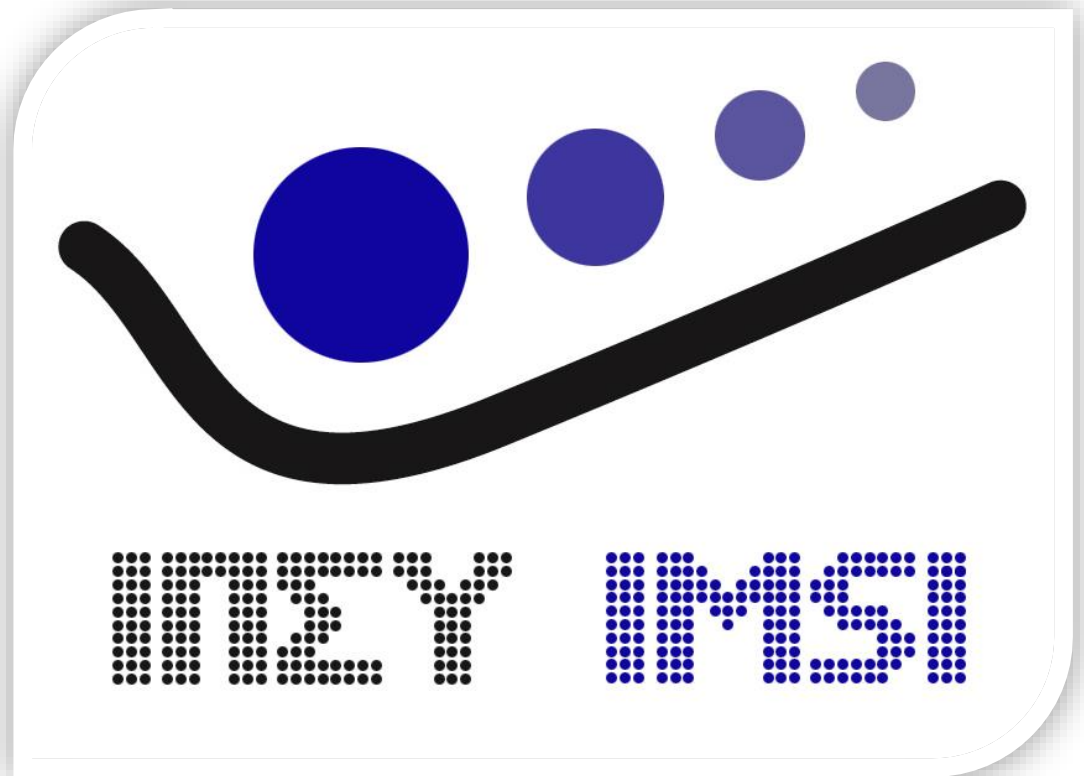
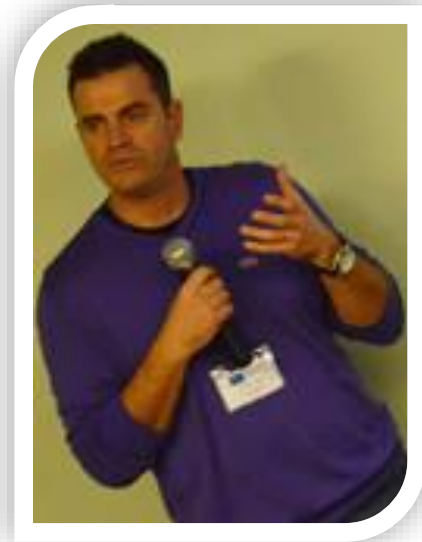


IMSI annual report 2023



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 2023, IMSI has significantly expanded the scope of its research efforts in a number of focus areas, including Big Data Analytics and Machine Learning, Big Data Research Infrastructures, Cloud Platforms and Data Services, Distributed and Web Information Systems, and User-centric Systems. 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, 12/3/2024



Contents

- Preface 1
- Profile 3
- Infrastructures..... 8
- Highlights 10
- Research Directions 13
- Projects 20
- Publications 62
- Dissemination Activities..... 66
- Systems 68
- Education 82
- Facts and Figures 85
- Contact 94



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, Amazon Web Services (AWS) cloud services are available to our researchers via the Open Clouds for Research Environment (OCRE) project and GRNET. Finally, 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

- Ioannis Emiris was invited to **The Cyprus Institute's Advisory Board** and to offer the institute's Colloquim, Nicosia, Cyprus; November 2023.
- UNDERPIN kickoff meeting was organized by IMSI and held in Athens in December 2023. UNDERPIN is a European project that started in December 2023, and is funded by the Digital Europe Programme under grant agreement No 101123179. It is led by Motor Oil Hellas, and ATHENA RC participates with the Information Management Systems Institute (IMSI).



From the UNDERPIN kick-off meeting in Athens, December 2023

The objective of UNDERPIN is to create a state-of-the-art platform that fosters dynamic asset management and predictive maintenance by leveraging DataSpace technologies in a way that is aligned with the vision of a unified European digital market.

- Jan. 2023 - ExtremeXP Kickoff. A new project, coordinated by ATHENA RC - IMSI, started in 01/2023. ExtremeXP (<https://extremexp.eu/>) proposes a new paradigm for experimentation-driven analytics, that aims at providing accurate, precise, fit-for-purpose, and trustworthy data-driven insights via evaluating different complex analytics variants, considering end users preferences and feedback in an automated way. Project Coordinator: George Papastefanatos.
- Nov. 2023 – Presentation of «Check4facts/Science» project in ERT (Hellenic Broadcasting Corporation) by George Papastefanatos. «Check4facts/Science» is a



platform for Public Discourse Fact Checking on four thematic sections: health, climate change, refugee/immigrant, and crime.

- A new collaboration between IMSI and INACCESS-POWERFACTORS started in 2023. IMSI has been contracted to offer its expertise in creating smart, AI tools for the predictive maintenance of photovoltaic parks. Project Coordinator from IMSI : Manolis Terrovitis
- Organized the Artificial Intelligence and Data Management applications for Renewable Energy Sources, collocated with ADBIS in Barcelona, September 2023. Co-organizers from IMSI: Manolis Terrovitis, George Papastefanatos, Giorgos Giannopoulos, Danae Pla-Karydi.
- Co-organized with INACCESS a rollout event for MORE project in Athens, with the participation of key stakeholders from the energy industry in Greece. Organizer from IMSI: Manolis Terrovitis.
- IMSI organized the “4th ACM Europe Summer School in Data Science” in Athens, Greece, 10-14 July, 2023. The school was attended by 68 PhD and MSc students from 12+ countries and featured 8 lectures and keynote talks from renowned researchers and industry leaders. Organization coordination from IMSI: Alkis Simitsis, Georgia Koutrika, Minos Garofalakis
<https://europe.acm.org/seasonal-schools/data-science/2023>
- IMSI participated in the organization of the “11th European Big Data Management & Analytics Summer School (eBISS 2023)”, in Barcelona, Spain, 3-7 July, 2023. The school was attended by 85+ MSc and PhD students, and featured 8 lectures and keynote talks. Organization coordination from IMSI: Alkis Simitsis
<https://cs.ulb.ac.be/conferences/ebiss2023/>

Awards

- **HBP Innovation Award to Y. Ioannidis for Medical Informatics Platform.** The HBP 2023 Innovation Awards presented during the final Human Brain Project Summit in Marseille (28-31 March 2023), honoured science to help patients. Prof. Yannis Ioannidis, University of Athens - Athena RC, and Prof. Philippe Ryvlin, University of Lausanne and their teams won the awards for developing the Medical Informatics Platform (MIP), capable of providing advanced analytics for diagnosis and research in clinical neuroscience. Embedded into EBRAINS, this open-source and free-of-use software allows sharing of health datasets and is distributed across hospitals and institutions in different countries. It aims to better understand brain disease, develop novel biomarkers and predictive models of brain disease useful for clinical practice, and thereby have a direct impact on patient management.
<https://www.athenarc.gr/el/news/hbp-innovation-award-ioannidis>
- **Highly Cited Researchers 2023:** Theodore Dalamagas was recognised as a Highly Cited Researcher from Clarivate Analytics for 2023. 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.

- **Best paper award:** Vayianos Pertsas and Panos Constantopoulos received the best paper award at the 15th International Joint Conference on Knowledge Discovery (KEOD) 2023.
- Alkis Simitsis received the ACM DOLAP 2023 **Test-Of-Time Award**.
- Alkis Simitsis received the VLDB 2023 **Distinguished Associate Editor Award**.
- Alkis Simitsis received the IEEE ICDE 2023 **Industry & Applications, Best Metareviewer Award**.
- Alkis Simitsis received the EDBT 2023 **Distinguished PC Member Award**.



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 and Biotechnology. The practice of life sciences is continuously becoming more data-driven. The group has a leading role in RIs (ELIXIR-GR, IBISBA.EU, 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. IBISBA.EU is a pan-European research infrastructure dedicated to Industrial Biotechnology. Inspired-RIs focuses on integrated structural biology, drug screening and target functional characterization.

Digital Curation. Digital curation encompasses a set of activities aiming at the production of high quality, dependable digital assets; their organization, archiving and long-term preservation; and the generation of added value from digital assets by means of resource-based knowledge elicitation. To ensure the adequate capture of the context of digital resources and their subsequent creative and effective use, we adopt a multidisciplinary approach that considers the full lifecycle of digital assets, such as records, digital surrogates and scholarly/ scientific datasets. A major ongoing line of research deals with the modeling and analysis of scholarly working practices and of workflows underpinning the related information processes. The originally formulated ESF NeDiMAH methods ontology (NeMO) has been generalized to become the



Scholarly Ontology, which in turn is being used to drive the automatic extraction of research processes from publications with significant success (see, for instance, the KEOD 2023 best paper), as well as the definition specialized information task models and workflows.

Digital research infrastructures. A strategic action line is the development of digital research infrastructures for the humanities at the national and european levels. At the national level, IMSI leads the APOLLONIS Greek Infrastructure for Digital Arts, Humanities and Language Research and Innovation (P. Constantopoulos, coordinator) resulting from the unification of CLARIN-EL and DARIAH-GR. At the european level, IMSI actively participates in the European Digital Research Infrastructure for the Arts and Humanities (DARIAH) since the preparatory phase, currently with leading roles in the directorship of the ERIC (A. Benardou is one of the three directors), VCC2, DARIAH's Virtual Competence Centre for Research and Education (M. Ilvanidou is VCC2 Co-Chair) and the Scientific Advisory Board (P. Constantopoulos, chair). IMSI has been heavily involved in building the ARIADNE infrastructure for archaeology. Currently, IMSI is participating in preparing 4CH, the Competence Centre for the Conservation of Cultural Heritage. Collaborations with other research infrastructures, also in areas besides the humanities, are actively pursued.

Cloud Platforms and Data Services

Data processing and analytics on the cloud. The key focus of the group is on technologies that enable scalable and energy efficient 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 measurement 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, as well as for predicting and reducing resource utilization and energy consumption in edge data centers. 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 pipelines, etc.) need to provide accurate, precise, fit-for-purpose, and trustworthy data-driven insights via evaluating different complex analytics variants, considering end users preferences and feedback in an automated way. They also need to operate in an energy efficient manner and be close to where data is generated, e.g., for reducing data transfer overhead, or when sensitive data cannot move out of the production system. In this context, the group has been active in developing scalable methods for users to visually interact and collaborate with complex data services and workflows. The group develops a variety of tools that offer scalable visual analytics capabilities for different types of data, e.g., spatial, timeseries.

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, with the approval of the AI act, 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. We further examine explainability beyond the AI model level, researching methods for explaining ML pipelines.

