ISSN 2724-6566

ISSUE 10 - 2021

Institute of Information Science and Technologies “A. Faedo”

ISTI NEWS

Digital innovation for next generation health and care

The transformation of health and care in the digital world has the potential to benefit people, health-care systems and the economy at a large.

Digital technologies such as mobile communications and Artificial intelligence (AI) offer new opportunities to transform the way we receive and provide health and care services.

They enable innovative approaches to the sustained independent living of the ageing and frail population as well as integrated health and social care that can improve long-term wellbeing and quality of life.

[continues on page 5]

Editorial

Still here - and standing strong

We are now in month 23 of the Covid19 epidemic. This unprecedented event has changed our lives and our way of working.

[continues on page 3]

“Ecce Homo” by Antonello da Messina, from non-invasive investigations to data fusion and dissemination

The investigation on Antonello da Messina’s Ecce Homo, the precious fifteenth-century painting kept at the Alberoni College in Piacenza, on the journal “Scientific Reports” of “Nature”.

[page 28]

EcoScope

Funded by Horizon 2020

The EcoScope project will develop an interoperable platform and a robust decision-making toolbox, available through a single public portal, to promote an efficient, ecosystem-based approach to the management of fisheries. It will be guided by policy makers and scientific advisory bodies, and address ecosystem degradation and the anthropogenic impact that cause fisheries to be unsustainably exploited across European seas.

[page 10]

National Research Council of Italy

Contents

Still here - and standing strong

[continued]

The last 6 months have not been easy, not for the country, not for the world - and certainly not for the scientific community. After a certain relief, owing to falling infection figures that justified a gradual return to the office in the autumn, we are currently being battered by an outbreak of the latest Covid variant and a general worsening in the situation. This means that, even while in the office, we are mostly working with digital tools and keeping in touch remotely. However, despite everything, we are still standing strong. A sense of humour and optimism have helped - especially when the news hasn't always been too good. Life goes on, with the expected and unexpected ups and downs. Happily, our research is progressing well and providing promising results. Maybe, we're doing better than we feel.

Information on our activities, including details of new projects and studies, can be found in this number. In particular, it has been particularly rewarding to see the quality and impact of past work being recognized internationally: see, on page 35, the Test of Time Award presented by IEEE to our colleagues Paolo Cignoni and Claudio Montani, for a result produced 15 years ago.

This editorial covers some of the main events at ISTI during the last semester: a member of staff elected to Scientific Board of CNR; the holding of the annual ISTI Day; activation of ISTI social media; an internal assessment of our scientific production.

Antonia Bertolino elected to the CNR Scientific Board

We are proud to announce that our colleague, Antonia Bertolino, is one of three members elected to the Scientific Board of the Italian National Research Council (CNR) by CNR researchers and technologists. The other two are Vito Mocella (ISASI-CNR) and Antonio Moretti (ISPA-CNR). The CNR Scientific Board is a collegial body of ten members. In addition to the three elected representatives, it will include seven members, still to be selected by the CNR Council of Administration from a shortlist of highly-qualified candidates drawn up by a committee of experts. The statutory role of the Scientific Board is to provide scientific advice to the CNR President and to promote proposals of strategic vision.

Of the seven internal candidates, Antonia was the only female. At a time when the EU is asking all research institutions to ensure fair opportunities and equal gender representation, CNR still has a long way to go. The first woman President, Maria Chiara Carrozza, was appointed only a few months ago, almost a century after its foundation. An important signal for the future.

Antonia has been a researcher at CNR for more than thirty years. Her area of expertise is in Software Engineering, in particular investigating program validation methodologies with the goal of increasing security and reliability. Her work is highly cited and has gained several international recognitions, including a distinguished award at ICSE, the flagship conference in software engineering, and a test-of-time award at ICST, the most prestigious conference in software testing, as well as two Facebook Research awards in 2019 and 2021.

Antonia already served as an elected member of the Scientific Board, from 2006 to 2011. At that time there were 20 members. Her objectives have not changed. She aims to contribute to improving the way research is conducted at CNR, increasing the effectiveness of recruiting and training procedures, and emphasizing international competitiveness. After having served in recent years on several councils and committees, with similar roles, for other international bodies, Antonia believes that it is now time to use this experience for her own institution. In particular, she hopes to play a role in deciding how the funds for education and research provided by the National Recovery and Resilience Plan (PNRR) will be employed.

Unfortunately, in recent years, we have witnessed a progressive distancing between central governance and the scientific network, and this has complicated and hindered research activity. The three elected members have the important responsibility of finally making the voice of all the personnel (researchers, technologists, technicians, administrators) heard, with the aim of facilitating work, enabling scientific activity and enhancing impact. Now, with a new President and a new scientific board, it is time to open a new page in our history.

