a service (FaaS) Platform-as-a-Service (PaaS), Software-as-a-Service (SaaS)) for hosting big data platforms, data and service catalogues and markets, secure processing and outsourcing of data and services to third parties, end-to-end big data analysis services in the cloud and at the edge, cloud native end-user applications for data analytics, as well as related tools and technologies, are covered.

- **Department on Distributed and Web Information Systems**

  The mission of the department is to address research, technological, and innovation challenges in the field of data management and processing in Web applications, and in distributed environments in general. In this context, we cover issues related to data modelling and management for the Data Web and the Semantic Web, Knowledge Graphs and Ontologies, integration of heterogeneous data sources, Web services, personalized information retrieval and recommendation, as well as sensor networks and peer-to-peer systems.

- **Department on User-centric Systems and Applications**

  The group's mission is research, development and innovation in the areas of user-centric and data-centric systems and applications. The group is specifically focusing on the development of algorithms and systems that (a) seamlessly learn from and adapt to users and data, (b) enable users to access data in more effective and human-like ways, (c) support fair and ethical data access and applications, and (d) promote data democratization in different real-life domains, from policy making to health and astrophysics. It naturally covers topics including data science, data exploration, intelligent data interfaces, recommender systems, conversational AI, personalization, fair, accountable and transparent algorithms, explainable systems, crowdsourcing, user analytics, visual analytics, and computer-assisted education.

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

- **IMSI co-organized ICDE 2021**. IMSI/ATHENA R.C, together with the Technical University of Crete (TUC), hosted the 37th IEEE International Conference on Data Engineering (ICDE), 19-22 Apr 2021. ICDE is the flagship IEEE conference addressing research issues in designing, building, managing, and evaluating advanced data-intensive systems and applications. For over three decades, ICDE has been a leading forum for researchers, practitioners, developers, and users to explore cutting-edge ideas and to exchange techniques, tools, and experiences. Prof. Minos Garofalakis, IMSI Director, was General (co)Chair of the 37th IEEE International Conference on Da-ta Engineering (ICDE). Moreover, five IMSI members participated in ICDE 2021 organization: George Papastefanatos (Web chair), Yannis Stavrakas (Finance chair), Alkis Simitsis (Registration chair), Georgia Koutrika (Sponsorship chair), Theodore Dalamagas (Local Arrangements chair).

- The EU (H2020-MSCA-ITN-2019) project **GRAPES** organized successfully the following three events during 2021:
    a. The **1st Doctoral school & Midterm meeting** that took place virtually on February 1st.
    b. The **Learning Week I: Academic skills and advanced topics in CAD** that took place in Barcelona, Spain on September 6th.
    c. The **1st Software and Industrial Workshop** that took place at the INRIA research center at Sophia Antipolis, France on December 6th.
        Project scientific coordinator: <u>Ioannis Z. Emiris.</u>

- Georgia Koutrika became a member of the ACM-RAISE Working Group.
- Georgia Koutrika became a member of ATHENA's Gender Equality Committee
- Georgia Koutrika is PC co-Chair, ACM Symposium on Cloud Computing, (SoCC) 2021
- A new collaboration between **IMSI and Intracom Telecom S.A.** started in March 2021. IMSI has been contracted for the provision of consulting services, design and development services for the redesign of the company's NFV-RI product with emphasis on the easy visualization of metric performance and resource optimization scenarios that perform Network Virtualized Functions within Cloud Environments. Project Coordinator from IMSI: <u>George Papastefanatos.</u>

- Organization of the <u>series of online courses</u> "Digital Collections Metadata Curation in the Arts and Humanities", 30 March, 6 and 13 April 2021.

- Organization of <u>"DH goes Viral" Digital Workshop</u>, 26 April 2021, with the support of a <u>DARIAH Theme grant,</u> in which the Twitter Conference 2020 participants were

reunited to assess the impact of the Covid-19 pandemic on DH research and pedagogy one year after.

- Organization of the [2nd APOLLONIS Summer School in Digital Humanities](#), 13-15 July 2021.
- Organization of the event "Creating a thematic action for the Greek 1940s in the APOLLONIS Infrastructure" at the Thessaloniki International Fair 2021, 18 September 2021.
- Launch of the new [APOLLONIS Services portal](#) that brings together all the digital services, tools and resources available through APOLLONIS, the Research Infrastructure for Digital Arts, Humanities and Language Research and Innovation coordinated by DCU.
- Publication of the book "DH goes Viral" funded by DARIAH Theme 2020.
- Organization of ["DH goes Viral" Digital Workshop](#), 26 April 2021, with the support of a [DARIAH Theme grant,](#) in which the Twitter Conference 2020 participants were reunited to assess the impact of the Covid-19 pandemic on DH research and pedagogy one year after.

## Awards
- ACM SIGMOD 2021: Alkis Simitsis was a Distinguished PC Member.

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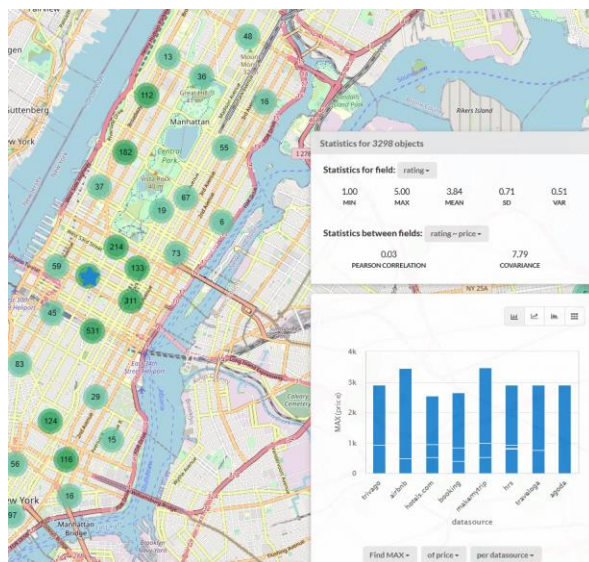
to integrate city data from disparate sources and provide a visual way to combine them and filter city areas.

- **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.

- **Defining and Detecting Complex Changes on RDF(S) Knowledge Bases**

  The dynamic nature of web data brings forward the need for maintaining data versions as well as identifying changes between them. In this work, we deal with problems regarding understanding evolution, focusing on RDF(S) knowledge bases, as RDF is a de-facto standard for representing data on the web. Revisiting past snapshots or the differences between them is not enough for understanding how and why data evolved. Instead, changes should be treated as first-class-citizens. In our view, this involves supporting semantically rich, user-defined changes, called complex changes, as well as identifying the relations between them. We focus on a declarative language for defining complex changes on RDF(S) knowledge bases. Moreover, on detecting complex change instances among dataset versions, which can be queried for analyzing evolution.

- **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.

- **Computational infrastructures for Life Sciences**
  We are building and maintaining Cloud-based computational infrastructures to support the work of life scientists. 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. Most of the work is being integrated into HYPATIA (https://hypatia.athenarc.gr/) the Cloud infrastructure that has been developed for the Greek Node of the ESFRI European RI ELIXIR (https://www.elixir-europe.org) and which supports the computational needs of the respective community of researchers.

- **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
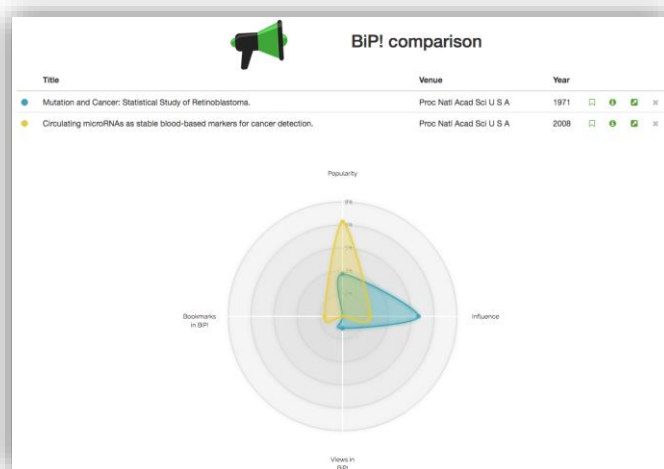
are working on citation network analysis technologies to identify high-impact research output, focusing on the study of impact aspects like long-term and short-term impact. We also apply text mining techniques on scientific texts to automatically extract useful knowledge. Finally, we develop powerful 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 ELIXRI-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.

- **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. This goal is actively pursued in MORE project (www.more2020.eu), which is funded by the EU. MORE aims to create a platform that will provide scalable Big Data analytics tools for sensor data coming from Renewable Energy Source, i.e., solar and wind parks. MORE will allow stakeholders in industry sectors with huge volumes of sensor data, especially the RES industry, to: a) scale the management of streaming and historical time series beyond an order of magnitude beyond the state-of-art and b) to perform forecasting, prediction and diagnostics using the whole data that is available to them with accuracy that outperforms existing approaches. In this context also plan to investigate the processing of deep learning operations (e.g., matrix factorization) in modern CPU architectures. Finally, we continue to work on optimization for modern architectures, especially for join operations.

- **Database Systems in the Era of Learning.**
  In this work, we revisit the traditional architecture components in database systems and depart from the 30+ year old concepts governing the design and implementation of critical database components, motivated by the recent, rapid progress on learning

technology. We investigate opportunities and explore ideas to bring the learning technology within the core of a database management system (DBMS) aiming at developing scalable, learning solutions to all challenging components of the database architecture, such as the optimizer, the execution engine, the process manager, the physical design, and so on. Our initial work includes (a) moving towards end-to-end learning optimizers, (b) employing learning techniques to manage and optimize workloads involving complex data, such as graph, spatial, and streaming, and (c) developing efficient and robust optimization techniques for distributed and cross-platform database systems.

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

- **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.
- **Domain specific Machine Learning and Analytics**
  Based on the experience we have obtained by implementing ML solutions for various disciplines, we have identified that, quite often, applying state of the art ML algorithms and workflows is not adequate to effectively solve specialized, but quite significant for real-world application, tasks. On the contrary, generic algorithms and processes often need to be significantly adapted/extended in order to handle specialized problems. This has become evident in several Earth Observation and analytics settings, as well in medical image analysis. Our aim is research how state of

the art ML/DL algorithms and methodologies can be properly extended, utilizing domain knowledge, in order to effectively solve real-world problems.

- **Explainability on AI, with emphasis on fairness of high-risk recommender system (RS) applications**

  In April 2021, the European Commission released a proposal for a harmonised regulatory framework on AI (the AIA), towards trustworthy AI. The AIA proposal differentiates between AI systems creating an unacceptable risk, a high risk, a transparency risk, and a minimal risk. AI systems posing an unacceptable risk are generally prohibited, whereas high-risk AI systems are subject to specific safeguards such as bias monitoring and correction; interpretability; and human oversight. With many real-world applications utilizing RS being categorized as high-risk (e.g., applications on Employment, Education and Credit-Worthiness), researching how explainability, fairness and user-interaction can be jointly handled and incorporated into such systems is a still open research problem with high value on a variety of applications. Our goal is to implement fairness aware explainability methods, as well as allow them to adapt and learn from user interaction-feedback.

## 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.

Components of Natural Language Database Interface are:

(a) A NL2SQL component takes as input a user question in natural language along with the database and produces a SQL query to be executed over this database.

(b) A SQL2NL component that does the inverse, and its natural language output can be used by the user to confirm the correctness of the query executed.

(c) A QueryResults2Text (QR2T) component takes a query, the results of the query, and a database, and returns a verbalisation of the results.

To bridge the gap between users and data, numerous NL2SQL systems have been developed that allow users to pose natural language questions over relational databases. Recently, novel NL2SQL systems are adopting deep learning methods with very promising results. At the same time, several challenges remain open making this area an active and flourishing field of research and development. We are working on leveraging recent advances of deep learning techniques for text-to-SQL translation in the context of different real cases, including policy making, astrophysics and cancer research.

Database interaction is often characterized as a non-trivial and time-consuming process due to user's inexperience with the data or the query language. Therefore, there is a need for the databases to be able to "talk back" in order to assist the users during data exploration and eventually lead them to the desired results. The SQL-to-Text problem has seen very little attention by the scientific community. Proposed solutions usually do not consider full relational databases and do not take advantage of the recent advances in the field of NLP.

We have developed a template-based approach that builds on Logos. Our novel extensions include improvements in terms of the system's translation capabilities and the fluency of the generated explanations. Furthermore, eQsplain is our deep learning SQL2NL system. Based on the T5 Transformer model, it can generate fluent and human-like query explanations. It is simultaneously trained for translating SQL and SPARQL. For "hard" scientific domains, we can incorporate information curated by experts, to improve explanations.