Fairness aware explainability. Implementing fairness auditing/bias detection algorithms that (a) can align with real world requirements and respective legal frameworks and (b) are transparent to the end users, comprises a significant research field with potentially high impact on real world systems. Our work produces state of the art work on explainable fairness auditing in various types of bias (statistical, difficulty of recourse, spatial bias). We further examine bias at the intersection of law and algorithms, researching how algorithmic fairness definitions and bias detection methods can be aligned with legal and ethical requirements in different application scenarios.

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. . Our current research focuses on leveraging pre-trained language models to automate entity extraction, entity linking and similarity search.

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. Our current research focuses on leveraging machine learning techniques to support similarity search and quality enhancement of geospatial data and time series.

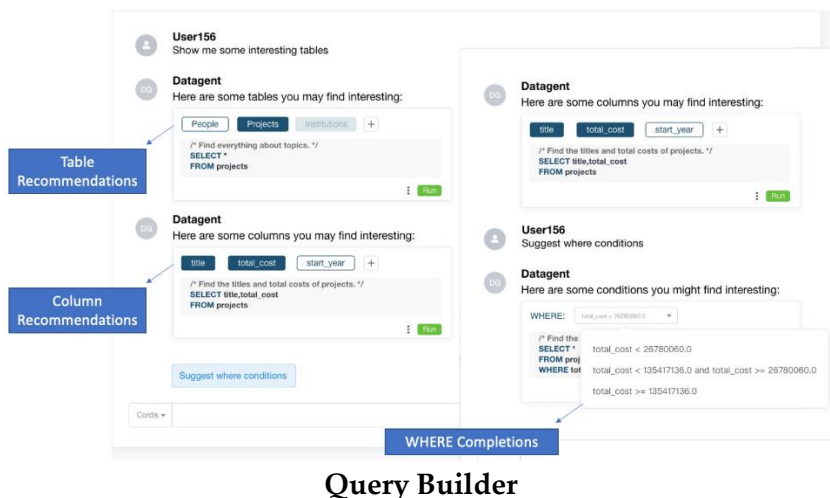
Leveraging Social Data. The availability of online data through social networks, especially Twitter, gives rise to several disparate and challenging problems: (a) how to leverage social data for obtaining new knowledge (data journalism, public opinion trends, brand monitoring), (b) how to use knowledge 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

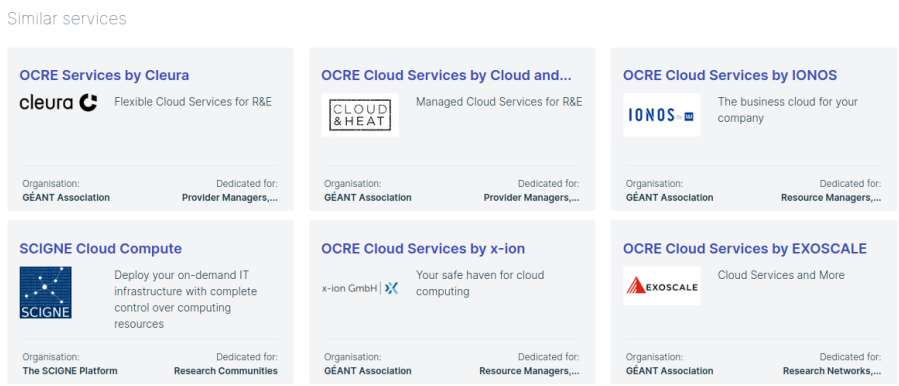
Futuristic Data Interfaces. The volume of data and the need for data democratization call for shift from the classical «query-based information access paradigm» popularized by existing systems to a novel information access paradigm, where the system takes upon a more conversational and active role in 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.



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



UNDERPIN: Pan-European data space for holistic asset management in critical manufacturing industries

Project manager	Stelios Sartzetakis
Coordinator	Motor Oil Hellas
IMIS - funding	195 KEuros
Programme	Digital Europe
Start date	1/12/2023
Duration	24 months
Website	https://underpinproject.eu/

The objective of UNDERPIN is to create a state-of-the-art platform that fosters dynamic asset management and predictive maintenance by leveraging DataSpace technologies in a way that is aligned with the vision of a unified European digital market. Specifically tailored for European manufacturers in the refinery and renewable energy domains, the platform is poised to revolutionize how industries approach data-driven solutions. The UNDERPIN DataSpace will ensure cross-organizational and cross-use-case data sharing and exchange, emphasizing data sovereignty. With a strong focus on fostering collaboration between SMEs and large industry players, the UNDERPIN DataSpace aims to drive innovation, enhance products and services, and elevate the industry to new heights of efficiency.



BIOINDUSTRY 4.0 - RI Services to Promote Deep Digitalization of Industrial Biotechnology - towards Smart Biomanufacturing



Project manager	Theodore Dalamagas
Coordinator	IMSI
IMIS - funding	424,000 euros
Programme	HORIZON-INFRA-2022-TECH-01
Start date	1/2023
Duration	48 months
Website	https://bioindustry4.hub.inrae.fr

Industrial biotechnology harnesses the power of microorganisms to produce chemicals and materials for a wide range of sectors including fuels, cosmetics, pharmaceuticals, and sustainable packaging. Accordingly, industrial biotechnology is considered to be a cornerstone of the circular bioeconomy. As a rapidly developing field, industrial biotechnology is data-intensive and thus an ideal candidate for the use of advanced digital methods. However, both industrial biotechnology and advanced digital technologies for industry are emerging areas. Therefore, the integration of these two relatively immature, fast-evolving technology families is a [challenge in itself](#). Bioindustry 4.0 is a cutting-edge initiative [bringing together academia and industry](#) in Europe to shape the future of industrial biotechnology. Spearheaded by [6 major European research infrastructures](#) and supported by a [consortium of 25 organizations](#), our mission is to enhance the competitiveness of the European biotech sector globally. By refining technologies such as AI, digital twin, and cloud technologies, we want to empower professionals from academia and industry to collaborate and push the boundaries of bioprocess design and control – making these technologies faster, cheaper, and more sustainable. Through Bioindustry 4.0 we intend to devise 5 new services that will be offered by European research infrastructures. These services will support the needs of academia and industry alike, providing them with access to some of the most advanced technologies available. These services include: (a) Fabric: A data fabric – a platform that helps users to structure and manage their datasets in ways compliant with international standards and suitable for use in algorithm training, (b) Sensors: Advanced process analytical technology (PAT) devices – these are cutting-edge sensor devices that will offer novel in-line bioprocess monitoring capabilities. (c) Strain: A strain discovery tool and decision support system – these tools will enhance the experience of microbial culture collection users, (d) Shadow: Digital shadows for bioprocess design – using physical and biological models, and a wealth of legacy and bespoke datasets, bioreactor digital shadows will be designed and trained, (e) Twin: Digital twins for bioprocess control – by using robust bioreactor models, Bioindustry 4.0 will develop digital twins suitable for real-time process control.





ExtremeXP: EXPerimentation driven and user eXPerience-oriented analytics for eXtremely Precise outcomes and decisions

Project manager	George Papastefanatos
Coordinator	IMSI / ATHENA RC
IMIS - funding	782.000 Euros
Programme	HORIZON-CL4-2022-DATA-01-01
Start date	01/01/2023
Duration	3 years
Website	http://extremexp.eu/

Extreme data characteristics (volume, speed, heterogeneity, distribution, diverse quality, etc.) challenge the state-of-the-art data-driven analytics and decision-making approaches in many critical domains such as crisis management, predictive maintenance, mobility, public safety, and cyber-security. At the same time, data-driven insights need to be extremely timely, accurate, precise, fit for purpose, and trustworthy, so that they can be useful.

ExtremeXP will handle the complexity of matching extreme needs with complex analytics processes (i.e., processes that involve and combine ML, data analysis, simulation, and visualization components) by placing the end user at the centre of complex analytics processes and relying on user intents and running experiments (i.e., trial and error) to prune the vast solution space of possible analytics workflows and configurations i.e., “variants”. Its main goal is to create a next generation decision support system that integrates novel research results from the domains of data integration, machine learning, visual analytics, explainable AI, decentralised trust, knowledge engineering, and model-driven engineering into a common framework. **The overarching idea of the framework is to optimise the properties of a complex analytics process that the end user cares about (e.g., accuracy, time-to-answer, specificity, recall, precision, resource consumption) by associating user profiles to computation variants.** The framework is envisioned as modular and extensible, orchestrating different services around an Experimentation Engine: Analysis-aware Data Integration, Extreme Data & Knowledge Management, User-driven AutoML, Transparent & Interactive Decision Making, and User-driven Optimization of Complex Analytics. The framework will be validated in five pilot demonstrators.



**SciLake: SCIENTIFIC KNOWLEDGE GRAPHS:
Democratising & making sense out of heterogeneous
scholarly content**

Project manager	Thanasis Vergoulis
Coordinator	IMSI
IMIS - funding	548k euros
Programme	HORIZON-INFRA-2021-EOSC-01-04
Start date	01-01-2023
Duration	36 months
Website	https://scilake.eu/

SciLake is a 3-year Horizon Europe research project that aims to extend the technical work in the field of Science Knowledge Graphs (SKGs) leveraging them as the foundation to establish the concept of the scientific lake: a research ecosystem to facilitate creating, combining, and querying cross-domain and domain-specific SKGs. This ecosystem, among others, will comprise tools that are capable to extract knowledge from unstructured (e.g., textual) information, facilitate the interoperability among SKGs, support various types of knowledge transformation, unify and simplify the way SKGs can be queried, and accelerate graph processing and analysis for SKGs.

The project will build a prototype of this concept delivering also, on top of it, an additional tier of smart services to assist in discovering scientific knowledge and improving research reproducibility. The work will be done in close consultation with four research communities (Neuroscience, Cancer research, Transportation research, and Energy research) and all the services will be tailored to take into consideration domain-specific requirements. Each of these communities will also demonstrate and evaluate selected SciLake services in the context of a domain-specific pilot. Finally, the project will leverage EOSC functionalities (e.g., comply with the EOSC Interoperability Framework for monitoring, accounting, and AAI) and will integrate its open-source services into the portfolio of OpenAIRE and EOSC Core services.

SciLake brings together a competent consortium comprising 13 partners from 9 different countries including partners with significant technological expertise in knowledge management and discovery to guarantee successful implementation of the scientific lake and its services but also domain experts from the four scientific disciplines that participate in the piloting activities.





GraspOS: Next Generation Research Assessment to promote Open Science

Project manager	Yannis Ioannidis (PI), Thanasis Vergoulis (technical manager)
Coordinator	IMSI
IMIS - funding	357k euros
Programme	HORIZON-INFRA-2022-EOSC-01-01
Start date	01-01-2023
Duration	36 months
Website	https://graspos.eu/

GraspOS is a 3-year Horizon Europe research project that sets out the ambitious goal to develop, assess and put into operation an open and trusted federated infrastructure offering data, tools, services and guidance to support and enable policy reforms for Open-Science-aware, responsible research assessment at researcher (individual/group), institutional, organisational and country level. Its key results include:

- The Open Science Assessment Framework (OSAF), a living and collaborative guide for responsible research assessment that aligns to Open Science practices and ensures that Open Science activities of researchers and research organisations are taken into consideration in research assessment events.
- A set of Assessment Portfolios to support the collection of inputs for research assessment and an online Assessment Registry to facilitate publishing assessment protocols developed for responsible research assessments.
- A set of metadata enrichment tools and services for enhancing missing attributes, semantics, context, and novel indicator values on research outputs and enriching the links among them with appropriate semantics.
- A set of monitoring tools and services offering Open Science indicators, contextual factors, qualitative information (e.g., narratives), and relevant evidence to support assessment processes and relevant types of analysis.
- A Federated Open Research Assessment Infrastructure that offers APIs to provide easy access to data assets coming from well-established providers of scholarly metadata information, which are valuable for the implementation of Open-Science-aware research assessment processes, aiming to also facilitate the development and provision of enrichment, monitoring, and other added-value services on top of these assets, paving the way for the creation of a research assessment data space.
- A Community of Practice of responsible research assessment experts from relevant networks, to facilitate knowledge sharing and capacity building.

