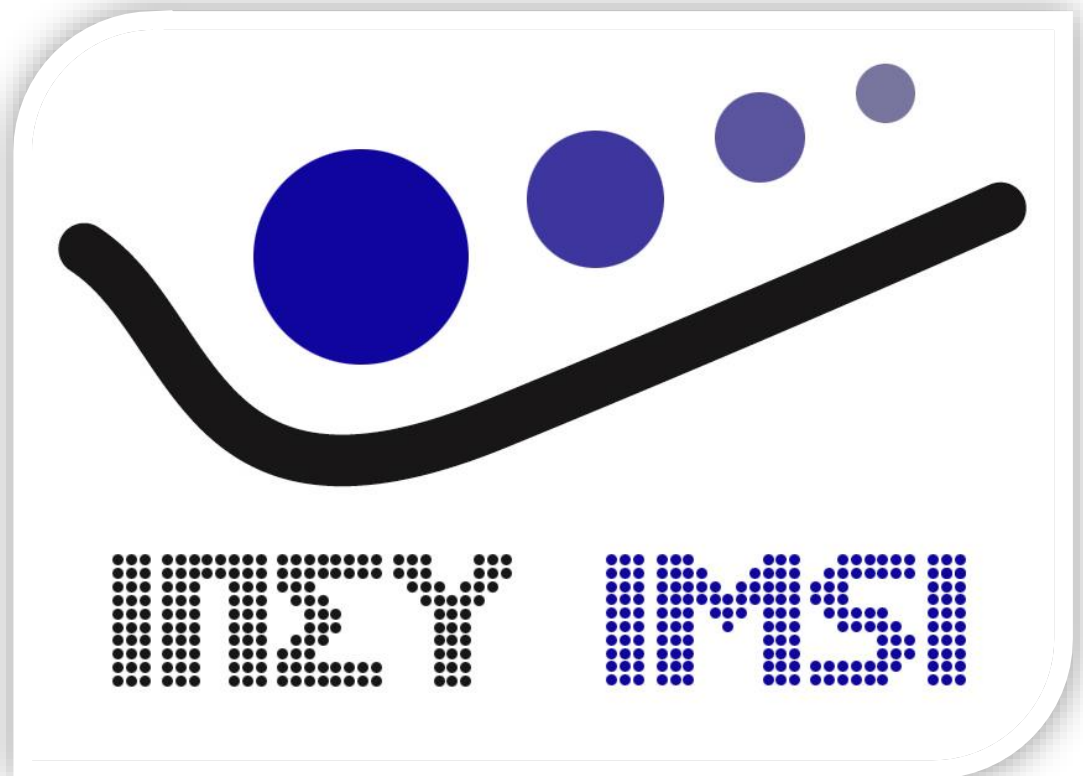
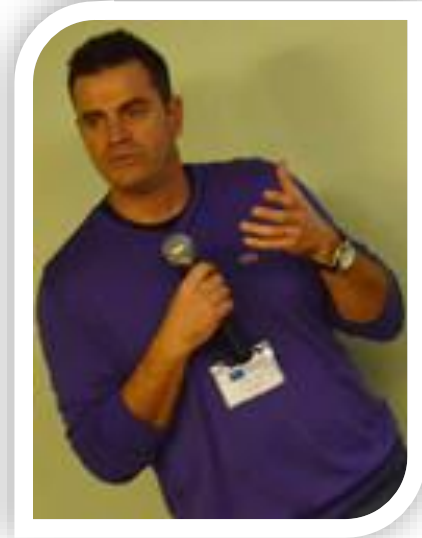


IMSI annual report 2022



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 2022, 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, 20/11/2023



Contents

Preface	1
Profile	3
Infrastructures.....	8
Highlights	10
Research Directions	12
Projects	18
Publications	53
Dissemination Activities.....	57
Systems	59
Education.....	73
Facts and Figures	76
Contact	85



Profile

The **Information Management Systems Institute (IMSI)** is a leading research institute in Greece and a Centre of Excellence for research, development, and innovation in the areas of large-scale information systems and Big Data management. IMSI lies within the **“ATHENA” Research and Innovation Center in Information, Communication and Knowledge Technologies**, which is the first research center in Greece with a focus exclusively on Information Society.

ATHENA R.C. was established in Athens in 2003, and it is a research and technology organization supervised by the General Secretariat for Research and Technology of the Ministry of Education, Research and Religious Affairs. The “Institute for the Management of Information Systems (IMIS)” was founded in 2007 with the mission to conduct research in the areas of data management and information systems. In 2017, the former name of the Institute changed to the current name “Information Management Systems Institute”, as decided by the Management Board of the ATHENA R.C. Accordingly, the short name became IMSI, replacing the former short name IMIS.

IMSI is today one of Greece’s premier research institutes 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 and 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 platforms and tools in both local industry and various Greek government organizations.

IMSI has become a reference point and a pole of attraction not only for highly qualified scientists and students working in the areas of information systems, databases, and Big Data technologies, but also for the dynamic Greek and European private sector in information technologies. IMSI continuously seeks to (a) engage in first-class, novel, internationally competitive research activities by successfully combining fundamental work and an applied orientation, (b) pursue collaborations with research teams of complementary expertise in its areas of interest, and (c) maintain societal and industrial relevance.

The **main scientific areas** of focus for IMSI include:



- Big Data and Scalable Data Analytics
- Machine Learning and Knowledge Discovery
- Large-Scale Information Systems
- Data Privacy and Data Marketplaces
- Blockchain Systems
- User-Centric Systems and Applications
- Cloud-based Platforms and Data Services
- Big Data Research Infrastructures

Research at IMSI ranges from basic to applied, and has a strong collaborative aspect, as it is typically conducted with national and international partners from industry as well as academia, in the context of innovative R&D 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.

IMSI is creating knowledge and technologies in some of the most aggressively developing sectors of the economy, such as Big Data Analytics and Machine Learning. Apart from scientific research, IMSI is also contributing to society in the following ways:

- It participates in high-profile EU and national infrastructures (including, OpenAIRE, ELIXIR, EOSC, and Apollonis) that support scientific research at the national and international level.
- It is a leading partner in flagship research efforts at both the EU and national level – prominent examples include The Human Brain Project (HBP) and the Greek Precision Medicine Initiative.
- It transfers knowledge to the industry through several EU and national projects.
- It contributes to educating and training new scientists and reversing brain drain, by participating in graduate education programs as well as funding and supervising MSc and PhD candidates.
- It attracts large amounts of competitive funding from national and international sources compared to its basic state funding.

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

- **Big Data Analytics and Machine Learning**
The department conducts 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 research directions include efficient and 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.
- **Big Data Research Infrastructures**
The department conducts 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, IMSI 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).

- **Cloud Platforms and Data Services**

The department conducts 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.

- **Distributed and Web Information Systems**

The department conducts 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.

- **User-centric Systems and Applications**

The department conducts research, development, and innovation in the areas of user-centric and data-centric systems and applications. IMSI 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 **Department of Coordination and Administration**, as well as 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
- Compute Nodes for research projects
- GPU Nodes for research projects

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

- **AutoFair project** (HE, GA No. 101070568) **started in October 2022**, comprises a strong consortium performing research on bias detection, correction and explanation in AI on use cases drafted by large companies like IBM and Workable. IMSI leads the development of the fairness-aware explainability methods, being in line with the latest developments in the EU, which pursue a more ethical, reliable, transparent, trustworthy and fair AI landscape.
- The EU (H2020-MSCA-ITN-2019) project **GRAPES**, coordinated by I. Emiris and focusing on 3D Shape Optimization and Manipulation by geometric modeling and deep learning methods, organized successfully the **2nd Doctoral school** that took place from Monday, June 13 to Friday, June 17 at the Università della Svizzera italiana (USI) in Lugano, Switzerland.
- A new **collaboration between 6 Research Centers** (ATHENA Research Center, National Centre for Social Research, National Observatory of Athens, Hellenic Centre for Marine Research, Biomedical Research Foundation of the Academy of Athens, Institute of Communication and Computer Systems) and 2 Universities (Department of Communication and Media Studies of the National and Kapodistrian University of Athens, Communication and Digital Media Department of the University of Western Macedonia) has started in 2022 around a project called Check4Facts \ Science. It will study fact-checking in media, in relation to news for climate change and public health \ pandemic. Project Coordinator from IMSI: George Papastefanatos.
- The Horizon Europe project **DataBri-X** (Data Process and Technological Bricks for expanding digital value creation in European Data Spaces) started in October 2022. The project is coordinated by ARC/IMSI (Stelios Sartzetakis). The project aspires to contribute in making Europe the most successful area in the world in terms of data sharing and data re-use.
- The H2020 project **SmartDataLake** (Sustainable Data Lakes for Extreme-Scale Analytics) was successfully completed, with the final review of the project taking place in February 2022. The project was coordinated by IMSI (Dimitris Skoutas) and involved designing novel methods and tools for data discovery, exploration and mining in data lakes.
- The Horizon Europe project **STELAR** (Spatio-Temporal Linked data tools for the AgRi-food data space) started in September 2022. The project is coordinated by IMSI (Dimitris Skoutas) and its goal is to develop an innovative Knowledge Lake

Management System for FAIR and AI-ready data with particular applications in the agrifood sector.

- Organization of the DARIAH Annual Event 2022, Athens, 31 May-03 June 2022.
- Georgia Koutrika is elected as a **member of the VLDB Endowment Board of Trustees**
- 1st Greek ACM-W Chapter **Winter School on Fairness in AI** co-organized by Georgia Koutrika and Evi Pitoura
- Georgia Koutrika was a panelist in the "**Futuristic Data Interfaces**" panel at DASFAA, April 12 2022 moderated by K. Stockinger with Jaydeep Sen, Immanuel Trummer, Lei Zou.
- Georgia Koutrika was a panelist in the "**#BreakTheBias in Science and Technology**" panel, panelist, moderated by Elina Makri with T. Alissafi, P. Nomikou, V. Kotroni, L. Panayotopoulos. March 8, 2022
- IMSI organized the "**1st DEDS Winter School on Ethical and Legal Aspects of Data**" in Athens, Greece, 2-8 April, 2022. The school was attended by 50+ PhD students and industry professionals from 10+ countries and featured 12 lectures and keynote talks from renowned researchers and industry leaders. Organization coordination from IMSI: Alkis Simitsis
- IMSI participated in the organization of the "**10th European Big Data Management & Analytics Summer School (eBISS 2022)**", in Cesena, Italy, 4-8 July, 2022. The school was attended by 70+ MSc and PhD students, and featured 10 lectures and keynote talks. Organization coordination from IMSI: Alkis Simitsis

Awards

- Manolis Terrovitis received IEEE MDM **Test-of-Time** (2008-2022) Award for the paper: Privacy Preservation in the Publication of Trajectories, Manolis Terrovitis and Nikos Mamoulis, IEEE MDM 2008
- **Highly Cited Researchers 2022:** Theodore Dalamagas was recognised as a Highly Cited Researcher from Clarivate Analytics for 2022. Clarivate identifies scientists who have demonstrated significant and broad influence, reflected in the publication of multiple papers frequently cited by their peers during the last decade (Web of Science). Researchers are selected for their exceptional influence and performance in one or more of 21 fields (those used in Essential Science Indicators™ or ESI) or across several fields.

Research Directions

IMSI research activities fall into the following areas.

Big Data Analytics and Machine Learning

Scalable, interactive Big Data analytics. One key direction of the group is to address a number of challenges relating to the data itself, the infrastructure and the users. Data challenges include its scale, heterogeneity, structure or lack thereof, dynamic nature and privacy. In this context, the group focuses on exploration and analysis of noisy data, including overlapping, incomplete or contradicting data from multiple sources, which is common due to the emergence of data aggregators for example. Approaches include query-time cleaning, repairing, deduplicating, clustering and exploration. Infrastructure challenges include the structure of the hardware (edge devices, distributed platforms, supercomputers), the physical storage vs. processing nodes and network structure, as well as application-specific data workflows. The group focuses on extreme-scale analytics by physical optimization over several criteria including runtime, throughput, latency, scheduling, system and monetary resources. The main objective is to bring computation closer to data, for example by in-situ data processing, leveraging hardware specificities without affecting the interface to applications and decoupling engine primitives from the underlying data store platform. The user challenge is essentially to make technologies accessible to non-expert data analysts. In this sense, novel algorithms are investigated to support exploring, processing, visualizing and extracting insights from data on the fly, guided by user interaction. Besides exploring data technology to assist machine learning, the group also investigates learning techniques to assist or even replace traditional data engine functions, such as query optimization, scheduling and workload management.