- **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.

  Towards this direction, we are working on algorithms that can provide query recommendations based on data analysis without the use of query logs. We deal with query completion recommendations in a cold-start setting, i.e., without relying on query logs, an inherently challenging problem. We present a new query

recommendation tool, PyExplore, that takes as input a SQL query and returns a set of ranked queries, each of them focusing on a subset of the initial query results that can help the user navigate the data space. PyExplore is based on the observation that people better perceive fewer dimensions, typically 2-3, and understand patterns in low-dimensional spaces. PyExplore first finds 'interesting' subsets of query attributes, i.e., that they exhibit correlation or diversity. Then, for each such subset, it clusters the data based on the values of the respective attributes.

- **Fair and Ethical Algorithmic Systems**
  Algorithmic systems, driven by large amounts of data, are increasingly being used in all aspects of society to assist people in forming opinions and taking decisions. For instance, search engines and recommender systems amongst others are used to help us in making all sorts of decisions from selecting restaurants and books, to choosing friends and careers. Other systems are used in school admissions, housing, pricing of goods and services, job applicant selection, and so forth. Such algorithmic systems offer enormous opportunities, but they also raise concerns regarding how fair they are. Hence, beyond efficiency and effectiveness of systems, the group investigates models and methods for fairness in algorithmic systems. Fairness, explainability, transparency are different sides of the same problem: how to make systems trustworthy.

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

- **User-Driven Data Management**
  The group is taking a holistic approach to user-centric systems and applications: building algorithms, systems, interfaces, and evaluation methodologies. In the effort to enable user-centric approaches, the group's research often focuses at the level of how to understand data and queries, and learn how to best process user queries, developing algorithms that leverage the best of both worlds, data management and deep learning, to build systems that can learn from user queries and from data to not only process queries more efficiently but also to understand user intention, adapt to users, and help the user achieve their information goals more effectively.
- **Rich and Highly Performant User-Defined Functions in Relational Databases.**
  The diversity and complexity of modern data management applications have led to the extension of the relational paradigm with syntactic and semantic support for User-Defined Functions (UDFs). Although well-established in traditional DBMS settings, UDFs have become central in many application contexts as well, such as data science, data analytics, and edge computing. Still, a critical limitation of UDFs is the impedance mismatch between their evaluation and relational processing. In this project, we design and develop an SQL extension with rich UDF support along with a pluggable architecture to easily integrate it with either server-based or embedded database engines. Our current focus is on Python UDFs that are fully integrated with relational queries as scalar, aggregator, or table functions. Key novel characteristics

include easy implementation of complex algorithms and several performance enhancements, including tracing JIT compilation of Python UDFs, parallelism and fusion of UDFs, stateful UDFs, and seamless integration with a database engine.

## Geospatial Data Management

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

- **Geolocated Time Series**

  Geolocated time series are time series that correspond to specific locations. They can represent, for example, visitor check-ins at certain venues or readings of sensors installed at various places. The amount and significance of such time series have increased in many domains over the last years. However, although several works exist for time series visualization and visual analytics in general, there is a lack of efficient techniques for visual exploration and analysis of geolocated time series in particular. In our work, we have developed two approaches that rely on hybrid spatial-time series indices to allow for efficient map-based visual exploration and summarization of geolocated time series data. In particular, we use the BTSR-tree index and we introduce a new variant of the iSAX index, called geo-iSAX. The former is a spatial-first hybrid index that extends the R-tree by maintaining bounds for the time series indexed at each node. Following a similar rationale, geo-iSAX is a time series-first hybrid index that maintains spatial MBRs of the geolocated time series indexed in each node. Based on these indices, we show how to efficiently produce map-based visualizations of geolocated time series at different levels of granularity.
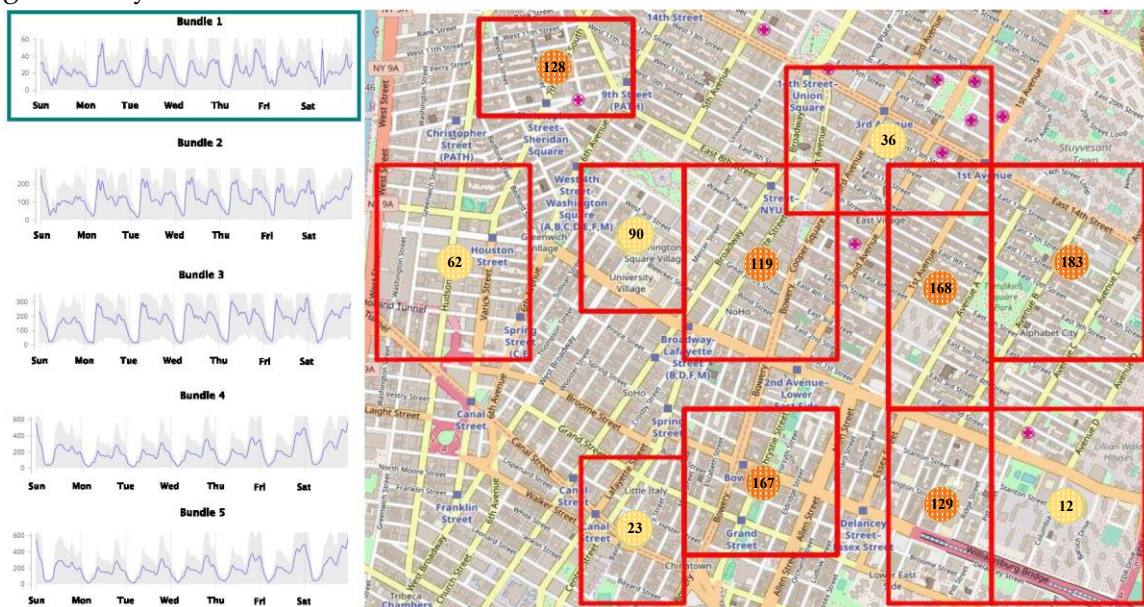


*New (emerging) research directions in this area include:*

- **Heterogeneous Information Networks**

  Heterogeneous Information Networks (HINs) are graphs comprising different types of nodes (entities) and edges (relationships). HINs offer an intuitive and generic model for representing complex information in various domains. A core concept for analyzing HINs is that of metapath, which is a path defined on the schema of the HIN. Metapaths represent relationships of different semantics between entities of the same or different type, providing a mechanism for exploring and analyzing a HIN from multiple perspectives. Thus, they are fundamental for several types of analyses in HINs, ranging from similarity joins to HIN embeddings and recommendations. These tasks become even more complex in the presence of entity types that are additionally associated with spatial or temporal properties. Spatial and temporal proximity are important factors in several analyses. Yet, since spatial and temporal relationships are typically not represented explicitly in the network structure, they cannot be captured by metapaths. Thus, analysis methods that are purely metapath-based will inevitably overlook these aspects, resulting in significant loss of information that is present in the data. Our works aims at finding the top-k most similar entities to a query with respect to different weighted combinations of both metapaths and/or spatio-temporal attributes.
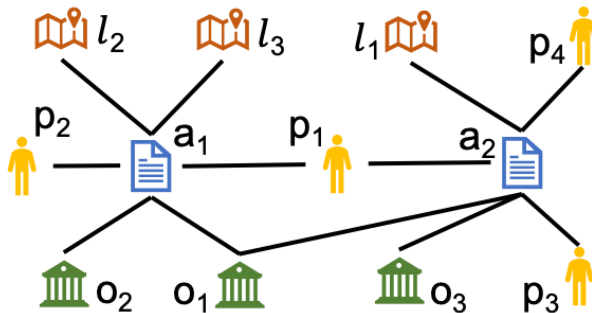
  

- **Extracting value from (underutilized) Big Geospatial Data via Deep Learning**

  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 modelling 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.

- **European Open Science Cloud.**
  The European Open Science Cloud (EOSC) is an environment for hosting and processing research data to support EU science, to provide European researchers, innovators, companies and citizens with a federated and open multi-disciplinary environment where they can publish, find and re-use data, tools and services for research, innovation and educational purposes. IMSI participates in the development of Core Services of the EOSC ecosystem, and more specifically it develops the catalogue of services of the EOSC portal, that offers research-infrastructures and commercial service providers to integrate and list their research services in uniform way in the EOSC portal.

# Projects

## EU R&D Projects
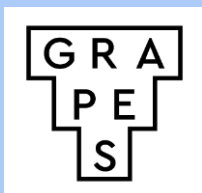
**GRAPES -  Learning, processing and optimising shapes**

| | |
|---|---|
| **Project manager** | Ioannis Z. Emiris |
| **Coordinator** | IMIS |
| **IMIS - funding** | 486,035 euros |
| **Programme** | Marie Sklodowska-Curie Innovative Training Networks |
| **Start date** | 1/11/2019 |
| **Duration** | 4,5 years |
| **Website** | http://grapes-network.eu/ |

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.

# XMANAI - Explainable Manufacturing Artificial Intelligence

| | |
|---|---|
| **Project manager** | Theodore Dalamagas |
| **Coordinator** | IMSI |
| **IMIS - funding** | 318,000 euros |
| **Programme** | H2020-ICT-2018-20, ICT-38-2020 |
| **Start date** | 11/2020 |
| **Duration** | 40 months |
| **Website** | https://ai4manufacturing.eu |

Despite the indisputable benefits of AI, humans typically have little visibility and knowledge on how AI systems make any decisions or predictions due to the so-called "black-box effect" in which many of the machine learning/deep learning algorithms are not able to be examined after their execution to understand specifically how and why a decision has been made. The inner workings of machine learning and deep learning are not exactly transparent, and as algorithms become more complicated, fears of undetected bias, mistakes, and miscomprehensions creeping into decision making, naturally grow among manufacturers and practically any stakeholder In this context, Explainable AI (XAI) is today an emerging field that aims to address how black box decisions of AI systems are made, inspecting and attempting to understand the steps and models involved in decision making to increase human trust. XMANAI aims at placing the indisputable power of Explainable AI at the service of manufacturing and human progress, carving out a "human-centric", trustful approach that is respectful of European values and principles, and adopting the mentality that "our AI is only as good as we are". XMANAI, demonstrated in 4 real-life manufacturing cases, will help the manufacturing value chain to shift towards the amplifying AI era by coupling (hybrid and graph) AI "glass box" models that are explainable to a "human-in-the-loop" and produce value-based explanations, with complex AI assets (data and models) management-sharing-security technologies to multiply the latent data value in a trusted manner, and targeted manufacturing apps to solve concrete manufacturing problems with high impact.

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

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

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**

| | |
|---|---|
| **Project manager** | prof Yannis Ioannidis |
| **Coordinator** | UNIPI |
| **IMIS - funding** | 347.208,75 € |
| **Programme** | H2020-RUR-2018-2020 (Rural Renaissance) |

| | |
|---|---|
| **Start date** | 01/06/2019 |
| **Duration** | 4 years |
| **Website** | https://desira2020.eu/ |

The project aims to improve the capacity of society and political bodies to respond to the challenges that digitalisation generates in agriculture, forestry and rural areas. The project approaches all activities considering digitalisation as a process of social transformation driven by digital technologies. It acknowledges the transformative nature of digitalisation and the impacts it can have in shaping the way rural communities and actors learn, work, travel, interact, etc., acting as a 'game changer' for rural territories and sectors. The digital transformation of rural areas generates winners (who benefit from the change), but also losers (who are marginalised by the change), as well as opponents (who resist to change) and proponents (who support or advocate for the change). To reap the benefits and reduce the risks associated with digitalisation, rural communities need to improve their capacity of understanding the impacts and changes in their context and to develop and plan appropriate actions in order to adapt. Sustainable digitalisation is key to minimise the costs and maximise the benefits of digital transformation, contributing to the Sustainable Development Goals (SDGs). The analytical work is conceptualised on the basis of the interconnection between society, plant and/or animals (Social, beings), data (Cyber, software) and things (Physical, Hardware) within the rural systems or contexts.