- A set of carefully designed Training material relevant to the subjects of research assessment and Open Science.

Finally, the project has initiated nine pilots encompassing a diverse array of use cases spanning various levels, scopes, and values. These pilots serve as platforms to showcase and practically test the assessment framework, infrastructure, tools, and services developed within the project. Furthermore, they adhere to a co-design approach, enabling stakeholders to provide feedback on implementation and ensure alignment with real-world needs and requirements.



TIER2: ENHANCING TRUST, INTEGRITY AND EFFICIENCY IN RESEARCH THROUGH NEXT-LEVEL REPRODUCIBILITY IMPACT PATHWAYS

Project manager	Thanasis Vergoulis
Coordinator	KNOW-CENTER GMBH RESEARCH CENTER FOR DATA-DRIVEN BUSINESS & BIG DATA ANALYTICS
IMIS - funding	162k euros
Programme	HORIZON-WIDERA-2022-ERA-01-41
Start date	01-01-2023
Duration	36 months
Website	https://tier2-project.eu/

TIER2 focuses on epistemic diversity by selecting three broad research areas (social, life and computer sciences) and two cross-disciplinary stakeholder groups (research publishers and funders) to systematically investigate reproducibility across contexts and to design, implement and assess systematic interventions addressing key levers of change. TIER2 targets to boost reproducibility knowledge, create tools, engage communities, and implement interventions across different contexts.



Co-Creating Positive and Sustainable Lifestyle Tool with and for European Citizens (PSLifestyle)

Project manager	Eleni Petra
Coordinator	SUOMEN ITSENAISYYDEN JUHLARAHASTO
IMIS - funding	253.881,25 Euros
Programme	H2020-EU.3.5
Start date	01/10/2021
Duration	3.5 years
Website	https://pslifestyle.eu/

Through our consumption behaviour we, individuals, are responsible for 72 per cent of the global greenhouse gas emissions. These are created by the way we live, travel, and eat and by what we buy. So far, the attempts to change people's habits by appealing to their rationale have not produced significant behaviour change. Aiming to help close the action gap between climate awareness and individual action, and to increase citizen participation in sustainability topics, the project builds a data-driven momentum for sustainable behaviour change across eight European countries. It does this by engaging citizens through a digital application to co-research, codevelop and uptake everyday life solutions for climate change, providing tools for the collection, monitoring and analysis of their environmental and consumption data. In the application, the citizens find out how their lifestyle and habits impact the environment. They will also see a list of smart everyday actions that are relevant to reduce their environmental impact. The application enables citizens to actively create data while monitoring their lifestyle induced climate impact and behaviour. The collected data will be used to improve the application itself, and, importantly to empower citizens to gain agency in research and policymaking. The ambition is to engage a total of four million EU citizens. The figure is based on the previous national success with a web-based carbon emission calculator, which is innovated further with the use of citizen science, by co-research at living labs, co-creating personalized sustainability plans, as well as with the integration of behaviour science knowledge into the co-research and design process. The overarching goal of PSLifeStyle is to co-create a web- and mobile-based sustainable lifestyle tool and open platform, that will enable, empower and encourage European citizens to take their personalized steps towards more sustainable lifestyles.



**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 Paraclite) 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 Intrasoftware 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.



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**

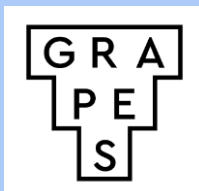
Project manager Manolis Terrovitis
Coordinator IMSI
IMIS - funding 886K Euro
Programme H2020-ICT-2018-20 / H2020-ICT-2020-1
Start date 1-10-2020
Duration 39 months
Website <https://www.more2020.eu/>

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 misconceptions 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.



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.

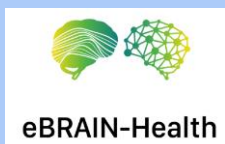


NI4OS-Europe 2020 - National Initiatives for Open Science in Europe

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.



eBRAIN-Health Actionable Multilevel Health Data

Project manager	Prof. Yannis Ioannidis
Coordinator	CHARITE - UNIVERSITAETS MEDIZIN BERLIN
IMIS - funding	485.625 €
Programme	HORIZON-INFRA-2021-TECH-01
Start date	1/1/2022
Duration	48 months
Website	https://www.ebrain-health.eu/

The Project eBrain-Health will deliver a distributed research platform for modeling and simulating complex neurobiological phenomena of human brain function and dysfunction in a data protection compliant environment. It will provide thousands of multilevel virtual brains from patients and healthy human controls for research and innovation. Brain data from multiple sources will be pre-processed. Solving the societal grand challenge of dementia is a big task. Yet it appears feasible in a collective approach. Therefore, we will build an interdisciplinary digital twin for dementia for modeling and simulating complex phenomena at the service of research infrastructure communities. EBRAINS-Health-Cloud will offer end-to-end services for personalized complex brain modeling and simulations in distributed e-infrastructures with data protection by design and by default and simulation-ready human multiscale brain data that range from molecular (genomics, proteomics, metabolomics) and cellular to electrophysiology and imaging to behavioural, clinical, life-style and environmental data as well as data from wearables. Brain data are pre-processed and annotated such that they all relate to a common reference 3D brain space. EBRAINS-Health-Cloud constitutes a blend of three large-scale research programs: the FET Flagship Human Brain Project with its EBRAINS Research Infrastructure, the EOSC project Virtual Brain Cloud with its Virtual Research Environment for sensitive data and the H2020 project AI-MIND with intelligent tools for dementia risk estimation. The project will have synergies to topics of the Digital Europe Program, such as artificial

intelligence, cybersecurity and supercomputing and the Health Data Space. EBRAINS-Health-Cloud offers a next generation clinical research infrastructure and creates an open yet protected space for groundbreaking digital health innovation by the research infrastructure communities comprising academia and the private sector.



EDITH: An ecosystem for digital twins in healthcare

Project manager	Prof. Yannis Ioannidis
Coordinator	VIRTUAL PHYSIOLOGICAL HUMAN INSTITUTE FOR INTEGRATIVE BIOMEDICAL RESEARCH VZW - VPH INSTITUTE
IMIS - funding	427.679€
Programme	DIGITAL-2021-DEPLOY-01
Start date	1/10/2022
Duration	24 months
Website	https://www.edith-csa.eu/

The overall aim of the EDITH project is to foster an inclusive ecosystem for Digital Twins in healthcare in Europe and to prompt the convergence of such an ecosystem towards a common strategy conducive to its further development. This is achieved by mapping and analysing the status of the fields which are crucial for the growth, uptake and use of digital twins in healthcare, including in silico medicine, health data interoperability, high performance computing, ethical and legal regulations etc. A vision for the integrated human digital twin will be developed, based on standardised (meta-)data and models, and a roadmap to realise that vision will be articulated. Additionally, a federated cloud-based repository will be established, to bring together currently available resources and best practices. The ecosystem will be leveraged to create a repository catalogue with available resources and recruit resources into the repository during the project. Conditions for integration in the repository in terms of required standards, regulations, meta-data, and others. will be identified. Finally, building on available infrastructure, a framework for a simulation platform will be put forward with pre-selected prototypes demonstrating a proof of concept. User communities (healthcare professionals, patients, industry and academia) will be actively involved in the process to ensure their needs are built into the architecture. Several activities will focus on the exploitation of parts of the repository and simulation

platform. Throughout the entire EDITH action, the community, its stakeholders and relevant international partners will be consulted via advisory boards, public meetings, community challenges and other public activities in order to firmly establish a sustainable ecosystem allowing to realise the vision of the integrated digital twin for personalised healthcare.



EBRAINS PREP: EBRAINS Preparatory Phase

Project manager Prof. Yannis Ioannidis

Coordinator EBRAINS

IMIS - funding 90.425€

Programme HORIZON-INFRA-2021-DEV-02 submitted for HORIZON-INFRA-2021-DEV-02 / 20 Jan 2022

Start date 1/9/2022

Duration 30 months

Website <https://www.ebrains.eu/projects/ebrains-prep>

This ESFRI Preparation Phase Project will place the EBRAINS research infrastructure (RI) developed by the Human Brain Project (HBP) on a sustainable footing. The new organisational framework will allow EBRAINS to evolve from the largely EC-funded, HBP-developed RI into a sustainable “Hub-and-Node” RI on the ESFRI Roadmap.

The Operation Phase EBRAINS RI will serve the whole brain research community, rather than the needs of a subset of users within that community; this will be reflected in a broader offering of tools and services by the RI. The Operation Phase RI will focus more on the processing, managing and sharing of data of all sorts, even those generated outside the RI, but will retain an important role in providing cutting-edge data-generating tools and services. While staying true to its roots in supporting basic research, it will expand its activities to support clinical research and innovative exploitation of brain-related technologies. The evolving RI offering will be determined with the help of a new governance structure, which focuses on determining the priority needs of the broad brain research community. The Operation Phase EBRAINS will be funded primarily by long-term commitments by the EU Member States and Associated Countries that are home to a Node in the EBRAINS RI.

Objective 1: To put in place a financial, legal and governance framework for

EBRAINS and prepare a plan to ensure the sustainable offering of the Service Capability by the Central Hub and the Nodes.

Objective 2: To obtain the legally binding commitment of EBRAINS Consortium Partners to offer services needed by the brain research community by setting up EBRAINS Nodes.

Objective 3: To deliver a complete overall technical design and associated cost estimate for the Service Capability.

Objective 4: To decide the legal structure that EBRAINS will adopt for the Operation Phase and produce draft statutes for EBRAINS in that structure.



PathOS: Open Science Impact Pathways

Project manager	Prof. Yannis Ioannidis
Coordinator	ATHINA-EREVNITIKO KENTRO KAINOTOMIAS STIS TECHNOLOGIES TIS PLIROFORIAS, TON EPIKOINONION KAI TIS GNOSIS - ATHENA - RESEARCH AND INNOVATION CENTER
IMIS - funding	285.625€
Programme	HORIZON-WIDERA-2021-ERA-01 submitted for HORIZON-WIDERA-2021-ERA-01 / 23 Sep 2021
Start date	1/9/2022
Duration	36 months
Website	https://pathos-project.eu/

PathOS aims to identify and quantify the Key Impact Pathways of Open Science relating to the research system and its interrelations with economic and societal actors. PathOS will enable a new understanding of OS impacts and their causal mechanisms through its workplan encompassing actions to synthesise and structure current evidence, development of new methods and tools for measuring impact, iterative pilot-testing via in-depth case studies, innovative dissemination and networking, and co-creation synthesis activities culminating in policy recommendations. This is pivotal in order to develop effective OS policy in the EU. It will do so by collecting concrete evidence of the causal effects of OS by studying the pathways of OS practices, from input to output, outcome and impact, including the consideration of enabling factors and key barriers. Impacts and pathways will be developed in particular in the three areas of science, society and economy. By investigating, measuring and comparing its costs and benefits

together with its pathways, PathOS will (i) bring a better understanding of the implications of open science for science, economy and society, (ii) provide recommendations to policy makers and other actors in the R&I ecosystem as to how and to what extent open science should be promoted in a balanced way, and (iii) develop innovative tools and methods using a big data to augment traditional ones for studying the causal effects of open science. This will enable evidence-based Open Science policy prioritisation, maximum OS impact, and increased R&I capacity in EU research systems.