A number of other events have characterized the last six months for us, at CNR in general and ISTI in particular.

Restructuring CNR

The first semester with our new President has just ended. She has been asked to address and work on solving an impressive number of problems. Most of these were not brand-new, but problematic and interrelated issues that have been impacting negatively on the organization and administration of CNR for too many years. The situation became increasingly complex when, in November, the Italian Government decided to include directives for a major restructuring of CNR in the Fiscal Law for 2022. This was unexpected and has opened the possibility for a comprehensive reform of CNR, at all possible levels (governance, central administration, staff recruitment, career progression, etc.).

Consequently, in a recent meeting with all the Directors of CNR Institutes, President Carrozza communicated that she intends to initiate a joint brainstorming effort aimed at redesigning CNR. This news was received

with enthusiasm. However, there are a number of pressing issues affecting our everyday activities and the idea of having to wait for the results of a complex reorganization increases the fear that we will have to cope with these problems for some time still.

On the positive side, a first important decision by our President, together with the General Director, regarded tenure for those included in a list of qualified researchers currently with contracts pro-tempore. Consequently, two ISTI researchers (Fabio Carrara and Francesca Pratesi) were given permanent contracts from December and, similarly, another three people should take up positions in February 2022.

ISTI Day

Last November we held the second virtual edition of our annual ISTI Day event. As Covid is still rampant and, unfortunately, the numbers are increasing (luckily, for the moment, at a lower rate than in other northern EU nations), encouraged by last year's successful experience, we decided to run the event in remote mode once again. Clearly, holding the event remotely has its pros and cons. We lose the vibe created through face-to-face communication but, at the same time, we profit from the wider diffusion of our results. In fact, all the talks given during ISTI Day 2021 were recorded, made available on our YouTube channel and publicised on our social media (https://www.isti.cnr.it/en/research/isti-day-2021).

ISTI Social

The setting up of a working group to establish a stronger presence of ISTI on social media is another recent action. ISTI is now present on LinkedIn, Twitter and Youtube, providing us with new tools to disseminate scientific results and activities. Social media are increasingly important in scientific communication (for example, and for what it's worth, a number of studies have found that articles with many tweets are more likely to be highly cited than less tweeted articles). Our social media will support our Labs in the dissemination and promotion of events, papers, projects, awards, achievements, jobs, seminars, and news in general.

ISTI Publications

Recently, we ran an analysis on the number and quality of our scientific production over the last two years. We found that the number of published papers has decreased (the number of journal papers shows a minor decrease, while the number of conference papers is more substantially reduced). This decrease in conference papers is probably due to two distinct factors: first, the results of recent competitive exams for permanent positions and the periodic research evaluation (VQR 2015-2019), which gave prominence to papers appearing in high quality publications, with a predominant preference for journal papers; second, the Covid19 epidemic, that caused either the cancellation or decrease in interest of a number of conferences. It is our impression that the long period spent working at home, in physical isolation, has had an impact on the more creative part of our activity. We believe that the loss of all those informal chances for discussing and brainstorming new ideas is a contributing factor in a reduced productivity in terms of new papers.

Let me conclude this editorial by expressing my gratitude to everyone, CNR in general, and our external national and international colleagues, for all their assistance in recent months.

But now a very special thank you to the staff of ISTI. If we are still standing strong, it entirely due to your hard work and collaboration throughout the year. Thank you, thank you, thank you.

Finally I wish you all a Happy and Prosperous NEW YEAR.

Contact: Roberto Scopigno, Director
direttore@isti.cnr.it

Digital innovation for next generation health and care

[continued]

The concept of "digital health" is situated here: at the crossroads of value-based, systems medicine and digital innovations. The goal is to harness the potentiality of the new technologies and make the patient the point of care; a complete paradigm shift moving towards a more predictive, preventative, personalised and participatory approach [1].

At the core of this vision, we find a combination of AI and machine learning, big data analytics, computer vision and graphics, as well as portable diagnostics, wearables and implantable sensors, and assistive technologies. The combination of these technologies with multiomics research is paving the way to a brand new understanding of the biological, social and environmental processes that underlie the onset of disease and is enabling the delivery of new models for disease risk prediction and early diagnosis.

This strategy is helping the health systems and the public health authorities to tackle the current Covid pandemic. While much has been done to increase vaccinations, perform more testing and tracing, and shift towards hybrid care, technological advances in standard techniques can be a major factor in helping to ease the burden of public health care institutions and caregivers, both now and in the future.