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

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

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 Secreteriat- The European Open Science Cloud for Research Secreteriat Project**

| | |
|---|---|
| **Project manager** | prof Yannis Ioannidis |
| **Coordinator** | Technopolis Consulting Group Belgium |
| **IMIS - funding** | 402.500,00 € |
| **Programme** | H2020-INFRAEOSC-2018-2020 (Implementing the European Open Science Cloud) |

| | |
|---|---|
| **Start date** | 01/01/2019 |
| **Duration** | 2,5 years |
| **Website** | https://www.eoscsecretariat.eu/ |

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**

| | |
|---|---|
| **Project manager** | prof Yannis Ioannidis |
| **Coordinator** | UPEM |
| **IMIS - funding** | 141.250,00 € |
| **Programme** | H2020-INFRAIA-2018-1 |

| | |
|---|---|
| **Start date** | 01/01/2019 |
| **Duration** | 4 years |
| **Website** | https://www.risis2.eu/ |

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**

| | |
|---|---|
| **Project manager** | prof Yannis Ioannidis |
| **Coordinator** | IMIS |
| **IMIS - funding** | 122.951,09 € |
| **Programme** | Fisheries and Maritime 2014-2020 |

**Start date**        04/05/2018

**Duration**          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).



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

**Project manager**   prof Yannis Ioannidis

**Coordinator**       PIN

**IMIS - funding**    136.875,00 €

**Programme**         H2020-INFRAIA-2018-1

**Start date**        01/01/2019

**Duration**          4 years

**Website**           https://ariadne-infrastructure.eu/

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.

**HBP SGA3 - Human Brain Project Specific Grant Agreement 3**

| | |
|---|---|
| **Project manager** | prof Yannis Ioannidis |
| **Coordinator** | PIN |
| **IMIS - funding** | 3.764.481,00 € |
| **Programme** | H2020-SGA-FETFLAG-HBP-2019 |
| **Start date** | 01/04/2020 |
| **Duration** | 3 years |
| **Website** | https://www.humanbrainproject.eu/en/ |

The last of four multi-year work plans will take the HBP to the end of its original incarnation as an EU Future and Emerging Technology Flagship. The plan is that the end of the Flagship will see the start of a new, enduring European scientific research infrastructure, EBRAINS, hopefully on the European Strategy Forum on Research Infrastructures (ESFRI) roadmap. The SGA3 work plan builds on the strong scientific foundations laid in the preceding phases, makes structural adaptations to profit from lessons learned along the way (e.g. transforming the previous Subprojects and Co-Design Projects into fewer, stronger, well-integrated Work Packages) and introduces new participants, with additional capabilities. The SGA3 work plan is built around improved integration and a sharpening of focus, to ensure a strong HBP legacy at the end of this last SGA. In previous phases, the HBP laid the foundation for empowering empirical and theoretical neuroscience to approaching the different spatial and temporal scales using state-

of-the-art neuroinformatics, simulation, neuromorphic computing, neurorobotics, as well as high-performance analytics and computing. While these disciplines have been evolving for some years, we now see a convergence in this field and a dramatic speedingup of progress. Data is driving a scientific revolution that relies heavily on computing to analyse data and to provide the results to the research community. Only with strong computer support, is it possible to translate information into knowledge, into a deeper understanding of brain organisation and diseases, and into technological innovation. In this respect, the underlying Fenix HPC and data e-infrastructure, co-designed with the HBP, will be key. The services offered by EBRAINS will be grouped in six Service Categories:

SC1: Curated and shared data: EBRAINS FAIR data services - neuroscience data publishing

SC2: Brain atlas services: navigate the brain in 3D - find, contribute and analyse brain data, based on location

SC3: Brain modelling and simulation workflows: integrated tools to create and investigate models of the brain

SC4: Closed loop AI and robotics workflows: design, test and implement robotic and AI solutions

SC5: Medical brain activity data platform: human intracerebral EEG database and analysis service

SC6: Interactive workflows on HPC or NMC: Europe-wide access to scalable and interactive compute services

**HumanE-AI-Net- HumanE AI Network**

| | |
|---|---|
| **Project manager** | prof Yannis Ioannidis |
| **Coordinator** | DFKI |
| **IMIS - funding** | 200.000,00 € |
| **Programme** | H2020-ICT-2019-3 |
| **Start date** | 01/09/2020 |
| **Duration** | 3 years |
| **Website** | https://www.humane-ai.eu/ |

The HumanE AI Net brings together top European research centers, universities and key industrial champions into a network of centers of excellence that goes beyond a narrow definition of AI and combines world leading AI competence with key players in related areas such as HCI, cognitive science, social sciences and complexity science. This is crucial to develop a truly Human Centric brand of European AI. We will leverage the synergies between the involved centers of excellence to develop the scientific foundations and technological breakthroughs needed to shape the AI revolution in a direction that is beneficial to humans both individually and societally, and adheres to European ethical values and social, cultural, legal, and political norms. The core challenge is the development of robust, trustworthy AI capable of what "understanding" humans, adapting to complex real-world environments, and appropriately interacting in complex social settings. The aim is to facilitate AI systems that enhance human capabilities and empower individuals and society as a whole while respecting human autonomy and self-determination. The HumanE AI Net project will engender the mobilization of a research landscape far beyond direct project funding, involve and engage European industry, reach out to relevant social stakeholders, and create a unique innovation ecosystem that provides a many fold return on investment for the European economy and society. We will make the results of the research available to the European AI community through the AI4EU platform and a Virtual Laboratory, develop a series of summer schools, tutorials and MOOCs to spread the knowledge, develop a dedicated innovation ecosystem for transforming research and innovation into an economic impact and value for society, establish an industrial Ph.D. program and involve key industrial players from sectors crucial to European economy in research agenda definition and results evaluation in relevant use cases.



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

| | |
|---|---|
| **Project manager** | Dimitris Skoutas |
| **Coordinator** | IMIS |
| **IMIS - funding** | € 853,125.00 |
| **Programme** | RIA, H2020-ICT-2018-2 |
| **Start date** | 1/1/2019 |
| **Duration** | 3 years |
| **Website** | https://smartdatalake.eu/ |

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**

| | |
|---|---|
| **Project manager** | Georgia Koutrika |
| **Coordinator** | ZHAW |
| **IMIS - funding** | 798.000 Euros |
| **Programme** | EU H2020 - H2020-EU.1.4.1.3. - Development, deployment and operation of ICT-based e-infrastructures, inode-project.eu |
| **Start date** | 01/11/2019 |
| **Duration** | 3.5 years |
| **Website** | https://www.inode-project.eu/ |

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**

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

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.

## LOCARD

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

| | |
|---|---|
| **Project manager** | Constantinos Patsakis |
| **Coordinator** | Athena Research Center |
| **IMIS - funding** | 700.000 |
| **Programme** | SU-FCT02 |
| **Start date** | 1/5/2019 |
| **Duration** | 36 months |
| **Website** | https://locard.eu/ |

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 crowdsource 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

| | |
|---|---|
| **Project manager** | Georgia Tsiliki |
| **Coordinator** | Heriot-Watt University |
| **IMIS - funding** | 124 037.5 Euros |
| **Programme** | H2020-NMBP-2017-two-stage |
| **Start date** | 1/1/2018 |
| **Duration** | 3.5 years |
| **Website** | https://www.h2020gracious.eu |

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**

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

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)**

| | |
|---|---|
| **Project manager** | George Papastefanatos |
| **Coordinator** | National Kapodistrian University of Athens |

| | |
|---|---|
| **IMIS - funding** | 470.625 Euros |
| **Programme** | H2020-EU.1.4.1.3 |
| **Start date** | 01/11/2019 |
| **Duration** | 3 years |
| **Website** | https://www.neanias.eu/ |

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**

| | |
|---|---|
| **Project coordinator** | George Papastefanatos |
| **Coordinator** | University of Greenwich |
| **IMSI - funding** | - |
| **Programme** | COST ACTION |
| **Start date** | 23/7/2017 |
| **Duration** | 4 years |
| **Website** | https://wise-act.eu/ |

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.



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

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

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.

**ELIXIR CONTAINERS2 - Making container services integratable, sustainable and widely adopted**

| | |
|---|---|
| **Project manager** | Thanasis Vergoulis |
| **Coordinator** | EMBL |
| **IMIS - funding** | € 3,231.25 |
| **Programme** | Strategic Implementation Study (ELIXIR Commissioned Services) |
| **Start date** | 31/8/2021 |
| **Duration** | 24 months |
| **Website** | https://www.imsi.athenarc.gr/en/projects/project/76 |

The aim of this project is to build on the current progress made through the ELIXIR-CONTAINERS project to enable adoption and deployment of protocols and services by the broader ELIXIR community at scale. It aims to coordinate existing efforts across ELIXIR, identify opportunities, contribute in a targeted and limited way with specific developments to connect relevant components and propose mechanisms for sustaining this effort over time.

**ELIXIR-CONVERGE: Development and long-term sustainability of new pan-European research infrastructures**

| | |
|---|---|
| **Project manager** | Thanasis Vergoulis |
| **Coordinator** | EMBL |
| **IMIS - funding** | € 50,047.50 |
| **Programme** | H2020-INFRADEV-2018-2020 (RIA) |
| **Start date** | 1/2/2020 |
| **Duration** | 43 months |
| **Website** | https://elixir-europe.org/about-us/how-funded/eu-projects/converge |

ELIXIR-CONVERGE is a project funded by the European Commission to help standardise life science data management across Europe. To achieve this standardisation, the project will develop a data management toolkit for life scientists. The toolkit will help ensure more research data is in the public domain, which will give scientists access to more data. This will allow them to discover new insights into the challenges facing society, such as food security and health in old age, and help stimulate innovation in biomedicine and biotechnology.

| | **OpenAIRE Nexus- OpenAIRE-Nexus Scholarly Communication Services for EOSC users** |
|---|---|

| | |
|---|---|
| **Project manager** | prof Yannis Ioannidis |
| **Coordinator** | OPENAIRE AMKE |
| **IMIS - funding** | 740.442,00 € |
| **Programme** | H2020-INFRAEOSC-2020-2 |
| **Start date** | 01/01/2021 |
| **Duration** | 2,5 years |
| **Website** | https://www.openaire.eu/openaire-nexus-project |

OpenAIRE-Nexus brings in Europe, EOSC and the world a set of services to implement and accelerate Open Science. To embed in researchers workflows, making it easier for them to accept and uptake Open Science practices of openness and FAIRness. To give the tools to libraries, research communities to make their content more visible and discoverable. To assist policy makers to better understand the environment and ramifications of Open Science into new incentives, scientific reward criteria, impact indicators, so as to increase research and innovation potential. To foster innovation, by providing SMEs with open data about scientific production. To this aim, OpenAIRE-Nexus onboards to the EOSC fourteen services, provided by public institutions, einfrastructures, and companies, structured in three portfolios: PUBLISH, MONITOR and DISCOVER. The services are widely used in Europe and beyond and integrated in OpenAIRE-Nexus to assemble a uniform Open Science Scholarly Communication package for the EOSC. The project aims at forming synergies with other INFRAEOSC-07 awarded projects, the INFRAEOSC-03 project, research infrastructures, einfrastructures, and scholarly communication services define a common Open Science interoperability framework for the EOSC, to facilitate sharing, monitoring, and discovery of EOSC resources across disciplines.

**EOSC Future**

| | |
|---|---|
| **Project manager** | prof Yannis Ioannidis |
| **Coordinator** | IMIS |
| **IMIS - funding** | 1.026.250,00 € |
| **Programme** | H2020-INFRAEOSC-2020-2 |
| **Start date** | 01/04/2021 |
| **Duration** | 2,5 years |
| **Website** | https://eoscfuture.eu/ |

is an EU-funded H2020 project that is implementing the European Open Science Cloud (EOSC). EOSC will give European researchers access to a wide web of FAIR data and related services. EOSC Future responds to INFRAEOSC-03-2020 call in order to integrate, consolidate, and connect e-infrastructures, research communities, and initiatives in Open Science to further develop the EOSC Portal, EOSC-Core and EOSCExchange of the European Open Science Cloud (EOSC). EOSC Future is structured around six thematic pillars that strategically group the work package tasks and activities and coherently present the key objectives and activities of the project to stakeholders: Pillar 1 on Policy & Strategy will coordinate and support a strategic vision for the future EOSC; Pillar 2 on Connection & Integration will connect and integrate EOSC infrastructures, data, and services; Pillar 3 on Excellent Science & Interdisciplinarity will extend the EOSC value Chain with scientific use cases; Pillar 4 on Growth & Innovation will grow and innovate EOSC with value-added services; Pillar 5 on Skills & Training will train users and providers of the EOSC ecosystem; and Pillar 6 on Engagement & Communications will engage the wider EOSC Community at a global level.