Project manager	Prof. Yannis Ioannidis
Coordinator	ARISTOTELIO PANEPISTIMIO THESSALONIKIS (AUTH)
IMIS - funding	181.375,00€
Programme	HORIZON-INFRA-2021-EOSC-01
Start date	01/10/ 2022
Duration	3,3 years
Website	https://raise-science.eu/

The mission of RAISE is to provide the infrastructure for a distributed crowdsourced data processing system, moving from open data to open access data for processing. RAISE will provide the mechanism for sending the algorithm to the dataset instead of sending the data to the algorithm. The real value of open data for the research community is not to access them but to process them as conveniently as possible in order to reduce time-to-result and increase productivity. RAISE aims at promoting a transparent way of sharing and processing data, enabling the research community to publish their work with evidence-based authenticity of the data-analysis performed, ensuring at the same time the accreditation of their work. RAISE will be grounded on the fundamental principle defined in the FAIR Guiding Principles for scientific data management and stewardship (Findability, Accessibility, Interoperability and Reusability). To do so, RAISE brings the processing algorithm (small size) to the dataset (large size) instead of downloading the dataset to the computer where the processing algorithm is. To increase the processing capacity of the dataset repositories, RAISE borrows the crowdsourcing concept where researchers can easily integrate in the existing workflows computers serving both their datasets and the processing capacity.

RAISE will produce the following Outputs: 1. A trustworthy crowdsourced network of RAI Certified nodes offering data storing and processing resources, 2. The RAI Cloud platform to orchestrate the data sharing, processing and finding, 3. The Research Analysis Identifier ? RAI , a unique identifier of any result along with the dataset information and the processing script, without disclosing any source code or raw data, 4. Dataset plagiarism identification and dataset proof-of-origin services, to maximise the level of trust of the RAISE system, 5. The RAI Synthetic Data Generator.



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).



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.





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.



EOSC Future

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

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 EOSC-Exchange 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. EOSC Future will unlock the potential of European research via a vision of Open Science for Society by (1) bringing all major stakeholders in the EOSC ecosystem together under one project umbrella to break the disciplinary and community silos and consolidate key EOSC project outputs, (2) developing scientific use cases in collaboration with the thematic communities showcasing the benefits and societal value of EOSC for doing excellent and interdisciplinary research, (3) engaging the wider EOSC community and increasing the visibility of EOSC through communications campaigns, marketing strategies, and physical and online engagement events, and (4) including the EOSC community in developing the EOSC Portal (including the long tail of science, public and private sectors, and international partners) via co-creation open calls.



4CH - Competence Centre for the Conservation of Cultural Heritage

Project coordinator	Panos Constantopoulos
Coordinator	ISTITUTO NAZIONALE DI FISICA NUCLEARE (INFN)
IMSI - funding	84.735 Euros
Programme	H2020-EU.3.6. - SOCIETAL CHALLENGES - Europe In A Changing World - Inclusive, Innovative And Reflective Societies
Start date	1/1/2021
Duration	36 months
Website	https://www.4ch-project.eu/

The 4CH project will design and prepare for a European Competence Centre on Cultural Heritage (CC) which will work proactively for the preservation and conservation of cultural heritage (CH). The project aims to start implementing the structure, organization and services of the CC which will operate as a virtual infrastructure providing expertise, advice and services using state-of-the-art ICT with a special focus on 3D technology. A Strategic Advisory Board, including representatives of major national and international CH bodies, will be established to advise on cultural, scientific, technological, financial, strategy and policy areas. 4CH will design and implement the ICT infrastructure of the CC, based on a Cultural Heritage Cloud, compliant with the EOSC vision, to provide an open collaborative digital space for cultural heritage conservation giving access to repositories of data, metadata, standards and guidelines. Access to High Performance Computing service providers will be supported.

The core results of 4CH work will comprise of:

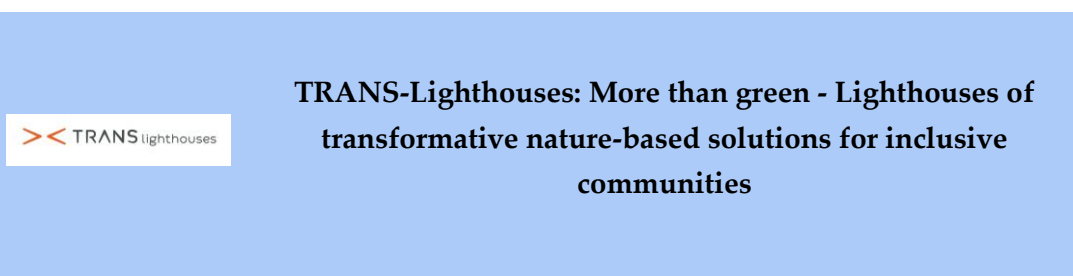
- a holistic interdisciplinary approach to state-of-the-art 3D documentation



of monuments and sites for conservation, preservation, access and exploitation

- guidance on policies and strategies
- a catalogue of well-documented standards
- an inventory of software and tools
- guidelines for data acquisition, management and storage
- a set of success stories on advanced digitization
- advice on funding opportunities
- training and education plans for professionals and managers
- communication and dissemination of 4CH results
- a complete business model and a sustainability plan for the Competence Centre.

4CH will foster the implementation of actions to substantially increase the number and the quality of digitized sites, documents and monuments.



Project manager	Panos Constantopoulos
Coordinator	CENTRO DE ESTUDOS SOCIAIS
IMSI - funding	100.000 Euros
Programme	HORIZON.2.6 - Food, Bioeconomy Natural Resources, Agriculture and Environment - 2.6.2 - Biodiversity and Natural Resources - Assessing the socio-politics of nature-based solutions for more inclusive and resilient communities
Start date	01/05/2023
Duration	42 months
Website	

The TRANS-lighthouses project aims to understand the strengths and limitations in the design and implementation of nature-based solutions. Based on material and immaterial evidence, it proposes to contribute to rethinking and reframing the main elements that compose the complexity of creating socially and

ecologically just solutions. As a project funded by the European Union, lasting from May 2023 to October 2026 and with a budget of almost 6 million euros, TRANS-lighthouses strengthens socio-politics as part of the public agenda for nature-based solutions towards systemic change.

TRANS-lighthouses also integrates a network of “lighthouses” in urban, rural, coastal and forest areas. The “lighthouses” are a metaphor for a set of local governance arrangements and instruments, within multi-stakeholder networks and concerted groups. They are aimed at improving the contributions of nature-based solutions and achieving, in an integrated way, ecological, social and economic objectives. To this end, new governance models will be tested, as well as approaches and tools for co-creation in small scale but big picture projects that can be upscaled over time.

Accordingly, each lighthouse is composed of living knowledge labs, assessment cases, pilot cases and international associated partners. In these spaces, the interaction of different knowledges, experiences and roles will support the assessment of ongoing solutions and the testing of new ones. In this way, it is intended to prioritise the perspectives of citizens, in dialogue with other interested actors for their co-creation.

The consortium of TRANS-lighthouses project comprises research and innovation performing organisations, policy-making institutions and civil society organisations, with 19 European partners from 10 countries. In terms of international cooperation, TRANS-lighthouses also integrates 9 associated partners from 7 countries in the Americas, Africa and Asia.

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.

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	24 months
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.



Digital Curation Services

Project manager	Panos Constantopoulos
Coordinator	IMSI
IMSI - funding	87.531,41 Euros
Programme	Internal project
Start date	1/2/2016
Duration	8 years
Website	http://www.dcu.gr/en/vπηροεσίες-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.



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	5 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.

Industry – Funded R&D Projects



TraMPaR: Trajectory Mining and Payment Routing

Project manager	Ioannis Z. Emiris
Coordinator	HSBC, Λονδίνο
IMIS - funding	12.000 euros
Programme	Industrial contract
Start date	01/11/2023
Duration	6 months
Website	-

The project develops an innovative way to implement payment routing, a problem encountered in several financial transactions. The problem will be tackled by new optimization methods based on path search and mining and adapted to the use case.



Smart application for enhancing photovoltaic park performance

Project manager Manolis Terrovitis

Coordinator IMSI

IMIS - funding 15000 Euros

Programme -

Start date 6/2023

Duration 12/2023

Website

The goal of the project is to provide technical and consulting services to INACCESS for enhancing applications that monitor photovoltaic parks with AI tools. The additional capabilities of the applications that will be developed will assist in the predictive maintenance of the park.



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 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, 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.



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	18 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

Book chapters

- Agiatis Benardou and Anna Maria Droumpouki: **Difficult Heritage and Immersive Experiences**. Book chapter in “Difficult Heritage and Immersive Experiences”, Benardou, A., and Droumpouki, A.M. (eds.), Routledge (2023)

Journal Publications

- Alexandros Zeakis, George Papadakis, Dimitrios Skoutas, Manolis Koubarakis: **Pre-trained Embeddings for Entity Resolution: An Experimental Analysis**. Proceedings of the VLDB Endowment. 16(9): 2225-2238 (2023)
- Yannis Foufoulas, Alkis Simitsis: **Efficient Execution of User-Defined Functions in SQL Queries**. Proc. VLDB Endow. 16(12): 3874-3877 (2023)
- Georgios Chatzigeorgakidis, Dimitrios Skoutas, Kostas Patroumpas, Themis Palpanas, Spiros Athanasiou, Spiros Skiadopoulou: **Efficient Range and kNN Twin Subsequence Search in Time Series**. IEEE Transactions on Knowledge and Data Engineering. 35(6): 5794-5807 (2023)
- Xiaoying Wu, Dimitri Theodoratos, Dimitrios Skoutas, Michael Lan: **A novel framework for the efficient evaluation of hybrid tree-pattern queries on large data graphs**. Information Systems. 117: 102249 (2023)
- S. Maroulis, N. Bikakis G. Papastefanatos, P. Vassiliadis, Y. Vassiliou: **Resource-Aware Adaptive Indexing for In-situ Visual Exploration and Analytics**. In VLDB Journal. 32(1): 199-227 (2023)
- Maria Ilvanidou, Vicky Dritsou, Maria Gavriilidou, Kanella Pouli, Yiorgos Tzedopoulos, & Irakleitos Souyioultzoglou (2023). **The "Digital Landscape in Greece" Web Survey**. DARIAH Annual Event 2023: Cultural Heritage Data as Humanities Research Data?, Budapest, Hungary. Zenodo. <https://doi.org/10.5281/zenodo.7973681>
- Vicky Dritsou, Achille Felicetti, & Panos Constantopoulos (2023). **Working with Flow: Workflows for Accessing 4CH Services**. DARIAH Annual Event 2023: Cultural Heritage Data as Humanities Research Data?, Budapest, Hungary. Zenodo. <https://doi.org/10.5281/zenodo.7973720>
- Pertsas, V. and Constantopoulos, P. (2023). **“Ontology-Driven Extraction of Contextualized Information from Research Publications”**. In Proc. 15th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - Volume 2: KEOD, ISBN 978-989-758-671-2, ISSN 2184-3228, pp. 108-118. *Best paper award*.