Large-scale Machine Learning. Beyond data analytics, a second direction is to investigate Machine Learning models to make predictions at large scale. This includes learning new representations from raw data, which can be used to solve new tasks. The key challenges are the scale and diversity of data, the missing, noisy or inconsistent supervision, as well as the dynamic nature of both. Besides the massively parallel processing on distributed platforms, learning on large-scale data is facilitated by continual learning on streaming data and non-parametric models that can be adapted easily. The group studies and builds on recent advances of self-supervised learning to compensate for the flaws or lack of supervision and extends the state of the art towards learning compact representations to enable scaling up. The focus is on strong mathematical foundations and interdisciplinary research to handle data of multiple modalities including vision, language, time series as well as structured and high-



dimensional data. Several application domains are considered, including geometric modeling and CAD/CAM, scientific databases and publications, information retrieval, bioinformatics. The results are applicable to several sectors including health, education, environment, transportation, finance, materials, food and agriculture.

Big Data Research Infrastructures

Generic Data infrastructures. Such infrastructures provide generic scalable data processing services for very large and heterogeneous scientific data, ready to be used as a building software block for other RIs. The group has a leading role in HELIX, a horizontal digital RI for data-intensive research, handling the data management, analysis, sharing, and reuse needs of Greek scientists and innovators in a cross-disciplinary, scalable, and low-cost manner. HELIX provides its services also as an autonomous RI in support of data sharing, open access publishing, and data experimentation.

Open Science. A critical mission of the European Commission is to provide unlimited, barrier free, Open Access to research outputs financed by public funding in EU. The group has a leading role in OpenAIRE, an RI whose mission is to fulfill the European Open Science Cloud (EOSC) vision, but has also a global outreach. Its operations already provide the glue for many user- and research-driven functionalities, whether these come from the long tail of science (repositories and local support) or domain disciplined research communities or other RIs.

Health. The practice of life sciences is continuously becoming more data-driven. The group has a leading role in RIs (ELIXIR-GR, Inspired-RIs, Oncopmnet) serving a range of domains from genomics and structural biology to medicine. ELIXIR-GR is the Greek node of ELIXIR, the distributed ESFRI RI for data, tools, standards, and training, serving the life science community for open, integrated, and state-of-the-art bioinformatics and biocomputing resources. Inspired-RIs focuses on integrated structural biology, drug screening and target functional characterization. Oncopmnet implements the Hellenic Precision Medicine Network on Oncology, providing digital tools and systems for organization, processing, and analysis of cancer data.

Humanities and Digital Curation. Humanities also goes digital. Thus, a strategic action line is the **development** of digital research infrastructures for humanities at the national and European levels with an emphasis on the lifecycle of curated data. The action line has been undertaken by the Digital Curation Unit (DCU) (a department of IMSI since 2009, led by Prof. Panos Constantopoulos, but now carrying out its activities under the Big Data Research Infrastructures Group). 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.

Cloud Platforms and Data Services

Data processing and analytics on the cloud. The key focus of the group is on technologies that enable scalable data analytics on the cloud. Our research is focused on the data analytics services layer, addressing any type of scalability problem using solutions based on the distribution of computation to multiple cores, VMs, or containers. Parallel in-memory data analytics operators, for complex data e.g., spatial, intervals, time series, incomplete and heterogenous data, are some of the most active research efforts. The group is also working on developing cloud-based data analytics services in the context of different disciplines and sectors. E.g., for energy analytics, for life sciences and medical data, data analysis operations tailored for pattern detection and extraction for energy monitoring data (timeseries) are optimized for accuracy and scalability in cloud environments. For telco data, research has focused on end-to-end big data solutions for stream analytics on network quality data coming from IoT devices, such as drones and autonomous cars. For scholarly data, research concerns performance and accuracy optimization of entity resolution and entity interlinking in data integration workflows.

Privacy-based processing of data in the cloud. A key concern with cloud-based analytics is privacy concerns and restrictions when personal data are involved. There are several privacy preserving strategies that can be employed to protect personal data including, designing principles, encryption, differential private algorithms and data anonymization techniques. The group has been active in most of these aspects, providing designing and governance principles for health information systems, data anonymization techniques and tools etc. The group supports the public open-source data anonymization tool Amnesia (<https://amnesia.openaire.eu>).

Data services. In modern cloud environments, data services (database-as-a-service, ML, etc.) often need to operate next to where data is generated, e.g., for reducing data transfer overhead or when sensitive data cannot move out of the production system, etc. In this context, new end-user data services and applications require in situ analysis, i.e., analysis is performed directly on the data residing at the edge, without the need to move and load the data on a cloud database. In this context, the group has been active in developing in-situ techniques, such as scalable interactive visualization techniques for in-situ visual analysis of data. The group has developed a public open-source visualization tool VisualFacts (<https://visualfacts.imsi.athenarc.gr/>).

Domain specific and explainable AI services. It is often the case that applying generic state of the art ML algorithms and workflows is not adequate to effectively solve specialized, but quite significant for real-world application, tasks. This has become

evident in various scenarios, including Earth Observation and analytics settings, as well as in medical image analysis. Our aim is to research how state of the art ML/DL algorithms and methodologies can be properly extended, utilizing domain knowledge, to effectively solve real-world problems. In parallel, explainability is to become a de-facto requirement for several types of AI systems and services. In this context, the group implements model agnostic explainability services, with emphasis on user interactivity and the explainability of fairness of AI systems, and their deployment in the form of Functions-as-a-Service.

Distributed and Web Information Systems

Web of Data. The Semantic Web is a collection of technologies that enable the linking and semantic annotation of various types of data from heterogeneous sources, leveraging information from standard vocabularies and ontologies. Linked Data, i.e., interrelated datasets, can boost knowledge discovery and data-driven analytics. Entity resolution and similarity joins lie at the heart of the interlinking process, as well as data integration in general. Addressing these problems raises challenges both in terms of efficiency and effectiveness. Regarding the former, scaling to very large collections of entities requires elaborate techniques for candidate selection and filtering. Achieving high accuracy is also challenging, due to the presence of various types of attributes, similarity measures and linking criteria, which leads to a large parameter space, involving different tradeoffs with respect to precision and recall.

Dynamics and Evolution of the Data Web. The management of evolving information in a decentralized setting introduces problems related to the archiving and preservation of interlinked information, temporal modelling & evolution management (change detection and propagation) as well as benchmarking techniques in this area. In our view, changes are discrete objects that have complex structure and retain their semantic and temporal characteristics, rather than being isolated low-level transformations on data.

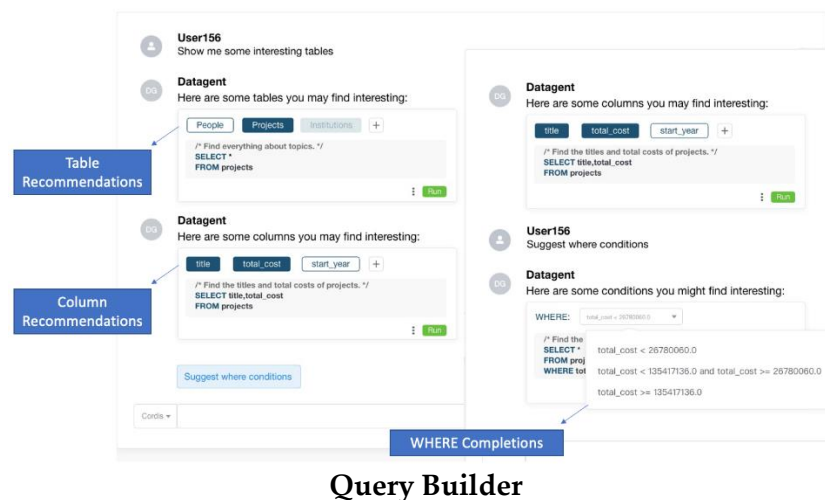
Geosocial networks. An increasingly large portion of data on the Web is associated with a spatial and/or temporal dimension. Also, spatial and temporal attributes are often inherently present in information generated by sensor networks and peer-to-peer systems. Location data and location-based services have a significant and widely recognized value in most, if not all, sectors of the data economy. Searching, integrating and mining geospatial data and time series is an active field of research with numerous new challenges.

Leveraging Social Data. The availability of online data through social networks, especially Twitter, gives rise to several disparate and challenging problems: (a) how to leverage social data for obtaining new knowledge (data journalism, public opinion trends, brand monitoring), (b) how to use knowledge graphs to create meaningful associations and recommendations between tweets and users, and (c) how to use diffusion patterns in Twitter to detect fake news.

User-centric Systems and Applications

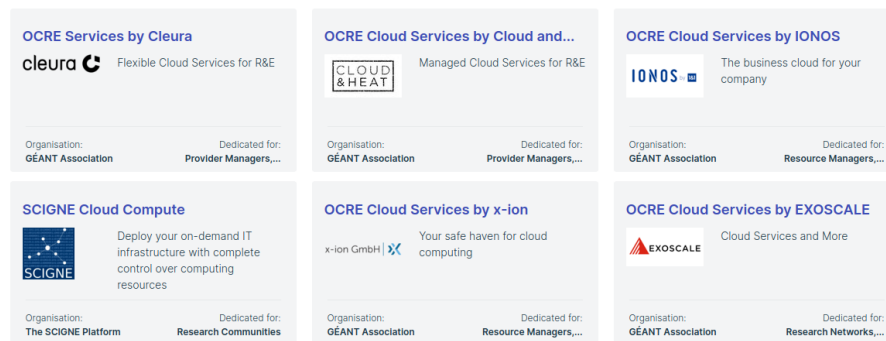
Futuristic Data Interfaces. The volume of data and the need for data democratization call for shift from the classical «query-based information access paradigm» popularized by existing systems to a novel information access paradigm, where the system takes upon a more conversational and active role in helping users effectively explore data of varying quality, complexity, and relevance. We are looking into text-to-SQL, SQL-to-Text and Data-to-Text systems, i.e., systems that allow users to ask queries using natural language, receive explanations of SQL queries and results, respectively. We are developing a SQL-to-Text system, called eQsplain that can learn to generate explanations for multiple query languages, such as SQL and SPARQL. We are also developing a Data-to-Text system, called QR2T, for explaining query results.

Intelligent Interactive Data Exploration. We are working on algorithms that allow the system to actively guide the user through their information access endeavor by offering query recommendations, and data exploration options as needed. Towards this direction, we have implemented recommendation algorithms in the context of the INODE platform as well as for the EOSC platform. In the former case, we have implemented a Query Builder that allows users to explore unseen data by helping them to build queries. The system learns in an online way by incorporating user feedback and adapting to user selections and preferences on the fly. The underlying algorithms leverage multi-armed bandits.



For the EOSC platform, we have developed a semantic similarity-based recommender engine that enables recommendations of services to EOSC consumers.

Similar services



EOSC Similar Services

Our research on futuristic data interfaces and intelligent interactive data exploration is brought together in our system Datagent 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.

Fair and Ethical Algorithmic Systems. Increasingly more aspects of our everyday lives are influenced by automated decisions made by systems that statistically analyze traces of our activities. It is thus natural to question whether such systems are trustworthy, particularly given the opaqueness and complexity of their internal workings. We are looking into how fairness is defined and achieved in ML systems (e.g., machine-generated recommendations and rankings). We are particularly developing new fairness-aware methods.

User-Driven Data Management. Our goal is to 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 goal more effectively.

Projects

EU R&D Projects



DataTools4Heart: A European Health Data Toolbox for Enhancing Cardiology Data Interoperability, Reusability and Privacy

Project manager	Minos Garofalakis
Coordinator	University of Barcelona (UB)
IMSI - funding	556.000 Euros
Programme	HORIZON-HLTH-2021-TOOL-06
Start date	1/10/2022
Duration	3 years
Website	https://www.datatools4heart.eu/

DataTools4Heart will create a comprehensive cardiology data toolbox for clinicians, researchers, and data scientists. Tools will allow **data ingestion and harmonisation, Natural Language Processing** in multiple languages, **federated machine learning and data synthesis**. Virtual assistants will aid users in navigating large multi-source cardiology data while adhering to European regulations and data standards.