**4CH - Competence Centre for the Conservation of Cultural Heritage**

| | |
|---|---|
| **Project coordinator** | Panos Constantopoulos |
| **Coordinator** | ISTITUTO NAZIONALE DI FISICA NUCLEARE |
| **IMSI - funding** | 84.375 euros |
| **Programme** | H2020-SC6-TRANSFORMATIONS-2018-2019-2020 |
| **Start date** | 1/1/2021 |
| **Duration** | 36 months |
| **Website** | https://www.4ch-project.eu/ |

The project aims to set up the methodological, procedural, and organizational framework of a Competence Centre, an infrastructure dedicated to knowledge organization and transfer through means such as training, standardization and inter-disciplinary collaboration. This European Competence Centre will need to adapt to many different conditions such as how cultural heritage is managed, the risks that may affect its conservation and how it may be valorised. Using a holistic and multidisciplinary approach to the conservation of cultural heritage, it will facilitate coordination between cultural heritage Institutions across Europe and provide services and tools to enable preservation and conservation of historical monuments and sites using the latest, most effective technologies with special attention to 3D. 4CH will benefit a range of institutions and other bodies, both public and private, responsible for managing European CH, service providers (CH professionals and SMEs), the creative industries and hospitality sector as well as heritage agencies, public bodies such as ministries and decision-makers who inform policies and strategies for conservation, preservation and digitization. The current Consortium includes leading institutions from the academia, industry, SMEs and research centres with complementary expertise and a wide geographic coverage of Europe who are guided by an Advisory Board consisting of high-level experts.

## National R&D Projects

**i4metal - Innovative Data Science Technologies for Scrap Valorization**

| | |
|---|---|
| **Project manager** | Theodore Dalamagas |
| **Coordinator** | IMSI |
| **IMIS - funding** | 250,000 euros |
| **Programme** | RESEARCH - CREATE - INNOVATE, Operational Programme Competitiveness, Entrepreneurship and Innovation 2014-2020 |
| **Start date** | 6/2020 |
| **Duration** | 30 months |
| **Website** | https://i4metal.gr |

The i4metal project will design and develop innovative ICT tools for organizing, processing and analyzing the data being collected during the operation of scrap processing facilities. Scrap (recyclable materials left over from product manufacturing and consumption, such as parts of vehicles, building supplies, and surplus materials) are an important source of secondary raw materials, with a key role in successfully transitioning to Circular Economy. These ICT tools, developed by the i4metal project, will support the creation of a data value chain in the scrap processing cycle to: (a) improve quality control of these materials; (b) increase accuracy standardization of scrap on the basis of their qualitative characteristics; (c) identifying and limiting cases of malicious practices; (d) optimizing the operation of scrap valorization processes; and (e) distributing open data regarding the collected scrap and the extracted secondary raw materials.

**URBANA**

**URBANA - Innovative Data Science Technologies for Scrap Valorization**

| | |
|---|---|
| **Project manager** | Theodore Dalamagas |
| **Coordinator** | IMSI |
| **IMIS - funding** | 143,000 euros |

| | |
|---|---|
| **Programme** | RESEARCH - CREATE - INNOVATE, Operational Programme Competitiveness, Entrepreneurship and Innovation 2014-2020 |
| **Start date** | 5/2020 |
| **Duration** | 30 months |
| **Website** | http://www.imsi.athenarc.gr/en/projects/project/69 |

In recent decades, the population percentage in most of the developed EU countries has been rising constantly, which has a negative impact both on the natural and productive resources of the urban centers and large provincial cities. As a consequence, the failure of the aforementioned to effectively respond to the demographic reality affects not only the urban landscape quality but also the environment in general, as the management of urban waste and sewage disposal is becoming increasingly difficult and the quality of air and water resources is constantly deteriorating and on the other hand the development of the economy, as insufficient employment opportunities can be created, especially for disadvantaged groups. The proposed project (URBANA) will create a technologically innovative platform to build a bridge of knowledge exchange between Agricultural Advisors and Urban Cultivators, fostering social innovation and collective awareness of the sustainability of cities. In particular, through URBANA, users will be able to present small-scale urban cultivation as well as other activities that they have implemented or implement on urban agriculture in the form of Do It Yourself (DIY) projects, while at the same time they will be able to make use of information and advice collected from crowdsourcing, IoT and social networks.



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

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

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

| | |
|---|---|
| **Project coordinator** | Alkis Simitsis |
| **Coordinator** | National Observatory of Athens |
| **IMSI - funding** | - |
| **Programme** | GSRT project |
| **Start date** | 2019 |
| **Duration** | 3 years |
| **Website** | - |

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.



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

| | |
|---|---|
| **Project manager** | George Papastefanatos |
| **Coordinator** | IMSI - Athena RC |
| **IMIS - funding** | 172,000 euros |
| **Programme** | 1st Call for Postdoc Research Projects - Hellenic Foundation for Research and Innovation (ELIDEK) |
| **Start date** | 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.

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

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

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.



**«Προσεγγιστικοί γεωμετρικοί αλγόριθμοι και συσταδοποίηση με εφαρμογές στα χρηματοοικονομικά»**

| | |
|---|---|
| **Project manager** | Ioannis Z. Emiris |
| **Coordinator** | IMIS |
| **IMIS - funding** | 41,041 euros |
| **Programme** | Υποστήριξη ερευνητών με έμφαση στους νέους ερευνητές – κύκλος Β |
| **Start date** | 01-03-2020 |
| **Duration** | 16 Months |

**Website**

Αντικείμενο της δράσης είναι η υλοποίηση της ερευνητικής πρότασης «Προσεγγιστικοί γεωμετρικοί αλγόριθμοι και συσταδοποίηση με εφαρμογές στα χρηματοοικονομικά». Στόχος της παρέμβασης είναι η ενίσχυση του ακαδημαϊκού βιογραφικού των ερευνητών και των ερευνητικών δεξιοτήτων τους, προκειμένου να βελτιωθούν οι προοπτικές ακαδημαϊκής / ερευνητικής καριέρας τους.

## Other Projects

|  | Europeana DSI-4 |
|---|---|

| | |
|---|---|
| **Project coordinator** | Dimitris Gavrilis |
| **Coordinator** | STICHTING EUROPEANA(EF) |
| **IMSI - funding** | 97,372 euros |
| **Programme** | Service Contract-"Deployment and Maintenance of Europeana DSI Core Services-SMART-2017/1136" |
| **Start date** | 1/9/2018 |
| **Duration** | 24 months |
| **Website** | https://pro.europeana.eu/project/europeana-dsi-4 |

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.

## Industry – Funded R&D Projects

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

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

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 the NFV-RI

| | |
|---|---|
| **Project manager** | George Papastefanatos |
| **Coordinator** | Intracom Telecom S.A. |
| **IMIS - funding** | 13,020 euros |
| **Programme** | Contract |
| **Start date** | 01/03/2021 |
| **Duration** | 3 months |
| **Website** | - |

The project concerns the provision of consulting services, design and

development services for the redesign of the company's NFV-RI product with emphasis on the easy visualization of metric performance and resource optimization scenarios that perform Network Virtualized Functions within Cloud Environment.

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

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

The project eLib aims at developing a digital library for the Independent Autority ofor 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.

### GeCoInt – Geometric Computation for the design of Integrated Circuit

| | |
|---|---|
| **Project manager** | Ioannis Z. Emiris |
| **Coordinator** | IMIS (in collaboration with ILSP) |

| | |
|---|---|
| **IMIS - funding** | 21,000 euros |
| **Programme** | ANSYS S.A. |
| **Start date** | 01-01-2020 |
| **Duration** | 12 Months |
| **Website** | https://www.imsi.athenarc.gr/el/projects/comproject/30 |

A joint project between Athena RC and ANSYS Hellas, that is focused on: spatial indexing data structures for facilitating region queries in the context of random walk-based capacitance extraction, enhancement, and extension of the solutions for optimal insertion order into the spatial index and lastly, evaluation of data-driven data structure for MEC queries. Our 5 collaborators (Master students, PhD candidates and Post-Docs) shall benefit from both top-notch research as well as a strong innovation component through a nexus of intersectoral secondments and the interaction with the highly qualified ANSYS Hellas R&D personnel.



**ΔΡΟΜΟ – Αυτόματη Δρομολόγηση βυτιοφόρων οχημάτων μεταφοράς καυσίμων**

| | |
|---|---|
| **Project manager** | Ioannis Z. Emiris |
| **Coordinator** | IMIS |
| **IMIS - funding** | 12,100 euros |
| **Programme** | EMDOT A.E. |
| **Start date** | 15-04-2021 |
| **Duration** | 2,5 Months |
| **Website** | https://www.imsi.athenarc.gr/el/projects/comproject/32 |

Στο έργο θα σχεδιαστούν αλγόριθμοι και θα αναπτυχθεί λογισμικό για την βέλτιστη δρομολόγηση στόλου βυτιοφόρων οχημάτων με σκοπό την όσο το δυνατόν οικονομικότερη μεταφορά παραγγελιών καυσίμων στα πρατήρια της επικράτειας, μεγιστοποιώντας ταυτόχρονα την χρήση των διαμερισμάτων των οχημάτων.

# Publications

## Book chapters

- Kanellos, T. Vergoulis, D. Sacharidis: "Ranking papers by expected short-term impact" Predicting the Dynamics of Research Impact (Springer, 2021)

## Journal Publications

- Paolo Manghi, Andrea Mannocci, Francesco Osborne, Dimitris Sacharidis, Angelo A. Salatino, Thanasis Vergoulis: New trends in scientific knowledge graphs and research impact assessment. Quant. Sci. Stud. 2(4): 1296-1300 (2021)
- Thanasis Vergoulis, Ilias Kanellos, Serafeim Chatzopoulos, Danae Pla Karidi, Theodore Dalamagas: BIP4COVID19: Releasing impact measures for articles relevant to COVID-19. Quant. Sci. Stud. 2(4): 1447-1465 (2021)
- Serafeim Chatzopoulos, Thanasis Vergoulis, Ilias Kanellos, Theodore Dalamagas, Christos Tryfonopoulos: Further improvements on estimating the popularity of recently published papers. Quant. Sci. Stud. 2(4): 1529-1550 (2021)
- Ilias Kanellos, Thanasis Vergoulis, Dimitris Sacharidis, Theodore Dalamagas, Yannis Vassiliou: Impact-Based Ranking of Scientific Publications: A Survey and Experimental Evaluation. IEEE Trans. Knowl. Data Eng. 33(4): 1567-1584 (2021)
- Evangelos Bartzos, Ioannis Z. Emiris, Jan Legerský, Elias P. Tsigaridas: On the maximal number of real embeddings of minimally rigid graphs in $R^2$, $R^3$ and $S^2$. J. Symb. Comput. 102: 189-208 (2021), https://doi.org/10.1016/j.jsc.2019.10.015
- Ioannis Z. Emiris, Angelos Mantzaflaris, Elias P. Tsigaridas: Multilinear polynomial systems: Root isolation and bit complexity. J. Symb. Comput. 105: 145-164 (2021), https://doi.org/10.1016/j.jsc.2020.06.005.
- Theodora Galani, George Papastefanatos, Yannis Stavrakas and Yannis Vassiliou: **Defining and Detecting Complex Changes on RDF(S) Knowledge Bases**. In Journal on Data Semantics (2021). https://doi.org/10.1007/s13740-021-00136-9
- Nikos Bikakis, Stavros Maroulis, George Papastefanatos, Panos Vassiliadis: **In-Situ Visual Exploration over Big Raw Data**. In Inf. Syst. 95: 101616 (2021)
- Theofilos Mailis, Yannis Kotidis, Stamatis Christoforidis, Evgeny Kharlamov, Yannis Ioannidis: "View selection over knowledge graphs in triple stores", Proceedings of the VLDB Endowment, Vol. 14, pp. 3281-3294, September 2021.
- K. El Raheb, M. Stergiou, A. Katifori, Y. Ioannidis: "Moving in the cube: A motion-based playful experience for introducing Labanotation to beginners", ACM Journal on Computing and Cultural Heritage (JOCCH), Vol. 14, No 2, pp. 1-26 June 2021.

- K. El Raheb, Y. Ioannidis: "Annotating the captured dance: reflections on the role of tool-creation", Int'l Journal of Performance Arts and Digital Media, Vol. 17, No. 1, pp. 118-137, May 2021.

- Maria Vayanou, Akrivi Katifori, Yannis Ioannidis: "Perspective Sharing in Culture Group Games: Passing Around the Social Mediator Role", Proceedings of the ACM on Human-Computer Interaction, Vol. CHI PLAY, pp. 1-24, June 2021.

- Angeliki Antoniou, Maria Vayanou, Akrivi Katifori, Angeliki Chrysanthi, Filippia Cheilitsi, Yannis Ioannidis: "Real Change Comes from Within!": Towards a Symbiosis of Human and Digital Guides in the Museum", ACM Journal on Computing and Cultural Heritage (JOCCH), Vol. 15, pp. 1-19, December 2021.