- Alexandros Kalimeris, Ioannis Psarros, Giorgos Giannopoulos, Manolis Terrovitis, George Papastefanatos, Gregory Kotsis: **Data-driven soiling detection in PV modules**. IEEE Journal of Photovoltaics 13 (3), 461-466, 2023.
- Kathryn Whitman, et al: **Review of Solar Energetic Particle Prediction Models**. Advances in Space Research, Volume 72, Issue 12, 2023.
- I.Z. Emiris, V. Margonis, and I. Psarros. **Near-neighbor preserving dimension reduction via coverings for doubling subsets of L1**. Theoretical Computer Science, 942:169–179, 2023.
- L. Calès, A. Chalkis, I.Z. Emiris, and V. Fisikopoulos. **Practical volume approximation of high-dimensional convex bodies, applied to modeling portfolio dependencies and financial crises**. Computational Geometry: Theory & Applications, 109, 2 2023.
- A. Chalkis, I.Z. Emiris, and V. Fisikopoulos. **A practical algorithm for volume estimation based on billiard trajectories and simulated annealing**. J. Experimental Algorithmics, 28(1.3):1–34, 2023.
- A. Chalkis, I.Z. Emiris, V. Fisikopoulos, E. Tsigaridas, and H. Zafeiropoulos. **Geometric algorithms for sampling the flux space of metabolic networks**. J. Computational Geometry, 14(1), 2023.
- P. Kakoulidis, I. Vlachos, D. Thanos, G.L. Blatch, I.Z. Emiris, and E. Anastasiadou. **Identifying and profiling structural similarities between spike of SARS-CoV-2 and other viral or host proteins with Machaon**. Communications Biology, 6(1), 2023.
- Antonios Kontaxakis, Nikos Giatrakos, Dimitris Sacharidis, Antonios Deligiannakis: **And synopses for all: A synopses data engine for extreme scale analytics-as-a-service**. Inf. Syst. 116: 102221 (2023)
- Y Fofoulas, E Zacharia, H Dimitropoulos, N Manola, Y Ioannidis, “**DETEXA: declarative extensible text exploration and analysis through SQL**”, International Journal on Digital Libraries, 1-13, May 2023

International Conference / Workshop Publications

- Daniele Lunghi, Alkis Simitsis, Olivier Caelen, Gianluca Bontempi: **Adversarial Learning in Real-World Fraud Detection: Challenges and Perspectives**. DEC@SIGMOD 2023: 27-33
- Alkis Simitsis, Spiros Skiadopoulos, Panos Vassiliadis: **The History, Present, and Future of ETL Technology** (invited). DOLAP 2023: 3-12
- Yannis Fofoulas, Alkis Simitsis: **User-Defined Functions in Modern Data Engines**. ICDE 2023: 3593-3598
- Daniele Lunghi, Alkis Simitsis, Olivier Caelen, Gianluca Bontempi: **Adversarial Learning in Real-World Fraud Detection: Challenges and Perspectives**. CoRR abs/2307.01390 (2023)

- Vasilis Gkolemis, Theodore Dalamagas, Eirini Ntoutsis, Christos Diou: **RHALE: Robust and Heterogeneity-Aware Accumulated Local Effects**. ECAI 2023: 859-866.
- Andra Ionescu, Kostas Patroumpas, Kyriakos Psarakis, Georgios Chatzigeorgakidis, Diego Collarana, Kai Barenscher, Dimitrios Skoutas, Asterios Katsifodimos, Spiros Athanasiou: **Topio: An Open-Source Web Platform for Trading Geospatial Data**. ICWE 2023: 336-351
- Andra Ionescu, Alexandra Alexandridou, Leonidas Ikononou, Kyriakos Psarakis, Kostas Patroumpas, Georgios Chatzigeorgakidis, Dimitrios Skoutas, Spiros Athanasiou, Rihan Hai, Asterios Katsifodimos: **Topio Marketplace: Search and Discovery of Geospatial Data**. EDBT 2023: 819-822 (Best Demo Award)
- Dimitra Paranou, Angelos Pentelas, Dimitris Katsiros, Konstantinos Maidatsis, George Giannopoulos, Evangelos Angelou, Nikos Anastopoulos, George Papastefanatos : **Forecasting Resource Demand for Dynamic Datacenter Sizing in Telco Infrastructures**. 2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy, 2023, pp. 3813-3822.
- Konstantinos Tertikas, Despoina Paschalidou, Boxiao Pan, Jeong Joon Park, Mikaela Angelina Uy, Ioannis Emiris, Yannis Avrithis, Leonidas Guibas: **Generating Part-Aware Editable 3D Shapes Without 3D Supervision**. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4466-4478. 2023.
- L. Kavouras, K. Tsopelas, G. Giannopoulos, D. Sacharidis, E. Psaroudaki, N. Theologitis, D. Rontogiannis, D. Fotakis, and I.Z. Emiris. **Fairness aware counterfactuals for subgroups**. In Proc. NeurIPS, Nov 2023. Poster.
- Paris Koloveas, Serafeim Chatzopoulos, Christos Tryfonopoulos, Thanasis Vergoulis: **BIP! NDR (NoDoiRefs): A Dataset of Citations from Papers Without DOIs in Computer Science Conferences and Workshops**. TPD L 2023: 99-105
- Serafeim Chatzopoulos, Kleanthis Vichos, Ilias Kanellos, Thanasis Vergoulis: **Piloting Topic-Aware Research Impact Assessment Features in BIP! Services**. ESWC (Satellite Events) 2023: 81-85
- Serafeim Chatzopoulos, Thanasis Vergoulis, Dimitrios Skoutas, Theodore Dalamagas, Christos Tryfonopoulos, Panagiotis Karras: **Atrapos: Real-time Evaluation of Metapath Query Workloads**. WWW 2023: 2487-2498
- Dimitris Sacharidis, Giorgos Giannopoulos, George Papastefanatos, Kostas Stefanidis: **Auditing for Spatial Fairness**. EDBT 2023.
- Giorgos Giannopoulos, Dimitris Sacharidis, Nikolaos Theologitis, Loukas Kavouras, Ioannis Z. Emiris: **FALE: Fairness-Aware ALE Plots for Auditing Bias in Subgroups**. Uncertainty meets Explainability Workshop @ ECML-PKDD 2023.
- Konstantinos Alexis, Stella Girtsou, Alexis Apostolakis, Giorgos Giannopoulos, and Charalampos Kontoes: **Next day fire prediction via semantic segmentation**. Workshop on Machine Learning for Earth Observation @ ECML-PKDD 2023.
- S Girtsou, A Apostolakis, G Giannopoulos, H Kontoes: **FireRisk: A Web Platform for next day fire forecasting**. EGU23.

- Vassilis Stamatopoulos, Stavros Maroulis, Konstantinos Kozanis, Ioannis Psarros, George Papastefanatos, Giorgos Giannopoulos, Manolis Terrovitis: **A Tool for Visual Exploration and Analysis of Solar Photovoltaic Module Data**. EDBT/ICDT Workshops 2023.
- A. Karvelas, Y. Foufoulas, A. Simitsis, Y. Ioannidis, “**Toulouse: Learning Joint Order Optimization Policies for Rule-based Data Engines**”, Proc. of the 2nd Int’l Workshop on Data Platform Design, Management, and Optimization (DataPlat’23), Ioannina, Greece, March 2023.
- Nikos Bikakis, Giorgos Giannopoulos, Nikolaos Sidiropoulos, Christina Flouda, Athanasios Doupas, Voula Giouli, Panagiotis Kariotis, Paraskevi Botini, Anna Vacalopoulou, Gregory Stainhaouer: **Exposing Geospatial Cultural Heritage Content in Map-based Applications**. EDBT/ICDT Workshops 2023.
- Achilleas Michalopoulos, Dimitrios Tsitsigkos, Panagiotis Bouros, Nikos Mamoulis, Manolis Terrovitis: **Efficient Nearest Neighbor Queries on Non-point Data**. SIGSPATIAL/GIS 2023: 33:1-33:4

Other Publications / Technical Reports

- Manolis Terrovitis, Dominik Slezak: **Data Infrastructures and Management - Introduction to the Special Theme**. ERCIM News 2023(133), 2023.
- Vasilis Gkolemis, Anargiros Tzerefos, Theodore Dalamagas, Eirini Ntoutsis, Christos Diou: **Regionally Additive Models: Explainable-by-design models minimizing feature interactions**. CoRRabs/2309.12215 (2023).



Dissemination Activities

Invited / Keynote Talks

IMSI members participated in the following invited / keynote talks:

- 07/2023, “Transitioning towards Open Scientometrics with Open Science Graphs”, Thanasis Vergoulis co-presented a tutorial at the ISSI 2023 international conference.
- 09/2023, “Open Science Knowledge Graphs: Transforming the Way we Manage, Explore, and Analyze Scientific Knowledge” Thanasis Vergoulis co-presented a workshop at the Open Science Fair 2023.
- Nov. 2023 – Participation in Press conference for the presentation of «Check4facts/Science» project, [George Papastefanatos](#).

Scientific Community Service

IMSI members have served in the **Program Committee** of more than 30 International Conferences and Workshops in 2023, including well-known Conferences like VLDB, ICDE, SIGMOD, EDBT, TPD, IEEE Big Data, ACM SIGSPATIAL, ACM CICM, IEEE DSAA, and more.

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

- IMSI co-organized the 6th International Workshop on Big Data Visual Exploration and Analytics (BigVis23) on March 28th 2023, held jointly with EDBT/ICDT2023 in Ioannina, Greece. [George Papastefanatos](#), [Nikos Bikakis](#) (Organizing committee)
- IMSI co-organized the workshop on “Artificial Intelligence and Data Management applications for Renewable Energy Sources” AID4RES, on September 4th, 2023, held jointly with ADBIS 2023, in Barcelona, Spain. [Manolis Terrovitis](#), [George Papastefanatos](#), [Giorgos Giannopoulos](#), [Danae Pla Karidi](#) (Organizing committee)
- [Thanasis Vergoulis](#) was in the organizing committee of the SciK 2023 workshop, co-located with the Web Conference 2023.
- [Alkis Simitsis](#) participated in the organization of the ACM International Workshop on Data Warehousing and OLAP (ACM DOLAP) as member of the Steering Committee.
- [Alkis Simitsis](#) participated as Associate Editor in the Proceedings of the VLDB Endowment for PVLDB Volume 16, as Associate Editor in the ACM/IMS Journal of Data Science, and as Editorial Board Member in Data & Knowledge Engineering, Elsevier.

- Manolis Terrovitis served as Editor for **Ercim News**.



Systems

- **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>

The screenshot displays the MAGE web interface, which is divided into three main sections: Load, Preprocess, and Discover. The Discover section is currently active and contains several configuration parameters:

- Max size: 100
- Size weight: 0.1
- Time budget (sec): 30
- Entropy mode: high
- Method: AdaptiveHybrid
- Initial seeds (%): 1
- Max overlap (%): 20
- Top k: 10

Below the configuration options is a button labeled "Compute Regions". The main area of the interface is a map of a city street grid. Several clusters of colored dots (representing geolocated entities) are visible on the map, each enclosed in a blue polygon. A tooltip is displayed over one of these clusters, showing the following information:

- rank: 1
- points: 100
- score: 0.9839559003083246

At the bottom of the map, there is a small text attribution: "Leaflet | Map tiles by Stamen Design, under CC BY 3.0. Data by © OpenStreetMap, under ODbL."

- **SPHINX**

SPHINX is a system for metapath-based entity exploration in Heterogeneous Information Networks (HINs), developed in the context of the EU funded project SmartDataLake. SPHINX allows users to define different views over a HIN based on both automatically selected and user-defined metapaths. Then, entity ranking and

similarity search can be performed over these views to find and explore entities of interest, taking also into account any spatial or temporal properties of entities. A Web-based user interface is provided to facilitate users in performing the various functionalities supported by the system, including metapath-based view definition, index construction, search parameters specification, and visual comparison of the results.

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

1. Select source entity type

Article

2. Select search preferences

Determine search preferences based on existing indices.