Data ingestion and harmonization. DataTools4Heart will develop a common data extraction tool to improve metadata and data interoperability while addressing data heterogeneity across European regions and cardiology units. This tool will be developed and validated through a modular and flexible Data Ingestion Suite deployed in 7 European sites. Interoperability of the Data Ingestion will be guaranteed with at least 4 standard-based data models (HL7 V2, HL7 CDA, OMOP CDM, and i2B2) and tested in 3 different use cases for AI modelling.

Natural Language Processing. DataTools4Heart will introduce a multilingual Natural Language Processing (NLP) suite to standardise the structuring of cardiology reports across European regions, including cardiology-specific entity recognition and machine translation. Such suite will include adaptation of 7 language models to the cardiology domain in English, Spanish, Italian, Romanian, Czech, Swedish, and Dutch using EHR data from clinical site partners. The project will include the release of clinical multilingual corpora (CardioSynth and

Paraclete) in 7 languages, with over 50% being low-resource and containing more than 500,000 words of clinical text.

Federated machine learning and data synthesis. With the aim to develop innovative methods for synthesising data, DataTools4Heart will build a privacy-preserving cardiology data toolbox to improve data reusability, while adhering to ethical and legal standards. A secure and federated network for cardiology data will be established in 7 European locations across all regions, as result from the cooperation of different stakeholders. Differentially private synthetic data will allow to handle data representative of a target population, scalable, shareable for research purposes, and able to reduce bias in algorithmic development. The legacy will be the creation of an open-source privacy-conscious synthetic dataset, CardioSynth. The process and the quality of synthetic data generation will be thoroughly evaluated over the course of the project.



DT4GS: Open collaboration and open Digital Twin infrastructure for Green Smart Shipping

Project manager	Ioannis Z. Emiris
Coordinator	INLECOM GROUP
IMSI - funding	142,500 euros
Programme	HORIZON Research and Innovation Actions (HORIZON-CL5-2021-D5-01-13)
Start date	01/06/2022
Duration	36 months
Website	https://dt4gs.eu/

DT4GS is aimed at delivering an "Open Digital Twin Framework" for both shipping companies and the broader waterborne industry actors to tap into new opportunities made available through the use of Digital Twins (DTs). The project will enable shipping stakeholders to embrace the full spectrum of DT innovations to support smart green shipping in the upgrade of existing ships and new vessels. DT4GS will cover the full ship lifecycle by embracing federation of DT applications as well as utilising DTLF policies and related shared-dataspace developments for the sector. DT4GS applications will focus on shipping companies but will also provide decarbonisation decision support system for shipyards, equipment manufacturers, port authorities and operators, river commissions, classification societies, energy companies and transport/corridor infrastructure companies. DT4GS's objectives are to: 1. Support shipping

companies in achieving up to 20% reduction in CO₂e with a 2026 horizon, by developing and deploying real-time configurable DTs for ship and fleet operational performance optimisation in 4 Living Labs involving shipping companies, with different vessel types, and establishing fully validated industry services for Green Shipping Operational Optimisation DTs expected to be adopted by 1000+ ships by 2030. 2. Establish a comprehensive zero-emission shipping methodology and support Virtual Testbed and Decision Support Systems that address both new builds and retrofits comprising: a. A DT4GS (Green Shipping) Dataspace for the broader shipping sector contributing to GAIA-X by establishing a core European industry resource that accelerates the green and digital transition of waterborne shipping and transport value chains. b. Simulation based solutions to retrofit ships, targeting 55% reduced CO₂e reduction by 2030. c. A smart green "new-build" reference design per vessel type. d. Virtual Testbed services for reducing the cost of physical testing of GS solutions by 20%.



DataBri-X: Data Process and Technological Bricks for expanding digital value creation in European Data Spaces

Project manager	Stelios Sartzetakis
Coordinator	IMSI / ATHENA RC
IMSI - funding	615K Euros
Programme	HORIZON-CL4-2021-DATA-01
Start date	01-10-2022
Duration	36 months
Website	https://databri-x.eu/

Through DataBri-X, European Data Spaces, platforms and marketplaces and their wide range of business, governmental and public, research and civil society stakeholders will be equipped with a holistic and flexible data governance process and a seamless integrated standards based toolbox for data- and metadata management which can be assembled along relevant requirements, provides open source as well as commercial tools (the bricks / bri-X), and mechanisms to load 3rd party resources like language resources or AI models, and can be easily deployed into Data Spaces and thereby will contribute to make Europe the most successful area in the world in terms of data sharing and data re-use, to gain the full benefit from the value of data, while respecting the legal framework relating to security and privacy.



STELAR: Spatio-TEmporal Linked data tools for the AgRI-food data space

Project manager	Dimitris Skoutas
Coordinator	IMSI
IMIS - funding	€ 948,125.00
Programme	Horizon Europe
Start date	1/9/2022
Duration	36 months
Website	https://stelar-project.eu/

STELAR will design, develop, evaluate, and showcase an innovative Knowledge Lake Management System (KLMS) to support and facilitate a holistic approach for FAIR (Findable, Accessible, Interoperable, Reusable) and AI-ready (high-quality, reliably labeled) data. The STELAR KLMS will allow to (semi-)automatically turn a raw data lake into a knowledge lake. This is achieved by (1) enhancing the data lake with a knowledge layer, and (2) developing and integrating a set of data management tools and workflows. The knowledge layer will comprise: (a) a data catalog offering automatically enhanced metadata for the raw data assets in the lake, and (b) a knowledge graph that semantically describes and interlinks these data assets using suitable domain ontologies and vocabularies. The provided tools and workflows will offer novel functionalities for: (a) data discovery and quality management; (b) data linking and alignment; and (c) data annotation and synthetic data generation. The KLMS will combine both human-in-the-loop and automatic approaches, to leverage background knowledge of domain experts while minimizing their involvement. To reduce manual effort and time, it will increase the automation of finding and selecting relevant data sources, configuring, and tuning the involved data management tools, and designing, executing, and monitoring end-to-end data processing workflows adapted to different user needs. The KLMS will include specialized tools and functions for geospatial, temporal, and textual data. An organization, ranging from a data-intensive SME to the operator of a data marketplace, will be able to use the STELAR KLMS to increase the readiness of its data assets for use in AI applications and for being shared and exchanged within a common data space. The STELAR KLMS will be pilot tested in diverse, real-world use cases in the agrifood data space, one of the nine data spaces of strategic societal and economic importance identified in the European Strategy for Data.



EVENFLOW – Robust Learning & Reasoning for Complex Event Forecasting

Project manager	Nikos Giatrakos
Coordinator	Intrasoft International
IMIS - funding	500,000 €
Programme	HORIZON-CL4-2021-HUMAN-01 — A Human-Centered and Ethical Development of Digital and Industrial Technologies 2021
Start date	01/10/2022
Duration	3 Years
Website	https://evenflow-project.eu/

A growing number of applications rely on AI-based solutions to carry-out mission-critical tasks, many of which are of temporal nature, dealing with ever-evolving flows of information. Crucial for mitigating threats and taking advantage of opportunities in such domains, is the ability to forecast imminent situations and critical complex events ahead of time. EVENFLOW will develop hybrid learning techniques for complex event forecasting, which combine deep learning with logic-based learning and reasoning into neurosymbolic forecasting models. The envisioned methods will combine (i) neural representation learning techniques, capable of constructing event-based features from streams of perception-level data with (ii) powerful symbolic learning and reasoning tools, that utilize such features to synthesize high-level, interpretable patterns of critical situations to be forecast. Crucial in the EVENFLOW approach is the online nature of the learning methods, which makes them applicable to evolving data flows and allows to utilize rich domain knowledge that is becoming available progressively, over time. To deal with the brittleness of neural predictors and the high volume/velocity of temporal data flows, the EVENFLOW techniques will rely on novel, formal verification techniques for machine learning, in addition to a suite of scalability algorithms for federated training and incremental model construction. The learnt forecasters will be interpretable and scalable, allowing for fully explainable insights, delivered in a timely fashion and enabling proactive decision making. EVENFLOW will be evaluated on three challenging use cases related to oncological forecasting in precision medicine, safe and efficient behavior of autonomous transportation robots in smart factories and reliable life cycle assessment of critical infrastructure.



ARCADIA - Autonomous Resource Allocation for Edge Infrastructures

Project manager	George Papastefanatos
Coordinator	Intrakom-Telecom S.A.
IMSI - funding	210.770,00€
Programme	Greece 2.0 - National Recovery and Resiliency Plan
Start date	01/07/2022
Duration	24 Months
Website	https://arcadia-project.gr/

The optimization of resource allocation in cloud computing environments is a crucial problem with research interest and direct application to a multitude of commercial applications. A prominent class of applications that strongly emphasize this problem pertain to the Network Function Virtualization (NFV) paradigm, which promotes the implementation of typical network functions, such as firewalls, load balancers, and intrusion detection systems, in the form of virtual machines or containers. This obviates the need for hardware-based network function deployments, effectively rendering NFV-enabled networks more scalable, reliable, and cost-effective. However, to fill up its promise, NFV shall incorporate sophisticated mechanisms that can promptly compute optimized resource allocations, subject to volatile user demands and energy consumption constraints. Along these lines, the main objective of ARCADIA is to investigate, design, and evaluate innovative computational frameworks and methods for optimized resource allocation in cloud computing environments. Concretely, ARCADIA focuses on a) systems exhibiting dynamic workload characteristics, and b) environments with high energy consumption requirements due to simultaneous and continuous operation of computer clusters and equipment. Both of these features can be found in edge systems, and specifically in edge data centers, which have become a pivotal computing part of next-generation networks.



Check4Facts \ Science

Project manager	George Papastefanatos
Coordinator	National Centre of Social Research
IMSI - funding	-
Programme	-
Start date	01/09/2022
Duration	12 Months
Website	http://check4facts.gr/

Check4Facts\Science is a self-funded project between 6 Research Centers (ATHENA Research Center, National Centre for Social Research, National Observatory of Athens, Hellenic Centre for Marine Research, Biomedical Research Foundation of the Academy of Athens, Institute of Communication and Computer Systems) and 2 Universities (Department of Communication and Media Studies of the National and Kapodistrian University of Athens, Communication and Digital Media Department of the University of Western Macedonia) in Greece. Check4facts aims to study fact-checking in the Greek public sphere and the media, in relation to news for climate change and public health \ pandemic. The Check4facts platform combines the automation of Machine Learning (ML) techniques, with the expertise of fact-checkers (researchers and scientific personnel from the consortium), to support a thorough and trustworthy workflow for fact checking and credibility assessment of news items. The results, realized as detailed assessment reports, are published to the general public via the Check4facts portal.



AutoFair - Human-Compatible Artificial Intelligence with Guarantees

Project manager	Ioannis Z. Emiris
Coordinator	Czech Technical University
IMIS - funding	496,875 euros
Programme	HORIZON-CL4-2021-HUMAN-01-01
Start date	1/10/2022
Duration	3 years
Website	https://humancompatible.org/

AutoFair deals with the design of fair, explainable and transparent AI algorithms. We address the matter of transparency and explainability of AI using approaches inspired by control theory. Notably, we consider a comprehensive and flexible

certification of properties of AI pipelines, certain closed loops and more complicated interconnections. At one extreme, one could consider risk averse a priori guarantees via hard constraints on certain bias measures in the training process. At the other extreme, one could consider nuanced communication of the exact tradeoffs involved in AI pipeline choices and their effect on industrial and bias outcomes, post hoc. Both extremes offer little in terms of optimizing the pipeline and inflexibility in explaining the pipeline's fairness-related qualities. Seeking the middle-ground, we suggest a priori certification of fairness-related qualities in AI pipelines via modular compositions of pre-processing, training, inference, and post-processing steps with certain properties. Furthermore, we present an extensive programme in explainability of fairness-related qualities. We seek to inform both the developer and the user thoroughly in regards to the possible algorithmic choices and their expected effects. Overall, this will effectively support the development of AI pipelines with guaranteed levels of performance, explained clearly. Three use cases (in Human Resources automation, Financial Technology, and Advertising) will be used to assess the effectiveness of our approaches.