- Manolis Koubarakis, Alexander Borgida, Panos Constantopoulos et al.: **A retrospective on Telos as a metamodeling language for requirements engineering**. Requirements Eng 26, 1–23 (2021). https://doi.org/10.1007/s00766-020-00329-x

- Agiatis Benardou, Michalis Kalamaras, Magdalini Sgouridi and Giannis Tsakonas: **The vision of an open, transferable, hybrid reading hub in urban Greece**. *Journal of Greek Media and Culture* 7(1), 125-136 (2021).

- Basei, G., Zabeo, A., Rasmussen, K., Tsiliki, G. and Hristozov, D., 2021. A Weight of Evidence approach to classify nanomaterials according to the EU Classification, Labelling and Packaging Regulation criteria. NanoImpact, 24, p.100359.

- Zabeo, A., Basei, G., Tsiliki, G., Peijnenburg, W. and Hristozov, D., 2022. Ordered weighted average based grouping of nanomaterials with Arsinh and dose response similarity models. NanoImpact, 25, p.100370.

- Jeliazkova, N., Bleeker, E., Cross, R., Haase, A., Janer, G., Peijnenburg, W., Pink, M., Rauscher, H., Svendsen, C., Tsiliki, G. and Zabeo, A., 2022. How can we justify grouping of nanoforms for hazard assessment? Concepts and tools to quantify similarity. NanoImpact, 25, p.100366.

- Antonis Mandamadiotis, Georgia Koutrika, Stavroula Eleftherakis, Apostolis Glenis, Dimitrios Skoutas, and Yannis Stavrakas. DatAgent: The Imminent Age of Intelligent Data Assistants. Proceedings of VLDB Endowment 14(12): 2815-2818, 2021.

- Kostis Pristouris, Harry Nakos, Yannis Stavrakas, Konstantinos I. Kotsopoulos, Theofanis Alexandridis, Myrto S. Barda, and Konstantinos P. Ferentinos. An Integrated System for Urban Parks Touring and Management. Urban Science, 5(4), 91, 2021.

## International Conference / Workshop Publications

- A. Mandamadiotis, S. Eleftherakis, A. Glenis, D. Skoutas, Y. Stavrakas, G. Koutrika. DatAgent: The Imminent Age of Intelligent Data Assistants. PVLDB, 14(12): 2815 – 2818

- A. Glenis, G. Koutrika. PyExplore: Query Recommendations for Data Exploration without Query Logs. ACM SIGMOD.

- O. Gkini, T. Belmpas, G. Koutrika, Y. Ioannidis: "An In-Depth Benchmarking of Text-to-SQL Systems", *Proc. 2021 Int'l ACM SIGMOD Conference*, Beijing, China, pp. 632-644, June 2021

- S. Chatzopoulos, T. Vergoulis, T. Dalamagas, C. Tryfonopoulos: VeTo-web: A Recommendation Tool for the Expansion of Sets of Scholars. JCDL 2021 (demo)

- G. Mouchakis, C. Kostopoulos, S. Konstantopoulos, I. Kanellos, A. Tzerefos, T. Vergoulis, T. Dalamagas: A Cloud-Native NGS Data Processing and Annotation Platform. VLDB DMAH 2021.

- S. Chatzopoulos, T. Vergoulis, P. Deligiannis, D. Skoutas, T. Dalamagas, C. Tryfonopoulos: SciNeM: A Scalable Data Science Tool for Heterogeneous Network Mining. EDBT 2021: 654-657

- T. Vergoulis, I. Kanellos, G. Giannopoulos, T. Dalamagas: Simplifying Impact Prediction for Scientific Articles. EDBT/ICDT Workshops 2021

- K. Zagganas, M. Lioli, T. Vergoulis, T. Dalamagas: Simplifying p-value calculation for the unbiased microRNA enrichment analysis, using ML-techniques. EDBT/ICDT Workshops 2021

- I. Kanellos, T. Vergoulis, D. Sacharidis, T. Dalamagas, Y. Vassiliou: Ranking Papers by their Short-Term Scientific Impact. ICDE 2021: 1997-2002

- T. Vergoulis, K. Zagganas, L. Kavouras, M. Reczko, S. Sartzetakis, T. Dalamagas: SCHeMa: Scheduling Scientific Containers on a Cluster of Heterogeneous Machines. SSDBM 2021: 243-247

- T. Vergoulis, I. Kanellos, C. Atzori, A. Mannocci, S. Chatzopoulos, S. La Bruzzo, N. Manola, P. Manghi: BIP! DB: A Dataset of Impact Measures for Scientific Publications. WWW (Companion Volume) 2021: 456-460

- P. Deligiannis, T. Vergoulis, S. Chatzopoulos, C. Tryfonopoulos: Visualising Scientific Topic Evolution. WWW (Companion Volume) 2021: 468-472

- Antonios Deligiannakis, Nikos Giatrakos, Yannis Kotidis, Vasilis Samoladas, Alkis Simitsis: Extreme-Scale Interactive Cross-Platform Streaming Analytics - The INFORE Approach. SEA-Data@VLDB 2021: 7-13

- Antonios Kontaxakis, Antonios Deligiannakis, Holger Arndt, Stefan Burkard, Claus-Peter Kettner, Elke Pelikan, Kathleen Noack: Real-time processing of geo-distributed financial data. DEBS 2021: 190-191

- Eleni Kougioumtzi, Antonios Kontaxakis, Antonios Deligiannakis, Yannis Kotidis: Towards creating a generalized complex event processing operator using FlinkCEP: architecture & benchmark. DEBS 2021: 188-189

- Nikos Giatrakos, Eleni Kougioumtzi, Antonios Kontaxakis, Antonios Deligiannakis, Yannis Kotidis: EasyFlinkCEP: Big Event Data Analytics for Everyone. CIKM 2021: 3029-303

- Marios Vodas, Konstantina Bereta, Dimitris Kladis, Dimitris Zissis, Elias Alevizos, Emmanouil Ntoulias, Alexander Artikis, Antonios Deligiannakis, Antonios Kontaxakis, Nikos Giatrakos, David Arnu, Edwin Yaqub, Fabian Temme, Mate

Torok, Ralf Klinkenberg: Online Distributed Maritime Event Detection & Forecasting over Big Vessel Tracking Data. IEEE BigData 2021: 2052-2057

- Graham Cormode, Minos N. Garofalakis, Michael Shekelyan: Data-Independent Space Partitionings for Summaries. PODS 2021: 285-298.

- Evangelos Bartzos, Ioannis Z. Emiris, Charalambos Tzamos: The m-Bézout Bound and Distance Geometry. CASC 2021: 6-20, https://doi.org/10.1007/978-3-030-85165-1_2

- Apostolos Chalkis, Emmanouil Christoforou, Theodore Dalamagas, Ioannis Z. Emiris: Modeling of Crisis Periods in Stock Markets. LION 2021: 55-65, https://doi.org/10.1007/978-3-030-92121-7_5

- Theodora Galani, Yannis Stavrakas, George Papastefanatos, Yannis Vassiliou: **Evo-Path: Querying Data Evolution Through Complex Changes**. In 10th International Conference on Data Science, Technology and Applications (DATA21). July 6-8, 2021, Online Event.

- Harry Nakos, Kostis Pristouris, and Yannis Stavrakas. **Assessing Visitor Activity in Urban Parks**. Urban Complex Systems 2021 (Conference on Complex Systems workshop), Lyon, France, October 26, 2021.

- Stavros Maroulis, Nikos Bikakis, George Papastefanatos, Panos Vassiliadis, Yannis Vassiliou: **RawVis: A System for Efficient In-situ Visual Analytics**. In ACM International Conference on Management of Data (SIGMOD/PODS 2021), Xi'an, Shaanxi, China, June 20 - 25, 2021 (Demo Paper)

- Giorgos Giannopoulos, George Papastefanatos, Dimitris Sacharidis, Kostas Stefanidis: **Interactivity, Fairness and Explanations in Recommendations**. In 6th International Workshop on Human Aspects in Adaptive and Personalized Interactive Environments (HAAPIE'21) in conjunction with ACM UMAP 2021, June 21-25, 2021, Utrecht, The Netherlands.

- Stavros Maroulis, Nikos Bikakis, George Papastefanatos, Panos Vassiliadis, Yannis Vassiliou: **Adaptive Indexing for In-situ Visual Exploration and Analytics**. In Proceedings of the 23rd International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data co-located with EDBT/ICDT 2021 Joint Conference, DOLAP@EDBT/ICDT 2021, Nicosia, Cyprus.

- Petros Galanis, Olympia Konstantakopoulou, Iliana Karagkouni, Parisis Gallos, Evangelos Tsampalas, Maria Lypiridou, Anastasia Gamvroula, Georgios Mavraganis, Efstathios Manios, George Ntaios, Efstathia Karagkiozi, Haralampos J. Milionis, Chara Evangelou, Argyro Tountopoulou, Ioanna Kouzi, Sofia Vassilopoulou, Athanasios Protogerou, Stamatina Samara, Asteria Karapiperi, Christos Savopoulos, Apostolos I. Hatzitolios, Athina Myrou, Ioannis Kalliontzakis, Aspasia Kouridaki, Stefanos Papastefanatos, George Papastefanatos, Kostas Vemmos, Panayiota Sourtzi, Thalia Bellali, Eleni Korompoki, Daphne Kaitelidou, Olga Siskou: **Assessing Staff's and Stroke Patients' Experiences in 8 Hospitals in Greece: Results from a Prospective Multi-Center Study ("SUN4Patients")**. In 20th

International Conference on Informatics, Management, and Technology in Healthcare (ICIMTH 2021): 392-396

- D. Petousi, A. Katifori, S. McKinney, S. Perry, M. Roussou, Y. Ioannidis: "Social bots of conviction as dialogue facilitators for history education: Promoting historical empathy in teens through dialogue", Proc. 20th Interaction Design and Children Conference, Athens, Greece, pp. 326-337, June 2021.
- Y. Foufoulas, L. Sidirourgos, E. Stamatogiannakis, Y. Ioannidis: "Adaptive Compression for Fast Scans on String Columns", Proc. 2021 Int'l ACM SIGMOD Conference, Beijing, China, pp. 554-562, June 2021.
- Ektor Vrettakis, Akrivi Katifori, Yannis Ioannidis: "Digital Storytelling and Social Interaction in Cultural Heritage-An Approach for Sites with Reduced Connectivity", Springer, Cham, pp. 157-171, December 2021.


## National Conference / Workshop Publications

- S. Eleftherakis, O. Gkini, G. Koutrika. Let the Database Talk Back: Natural Language Explanations for SQL. SEADATA@VLDB2021
- Eleni Lavasa, Giorgos Giannopoulos, Athanasios Papaioannou, Anastasios Anastasiadis, et al. **Assessing the Predictability of Solar Energetic Particles with the Use of Machine Learning Techniques**. Solar Physics volume 296, Article number: 107 (2021) (**Editors' Choice**)
- Stella Girtsou, Alexis Apostolakis, Giorgos Giannopoulos, Charalampos Kontoes: **A Machine Learning Methodology for Next Day Wildfire Prediction**. IGARSS 2021: 8487-8490
- Vassilis Kaffes, Dimitris Sacharidis, Giorgos Giannopoulos: **Model-Agnostic Counterfactual Explanations of Recommendations**. UMAP 2021: 280-285
- Giorgos Giannopoulos, George Papastefanatos, Dimitris Sacharidis, Kostas Stefanidis: **Interactivity, Fairness and Explanations in Recommendations**. UMAP (Adjunct Publication) 2021: 157-161
- Thanasis Vergoulis, Ilias Kanellos, Giorgos Giannopoulos, Theodore Dalamagas: **Simplifying Impact Prediction for Scientific Articles**. EDBT/ICDT Workshops 2021


## Other Publications / Technical Reports

- Ludovic Calès, Apostolos Chalkis, Ioannis Z. Emiris: The cross-sectional distribution of portfolio returns and applications. CoRR abs/2105.06573 (2021)
- Agiatis Benardou, Vicky Dritsou and Maria Ilvanidou (eds): **DH Goes Viral**. (2021) DOI: 10.5281/zenodo.5793151

# Dissemination Activities

## Invited / Keynote Talks

IMSI members participated in the following invited / keynote talks:

- A. Simitsis. Advanced Optimization Techniques in Infore. Infore meeting, Virtual Event, April 14, 2021.
- G. Papastefanatos. "Sustainability of the EOSC Portal" In Sustaining EOSC Session of the EOSC Symposium June 2021, Virtual Event.