+ Add new index

i) Preferences on source entities

Type	Preferences	Weights
<input checked="" type="radio"/> Temporal	Select Date	0 0.25 0.5 0.75 1

ii) Metapath preferences

Metapath	Preferences on target entities	Weights
<input checked="" type="radio"/> Article - Organisation	twitter x	0 0.25 0.5 0.75 1
<input type="radio"/> Article - Person	donald trump x barack obama x	0 0.25 0.5 0.75 1

3. Select value for k

50

Search

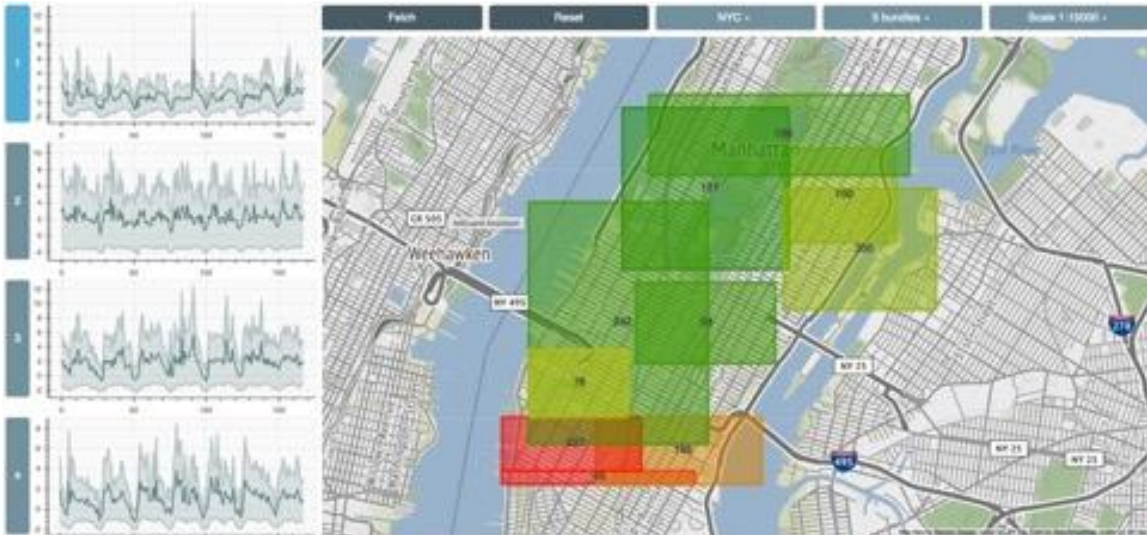
Results

#	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/smardatalake/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/>



Step 1 **User156**
Find projects that started before 2018

Step 2 **Datagent**
Best interpretation:

```
/* Find the titles and end years of projects whose end year is less than 2018. */
SELECT projects.title, projects.end_year
FROM projects
WHERE projects.end_year < 2018
```

NL Explanation

SQL Interpretation

Step 3

projects.title	projects.end_year
ALFRED - Personal Interactive Assistant for Independent Living and Active Ageing	2016
Microbial Biomarker Records in Tibetan Peats: Monsoon Variability and its Impact on Methane Biogeochem...	2016
Post-glacial recolonisation and Holocene anthropization impact on populations of shrews and hedgehog...	2016
Molecular Mechanisms Employed by the Newly Assigned RNA-binding Protein FASTKD2	2016
Identifying the targets and mechanism of action of the SUMO targeted ubiquitin ligase RNF4 in respon...	2017

Find similar queries

Get query recommendations

Cordis

Datagent
Best interpretation:

```
/* Find the titles and end years of projects whose end year is less than 2018. */
SELECT projects.title, projects.end_year
FROM projects
WHERE projects.end_year < 2018
```

Current Query

User156
Get query recommendations

Datagent
Here are some queries you might find interesting:

```
/* Find the titles and end years of projects whose end year is less than 2018 and end year is greater than or equal to 2014.5. */
SELECT projects.title, projects.end_year
FROM projects
WHERE projects.end_year < 2018 and end_year >= 2014.5
```

```
/* Find the titles and end years of projects whose end year is less than 2018 and end year is less than 2011.5. */
SELECT projects.title, projects.end_year
FROM projects
WHERE projects.end_year < 2018 and end_year < 2011.5
```

```
/* Find the titles and end years of projects whose end year is less than 2018, end year is greater than or equal to 2011.5 and end year is less than 2014.5. */
SELECT projects.title, projects.end_year
FROM projects
WHERE projects.end_year < 2018 and end_year >= 2011.5 and end_year < 2014.5
```

Query Recommendations

Cordis

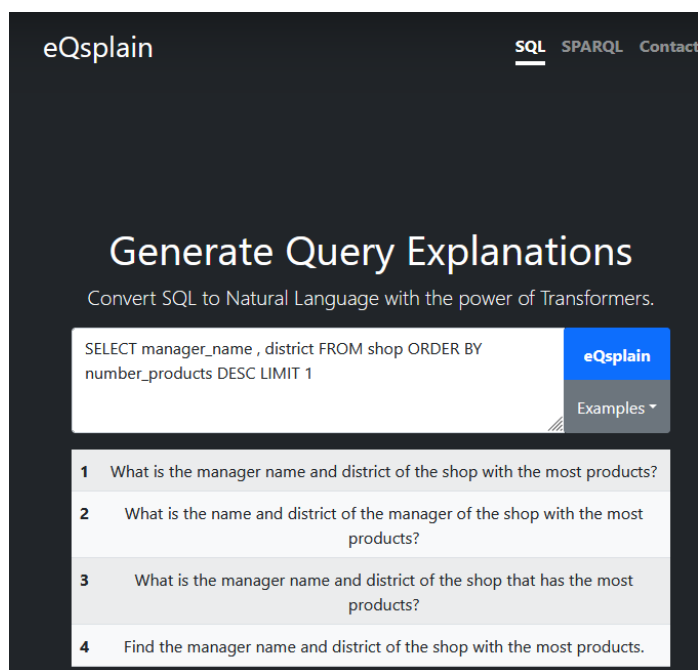
More information about DatAgent can be found here:
<http://vldb.org/pvldb/vol14/p2815-madamadiotis.pdf>

- **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 OAI-compliant and preserves the whole lifecycle of each digital object. All ingestion and curation actions create new versions of metadata streams that are stored and semantically annotated, thus allowing to view the entire history of changes associated with each digital resource. MORe is tuned to support massive imports in the order of about 0.8 million records per hour. Online services with the MORe system are provided to Europeana and the related communities formed and supported through the projects CARARE, 3-D ICONS, ARIADNE, LoCloud and CEF Europeana. A total of approximately 70 content providers in over 20 countries, are being regularly served by the MORe aggregation service to aggregate, enrich and deliver content to Europeana, with approximately 10 million heritage asset records processed so far.

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

- **NeMO and SO**

The NeDiMAH Methods Ontology (NeMO) is a comprehensive ontological model of scholarly practice in the arts and humanities, offered and maintained by DCU/IMSI, the development of which was undertaken through the ESF Research Network [NeDiMAH](#). NeMO is a CIDOC CRM - compliant ontology which explicitly addresses the interplay of factors of agency (actors and goals), process (activities and methods) and resources (information resources, tools, concepts) manifest in the scholarly process. It builds on the

results of extensive empirical studies and modeling of scholarly practices performed by DCU/IMSI in projects [DARIAH](#) and [EHRI](#). NeMO incorporates existing relevant taxonomies of scholarly methods and tools, such as TaDIRAH, the arts-humanities.net and Oxford taxonomies of ICT methods, DHCommons, CCC-IULA-UPF and DiRT, through appropriate mappings of the concepts defined therein onto a semantic backbone of NeMO concepts. It thus enables combining documentary elements on scholarly practices of different perspectives and using different vocabularies. NeMO was subsequently generalized to the domain-neutral Scholarly Ontology (SO).

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

- **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.

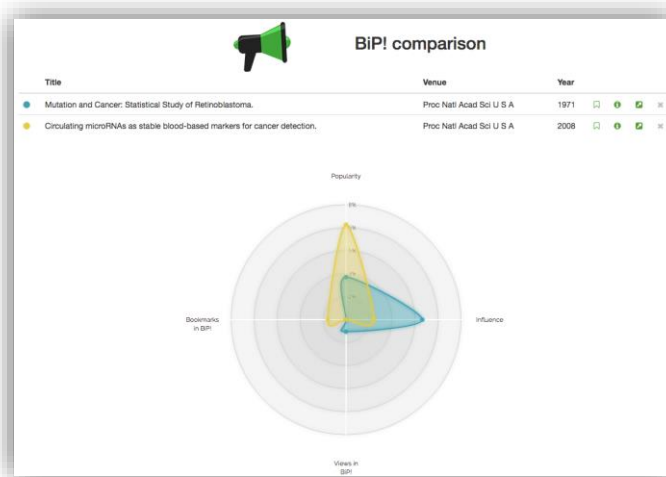
<https://github.com/DAIAD>

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

- **BIP! Services**

BIP! Services is a suite of tools that assist scientific discovery and research assessment. BIP! Finder tool supports ranking and comparing scientific articles based on different aspects of their impact. BIP! Scholar is a service that allows researchers to create their own academic profiles.

<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>

- **BIP! Ranker**

BIP! Ranker is an open source library containing implementations of paper ranking methods that have been proposed in the literature. Our implementations utilise a suite of Spark scripts and can be used either on a single machine, or a 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/athenarc/Bip-Ranker>

- **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).

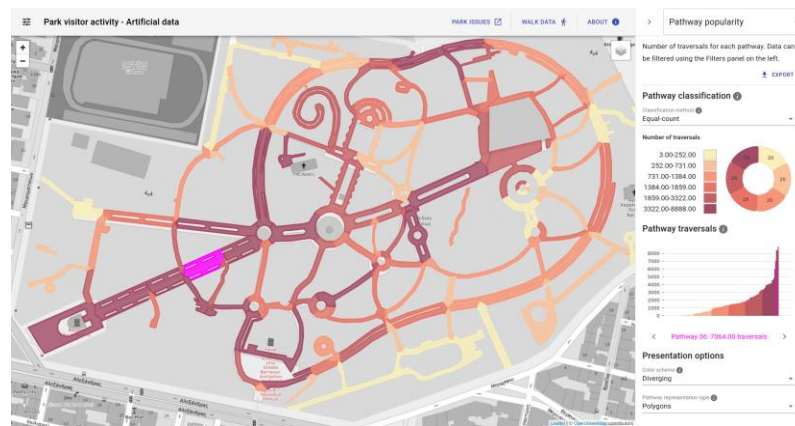
<https://github.com/athenarc/YeSQL>

- **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.



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

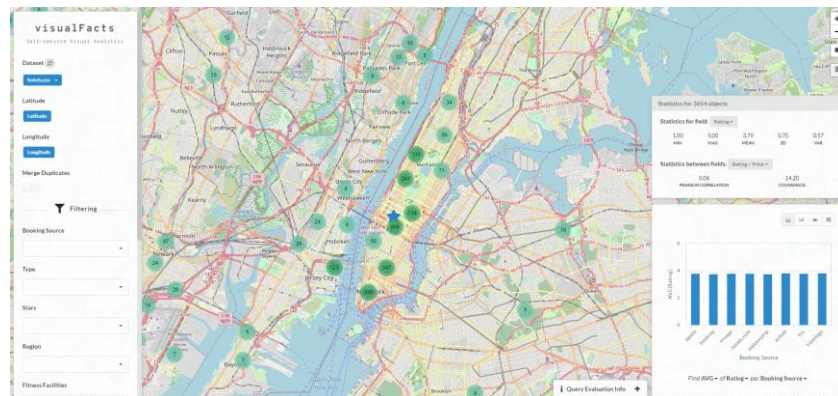
through faceted browsing, hierarchical representation of coded lists in charts, free keyword search of literal values, and capabilities for combining different datasets along common dimensions. Moreover, it makes knowledge reusable by making all data available for download in various formats including 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>

- **A Self Service Visualization Platform for Renewable Energy Analytics**



The tool is a visual analytics platform for big geo-located timeseries data, enriched with operations for the analysis of measurements coming from renewable energy sources, such as wind and solar parks. It offers functionality for the exploration and the application of complex and real-time visual analytics at the macroscale (e.g., park KPIs on a map) and microscale (e.g., analysis of multivariate timeseries from multiple sensors), such as, pattern extraction, outlier detection, timeseries forecasting, etc. Furthermore, the platform utilizes a novel caching approach for achieving interactive responses when user analyses large volumes of multivariate timeseries data, in real time.