MORE **Management of Real-time Energy Data**

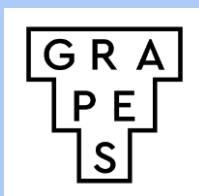
Project manager	Manolis Terrovitis
Coordinator	IMSI
IMIS - funding	886K Euro
Programme	H2020-ICT-2018-20 / H2020-ICT-2020-1
Start date	1-10-2020
Duration	39 months
Website	https://www.more2020.eu/

The widespread use of sensor and IoT devices is generating huge volumes of time series data in various industries like finance, energy, factories, medicine, manufacturing and others. Industries use these data for monitoring, but their main potential is still untapped. Existing techniques and software for time series management do not provide tools sufficiently scalable and sophisticated for managing the huge volumes of data or adequate forecasting, prediction and diagnostics.

MORE will create a platform that will address the technical challenges in time series and stream management, focusing on the RES industry. MORE's platform

will introduce an architecture that combines edge computing and cloud computing to be able to guarantee both responsiveness and provide sophisticated analytics simultaneously. This architecture will be combined with the usage of time series summarization techniques, or as we more accurately term them in MORE, modelling techniques for sensor data. Models are any compressed representations that allow the reconstruction of the original data points of a time series (e.g. a linear function) within a known error-bound (possibly zero). This approach has synergies with the edge computing approach, since summarization can be done at the edge, reducing the load in the whole data processing pipeline.

MORE will introduce advanced analytics tools for prediction, forecasting and diagnostics based on two technological directions: machine learning and pattern extraction, with emphasis to motifs, which is the state-of-the-art for time series. MORE will adjust these techniques to work directly on models of data, thus enabling them to scale beyond state-of-the-art. The ability to ingest huge volumes of data will have an important impact to the accuracy of the prediction and diagnostics models.



GRAPES - Learning, processing and optimising shapes

Project manager	Ioannis Z. Emiris
Coordinator	IMIS
IMIS - funding	486,035 euros
Programme	Marie Skłodowska-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	UNIFI
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.



**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.



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.



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 crowdsourcing module to collect citizen reports of selected violations, a crawler to detect and correlate online deviant behaviour, and a toolkit for investigators that will assist them in collecting online and offline evidence. This will be powered by an immutable storage and an identity management system that will protect privacy and handle access to evidence data using a Trusted Execution Environment. Blockchain technology will not only guarantee that information about the evidence cannot be tampered with, but allow interoperability without the need for a trusted third party.



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	13/10/2017
Duration	4.5 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 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, infrastructures, 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, infrastructures, 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.



DEDS

Data Engineering for Data Science

Project manager	Alkis Simitsis
Coordinator	-
IMIS - funding	685 KEuros
Programme	H2020-MSCA-ITN-2020
Start date	01-03-2021
Duration	36 months
Website	https://deds.ulb.ac.be/

Data is a key asset in modern society. Data Science, which focuses on deriving valuable insight and knowledge from raw data, is indispensable for any economic, governmental, and scientific activity. Data Engineering provides the data ecosystem (i.e., data management pipelines, tools and services) that makes Data Science possible. The European Joint Doctorate in "Data Engineering for Data Science" (DEDS) is designed to develop education, research, and innovation at the intersection of Data Science and Data Engineering. Its core objective is to provide holistic support for the end-to-end management of the full lifecycle of data, from capture to exploitation by data scientists. DEDS operates under the Horizon 2020 - Marie Skłodowska-Curie Innovative Training Networks (H2020-MSCA-ITN-2020) framework. It is jointly organised by Université Libre de Bruxelles (Belgium), Universitat Politècnica de Catalunya (Spain), Aalborg Universitet (Denmark), and the Athena Research and Innovation Centre (Greece). Partner organisations from research, industry and the public sector prominently contribute to the programme by training students and providing secondments in a wide range of domains including Energy, Finance, Health, Transport, and Customer Relationship and Support. DEDS is a 3-year doctoral programme based on a co-tutelle model. A complementary set of 15 joint, fully funded, doctoral projects focus on the main aspects of holistic management of the full data lifecycle. Each doctoral project is co-supervised by two beneficiaries and includes a secondment in a partner organisation, which grounds the research in practice and validate the proposed solutions. DEDS delivers innovative training comprising technical and transversal courses, four jointly organized summer and winter schools, as well as dissemination activities including open science events and a final conference.

Upon graduation, a joint degree from the universities of the co-tutelle will be awarded.



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.



CLIMPACT – Flagship Initiative for Climate Change

Project coordinator	Alkis Simitsis
Coordinator	National Observatory of Athens
IMSI - funding	115.1 KEuros
Programme	GSRT project
Start date	2019
Duration	3 years
Website	-

Climpact is a flagship initiative on climate change to coordinate a Pan-Hellenic network of institutions responsible for the integration, harmonization, and optimization of existing climate services, early warning systems and measurements from relevant national infrastructures in Greece. 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.



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	3.5 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.

Digital Landscape



The emerging landscape of digital work practices in the Humanities in the context of the European projects DARIAH and CLARIN

Project manager	Panos Constantopoulos
Coordinator	IMSI - Athena RC
IMIS - funding	364,999 Euros
Programme	HFRI
Start date	01/04/2022
Duration	2 years
Website	https://digital-landscape.gr/

Digital research infrastructures play a catalytic role in the digital transformation of research. The Greek Infrastructure for Digital Arts, Humanities and Language Research and Innovation “APOLLONIS” and the European research infrastructures DARIAH for the Arts and Humanities and CLARIN for Language Resources and Technology, to which APOLLONIS is connected, are important facilitating agents of this transformation in Greece and in Europe respectively. The recent pandemic and the restrictions it imposed brought about new needs in Humanities research and pedagogy, introducing new modi operandi, accelerating the digital transformation, but also demonstrating the gaps that still remain. In this emerging new landscape, the proposed project seeks to identify, analyze and model current trends of digital work practices in the Humanities in Greece and the possible impact of the pandemic on them, and support the penetration and consolidation of these practices using services of the APOLLONIS infrastructure and its synergies with DARIAH and CLARIN. The tools to collect evidence-based information will be a communities web-survey and focus groups, while the analysis of the collected data will be used to inform the planning of the next phase of APOLLONIS infrastructure, ensuring digital services remain timely and state-of-the-art. The project will also develop an ontology-based semantic representation of digital work processes in the Humanities and Social Sciences and provide workflow models that could serve as application guides. Training activities targeted to various communities will be launched to promote the use of digital methods, tools and practices in Humanities research in Greece and all resources will be made available online. Finally, the project will ensure uninterrupted collaboration with the European infrastructures, DARIAH and CLARIN.

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 big data solution and methods for stream analytics on network data

Project manager	George Papastefanatos
Coordinator	Intracom Telecom S.A.
IMIS - funding	-
Programme	Contract
Start date	16/10/2017

Duration 62 months

Website -

A new collaboration between **IMSI, Intracom Telecom and Ericsson** started in *October 2017*. IMSI has been contracted to design , develop and maintain 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 IoT devices, their scalable processing, analysis and storage and the visualization of several KPIs.



eLib GGDE - eLib of Independent Authority for Public Revenue

Project coordinator George Papastefanatos

Coordinator Remaco S.A.

IMSI - funding -

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



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	01-03-2021
Duration	12 Months
Website	https://www.imsi.athenarc.gr/el/projects/comproject/30

A joint R&D project between Athena RC and ANSYS Hellas, that is focused on: (a) spatial indexing data structures for facilitating region queries in the context of random walk-based capacitance extraction, (b) enhancement, and extension of the solutions for optimal insertion order of VLSI elements into the spatial index and (c) evaluation of data-driven data structure for Max Empty Cube queries. We develop prototype code and evaluate it experimentally on ANSYS real data in simulating VLSI designs for identifying defects before fabrication. The project's members shall benefit from both top-notch research as well as a strong innovation component through interaction with the highly qualified ANSYS Hellas R&D personnel.



PREFERRED - PREVENTING FIRE EVENTS BY REDISCOVERING AND EXTENDING DEEP LEARNING METHODS

Project manager	Giorgos Giannopoulos
Coordinator	GEOAPIKONISIS S.A.
IMIS - funding	95,000 euros
Programme	Synergies for Research and Innovation in the Prefecture of Attica
Start date	8/2022
Duration	17 months

PREFERRED develops novel ML and DL methods for the next-day fire risk prediction, as well as smart analytics and visualization interfaces on top of them. The project will incorporate its results into decision support platform that will be directly exploitable by disaster management stakeholders (fire service, civil protection, NGOs) in the prevention and management of fire events.

Other Projects



Digital Curation Services

Project manager	Panos Constantopoulos
Coordinator	IMSI
IMIS - funding	94.109,12 euros
Programme	Internal project
Start date	1/2/2016
Duration	7 years
Website	http://www.dcu.gr/en/υπηρεσίες-ψηφιακής-επιμέλειας/

The internal project Digital Curation Services supports the advancement of research and development efforts in the research directions and work programme of IMSI by using chiefly funds remaining from the execution of service provision projects. In view of parallel externally funded projects, this project on one hand ensures the continuity of our work programme, while on the other it enables actions of topical interest for which external funds may be insufficient at the time. It also allows better integrating and employing results of completed projects in sustained and new digital curation services, thus supporting the capitalization of those results.

Publications

Books

- Benardou, A., Droumpouki, A.M. (Eds.). (2022). **Difficult Heritage and Immersive Experiences** (1st ed.). Routledge. <https://doi.org/10.4324/9781003200659>, 2022

Book chapters

- Nikos Giatrakos, Antonios Deligiannakis, Konstantina Bereta, Marios Vodas, Dimitris Zisis, Elias Alevizos, Charilaos Akasiadis, Alexander Artikis: **Processing Big Data in Motion: Core Components and System Architectures with Applications to the Maritime Domain**. Technologies and Applications for Big Data Value 2022: 497-518
- Constantopoulos, P., Dritsou, V., Ilvanidou, M., Chroni, A. **Aggregation and Curation of Historical Archive Information**. In: Karagiannis D., Lee M., Hinkelmann K., Utz W. (eds) Domain-Specific Conceptual Modeling. Springer, Cham. https://doi.org/10.1007/978-3-030-93547-4_23, 2022.