## Scientific Community Service

IMSI members have served in the **Program Committee** of more than 16 International Conferences and Workshops in 2021, including well-known Conferences like ACM SIGMOD, ACM KDD, CIKM, CIDR, and more.

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

- **IMSI co-organized ICDE 2021**. Minos Garofalakis (General chair), George Papastefanatos (Web chair), Yannis Stavrakas (Finance chair), Alkis Simitsis (Registration chair), Georgia Koutrika (Sponsorship chair), Theodore Dalamagas (Local Arrangements chair).
- **ACM SIGMOD Record**. A. Simitsis, Associate Editor, Research Center Column.
- **Data & Knowledge Engineering**, Elsevier. A. Simitsis, Editorial Board Member.
- **ACM DOLAP**: ACM International Workshop on Data Warehousing and OLAP: A. Simitsis, Steering Committee
- VLDB Journal: Georgia Koutrika is co-editor-in-Chief
- Greek ACM-W Chapter: Georgia Koutrika is treasurer
- ACM SIGMOD:  Georgia Koutrika is Associate Information Director
- Thanasis Vergoulis co-organised the Sci-K 2021 international workshop (in conj. With The Web Conf 2021).
- George Papastefanatos co-organized the 5th International Workshop on Big Data Visual Exploration and Analytics (BigVis2021) in conjunction with EDBT/ICDT 2021
- George Papastefanatos served as Guest Editor in Elsevier Big Data Research Journal for the Special Issue on Interactive Big Data Visualization and Analytics.
- Agiatis Benardou served DARIAH-EU as VCC2 (Research and Education Liaison) Co-Chair.

- Agiatis Benardou served as Programme Committee member at the annual DARIAH-EU event, held virtually in September 2021.

# Systems

- **SPHINX**

SPHINX is a system for metapath-based entity exploration in Heterogeneous Information Networks (HINs), developed in the context of the EU funded project SmartDataLake. SPHINX allows users to define different views over a HIN based on both automatically selected and user-defined metapaths. Then, entity ranking and similarity search can be performed over these views to find and explore entities of interest, taking also into account any spatial or temporal properties of entities. A Web-based user interface is provided to facilitate users in performing the various functionalities supported by the system, including metapath-based view definition, index construction, search parameters specification, and visual comparison of the results.

https://www.vldb.org/pvldb/vol13/p2913-chatzopoulos.pdf

## 1. Select source entity type

Article

## 2. Select search preferences

Determine search preferences based on existing indices.

+ Add new index

### i) Preferences on source entities

| Type | Preferences | Weights |
|------|-------------|---------|
| ⬤ Temporal | Select Date | 0   0.25   0.5   0.75   1 |

### ii) Metapath preferences

| Metapath | Preferences on target entities | Weights |
|----------|-------------------------------|---------|
| ⬤ Article - Organisation | twitter × | 0   0.25   0.5   0.75   1 |
| ⬤ Article - Person | donald trump ×   barack obama × | 0   0.25   0.5   0.75   1 |

## 3. Select value for k

50

Q Search

## Results

| # | Article | Score | Article - Organisation | Article - Person |
|---|---------|-------|------------------------|------------------|
| 1 | 20190520_filtered.csv:20190520200000-1330 Q | 0.7400000000000017 | huawei, twitter, cnn, white house, facebook, united states score: 3.3382377953649984e-15 | barack obama, donald trump score: 1 |

- **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/smartdatalake/spaTScope
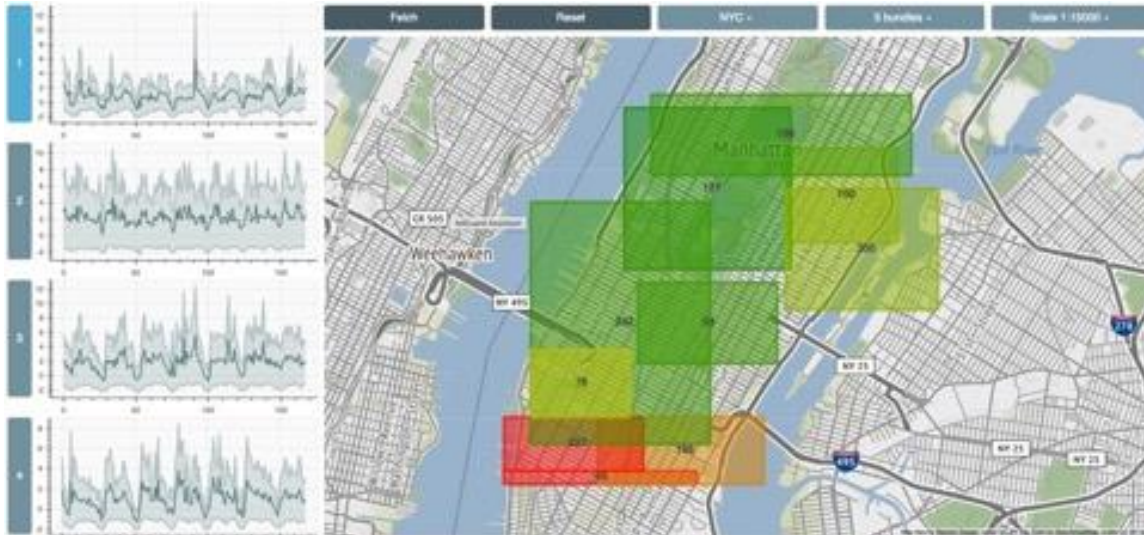


- **THOR**

Numerous search systems have been implemented that allow users to pose unstructured queries over databases without the need to use a query language, such as SQL. Unfortunately, the landscape of efforts is fragmented with no clear sight of which system is best, and what open challenges we should pursue in our research. To help towards this direction, we present THOR that makes 4 important contributions: a query benchmark, a framework for comparing different systems, several search system implementations, and a highly interactive tool for comparing different search systems.



https://darelab.imsi.athenarc.gr/thor/home

- **Datagent**

DatAgent is an intelligent data assistant system that allows users to ask queries in natural language, and can respond in natural language as well. Moreover, the

system actively guides the user using different types of recommendations and hints, and learns from user actions.

- **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 OAIS-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/

- **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

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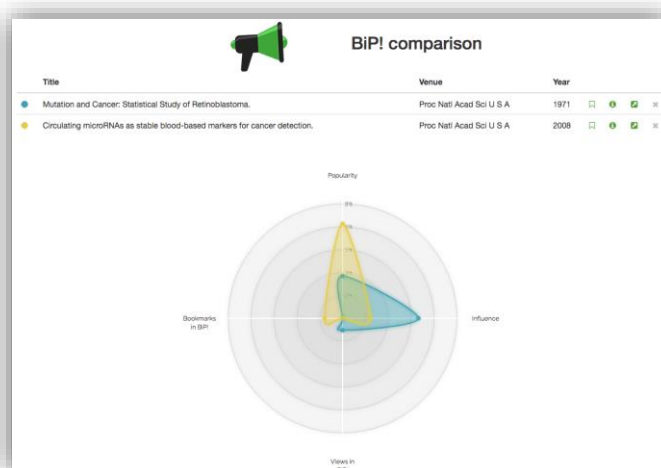
influence (i.e., their long-term impact in the discipline). Furthermore, the tool provides useful features like intuitive infographics for each article and a mechanism of bookmarks.

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

- **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/

- **SCHeMa**

SCHeMa (Scheduler for scientific Containers on clusters of Heterogeneous Machines) an open source platform to facilitate the execution and reproducibility of computational experiments on heterogeneous clusters. The platform exploits containerization, experiment packaging, and workflow management technologies to ease reproducibility,

while it leverages machine learning technologies to automatically identify the type of node that is more suitable to undertake each submitted computational task.
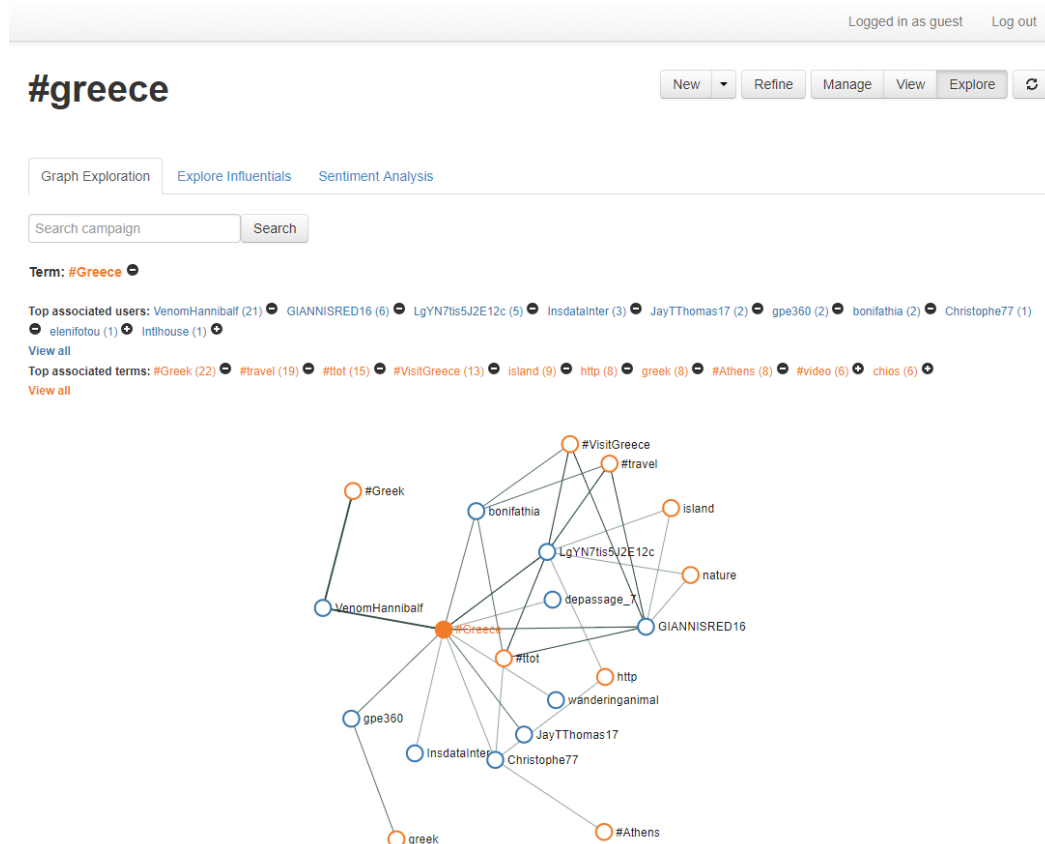https://github.com/athenarc/schema

- **SheerMP**

Optimizer for streaming analytics in cross-platform and cross-site environments.
https://infore-project.eu/deliverables.html

- **TwitHoard**

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.

http://twithoard.imis.athena-innovation.gr:8080/twithoard

- **eLib**

The project eLib aims at developing a digital library for the Independent Autority 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

- **PHONY: Automatic Dataset Generation 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.

- **Park Visitor Activity**

Urban parks are a vital part of cities around the world, accepting millions of visitors daily. However, in most cases urban parks do not enjoy the technological support that could



enhance visitor experience and facilitate park administrators. VR-Park is a project that

addresses the above issues and uses as a case study the "Pedion Areos" park, a prominent park in Athens, Greece. It comprises a mobile phone application used by the park visitor, and web-based applications used by the park administrator. Our work focuses on Park Visitor Activity, an innovative application used by the park administrator to collectively assess visitor movement and activity in general. The movement of visitors inside the park is collected and analyzed, to provide patterns of usage of the park areas: spots where people gather, pathways that are used the most, months of the year / time of day when visits have a peak, etc. Such information is invaluable for taking informed decisions about the management of the park. Comprehending how people tend to move and how they use an open area can be very useful in a wide spectrum of cases that go beyond urban park administrators.