- **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 use to create supportive material to the anonymization process, like generalization hierarchies. Amnesia offers k-anonymity and km-anonymity and a parallel scalable anonymization algorithm, it is available through the OpenAIRE infrastructure.

- **FAGI**

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

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



Education

PhD / MSc / Diploma Thesis Co-supervision

IMSI members actively co-supervise several undergraduate and graduate university students, who often conduct their work in IMSI premises. As a result, in 2023 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 PhD students. The following PhD dissertations have been completed in 2023:

- Serafeim Chatzopoulos. Topic: **Data mining for scholarly information networks**. Joint supervision with the University of the Peloponnese. Collaborating researcher: Thanasis Vergoulis.
- Stavros Maroulis. Topic: **Adaptive indexing for interactive visual exploration and analytics**. Joint supervision with the National Technical University of Athens. Collaborating researchers: George Papastefanatos.

IMIS members also co-supervise MSc students. The following MSc dissertations have been completed in 2023:

- Synodinou, Anna (2023). ETL workflow application study on archival data. Department of Informatics, Athens University of Economics and Business, http://www.pyxida.aueb.gr/index.php?op=view_object&object_id=10693. Supervised by: Vicky Dritsou.
- Kassapaki, Maria-Eleni (2023). Topic: Working practices knowledge base: Analysis and extraction of textual data using machine learning methods. Department of Informatics, Athens University of Economics and Business. Supervised by: Panos Constantopoulos, Vayianos Pertsas.
- Martoglou, Eleni (2023). Documentation and digitization of publications and audio-visual content on modern Greek history through the ASKI collections. Department of Informatics, Athens University of Economics and Business. Supervised by: Agiatis Benardou and Vicky Dritsou.

The following PhD and MSc students collaborated closely with / were supervised by IMSI members in their research during 2023:

- Giorgos Alexiou. Topic: **Query driven Entity Resolution**. Joint supervision with the National Technical University of Athens. Collaborating researcher: George Papastefanatos.



- Thomas E. Makrygiannis, Vasileios S. Stamatopoulos: **Leveraging the Matrix Profile for Real-Time Atrial Fibrillation Detection**. Msc Thesis, National Kapodistrian University of Athens. Supervisor: Theodore Dalamagas, Collaborating researcher: George Papastefanatos.
- Theodore Routsis: **Recommendations for Interactive data visualization of spatial - categorical data**. Msc Thesis, University of Piraeus. Supervisor: George Papastefanatos.
- Paris Koloveas. Topic: **Knowledge extraction from scientific texts and citation classification**. Joint supervision with the University of the Peloponnese. Collaborating researcher: Thanasis Vergoulis.
- 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.
- Christina Kakalou. Topic: **Digital-assisted Health Information Material Quality Assessment**. In collaboration with the University of Athens. Researcher-in-charge from IMSI: Theodore Dalamagas.
- 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
- 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.

Supervision of Internships in IMIS:

- Nikos Masouras: 10/2023-03/2024. Undergraduate student in the Department of Informatics, National Kapodistrian University of Athens. Supervisor: George Papastefanatos.
- Andreas Papageorgiou: 02/2023-08/2023. Undergraduate student in the Department of Informatics, National Kapodistrian University of Athens, Supervisor: George Papastefanatos.

Other Educational Activities

Other educational activities involving IMIS members include the following.

- Thanasis Vergoulis co-taught the **Big Data Systems Architecture** course of the Business Analytics MSc program of the Athens University of Economics and Business.
- Serafeim Chatzopoulos co-taught the **Data Management** course of the Data Science MSc program of the University of the Peloponnese.
- George Papastefanatos has been an Adjunct Lecturer, teaching Visual Analytics and Big Data Management courses at the Business and Data Analytics (BDA) track of the Cybersecurity and Data Science MSc program of University of Piraeus.



Facts and Figures

Financial report

In 2023, IMSI continued its participation in EC and national funded research and development projects. The key economic indicators regarding the expenses and revenues in 2023 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 74%, comes from EC funded projects. The revenues from the activities of IMSI (EC projects, National funded projects, Product and Service Sales, Other revenues) are more than 13 times the public expenditure received by IMSI.

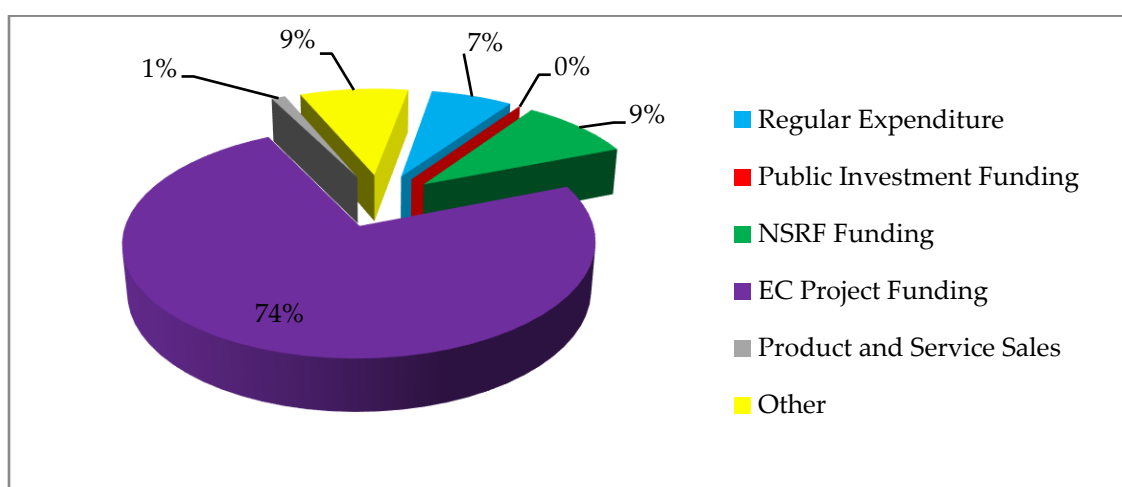


Figure 1. Distribution of revenues in 2023

Table 1. Expenses and Revenues for 2023

Expenses 2023 (in Euros)	
Travel Expenses	215.730,10
Operational Costs	168.722,03
Equipment	134.698,53
Other Expenses	817.189,16
Personnel fees and payments to third parties	4.908.934,77
Total	6.245.274,59

Revenues 2023 (in Euros)	
Regular Expenditure	371.205,00
Public Investment Funding	0,00
NSRF Funding	513.923,91
EC Project Funding	4.004.224,58
Product and Service Sales	64.009,17
Other	493.765,61
Total	5.447.128,27

Table 2 shows the revenues of IMSI since 2020 while a comparison of the revenues in the years 2020 - 2023 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 2020 - 2023. As expected, a part of the NSRF Funding within the Partnership Agreement 2014-2020 was paid off in the year 2020, with the revenues from the participation of IMSI in national funded projects reaching 0.9M euros.

Table 2. Revenues from 2020 to 2023

	Revenues			
	2020	2021	2022	2023
Regular Expenditure	330.874,24	358.173,78	362.550,34	371.205,00
Public Investment Funding	222.884,40	0,00	0,00	0,00
NSRF Funding	900.019,00	482.358,02	331.446,15	513.923,91
EC Project Funding	4.747.206,63	3.498.558,06	5.433.594,48	4.004.224,58
Product and Service Sales	143.597,65	190.461,54	100.464,37	64.009,17
Other	29.847,85	8.114,43	793.425,79	493.765,61
Total	6.374.429,77	4.537.665,83	7.021.481,13	5.447.128,27

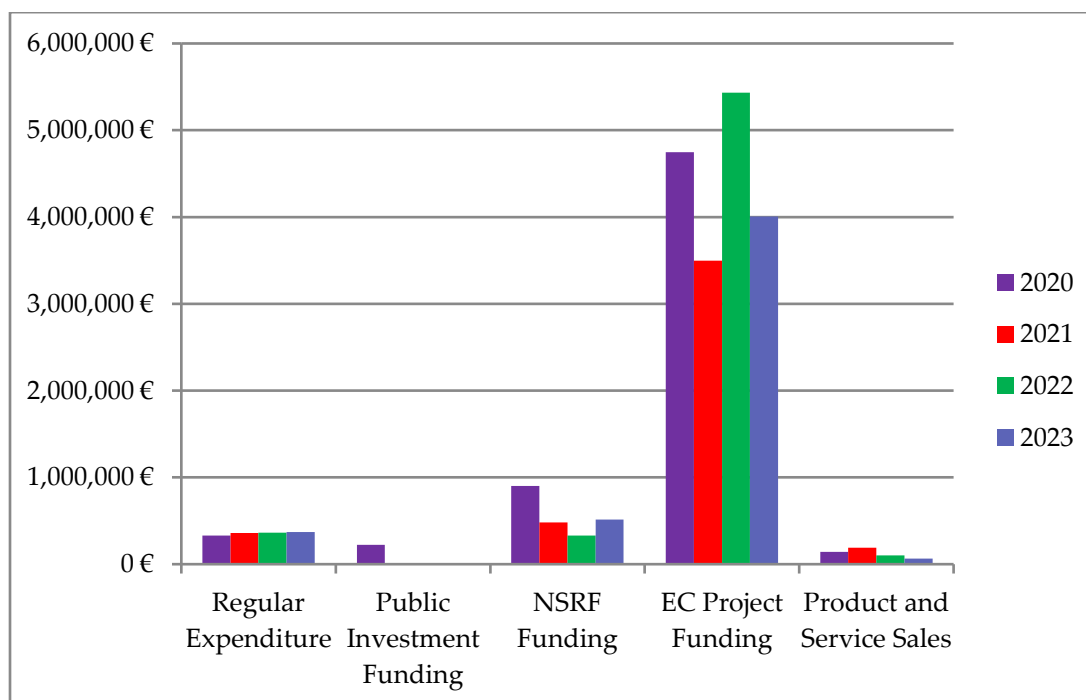


Figure 2. Comparison of revenues 2020 - 2023

Staff

Director & Researchers

- Garofalakis Minos, Professor, Technical University of Crete, Director
- Koutrika Georgia, Researcher A
- Sartzetakis Stylianos, Researcher A
- Simitsis Alkiviadis, Researcher A
- Dalamagas Theodore, Researcher A
- Stavrakas Yannis, Researcher A
- Skoutas Dimitris, Researcher 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, Emeritus Professor, AUEB
- Deligiannakis Antonis, Professor, Technical University of Crete
- Douligeris Christos, Professor, University of Piraeus
- Giatrakos Nikolaos, Assistant professor, Technical University of Crete
- El Raheb Katerina, Assistant Professor, University of Peloponnese
- Emiris Ioannis, Professor, University of Athens
- Fotakis Dimitrios, Professor, NTUA
- Ioannidis Yannis, Professor, University of Athens
- Kotidis Yannis, Professor, AUEB
- Mamoulis Nikolaos, Professor, University of Ioannina