Journal Publications

- Panagiotis Tampakis, Eva Chondrodima, Andreas Tritsarolis, Aggelos Pikrakis, Yannis Theodoridis, Kostis Pristouris, Harry Nakos, Panagiotis Kalampokis, Theodore Dalamagas: **i4sea: A Big Data Platform for Sea Area Monitoring and Analysis of Fishing Vessels Activity**. Geo spatial Inf. Sci. 25(2): 132-154 (2022)
- Serafeim Chatzopoulos, Thanasis Vergoulis, Theodore Dalamagas, Christos Tryfonopoulos: **VeTo+: improved expert set expansion in academia**. Int. J. Digit. Libr. 23(1): 57-75 (2022)
- A. Chalkis, I.Z. Emiris, V. Fisikopoulos, P. Repouskos, and E. Tsigaridas. **Efficient sampling in spectrahedra and volume approximation**. Linear Algebra & Appl., 648:204–232, 2022. <https://doi.org/10.1016/j.laa.2022.04.002>.
- E. Christoforou, H. Leontiadou, F. Noé, J. Samios, I.Z. Emiris, and Z. Cournia. **Investigating the bioactive conformation of Angiotensin II using hidden Markov state modeling revisited with web-scale clustering**. J. Chemical Theory & Computation, 18(9):5636–5648, 2022.
- E. Bartzos, I.Z. Emiris, and R. Vidunas. **New upper bounds for the number of embeddings of minimally rigid graphs**. Discrete & Computational Geometry, 68(3):796–816, 3 2022

- Alexis Apostolakis, Stella Girtsou, Giorgos Giannopoulos, Nikolaos S. Bartsotas, Charalampos Kontoes: **Estimating Next Day's Forest Fire Risk via a Complete Machine Learning Methodology**. Remote. Sens. 14(5): 1222 (2022)
- George Papastefanatos, Marios Meimaris, Panos Vassiliadis: **Relational schema optimization for RDF-based knowledge graphs**. Inf. Syst. 104: 101754 (2022)
- Nikos Bikakis, Panos K. Chrysanthis, George Papastefanatos, Tobias Schreck: **Special Issue on Machine Learning Approaches in Big Data Visualization**. IEEE Computer Graphics and Applications 42(3): 39-40 (2022)
- George Stamatakis, Antonis Kontaxakis, Alkis Simitsis, Nikos Giatrakos, Antonios Deligiannakis: **SheerMP: Optimized Streaming Analytics-as-a-Service over Multi-site and Multi-platform Settings**. EDBT 2022: 2:558-2:561
- Vasileios Stavropoulos, Elias Alevizos, Nikos Giatrakos, Alexander Artikis: **Optimizing complex event forecasting**. DEBS 2022: 19-30
- Alexandros Zeakis, Dimitrios Skoutas, Dimitris Sacharidis, Odysseas Papapetrou, Manolis Koubarakis: **TokenJoin: Efficient Filtering for Set Similarity Join with Maximum Weighted Bipartite Matching**. Proc. VLDB Endow. 16(4): 790-802 (2022)
- E. Pitoura, K. Stefanidis, **G. Koutrika**. Fairness in Rankings and Recommendations: An Overview, The VLDB J. 31(3): 431-458, 2022

International Conference / Workshop Publications

- Anargiros Tzerefos, Ilias Kanellos, Serafeim Chatzopoulos, Theodore Dalamagas, Thanasis Vergoulis: **SurvAnnT: Facilitating Community-Led Scientific Surveys and Annotations**. TPDJ 2022: 543-547
- Electra Athanatsiki, Ifigeneia-Maria Tsioutsia, Harry Nakos, Panagiotis Kalampokis, Theodore Dalamagas: **Urbana: An Innovative Platform for Collective Awareness and Enhancement of Urban Agriculture**. HAICTA 2022: 162-163
- Vasilis Gkolemis, Theodore Dalamagas, Christos Diou: **DALE: Differential Accumulated Local Effects for efficient and accurate global explanations**. ACML 2022
- C. Checa and I.Z. Emiris. **A greedy approach to the Canny-Emiris formula**. In Proc. Annual ACM Intern. Symp. on Symbolic and Algebraic Computation (ISSAC), pages 283–291, 2022. <https://doi.org/10.1145/3476446.3536180>.
- E. Bartzos, I.Z. Emiris, I.S. Kotsireas, and C. Tzamos. **Bounding the number of roots for multi-homogeneous systems**. In Proc. Annual ACM Intern. Symp. on Symbolic and Algebraic Computation (ISSAC), pages 255–262, Lille, France, July 2022. <https://doi.org/10.1145/3476446.3536189>.
- Seshu Tirupathi, Dhaval Salwala, Giulio Zizzo, Ambrish Rawat, Mark Purcell, Søren Keijser Jensen, Christian Thomsen, Nguyen Ho, Carlos E. Muñoz-Cuza, Jonas Brusokas, Torben Bach Pedersen, Giorgos Alexiou, Giorgos Giannopoulos, Panagiotis Gidararakos, Alexandros Kalimeris, Stavros Maroulis, George

Papastefanatos, Ioannis Psarros, Vassilis Stamatopoulos, Manolis Terrovitis:

Machine Learning Platform for Extreme Scale Computing on Compressed IoT Data. Big Data 2022: 3179-3185

- Suela Isaj, Vassilis Kaffes, Torben Bach Pedersen, Giorgos Giannopoulos: **A Supervised Skyline-Based Algorithm for Spatial Entity Linkage.** EDBT 2022: 2:220-2:233
- Voula Giouli, Anna Vacalopoulou, Nikolaos Sidiropoulos, Christina Flouda, Athanasios Doupas, Giorgos Giannopoulos, Nikos Bikakis, Vassilis Kaffes, Gregory Stainhaouer: **Placing multi-modal, and multi-lingual Data in the Humanities Domain on the Map: the Mythotopia Geo-tagged Corpus.** LREC 2022: 2856-2864
- George Papastefanatos, Giorgos Alexiou, Nikos Bikakis, Stavros Maroulis, Vassilis Stamatopoulos: **VisualFacts: A Platform for In-Situ Visual Exploration and Real-Time Entity Resolution.** In EDBT/ICDT Workshops 2022
- Kostas Patroumpas, Dimitrios Skoutas, Dimitris Sacharidis: **MAGE: Discovering Mixture-based Areas of Interest over Geolocated Entities.** EDBT 2022: 2:574-2:577
- Xiaoying Wu, Dimitri Theodoratos, Dimitrios Skoutas, Michael Lan: **Efficient In-Memory Evaluation of Reachability Graph Pattern Queries on Data Graphs.** DASFAA (1) 2022: 55-71
- G. Katsogiannis-Meimarakis, G. Koutrika. **Data Democratisation with Deep Learning: An Analysis of Text-to-SQL Systems,** TheWebConf 2022
- Konstantinos Theocharidis , Manolis Terrovitis, Spiros Skiadopoulos, Panagiotis Karras: **A Content Recommendation Policy for Gaining Subscribers.** SIGIR 2022: 2501-2506
- Yannis E. Foufoulas, Alkis Simitsis, Eleftherios Stamatogiannakis, Yannis E. Ioannidis: **YeSQL: "You extend SQL" with Rich and Highly Performant User-Defined Functions in Relational Databases.** Proc. VLDB Endow. 15(10): 2270-2283 (2022)
- Yannis E. Foufoulas, Alkis Simitsis, Yannis E. Ioannidis: **YeSQL: Rich User-Defined Functions without the Overhead.** Proc. VLDB Endow. 15(12): 3730-3733 (2022)
- George Stamatakis, Antonis Kontaxakis, Alkis Simitsis, Nikos Giatrakos, Antonios Deligiannakis: **SheerMP: Optimized Streaming Analytics-as-a-Service over Multi-site and Multi-platform Settings.** EDBT 2022: 2:558-2:561
- Dimitris Tsesmelis, Alkis Simitsis: **Database Optimizers in the Era of Learning.** ICDE 2022: 3213-3216
- Kleanthis Vichos, Michele De Bonis, Ilias Kanellos, Serafeim Chatzopoulos, Claudio Atzori, Natalia Manola, Paolo Manghi, Thanasis Vergoulis. **A preliminary assessment of the article deduplication algorithm used for the OpenAIRE Research Graph.** IRCDL 2022
- Thanasis Vergoulis, Serafeim Chatzopoulos, Kleanthis Vichos, Ilias Kanellos, Andrea Mannocci, Natalia Manola, Paolo Manghi. **BIP! Scholar: A Service to Facilitate Fair Researcher Assessment.** JCDL 2022, 2022

Other Publications / Technical Reports

- Giorgos Alexiou, George Papastefanatos, Vassilis Stamatopoulos, Georgia Koutrika, Nectarios Koziris: **QueryER: A Framework for Fast Analysis-Aware Deduplication over Dirty Data**. arXiv:2202.01546 (2022)
- Kakouti, G., Dritsou, V., Benardou, A. (2022) ‘**Modeling historical storytelling**’. Storytelling - DARIAH Annual Event 2022, Zenodo, 28 May. doi: 10.5281/zenodo.6589728
- Sihem Amer-Yahia, Yael Amsterdamer, Sourav S. Bhowmick, Angela Bonifati, Philippe Bonnet, Renata Borovica-Gajic, Barbara Catania, Tania Cerquitelli, Silvia Chiusano, Panos K. Chrysanthis, Carlo Curino, Jérôme Darmont, Amr El Abbadi, Avriella Floratou, Juliana Freire, Alekh Jindal, Vana Kalogeraki, Georgia Koutrika, Arun Kumar, Sujaya Maiyya, Alexandra Meliou, Madhulika Mohanty, Felix Naumann, Nele Sina Noack, Fatma Özcan, Liat Peterfreund, Wenny Rahayu, Wang-Chiew Tan, Yuanyuan Tian, Pinar Tözün, Genoveva Vargas-Solar, Neeraja J. Yadwadkar, Meihui Zhang: **Diversity and Inclusion Activities in Database Conferences: A 2021 Report**. SIGMOD Rec. 51(2): 69-73 (2022)

Dissemination Activities

Invited / Keynote Talks

IMSI members participated in the following invited / keynote talks:

- **G. Papastefanatos, November 2022.** “Onboarding and Integration of services and resources in the EOSC Future platform”. In European Open Science Cloud (EOSC) Symposium, 16 November 2022, Prague.
- **G. Papastefanatos, July 2022.** “Introducing the Check4facts platform and machine learning technologies to support public statements verification”, In 1st International Scientific one-day conference on fact-checking in Greece “Tackling Misinformation: The Contribution of Universities and Research Centers”. 04 July 2022, Athens.
- **G. Papastefanatos, April 2022.** “The Providers’ portal of the European Open Science Cloud”. In EOSC Core and Marketplace Providers’ Days Event, 26-27 April 2022, Virtual Event
- **Dritsou, V.** The Greek Landscape of Digital Humanities Initiatives. Fiesole Retreat 2022, 5 April, Athens, Greece.
- Georgia Koutrika gave a keynote on “Democratizing Data Access: What if we could just talk to our data?” in the International Conference on Theory and Practice of Digital Libraries, TPDL 2022, Padua, September 20 – 23, 2022
- Georgia Koutrika gave a talk on “Fairness in Algorithmic Systems: A Reality or a Fantasy?” as part of the Distinguished Lecture Series Talk, Hong Kong Baptist University, May 3, 2022
- Georgia Koutrika gave a talk on “Intelligent Data Assistants: Democratizing Data Access”, as part of the TU Berlin/BIFOLD Research Colloquium, April 4, 2022.
- Georgia Koutrika gave a talk on “Data Democratisation with Deep Learning: An Analysis of Text-to-SQL Systems”, in the RAIS Summer School and Workshop, Sweden, September 12-16, 2022
- Georgia Koutrika gave a talk on "Algorithmic Fairness", in the PhD-summer school 2022 on Big Data, Black Forest, Germany, June 20-24, 2022
- Georgia Koutrika gave a talk on “Fairness in Algorithmic Systems: A Reality or a Fantasy?”, in the 1st DEDSSchool on “Ethical and Legal Aspects on Data”, April 4 – 8, 2022

- Georgia Koutrika gave a talk on “Fairness in Ranking and Recommenders: An Overview” 1st ACM-W Greek Chapter Winter School on "Fairness in AI" (with E. Pitoura and K. Stefanidis), February 24-25, 2022.

Scientific Community Service

IMSI members have served in the **Program Committee** of more than 25 International Conferences and Workshops in 2022, including well-known Conferences like SIGMOD, VLDB, ICDE, EDBT, WISE, IEEE Big Data, CIDR, DATA, PVLDB, EDBT, SIGSPATIAL, and more.

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

- **G. Papastefanatos**, 5th International Workshop on Big Data Visual Exploration and Analytics (BigVis2022) in conjunction with EDBT/ICDT 2022, March 29 2022 , Edinburgh, UK [Online]
- **G. Papastefanatos**, Guest Editor in IEEE Computer Graphics and Applications. Special Issue on Machine Learning Approaches in Big Data. 2022.
- **Georgia Koutrika** is the VLDB 2023 PC co-chair, and PVLDB Vol 16 EiC.



- **MAGE**

MAGE is an open-source tool for mixture-based best region search over geolocated entities of different types, like Points of Interest, geotagged posts or photos. MAGE detects the top- k areas of arbitrary shapes exhibiting high or low mixture patterns. Through a graphical interface, users can specify their preferences, execute the selected algorithm, and visually inspect the results on the map to unveil interesting patterns.