- **Socioscope**

A visual analysis tool, used for visualization and exploration of social and political data (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.
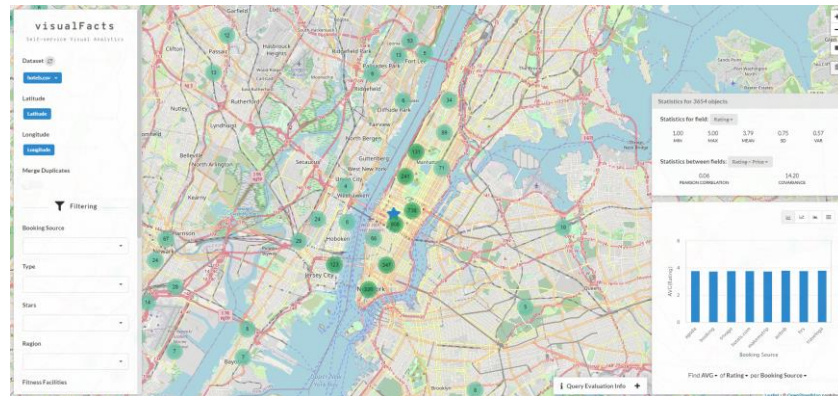www.socioscope.gr

- **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

- **Check4Facts. A platform for Public Discourse Fact Checking**

Check4facts aims to study fact-checking in the Greek public sphere, and particularly in relation to political personnel's discourse, addressing two major themes: (a) the immigration/ refugee issue, and (b) crime, raising facticity/truth issues on them. The Check4facts platform combines the automation of Machine Learning (ML) techniques, with the expertise of fact-checkers, to support a thorough and trustworthy workflow for political statement credibility assessment. The results of statement assessments, realized

as detailed assessment reports, are published to the general public via the Check4facts portal. http://check4facts.imsi.athenarc.gr/

- **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 uses 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 2021 several MSc and Diploma thesis have been co-supervised by IMSI members, who often serve as members in the respective examination committees.

IMIS members also co-supervise PhD students. The following PhD dissertations have been completed in 2021:

- 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.

The following PhD and MSc students collaborated closely with / were supervised by IMSI members in their research during 2021:

- Alexandros Zeakis. Topic: **Similarity Joins with multiple matching criteria**. Joint supervision with the University of Athens. Collaborating researcher: Dimitris Skoutas.
- Dimitris Tsesmelis. Topic: **Physical Optimization for large scale, data science workloads**. Collaborating researcher: Alkis Simitsis.
- Ibraheem Taha. **Interactive exploration & analytics on complex big data**. Collaborating researcher: Alkis Simitsis.
- Antheas Kapenekakis. **Privacy-aware data**. Collaborating researcher: Alkis Simitsis.
- Daniele Lunghi. **Scalable model selection in stream settings**. Collaborating researcher: Alkis Simitsis.
- Christos Papadopoulos. **A platform for prescriptive analytics**. Collaborating researcher: Alkis Simitsis.
- Antonis Kontaxakis. **End-to-end optimization for data science in the wild**. Collaborating researcher: Alkis Simitsis.
- Konstantinos Alexis. Topic: **Deep learning for computer vision with emphasis on noisy, heterogeneous and imbalanced datasets**. University of Athens. Collaborating researcher: Giorgos Giannopoulos.

- Eleni Lavasa. Topic: **Studying Solar Energetic Particle Events, contributing to their prognosis using machine learning techniques**. University of Athens. Collaborating researcher: Giorgos Giannopoulos.
- 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.
- Giorgos Alexiou, Topic: **Entity disambiguation and data interlinking.** Joint supervision National Technical University of Athens. Collaborating researcher: George Papastefanatos.
- Stavros Maroulis, Topic: **In Situ Data exploration and Visualization.** Joint supervision National Technical University of Athens. Collaborating researcher: George Papastefanatos.
- Christos Tsapelas (University of Athens) – Deep Learning for Query Optimization - supervisor: Georgia Koutrika.
- Antonis Mandamadiotis (University of Athens) – Query Recommendations using Multi-armed Bandits - supervisor: Georgia Koutrika.
- Mike Xydas (University of Athens) – Data to Text: Generation of Explanations for Query Results - supervisor: Georgia Koutrika.
- Anna Mitsopoulou (University of Athens) – NL-to-SQL Query Generation - supervisor: Georgia Koutrika.
- George Katsogiannis-Meimarakis (University of Athens) – SQL-to-NL Query Explanations - supervisor: Georgia Koutrika.

## Other Educational Activities

Other educational activities involving IMIS members include the following.

- Organization of the series of online courses "Digital Collections Metadata Curation in the Arts and Humanities", 30 March, 6 and 13 April 2021.
- Organization of "DH goes Viral" Digital Workshop, 26 April 2021, with the support of a DARIAH Theme grant, in which the Twitter Conference 2020 participants were reunited to assess the impact of the Covid-19 pandemic on DH research and pedagogy one year after.

- Organization of the [2nd APOLLONIS Summer School in Digital Humanities](#), 13-15 July 2021.
- Thanasis Vergoulis taught the "Data Management" course of the post-graduate program "Data Science" of NCSR Demokritos.
- Thanasis Vergoulis co-taught the "Big data systems and architecture" course of the post-graduate program "Business Analytics" of the Athens Univ. of Economics & Business.
- 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.
- George Papastefanatos was an Adjunct Lecturer for the course "Big Data Management" (In MSc in Cybersecurity and Data Science, Full-time Program – University of Piraeus).
- George Papastefanatos was an Adjunct Lecturer for the course "Advanced Topics in Data Engineering" (In MSc in Business Analytics, Full-time Program – Athens University of Economic and Business).
- George Papastefanatos supervised the following Internships.
  - D. Stoikou (Undergraduate - School of Electrical and Computer Engineering, NTUA), 2021.
  - K. Kozanis (Undergraduate - Department of Informatics, University of Thessaly), 2021.

# Facts and Figures

## Financial report

In 2021, IMSI continued its participation in EC and national funded research and development projects. The key economic indicators regarding the expenses and revenues in 2021 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 77%, 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 11 times the public expenditure received by IMSI.
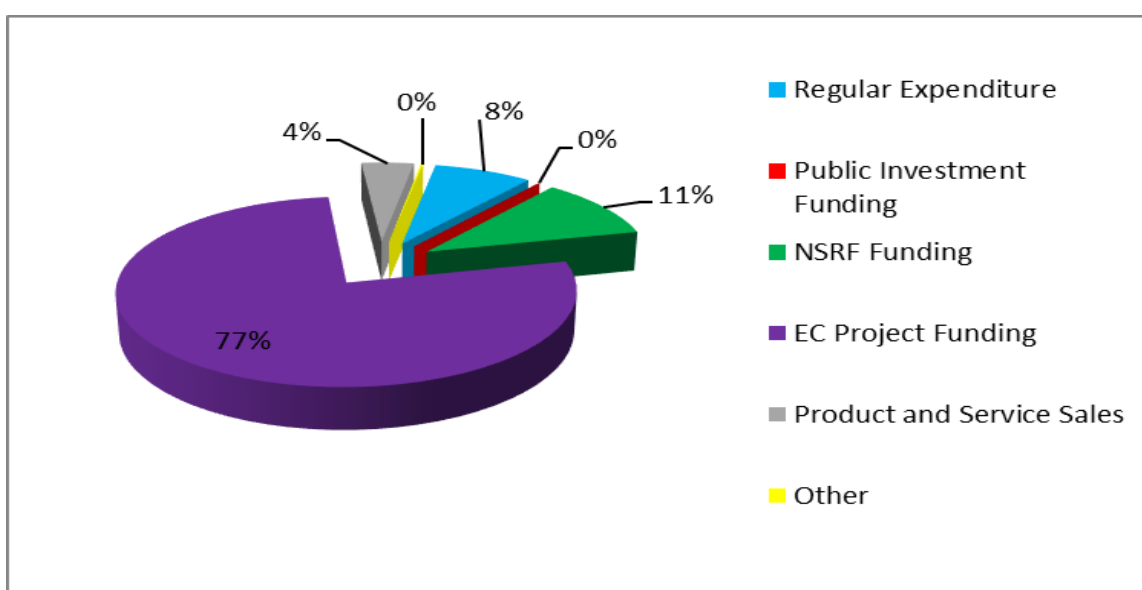


**Figure 1. Distribution of revenues in 2021**

**Table 1. Expenses and Revenues for 2021**

| Expenses 2021 (in Euros) | |
|---|---|
| Travel Expenses | 26.024,02 |
| Operational Costs | 46.608,82 |
| Equipment | 72.558,90 |
| Other Expenses | 242.695,69 |
| Personnel fees and payments to third parties | 4.474.648,17 |
| **Total** | **4.862.535,60** |

| Revenues 2021 (in Euros) | |
|---|---|
| Regular Expenditure | 358.173,78 |
| Public Investment Funding | 0,00 |
| NSRF Funding | 482.358,02 |
| EC Project Funding | 3.498.558,06 |
| Product and Service Sales | 190.461,54 |
| Other | 8.114,43 |
| Total | 4.537.665,83 |

Table 2 shows the revenues of IMSI since 2018 while a comparison of the revenues in the years 2018 - 2021 is illustrated in Figure 2. We can see that the revenues coming from participation in European projects reached between 3.5M and 4.7M within the period 2018 - 2021. As expected, a part of the NSRF Funding within the Partnership Agreement 2014-2020 was paid off in the years 2018&2019, with the revenues from the participation of IMSI in national funded projects reaching 1,2M euros.

**Table 2. Revenues from 2018 to 2021**

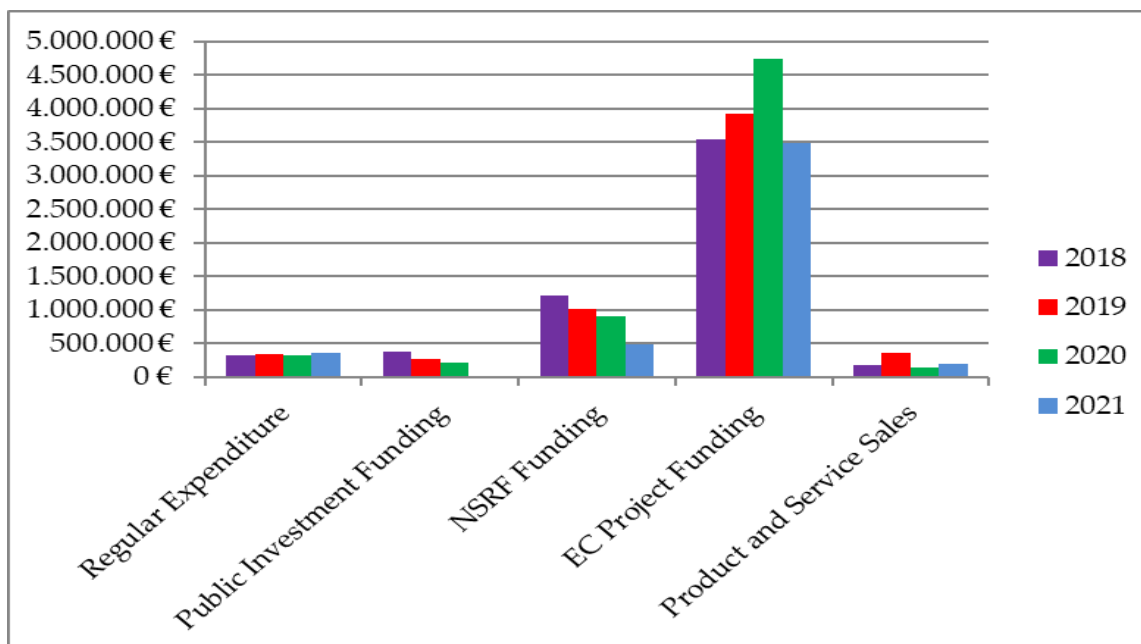| Revenues | | | | |
|---|---|---|---|---|
| | 2018 | 2019 | 2020 | 2021 |
| Regular Expenditure | 324.794,26 | 346.777,65 | 330.874,24 | 358.173,78 |
| Public Investment Funding | 383.261,82 | 267.080,07 | 222.884,40 | 0,00 |
| NSRF Funding | 1.208.807,34 | 1.023.287,09 | 900.019,00 | 482.358,02 |
| EC Project Funding | 3.536.570,82 | 3.926.562,69 | 4.747.206,63 | 3.498.558,06 |
| Product and Service Sales | 173.080,02 | 359.999,93 | 143.597,65 | 190.461,54 |
| Other | 31.198,95 | 59.553,25 | 29.847,85 | 8.114,43 |
| Total | 5.657.713,21 | 5.983.260,68 | 6.374.429,77 | 4.537.665,83 |

**Figure 2. Comparison of revenues 2018 - 2021**

## 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
- Avrithis Yannis, Researcher A
- Skoutas Dimitris, Researcher B
- Terrovitis Manolis, Researcher B
- Papastefanatos George, Researcher B