- Patsakis Konstantinos, Associate Professor, University of Piraeus
- Samoladas Vasilis, Associate Professor, Technical University of Crete
- Vassiliadis Panos, Professor, University of Ioannina
- Vergados Dimitrios, Professor, University of Piraeus
- Xenakis Christos, Professor, University of Piraeus

Postdoctoral Researchers / Scientific Collaborators

- Adamidi Eleni, BSc & MSc in Applied Mathematics and Physics NTUA, PhD NTUA
- Arkoumanis Constantinos, Elec. & Comp. Eng. NTUA, PhD NTUA
- Bartzos Evangelos, BSc in Mathematics, University of Patras, MSc in advanced mathematics, Université Lyon 1, PhD Dept Informatics & Telecom. UoA
- Benardou Agiatis, BA, MA & 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
- Chatzopoulos Serafeim, BA, Msc, Dept. of Inform. and Telecom. UoA, PhD University of Peloponnese
- Christodoulou Eleni, Elec. & Comp. Eng. NTUA, Mphil in Computational Biology, University of Cambridge, PhD Biotec, Technical University Of Dresden
- Dimitropoulos Charalampos, Elec. & Elec. Eng. City University London, PhD University of London
- Dritsou Vicky, BSc in Informatics, AUEB, MSc in Information Systems, AUEB, PhD in Informatics, AUEB
- Georgakilas Giorgos, Computer Engineering and Informatics Department, University of Patras, MSc in Bioinformatics, UoA, PhD, Electrical and Computer Engineering Department, University of Thessaly
- Giannopoulos Giorgos, Elec. & Comp. Eng. NTUA, PhD NTUA
- Grypari Ioanna, BSc in Applied Mathematics & Economics, Brown University, MSc LSE, PhD in Economics, University of Minnesota
- Kaklis Dimitrios, Dept. of Inform. and Telecom. UoA, PhD Dept. of Informatics and Telematics, HUA
- Kalampratsidou Vilelmini, Electronics and Computer Engineering, Technical University of Crete, PhD in Computer Science, Rutgers University
- Kanellos Ilias, Elec. & Comp. Eng. NTUA, PhD NTUA
- Katifori Akrivi, BSc, MSc & PhD, Dept. of Inform. and Telecom. UoA
- 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
- Kolidakis Stylianos, BSc in Electrical and Electronics Engineering, AUTH, MSc & PhD, D.U.Th
- Konaxis Christos, Dept. of Mathematics UoA, PhD . Dept Informatics & Telecoms, UoA
- Lykousas Nikolaos, Department of Informatics, University of Piraeus, MSc in Intelligent Interactive Systems, Universitat Pompeu Fabra, PhD University of Piraeus
- Mailis Theofilos, Elec. & Comp. Eng. NTUA, PhD NTUA
- Maroulis Stavros, Elec. & Comp. Eng. NTUA, PhD NTUA
- Pla-Karidi Danae, Elec. & Comp. Eng. NTUA, PhD NTUA
- Psarros Ioannis, BSc, MSc & PhD, Dept. of Inform. and Telecom. UoA
- Raouzaïou 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
- 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
- Zagganas Constantinos, Elec. & Comp. Eng. NTUA, PhD University of Peloponnese

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
- Gaitanidou Kalipso, School of Political Sciences AUTH, Administrative Support
- Paradissi Eva, Administrative Support
- Pikrammenos Ioannis, Department of Computer Engineering and Informatics, University of Patras, Msc, PhD NTUA

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
- Argiropoulos Vassilios, Department of Informatics, University Of Piraeus
- Avgeridis Lazaros, Dept. of Inform. and Telecom. UoA
- Athanasiou Spyros, Elec. & Comp. Eng. NTUA
- 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
- Checa Carles, BSc in Mathematics, Universitat Politècnica de Catalunya, MSc in Mathematics, Università degli Studi di Padova & Leiden University
- Deligiannis Panagiotis, Department of Informatics and Telecommunications, University of Peloponnese
- Dimakopoulos Nikolaos, Dept. of Inform. and Telecom. UoA, MSc in Data Science & Machine Learning, NTUA
- Dorgiakis Michail, BSc in Computer Science, University of Crete, MSc in Computer Science, Vrije University Amsterdam
- Eleftheraki Stavroula, BSc in Statistics, University of the Aegean, MSc in Data Science & Information Technologies, UoA
- Emmanouilidi Sofia, Deree, Business Administration (Accounting and Finance)
- Fabrici Eros, BSc in Computer Science, Università Degli Studi di Udine, MSc in Data Science and Scientific Computing, Università Degli Studi di Trieste

- Fergadis Aris, Elec. & Comp. Eng. University of Patras, MSc in Human Language Technology UoA/NTUA
- Filippopolitis Konstantinos, BSc in Informatics, University of Piraeus
- Fofoulas 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
- Georgomanolis Nikolaos, Department of Computer Engineering and Informatics, University of Patras
- Gentimis Athanasios, BSc in Electrical and Computer Engineering, NTUA, MSc in Cyber Communication & Interactive Media, UoA
- Genevaux Anna, BSc in Applied Mathematics for Engineering, Industry and Innovation, University Paul Sabatier Toulouse
- Georgopoulos Vassilios, Dept. of Inform. and Telecom. UoA
- Giannopoulou Roza, BA in Architecture, Kingston University, MSc in Architecture, University of Westminster
- Gidaracos 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
- Iliopoulou Antonia, BSc & MSc, Dept. of Inform. and Telecom. UoA
- Ilvanidou Maria, Faculty of History and Archaeology, UoA, Msc, University of Crete
- Kafasis Periklis, BSc in Electrical & Computer Engineering AUTH, MBA in Business Administration and Management, Hellenic Open University
- Kalampokis Panagiotis, BA, Computer Science, University of Ioannina, Msc in Mobile and Web Computing, IHU
- Kalimeris Alexandros, BSc & MSc, Dept. of Inform. and Telecom. UoA
- Karabatsis Thanasis - Michail, Dept. of Inform. and Telecom. UoA
- Karagiannis Asterios, Department of Electrical and Electronics Engineering, University of West Attica
- Karmas Athanasios, Elec. & Comp. Eng. NTUA, MSc NTUA
- Karvounari Sofia, BSc, MSc, Dept. of Inform. and Telecom. UoA
- Kasomoulis Aristotelis, BSc, MSc, Dept. of Inform. and Telecom. UoA
- Katrakazi Marianna, BA in Law, UoA, MA University of Strasbourg – CEIPI
- Katsiros Dimitrios, Elec. & Comp. Eng. NTUA
- 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
- Koloveas Paris, BSc & MSc, Dept. of Informatics and Telecommunications, University of Peloponnese
- Koltsida Panayota, BA, Msc, Dept. of Inform. and Telecom. UoA
- Konstantinidis Orestis, School of Applied Mathematical and Physics Science NTUA, MSc in Computer Science UoA
- Kontopidi Maria, BSc & MSc, Department of Geography, University of the Aegean

- Kosmidou Konstantina, Department of Production and Management Engineering, Democritus University of Thrace
- Kouli Vasiliki, Department of Business Administration AUEB, MSc In International Corporate Finance & Banking, University Of Glasgow
- Koulocheri Eleni, BSc in Mathematics, UoA, MSc in Information Systems, Hellenic Open University
- Kouri Lydia, BA in Social and Education Policy, University of Peloponnese, MA in Political Science and Sociology, UoA
- Kouvaras Ioannis, Elec. & Comp. Eng. NTUA
- Kozanis Konstantinos, Dept. Of Informatics & Telecommunications, University of Thessaly
- Lampea Theodora, School of Applied Mathematical and Physics Science 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
- Lentzos Konstantinos, BSc in Mathematics UoA, MSc UoA
- Maidatsis Konstantinos, Dept. of Inform. and Telecom. UoA
- Maili Evdokia, BA in Mathematics, University of Patras, Msc, Dept. of Inform. and Telecom. UoA
- Mandamadiotis Antonios, Department of Information and Communication Systems Engineering, University of the Aegean, MSc in Data Science and Information Technologies, UoA
- Mantas Andreas, Department of Computer Engineering and Informatics, University of Patras
- Mantes Athanasios, BSc, MSc, Department of Computer Engineering and Informatics, University of Patras, MSc in Geoinformatics, NTUA, MSc in Bioinformatics, UoA
- Manola Natalia, Dept. of Physics, UOA
- Marinos Kouris Christos, Department Of Business Administration, AUEB, MSc Strathclyde University of Glasgow
- Markou Melina Elena, BSc In Mathematics with Informatics, University Of Oslo, MSc in Data Science University Of Oslo
- 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
- Melekopoglou Maria Mairi, Elec. & Comp. Eng. NTUA, MSc in Computer science, University of Wisconsin-Madison
- Mexis Konstantinos, Dept. of Elec. & Comp. Eng. NTUA
- Mitsopoulou Anna, Dept. of Inform. and Telecom. UoA
- Nakos Charalampos, Elec. & Comp. Eng. NTUA
- Oikonomou Leonidas, Architecture, School of Architecture NTUA, MSc NTUA
- Palaiologou Theoni, , BSc & MSc, Dept. of Inform. and Telecom. UoA
- 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
- Papanikos Georgios, BA, MSc, Dept. of Inform. and Telecom. UoA
- Pappas Nikolaos, Department of Electrical and Computer Engineering, University of Thessaly
- 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
- Pitsilou Vasiliki, Dept. of Inform. and Telecom. UoA
- Pristouris Costis, Elec. & Comp. Eng. NTUA, MSc in Geoinformatics, NTUA
- Psalla Maria, School of Law, AUTH, MSc University of Piraeus
- Psaroudaki Eleni, Elec. & Comp. Eng. NTUA
- Renieris Loukas, BSc in Informatics, AUEB, MSC in Informatics, University of Piraeus
- Rodopoulos Charalampos, BSc in Mathematics, University of Patras
- Rontogiannis Dimitrios, Dept. of Inform. and Telecom. UoA
- Skoufis Petros, Dept. of Elec. & Comp. Eng. NTUA
- Smyrniaios Lampros, Dept. of Inform. and Telecom. UoA
- Spirou Konstantinos, Dept. of Inform. and Telecom. UoA
- Stamatakis Giorgos, Elec. & Comp. Eng., Technical University of Crete
- Stamatopoulos Vassilis, Dept. of Inform. and Telecom. UoA
- Stefou Thodoris, BSc, MSc, Dept. of Inform. and Telecom. UoA
- Streviniotis Errikos, BSc, MSc, Elec. & Comp. Eng. Technical University of Crete
- 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
- Theologitis Nikolaos, Elec. Comp. & Eng. University of Patras
- Thorpe Rowan, Newlands College
- Toli Eleni, Department of History and Archaeology, University of Ioannina, MSc in Labour Economics Ruhr University Bochum
- Tsilipakos Georgios, BSc in Physics, AUTH, MSc Dept. of Inform. and Telecom. UoA
- Tragkas Alexandros, Elec. & Comp. Eng. Technical University of Crete
- Triantafyllou Konstantinos, Dept. of Inform. and Telecom. UoA
- Triantos Konstantinos, School of Rural, Surveying and Geoinformatics Engineering, NTUA
- 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
- Tzerefos Anargiros, Department of Informatics and Telecommunications, University of Peloponnese
- Vagianou Maria, Department of Informatics, University of Piraeus, MSc UoA
- Vargiamis Michalis, Dept. of Inform. and Telecom. UoA
- Vichos Kleanthis, Department of Electrical and Computer Engineering, University of Patras

- Xagorari Katerina, Department of Informatics and Telecommunications, University of Peloponnese
- Xenios Stefanos, BSc in Chemical Engineering, NTUA
- 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
- 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>