<https://openproceedings.org/2022/conf/edbt/paper-147.pdf>



- **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**
Temporal

Preferences
Select Date

Weights
0 0.25 0.5 0.75 1

ii) Metapath preferences

☒ **Metapath**
Article - Organisation

Preferences on target entities
twitter x

☒ Article - Person

Weights
0 0.25 0.5 0.75 1


0 0.25 0.5 0.75 1

3. Select value for k

50

Search

Results

#	Article	Score	Article - Organisation	Article - Person
1	20190520_filtered.csv;20190520200000-1330 	0.74000000000000017	huawei, twitter, cnn, white house, facebook, united states score: 3.3382377953649984e-15	barack obama, donald trump score: 1

- **spaTScope**

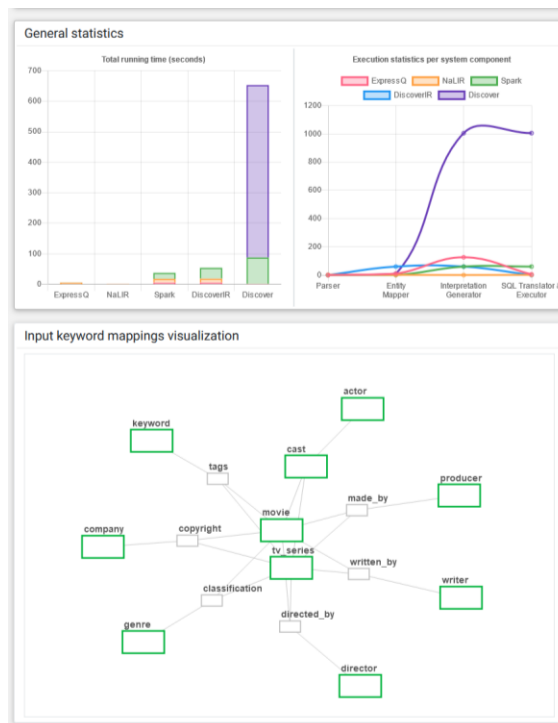
spaTScope is a web application for visual exploration of geolocated time series developed in the context of the EU-funded projects SLIPO and SmartDataLake. spaTScope allows users to visually explore large collections of geolocated time series and obtain insights about trends and patterns in their area 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.

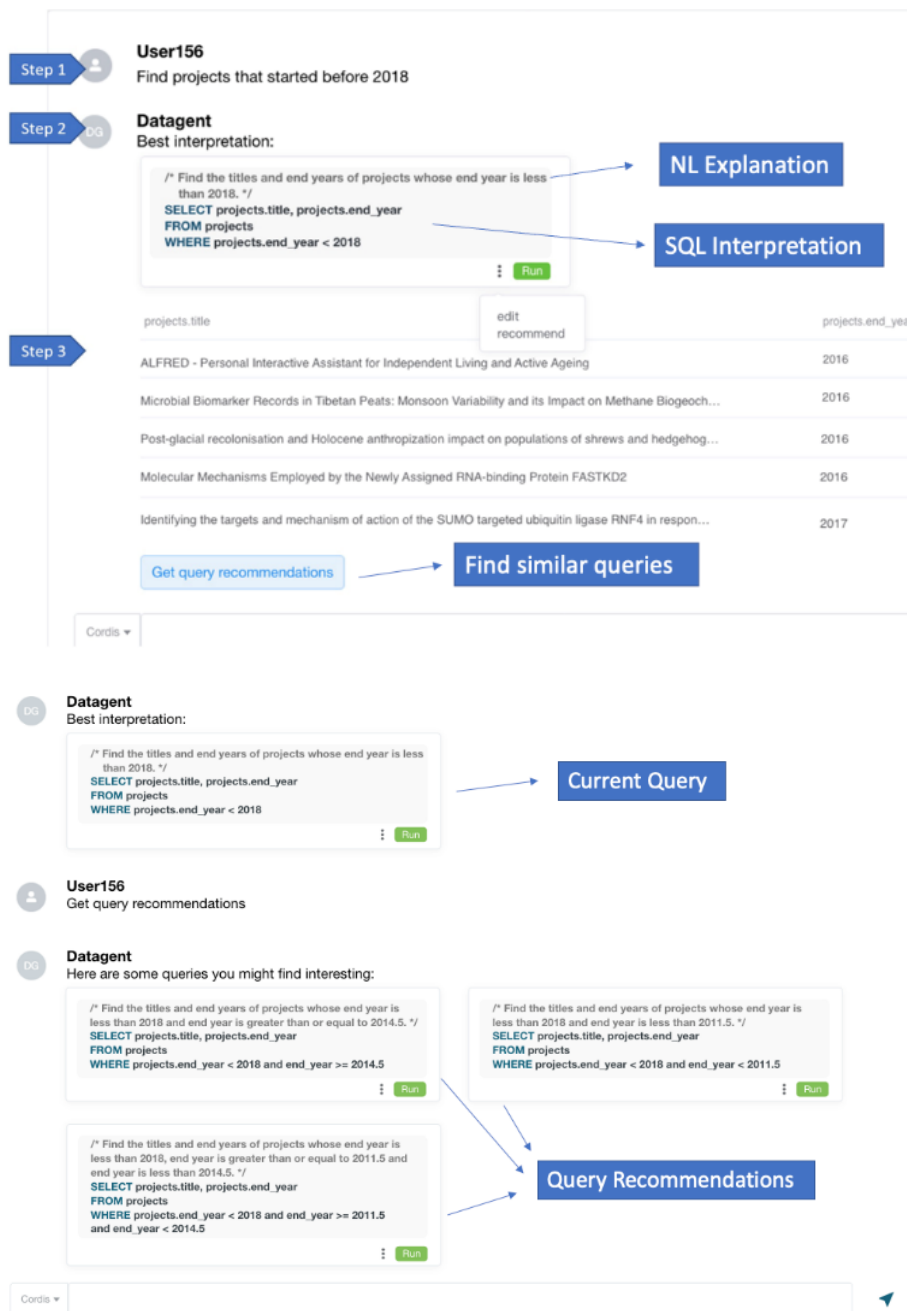


<https://darelab.imsi.athenarc.gr/thor/home>

- **DatAgent**

DatAgent 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.

<https://darelab.imsi.athenarc.gr/datagent/>



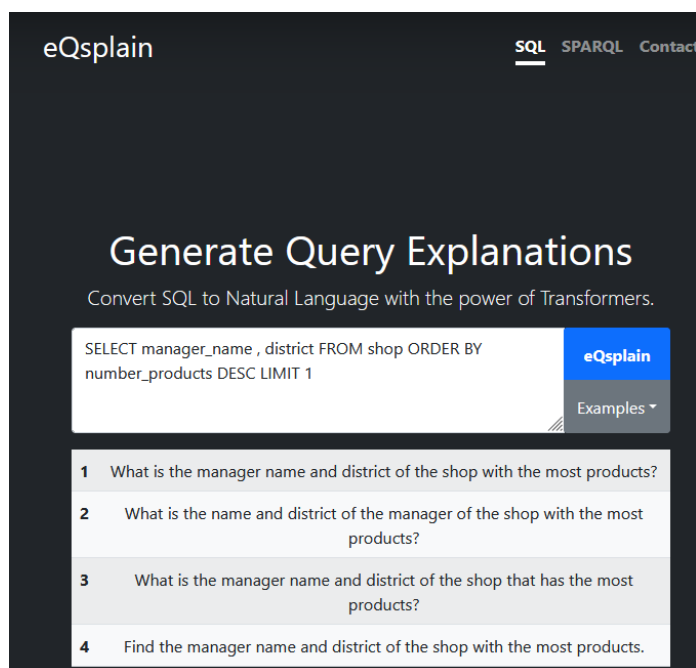
More information about Datagent can be found here:
<http://vldb.org/pvldb/vol14/p2815-madamadiotis.pdf>

- **eQsplain**

In the age of the data revolution, more and more human activities, from business operations to medical research are dependent on data. At the same time, immense amounts of information are stored in relational databases, which remain inaccessible without the knowledge of certain query languages is required. While many research efforts are being made towards the translation of text to SQL, this does not suffice in order to create a complete Natural Language Interface (NLI) for Databases. The reverse process (SQL-to-text) is of equal importance in order for the user to seamlessly use a

database with only their native language. The system is based on the T-5 Transformer-based Pre-trained Language Model, taking advantage of its capabilities to generate fluent and human-like natural language. eQsplain can learn to generate explanations for multiple query languages, such as SQL and SPARQL.

<https://darelab.athenarc.gr/eqsplain/sql>



- **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**

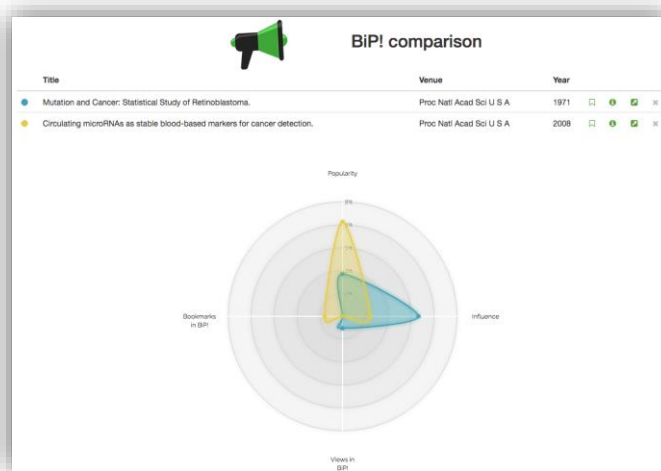
DAIAD is the first *integrated residential demand management* system for water. It applies Big Data and Machine Learning technologies to leverage smart water meter data, engaging and informing consumers to induce sustainable changes in consumption behaviour, as well as providing novel large-scale analytics 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>

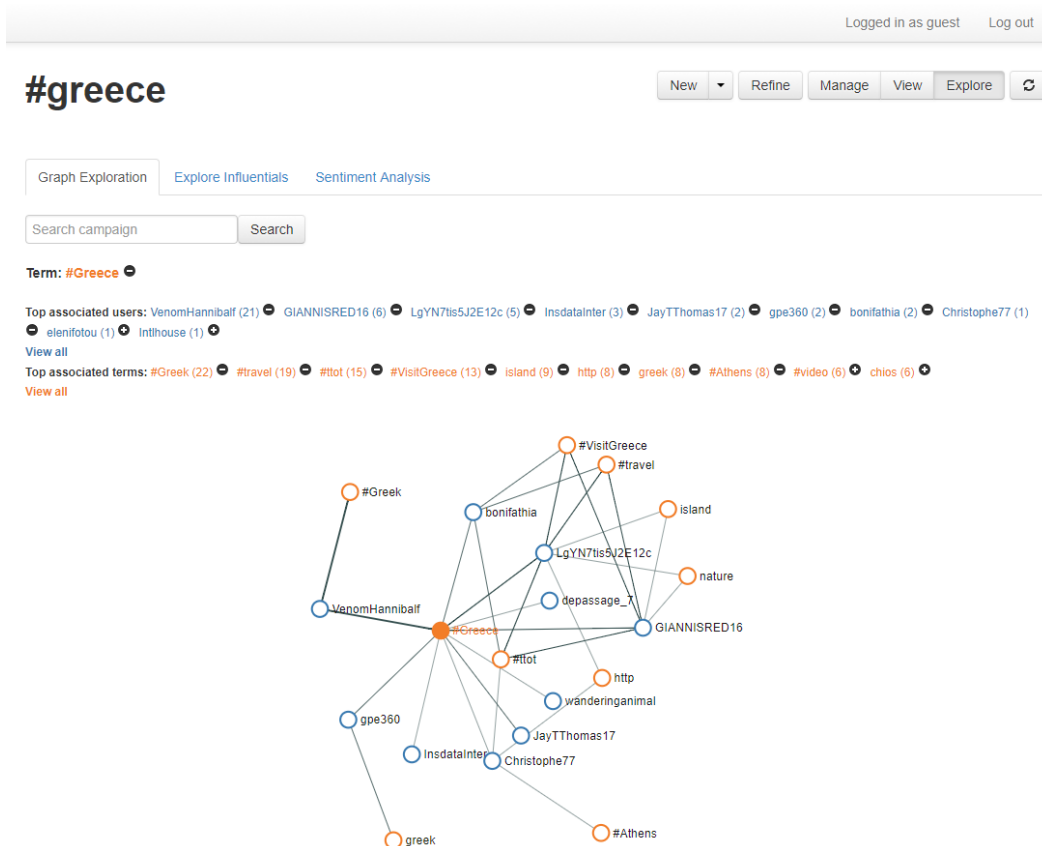
- **YeSQL**

YeSQL is an SQL extension that provides more usable, more expressive, and more performant Python UDFs and can be integrated into both server-based and embedded DBMSs. It enriches SQL with a functional syntax that unifies the expression of relational and user-defined functionality and optimizes the execution of both in a seamless fashion, assigning processing tasks to the DBMS or the UDF host language VM accordingly and employing efficient low-level implementation techniques. Key characteristics of the YeSQL language that enhance usability and expressiveness include (a) stateful, parametric, and polymorphic UDFs, (b) dynamically typed UDFs, (c) scalar and aggregate UDFs returning arbitrary table forms, and (d) UDF pipelining. Key performance characteristics include (a) seamless data exchange between the UDF and the DBMS, (b) JIT-compiled UDFs, (c) UDF parallelization, (d) stateful UDFs, and (e) UDF fusion. YeSQL is designed to work in synergy with existing systems. It is fully implemented on top of SQLite API, originally introduced in SQLite, but there is also a prototype implementation on top of a read optimised server database (MonetDB).