However, if we are to achieve the promises of digital health, vast amounts of high-quality data will need to be curated, and suitable regulations for privacy, data ownership, security and liability stipulated. The solutions proposed will have to be standardised and interoperable. Only those applications that guarantee reliability, end-users' trust and total patients' safety can be expected to have real impact and effective uptake.

Two main technologies exemplify the great potential of effective digital health systems: AI in Medical Imaging and Technologies for Remote Monitoring and Assistance. At ISTI, we are working in both areas.

ARTIFICIAL INTELLIGENCE and MEDICAL IMAGING

The successes of AI in computer vision have great potential. AI is at the basis of the most recent advances in computerised applications for medical image analysis and understanding. The main benefit is to reduce the workload of radiologists in intensive error-prone manual tasks, and exploit the rich content of imaging data to identify disease phenotypes for more accurate diagnoses and prognostic evaluations. Research activities at ISTI are at the forefront in this field, and include much involvement in standardization and regulation initiatives.

The Signals & Images Lab (SI-Lab) is a partner in the H2020 ProCAncer-I project, one of four large research and innovation actions funded by the European Commission in 2020 under the Call "Artificial Intelligence for Health Imaging". The project is working to deliver an AI Platform integrating imaging data and models to support precision care throughout the continuum of prostate cancer. The final goal is to develop advanced AI models to address crucial clinical needs, such as diagnosis, metastases detection and prediction of treatment response. To ensure the rapid clinical uptake of such models, ISTI-CNR and other partners are defining a governance framework for AI models to robustly monitor their performance, accuracy and reproducibility. The SI-Lab is also an active member of the Artificial Intelligence for Health Imaging (AI4HI) initiative, whose first outcome is the definition of the FUTURE-AI best practices for trustworthy AI in medical imaging. FUTURE-AI is an international, multi-stakeholder initiative aimed at defining and maintaining concrete guidelines that will facilitate the design, development, validation and deployment of trustworthy AI solutions based on six guiding principles: Fairness, Universality, Traceability, Usability, Robustness and Explainability. A large stakeholder survey has been launched based on the first version of the guidelines [2] to collect feedback from the relevant communities. The SI-Lab has also worked on designing and developing deep learning models to assess the aggressiveness of prostate cancer from Magnetic Resonance Imaging (MRI) data. The idea is to deliver methods that support the diagnostic process by limiting the need for invasive procedures such as biopsy [3]. Deep learning models [4] have been also successfully applied to the quantitative estimation of liver steatosis from ultrasound examinations. The goal is to replace more expensive methods, such as magnetic resonance spectroscopy, with comparable accuracy.

The Artificial Intelligence for Media and Humanities Lab (AIMH-Lab) has worked on the accurate diagnosis of behavioural variant frontotemporal dementia based on MRI analysis. The goal is to alleviate the tedious task of human-assisted feature crafting and to robustly handle imaging data acquired with different diagnostic devices. The research team designed and adapted simple (logistic regressor, multi-layer perceptron), advanced (convolutional networks), and modern (attention-based, transformer-like) neural architectures to work with 3D MR images, and have measured a constant improvement in the classification metrics, which reflects the underlying model complexity. In this way, they were able to improve on state-of-the-art manually engineered features and classical machine learning tools. The team also found that data intermixing, as in the case of the transformer-like architectures, produces a more robust and transferable model that appears oblivious

in terms of data acquisition devices [5]. This suggests that further investigation of such architectures opens may make it possible to advance the state-of-the-art in both classification and segmentation of medical imaging data.

The AIMH-Lab has also worked on optimising the task of detecting and quantifying biological structures in medical imaging, taking into account annotators' uncertainty. This is one of the most common tasks enabling basic research in this field, such as disease diagnosis via cell growth estimation. Since the manual detection of cells or complex structures on a large scale is often a tedious, time-consuming, and error-prone task, it was one of the earliest tasks tackled by data-driven artificial intelligence systems like neural networks. The AIMH team has devised a two-stage counting methodology (see Figure 1), which robustly handles the uncertainty of manual annotations. The methodology exploits largely available weakly-labelled single-annotator data plus a small set of multi-annotator data in which variability of labels occurs. In the first scoring stage, an initial cell localization is learnt by following standard practice. The second scoring stage is trained on multi-annotator data and assesses each object localized in the previous step. The proposed approach predicts scores associated with detected objects that are correlated with annotators in terms of agreement in detection. This improves the filtering of low-confidence detections. Experiments in counting perineuronal nets in microscopy images have shown a reduction of the counting error up to 70% when adopting our two-step methodology, independently of the specific neural architecture used.