### 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
- Delis Alexios, Professor, University of Athens
- Douligeris Christos, Professor, University of Piraeus
- Emiris Ioannis, Professor, University of Athens
- Fotakis Dimitris, Associate Professor, NTUA
- Ioannidis Yannis, Professor, University of Athens
- Karantzalos Konstantinos, Associate Professor, NTUA
- Kokosis Antonios, Professor, NTUA
- Kotidis Yannis, Associate Professor, AUEB

- Maravelakis Manolis, Associate Professor, Hellenic Mediterranean University
- Papatheodorou Christos, Professor, University of Athens
- Patsakis Konstantinos, Assistant Professor, University of Piraeus
- Samoladas Vasilis, Associate Professor, Technical University of Crete

## Postdoctoral Researchers / Scientific Collaborators

- Akinsete Ebun, BSc in Architectural Technology, University of Bolton, MSc in Project Management, University of Manchester, PhD in Urban Regeneration and Sustainable Communities, University of Bolton
- Arkoumanis Constantinos, Elec. & Comp. Eng. NTUA, PhD NTUA
- Benardou Agiatis, BA, PhD in Ancient History, King's College London
- Bikakis Nikos, Elec. & Comp. Eng., Technical University of Crete, PhD NTUA
- 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
- Georgakilas George, Department of Computer Engineering and Informatics, University of Patras, MSc UoA, PhD University of Thessaly
- 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
- Kalozoumi Georgia, Department of Biochemistry & Biotechnology, University of Thessaly, MSc in Molecular Medicine, UoA, PhD in Molecular Biology, UoA
- Kanellos Ilias, 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
- Michalopoulos Christos, Mech. Eng. NTUA, MSc NTUA, PhD NTUA
- Pierrakos Dimitris, Dept. of Physics UoA, PhD UoA
- Psarros Ioannis, BSc, MSc & PhD, Dept. of Inform. and Telecom. UoA
- Ragia Lemonia, Department of Rural and Surveying Engineering, Aristotle University of Thessaloniki, PhD University of Bonn
- Raouzaiou Amaryllis, Elec. & Comp. Eng. NTUA, MSc & PhD NTUA
- Sakellariou Iason – Apostolos, BSc in Physics, University of Crete, MSc in Physics, Paris Diderot University, PhD in Physics, Paris-Sud University
- Sidirourgos Lefteris, BSc in Mathematics, University of Crete, MSc in Computer Science, University of Crete, PhD in Computer Science, University of Amsterdam
- 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

- Aidonopoulos Orfeas, Comp. & Inform. Eng. Univ. of Patras, Msc Dept. of Inform. & Telecom, UoA
- Alexakis Michalis, Elec. & Comp. Eng. NTUA
- Alexiou George, University of Aegean, Msc AUTH
- Alexis Konstantinos, Dept. of Inform. and  Telecom. UoA, MSc in Data Science, AUEB
- Andreou George, BSc in Physics UoA
- Apostolopoulos Theodoros, BSc in Mathematics UoA, MSc University of Piraeus
- Athanasiou Spyros, Elec. & Comp. Eng. NTUA
- Athanasoulias Sotiris, BSc in Computer Science & Artificial Intelligence, University of Sussex, MSc in Artificial Intelligence & Adaptive Systems, University of Sussex
- Balasis Ioannis, Dept. of Inform. and  Telecom. UoA
- Betchavas Panagiotis - Ioannis, BSc in Informatics, AUEB, MSc in Data Science, AUEB
- Bodare Filippos - Andreas, BSc in Physics, UoA, MSc in Environmental sciences, UoA
- 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, <u>Harokopio University, Msc, Department of Informatics, AUEB</u>
- Chatzopoulos Serafeim, BA, Msc, Dept. of Inform. and Telecom. UoA
- Checa Carles, BSc in Mathematics, Universitat Politècnica de Catalunya, MSc in Mathematics, Universittà degli Studi di Padova & Leiden University
- Christoforou Emmanouil,  BSc in Computer Science, AUEB, MSc in Bioinformatics, UoA
- Chryssanthou Argiris, BSc in Informatics, AUEB, MSc in Information security and computer crime, University of Glamorgan
- Deligiannis Panagiotis, Department of Informatics and Telecommunications, University of Peloponnese
- 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
- Dioletis Minas, BSc in Informatics, University of Piraeus, MSc UoA
- Dorgiakis Michail, BSc in  Computer Science, University of Crete, MSc in Computer Science, Vrije University Amsterdam
- Eleftheraki Stavroula, BSc in Statistics, University of the Aegean, MSc in Data Science & Information Technologies, UoA
- Emmanouilidi Sofia, Deree, Business Administration (Accounting and Finance)
- Farao Aristeidis, BSc, MSc, Dept. of Digital Systems, University of Piraeus
- Filippopolitis Konstantinos, BSc in Informatics, University of Piraeus
- 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
- Georgopoulos  Vasilios, Dept. of Inform. and  Telecom. UoA
- Gialitsis Nikolaos, Dept. of Inform. and  Telecom. UoA, MSc UoA
- Giannopoulou Roza, BA in Architecture, Kingston University, MSc in Architecture, University of Westminster
- Giannoudis Asterios, Dept. of Inform. and  Telecom. UoA
- Gkamiliari Aggeliki, Elec. & Comp. Eng. NTUA, MSc NTUA
- Gkolemis Vasilis, BSc in Electrical and Computer Engineering, Aristotle University of Thessaloniki, MSc in Data Science, University of Edinburgh
- Glenis Apostolos, BSc in Computer Science, University of Piraeus, MSc in  Computer Science, University of Crete
- Guittard Alice, BA in Geography, University of Toulouse 2, MSc in in Geography & Land Planning, University of Montpellier 3, MSc in Physical Geography, Geomatic with Remote Sensing and GIS, Stockholm University
- 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
- Kakaletris Georgios, BA, Msc, Dept. of Inform. and Telecom. UoA
- Kalampokis Panagiotis, BA, Computer Science, University of Ioannina, Msc in Mobile and Web Computing, IHU
- Kalimeris Alexandros, Dept. of Inform. and Telecom. UoA
- Kanavou Kalliopi, Department of Library Science & Information Systems, Technological Educational Institute of Athens, MSc in Informatics, University of Piraeus
- Karabatsis Thanasis - Michail, Dept. of Inform. and  Telecom. UoA
- Karamanolis Georgios, Dept. of Inform, AUEB
- Karmas Athanasios, Elec. & Comp. Eng. NTUA, MSc NTUA
- Karvounari Sofia, BSc, MSc, Dept. of Inform. and  Telecom. UoA
- Karvounis Manos, BSc, MSc & PhD candidate, Dept. of Inform. and  Telecom. UoA
- Kasomoulis Aristotelis, BSc, MSc, Dept. of Inform. and  Telecom. UoA
- Katrakazi Elissavet Olga, BA AUEB, Msc in Applied Environmental Economics, University of Kent
- Katrakazi Marianna, BA in Law, UoA, MA University of Strasbourg – CEIPI
- Katsogiannis - Meimarakis George, Dept. of Inform. and  Telecom. UoA
- Katsouli Maria – Olympia, BSc in Informatics, University of Piraeus, MSc in Brain Sciences, University of Glasgow
- 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
- Kechagias Konstantinos, Dept. of Inform. and  Telecom. UoA
- Kokogiannaki Argiro, Dept. of Inform. and Telecom. UoA
- Koltsida Panayota, BA, Msc, 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
- Koulalis Antonios, BA Dept. of Inform. and Telecom. UoA
- Koulocheri Eleni, BSc in Mathematics, UoA,  MSc in Information Systems, Hellenic Open University

- Kouvaras Ioannis, Elec. & Comp. Eng. NTUA
- Kravvaritis Constantinos, Dept. of Inform. and Telecom. UoA, Msc, Dept. of Inform, AUEB
- Lavasa Eleni, Dept. of Physics, UoA, MSc, University of Peloponnese
- Lazaropoulou Polyxeni, Department of Production and Management Engineering, Democritus University of Thrace, MSc, Democritus University of Thrace
- Lebesis Antonis, BA, Msc, Dept. of Inform. and Telecom. UoA
- Leonidakis Nikiforos, Dept. of Inform. and Telecom. UoA
- Liatas Christos, Electrical and Electronic Engineering, University of Nottingham, MSc University of Bristol
- Lykokanellos Filopoimin, Department of Computer Engineering and Informatics, University of Patras, MSc in Computer Science, University of Patras
- 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
- Mandamadiotis Antonios, Department of Information and Communication Systems Engineering, University of the Aegean, MSc in Data Science and Information Technologies, UoA
- Mantes Athanasios, BSC, MSc, Department of Computer Engineering and Informatics, University of Patras, MSc in Geoinformatics, NTUA, MSc in Bioinformatics, UoA
- Manola Natalia, Dept. of Physics, UOA
- Marinos Kouris Christos, Department Of Business Administration, AUEB, MSc Strathclyde University of Glasgow
- Maroulis Stavros, Elec. & Comp. Eng. NTUA
- Martzios Alexandros, BSc in Information Technology, University of Hertfordshire, IST College
- Mathioulaki Eleni, Elec. & Comp. Eng. NTUA
- Mavrogiannakis Nikos, BA, Msc in Physics, University of Manchester
- Mihalenas Vasilis, School of Architecture, NTUA
- Nakos Charalampos, Elec. & Comp. Eng. NTUA
- Ntalidou Nikoletta, BA in Law, Aristotle University of Thessaloniki, LL.M University of Piraeus
- Oikonomou Leonidas, Architecture, School of Architecture NTUA, MSc NTUA
- Panagiotopoulos George, Dept. of Inform. and Telecom. UoA
- 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
- Papageorgiou-Mariglis Asimakis, BSc in Physics, University of Crete, MSc University of Edinburgh
- Papanikos Georgios, BA, MSc, Dept. of Inform. and Telecom. UoA
- Pappas Dimitris, BSc in Computer Science, University of Crete, MSc in Computer Science, AUEB
- Patroumpas Konstantinos, Comp. & Inform. Eng. University of Patras, Msc NTUA
- 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
- 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

- Renieris Loukas, BSc in Informatics, AUEB, MSC in Informatics, University of Piraeus
- Sifakaki Ilektra, BA in Mathematics, UoA, Msc, Department of Informatics and Telematics, <u>Harokopio University</u>
- Smyris Georgios, BSc in Physics, UoA, MSc UoA
- Smyrnaios Lampros, 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
- Spirou Konstantinos, Dept. of Inform. and Telecom. UoA
- Stamatakis Giorgos, Elec. & Comp. Eng., Technical University of Crete
- StamatopoulosVassilis, Dept. of Inform. and Telecom. UoA
- Stefou Thodoris, BSc, MSc, Dept. of Inform. and Telecom. UoA
- Stergiopoulou Lydia, Department of Economics, AUEB, MSc in Environmental Sciences, Policy and Management, University of Manchester
- Svingos Christoforos, Dept. of Inform. and Telecom. UoA, MSc UoA
- Tsesmelis Dimitris, Dept. of Inform. and Telecom. UoA, MSc in Big Data Management and Analytics, ULB, UPC, CS
- Tertikas Konstantinos, Elec. & Comp. Eng. NTUA, MSc University of Southampton
- Thorpe Rowan, Newlands College
- Toli Eleni, Department of History and Archaeology, University of Ioannina, MSc in Labour Economics Ruhr University Bochum
- Triantafyllou Konstantinos, Dept. of Inform. and Telecom. UoA
- Tsapatsari Mirto, BA in Law, UoA, LL.M in International Business Law, VU University Amsterdam
- Tsapelas Christos, Dept. of Inform. and Telecom. UoA
- Tsiakaliaris Christos, Elec. & Comp. Eng. NTUA, MSc in Information Systems Engineering, University of Manchester, Institute of Science and Technology
- Tsitsigkos Dimitris, Dept. of Inform. and Telecom. UoA
- Tzerefos Anargiros, Department of Informatics and Telecommunications, University of Peloponnese
- Tziotzios Diamantis, Elec. & Comp. Eng. DUTH
- Vargiamis Michalis, Dept. of Inform. and Telecom. UoA
- Vassiliou Anastasia, Department of History and Archaeology, UoA, MSc in Applied Geography and Spatial Planning, Harokopio University
- Vichos Kleanthis, Department of Electrical and Computer Engineering, University of Patras
- Xagorari Katerina, Department of Informatics and Telecommunications, University of Peloponnese
- Xiros Panagiotis, Elec. & Comp. Eng. NTUA
- Xydas Michail, Dept. of Inform. and Telecom. UoA, MSc in Data Science and Information Technologies, UoA
- 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

Most current information about activities and contacts for the staff are available on the IMSI website:
https://www.imsi.athenarc.gr/en