<https://github.com/athenarc/YeSQL>

- **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>

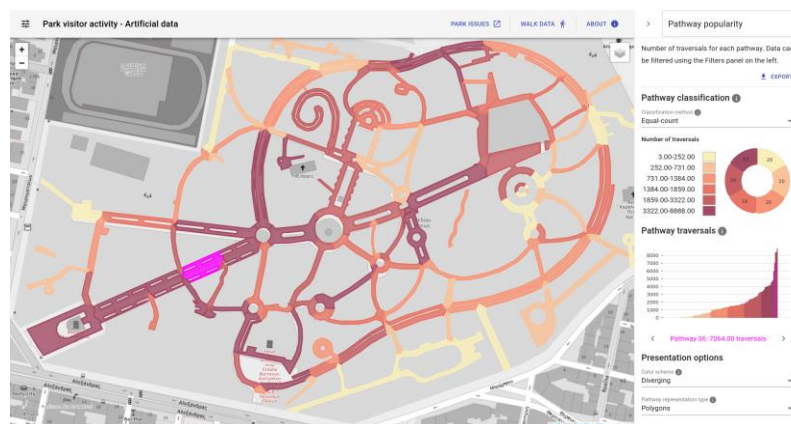
- **eLib**

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

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

- **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

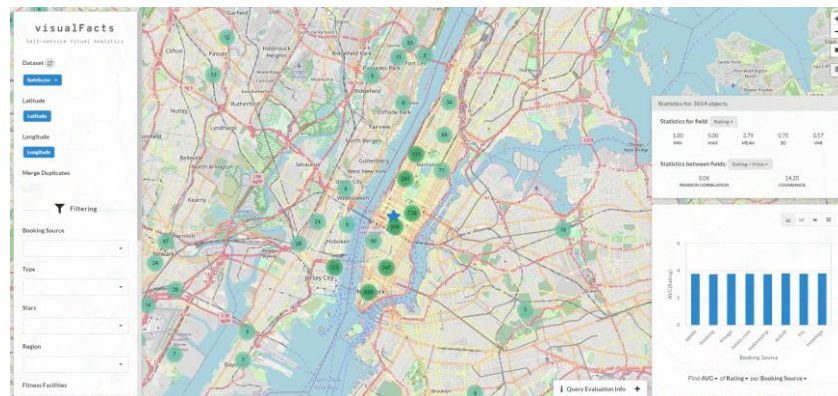
through faceted browsing, hierarchical representation of coded lists in charts, free keyword search of literal values, and capabilities for combining different datasets along common dimensions. Moreover, it makes knowledge reusable by making all data available for download in various formats including 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 users who are not IT experts to visually explore the data and the impact of different anonymization settings on them. It helps users 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 2022 several MSc and Diploma thesis have been co-supervised by IMSI members, who often serve as members in the respective examination committees.

IMIS members co-supervise MSc students. The following MSc dissertations have been completed in 2022:

- Panagiotis Leontaridis. Topic: **“Knowledge base for work practices: Text pre-processing and annotation for machine learning.”** Collaborating researchers: [Panos Constantopoulos](#), Vayianos Pertsas
- Kyriaki Hampilidou. Topic: **“Knowledge base for work practices: adaptation of knowledge extraction rules”**. Collaborating researchers: [Panos Constantopoulos](#), Vayianos Pertsas
- Eleni Makrygiorgou. Topic: **“Knowledge base for work practices: creation of training and evaluation sets for machine learning.”** Collaborating researchers: [Panos Constantopoulos](#), Vayianos Pertsas
- Sofia N. Karvounari, **Management of Scientific Analysis and Simulation Workflows over High Performance Computing Systems**. National and Kapodistrian University of Athens, Athens, Greece, May 2022. Collaborating researcher: [Alkis Simitsis](#).

The following PhD and MSc students collaborated closely with / were supervised by IMSI members in their research during 2022:

- Alexandros Zeakis. Topic: **Similarity Joins with multiple matching criteria**. Joint supervision with the University of Athens. Collaborating researcher: Dimitris Skoutas.
- 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
- Stavroula Eleftheraki, Phd Student. Topic: **Fairness in Recommendations**. Joint supervision with the University of Grenoble. Supervisor from IMSI: Georgia Koutrika.

- Antonis Mandamadiotris, Phd Student. Topic: **Intelligent Interactive Data Exploration**. Joint supervision with the University of Grenoble. Supervisor from IMSI: Georgia Koutrika.
- Christos Tsapelas, MSc student (University of Athens). Topic: **Deep Learning for Query Optimization** - supervisor: Georgia Koutrika
- George Katsogiannis-Meimarakis, MSc student (University of Athens). Topic: **Deep Learning for NL2SQL and SQL2NL Translation**. Supervisor: Georgia Koutrika
- Mike Xydas, MSc student (University of Athens). Topic: **Verbalising Query Results to Text**. Supervisor: Georgia Koutrika
- Anna Mitsopoulou, MSc student (University of Athens). Topic: **NL-to-SQL Query Generation**. Supervisor: Georgia Koutrika
- Ismimi Bouliari, MSc student (University of Athens). Topic: **Hybrid Recommendations for fighting Popularity Bias**. Supervisor: Georgia Koutrika
- 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: Minos Garofalakis.
- Eros Fabrici. **Privacy-preserving Data Integration**. Collaborating researcher: Minos Garofalakis.
- 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.
- Marcos Carvalho. **Transparent in-situ data processing**. Collaborating researcher: Alkis Simitsis.

Other Educational Activities

Other educational activities involving IMIS members include the following.

- 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 in the following MSc programmes:
 - “Visual Analytics” (In MSc in Cybersecurity and Data Science, Full-time Program – University of Piraeus)
 - “Big Data Management” (In MSc in Cybersecurity and Data Science, Full-time Program – University of Piraeus).

- “Advanced Topics in Data Engineering” (In MSc in Business Analytics, Full-time Program – Athens University of Economic and Business).
- Manolis Terrovitis co-taught the **Spatial Databases** course of the GeoInformatics MSc program of NTUA.



Facts and Figures

Financial report

In 2022, IMSI continued its participation in EC and national funded research and development projects. The key economic indicators regarding the expenses and revenues in 2022 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, Other revenues) are more than 18 times the public expenditure received by IMSI.

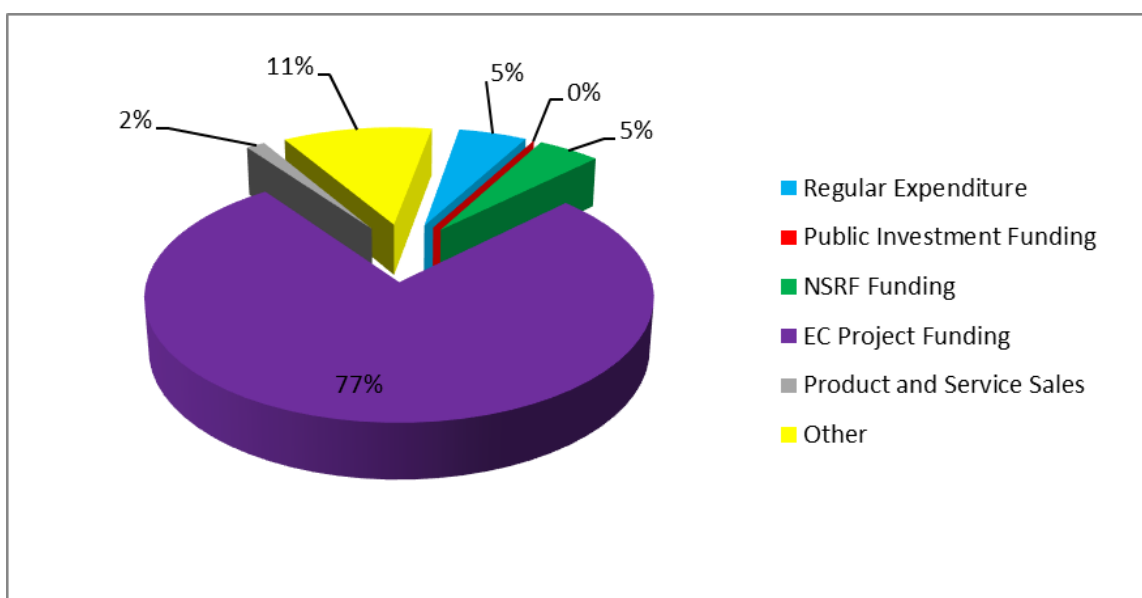


Figure 1. Distribution of revenues in 2022

Table 1. Expenses and Revenues for 2022

Expenses 2022 (in Euros)	
Travel Expenses	124.834,89
Operational Costs	97.394,67
Equipment	112.853,38

Other Expenses	1.412.523,05
Personnel fees and payments to third parties	4.453.352,43
Total	6.200.958,42
Revenues 2022 (in Euros)	
Regular Expenditure	362.550,34
Public Investment Funding	0,00
NSRF Funding	331.446,15
EC Project Funding	5.433.594,48
Product and Service Sales	100.464,37
Other	793.425,79
Total	7.021.481,13

Table 2 shows the revenues of IMSI since 2019 while a comparison of the revenues in the years 2019 - 2022 is illustrated in Figure 2. We can see that the revenues coming from participation in European projects reached between 3.5M and 5.4M within the period 2019 - 2022. As expected, a part of the NSRF Funding within the Partnership Agreement 2014-2020 was paid off in the year 2019, with the revenues from the participation of IMSI in national funded projects reaching 1M euros.

Table 2. Revenues from 2019 to 2022

	Revenues			
	2019	2020	2021	2022
Regular Expenditure	346.777,65	330.874,24	358.173,78	362.550,34
Public Investment Funding	267.080,07	222.884,40	0,00	0,00
NSRF Funding	1.023.287,09	900.019,00	482.358,02	331.446,15
EC Project Funding	3.926.562,69	4.747.206,63	3.498.558,06	5.433.594,48
Product and Service Sales	359.999,93	143.597,65	190.461,54	100.464,37

Other	59.553,25	29.847,85	8.114,43	793.425,79
Total	5.983.260,68	6.374.429,77	4.537.665,83	7.021.481,13

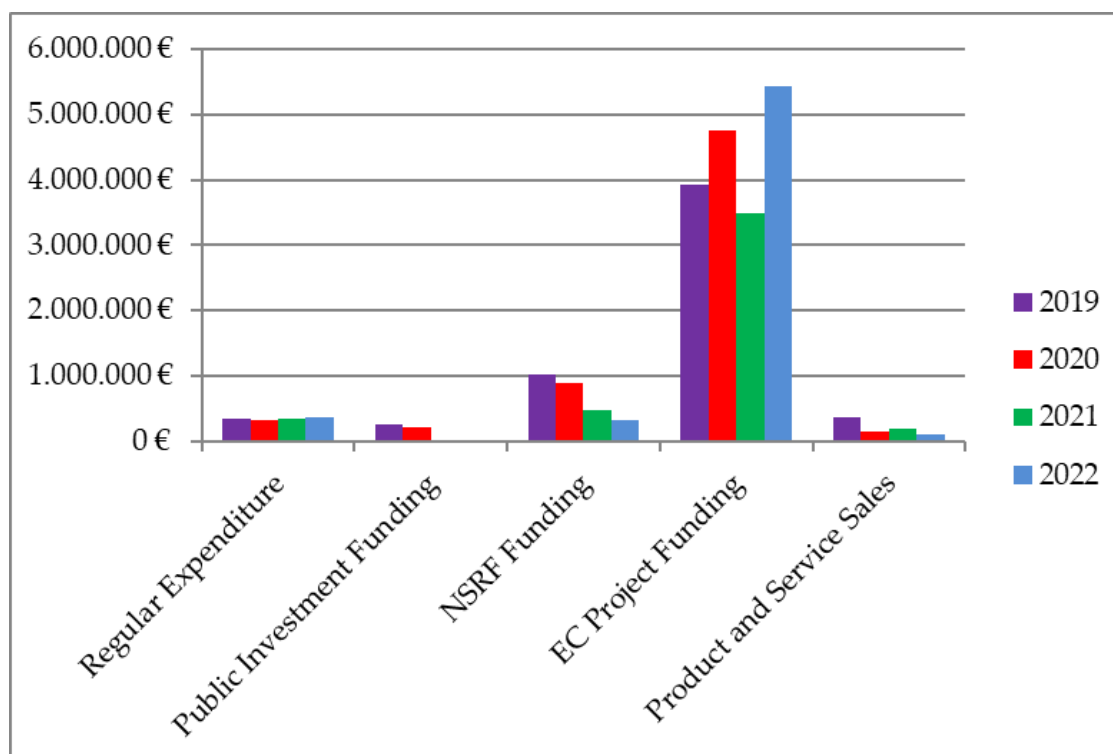


Figure 2. Comparison of revenues 2019 - 2022

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

- Dasaklis Thomas, Assistant professor, Hellenic Open University
- Deligiannakis Antonis, Assistant Professor, Technical University of Crete
- Douligieris Christos, Professor, University of Piraeus
- Emiris Ioannis, Professor, University of Athens
- Ioannidis Yannis, Professor, University of Athens
- Karantzaos Konstantinos, Associate Professor, NTUA
- Kokosis Antonios, Professor, NTUA
- Kotidis Yannis, Associate Professor, AUEB
- Mamoulis Nikolaos, Professor, University of Ioannina
- Patsakis Konstantinos, Assistant Professor, University of Piraeus

Postdoctoral Researchers / Scientific Collaborators

- Arkoumanis Constantinos, Elec. & Comp. Eng. NTUA, PhD NTUA
- Benardou Agiatis, BA, PhD in Ancient History, King's College London
- Casino Fran, BSc in Computer Science, URV, MSc in Computer Security and Intelligent Systems, URV, PhD in Computer Science, Rovira i Virgili University
- Chatzigeorgakidis Giorgos, Elec. & Comp. Eng., Technical University of Crete, MSc in Computer Science, Denmark's Technical University, PhD University of Peloponnese
- 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
- FilidouVasileia, BSc in Physics, University of Patras, MSc in Nonlinear Optics, University of Patras, PhD in Quantum Computing, Oxford University
- 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
- Konaxis Christos, Dept. of Mathematics UoA, PhD . Dept Informatics & Telecoms, UoA
- Kanellos Ilias, Elec. & Comp. Eng. NTUA, PhD NTUA
- Kavouras Lukas, School of Applied Mathematical and Physical Sciences, NTUA, MSc in Logic, Algorithms and Computation (NTUA, University of Athens, University of Patras), PhD NTUA
- Mailis Theofilos, Elec. & Comp. Eng. NTUA, PhD NTUA
- 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
- Raouzaoui 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
- 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
- Vutsinas Theodore, Department of Mechanical Engineering, University of Patras, PhD University of Piraeus
- 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
- Apostolopoulos Theodoros, BSc in Mathematics UoA, MSc University of Piraeus
- Argiropoulos Vasileios, Student, Dept. of Informatics, University of Piraeus
- Athanasiou Spyros, Elec. & Comp. Eng. NTUA
- Avgeridis Lazaros, Dept. of Inform. and Telecom. UoA
- Balasis Ioannis, Dept. of Inform. and Telecom. UoA
- Balasoglou Argyri, Department of Economics, University of Piraeus, MBA University of Piraeus
- Betchavas Panagiotis - Ioannis, BSc in Informatics, AUEB, MSc in Data Science, AUEB
- Bodare Filippos - Andreas, BSc in Physics, UoA, MSc in Environmental sciences, UoA
- Boulieris Petros, BSc in Statistics, AUEB, MSc in Data Science, AUEB
- Chasialis Konstantinos, Dept. of Inform. and Telecom. UoA
- Chatzidimitriou Theodoros, BSc in Chemical Engineering, NTUA
- Chatzilenas Christos, Department of Informatics and Telematics, Harokopio University, Msc, Department of Informatics, AUEB
- Chatzopoulos Serafeim, BA, Msc, Dept. of Inform. and Telecom. UoA
- Checa Carles, BSc in Mathematics, Universitat Politècnica de Catalunya, MSc in Mathematics, Università degli Studi di Padova & Leiden University
- Chrysanthou Argyris, BSc in Informatics, AUEB, MSc in Information security and computer crime, University of Glamorgan
- Deligiannis Panagiotis, Department of Informatics and Telecommunications, University of Peloponnese
- 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
- Eleftherakos Athanasios, Department of Automation Engineering, Piraeus University of Applied Sciences
- Emmanouilidi Sofia, Deree, Business Administration (Accounting and Finance)
- Evangelopoulou Eleni, BA & MA, Department of Communication and Media Studies, UoA
- Fabrici Eros, BSc in Computer Science, Università Degli Studi di Udine, MSc in Data Science and Scientific Computing, Università Degli Studi di Trieste
- 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
- Fourka Maria Niki, BSc in Statistics, AUEB, MSc in Data Science, AUEB

- Galouni Constantina, Dept. of Inform. and Telecom. UoA
- Georgopoulos Vasilios, Dept. of Inform. and Telecom. UoA
- Giannopoulou Roza, BA in Architecture, Kingston University, MSc in Architecture, University of Westminster
- Gidarakos Panagiotis, School of Applied Mathematical and Physics Science, NTUA
- 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
- Iatropoulou Katerina, BA, Msc, Dept. of Inform. and Telecom. UoA
- Ilvanidou Maria, Faculty of History and Archaeology, UoA, Msc, University of Crete
- Ioannidou Ifigeneia Ioanna, BSc in Chemistry, University of Crete
- Kalampokis Panagiotis, BA, Computer Science, University of Ioannina, Msc in Mobile and Web Computing, IHU
- Kallai Tundee, Faculty of Humanities, Lorand Eotvos University (ELTE)
- Kamperi Petro- Foti, 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
- Karantzas Georgios, Student, Dept. of Informatics, University of Piraeus
- Karvounari Sofia, BSc, MSc, 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
- Kechagias Konstantinos, Dept. of Inform. and Telecom. UoA
- Kokkalis George, Department of Computer Engineering and Informatics, University of Patras
- Kokogiannaki Argiro, Dept. of Inform. and Telecom. UoA
- Kolidakis Stylianos, BSc in Electrical and Electronics Engineering, AUTH, MSc & PhD, D.U.Th.
- Koltsida Panayota, BA, Msc, Dept. of Inform. and Telecom. UoA
- Korovesi Chrysoula, Department of Architectural Engineering, NTUA, MA in Digital Arts, ASFA
- Koulocheri Eleni, BSc in Mathematics, UoA, MSc in Information Systems, Hellenic Open University
- Kouvaras Ioannis, Elec. & Comp. Eng. NTUA
- Kozanis Konstantinos, Dept. of Inform. and Telecom., University of Thessaly
- Kouli Vasiliki, Bachelor in Business Administration, AUEB, MSc in International Corporate Finance and Banking, University of Glasgow
- Kravvaritis Constantinos, Dept. of Inform. and Telecom. UoA, Msc, Dept. of Inform, AUEB
- Lampea Theodora, Student, School of Applied Mathematical and Physical Sciences, NTUA
- 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
- Leonidakis Nikiforos, Dept. of Inform. and Telecom. UoA
- Lentoudi Maria, Department of Sociology, Panteion University, MSc in services Management, AUEB
- Lentzos Konstantinos, BSc in Mathematics UoA, MSc UoA
- Lykousas Nikolaos, Department of Informatics, University of Piraeus, MSc in Intelligent Interactive Systems, Universitat Pompeu Fabra
- Maidatsis Konstantinos, Student, Dept. of Inform. and Telecom. UoA
- Maili Evdokia, BA in Mathematics, University of Patras, Msc, Dept. of Inform. and Telecom. UoA
- Makrodimitris Georgios, BSc in Mathematics, University of Patras, MSc University of the Aegean, PhD Candiadate, University of Piraeus
- Mandamadiotis Antonios, Department of Information and Communication Systems Engineering, University of the Aegean, MSc in Data Science and Information Technologies, UoA
- Mantas Andreas, Department of Computer Engineering and Informatics, University of Patras
- Mantas Nikolaos, BSc in Informatics, University of Piraeus
- 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
- Mastoraki Aikaterina, Dept. of Inform. and Telecom. UoA
- Mathioulaki Eleni, Elec. & Comp. Eng. NTUA
- Mavrogiannakis Nikos, BA, Msc in Physics, University of Manchester
- Mazopoulou Falia, Department of Foreign Languages, Translation and Interpreting, Ionian University, MBA University of Piraeus
- Mexis Konstantinos, Dept. of Elec. & Comp. Eng. NTUA
- Mitsopoulou Anna, Dept. of Inform. and Telecom. UoA
- 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
- Pantis Georgios, BSc in Mathematics, University of Patras, MSc in Statistical Science, University of Oxford
- 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
- Papagiannopoulou Stavroula, School of Law, UoA
- Papanikos Georgios, BA, MSc, Dept. of Inform. and Telecom. UoA
- Pappas Dimitris, BSc in Computer Science, University of Crete, MSc in Computer Science, AUEB

- Paranou Dimitra, BSc in Informatics, AUTH
- 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
- Psalla Maria, School of Law, AUTH, MSc University of Piraeus
- Psallidas Michail, Comp. & Inform. Eng. University of Patras
- Renieris Loukas, BSc in Informatics, AUEB, MSc in Informatics, University of Piraeus
- Repouskos Panagiotis, Dept. of Inform. and Telecom. UoA
- Sifakaki Ilektra, BA in Mathematics, UoA, Msc, Department of Informatics and Telematics, Harokopio University
- Skoufis Petros, Dept. of Elec. & Comp. Eng. NTUA
- Smyrniaios 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
- Stamatopoulos Vassilis, Dept. of Inform. and Telecom. UoA
- Stefou Thodoris, BSc, MSc, Dept. of Inform. and Telecom. UoA
- Taha Ibraheem, Faculty of Engineering and Information Technology, An-Najah National University, MSc Dept. of Inform. and Telecom. UoA
- 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
- Tsesmelis Dimitris, Dept. of Inform. and Telecom. UoA, MSc in Big Data Management and Analytics, ULB, UPC, CS
- Tsilipakos Georgios, BSc in Physics, AUTH, MSc Dept. of Inform. and Telecom. UoA
- 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
- Tsopelas Konstantinos, Dept. of Elec. & Comp. Eng. NTUA
- Tzamos Charalampos, BSc & MSc, Dept. of Inform. and Telecom. UoA
- Tzerefos Anargiros, Department of Informatics and Telecommunications, University of Peloponnese
- Vagianou Maria, Department of Informatics, University of Piraeus, MSc UoA
- Vareltsis Georgios Evaggelos, BSc in Computer Science, University of Crete, MSc University College London
- 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
- Vlachakis Nikolaos, Student, Dept. of Elec. & Comp. Eng. NTUA

- 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
- Zafeirakopoulou Efrosini, BSc in Geology, University of Patras
- Zagganas Constantinos, Elec. & Comp. Eng. NTUA
- Zeakis Alexandros, Dept. of Inform. And Telecom. UoA, MSc in Data Science, AUEB



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>