The Formal Methods and Tools Lab (FMT-Lab) has worked on linking formal, unambiguous logical specifications to medical imaging processing. The research team involved has aimed at encoding expert knowledge in an executable form, thus ensuring that procedures are intelligible to domain experts (and not solely to programmers). This approach can be extended to other forms of analysis and may guarantee that Machine Learning procedures respect user-specified guidelines or protocols, akin to the idea of hybrid Artificial Intelligence. A domain-specific language called ImgQL ("Image Query Language") has been derived from Topological Spatial Logics, and from recent developments in the "Spatial Logic of Closure Spaces". ImgQL introduces declarative concepts such as *contact* between regions, *distance*, and *reachability*. The tool VoxLogicA is a "spatial model checker", capable of automatically computes ImgQL queries on (possibly large) image datasets. The tool is designed to support image analysis with an emphasis on simplicity, focused on producing explainable and implementation-independent results, but it is also highly optimized for performance. A ten-line-long ImgQL specification [6] has been used for the segmentation of Glioblastoma in circa 200 cases from the 2017 "Brain Tumour Segmentation (BRATS) challenge" dataset, scoring among the top ranking methods. VoxLogicA has been also used for skin lesion segmentation, the first task in melanoma diagnosis, obtaining results, in terms of accuracy and computational efficiency, in line with the state-of-the-art. The FMT-Lab team has also delivered a GPU-based version of VoxLogicA, exploiting the computational power of modern GPU devices. For more details, the reader is encouraged to see [6] and the references included.

INNOVATIVE TECHNOLOGIES for REMOTE MONITORING and ASSISTANCE

The use of innovative technologies to support remote- and self-care is considered to be the holy grail of any sustainable health service. This is particularly important in assistive scenarios for frail target population groups.

At ISTI, smart solutions for remote monitoring are being investigated with the aim of anticipating and avoiding adverse events, managing dangerous situations, and improving welfare while reducing costs and reallocating time, people and money, more conveniently. For example, in the TiAssisto Project, funded by the Tuscan Region, the SI-Lab group focuses on people with Covid-19 who may also have other serious conditions such as diabetes or heart failure. The main research interest of the project lies in the design and development of an integrated ecosystem of artificial intelligence models connected to a decision support system capable of analysing both data sent by patients and the output of diagnostic procedures performed by physicians, such as thoracic ul-

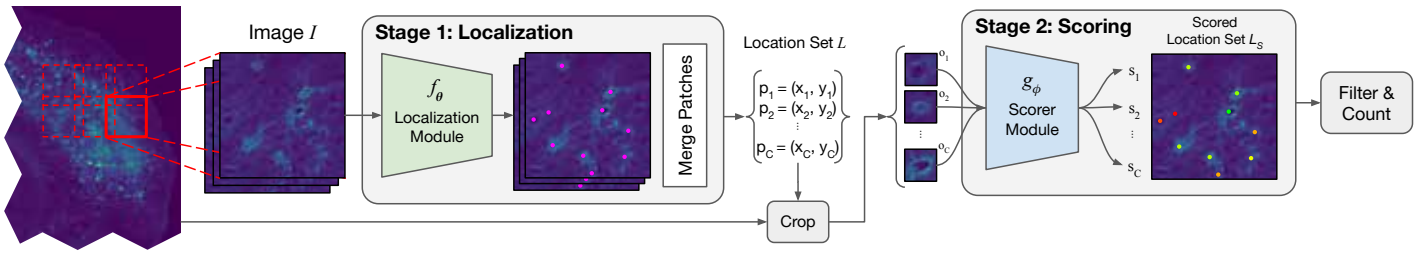


Figure 1. The two-stage cell counting pipeline proposed by the AIMH team for identifying and counting cells in microscopy images. The first stage detects the objects by exploiting a localization model trained on dot-annotated images that may have weak labels. The second stage consists of a scorer model that assigns an "objectness" score to the objects localized in the previous step. Predicted scores correlate with the agreement among annotators; higher scores tend to be assigned to patterns found by more annotators

trasound, electrocardiograms and blood gas analysis. (See Figure 2). The ultimate goal is pre-diagnosis plus the provision of valuable suggestions for supporting medical decision making.

When coping with more specific issues, for example assisting frail target populations such as older adults or children with neuro-behavioural disorders, research activities at ISTI are conducted as part of national and European initiatives.

An ageing population presents one of the most significant socio-economic challenges of the 21st century. According to estimates, more than 20% of Europeans will be 65 or older by 2025. For this reason, the EU has devoted considerable resources to ICT projects in the AHA (Active and Healthy Ageing) domain and has fostered the creation of a plethora of platforms to respond to the needs of this ageing population. Although these projects have stimulated fundamental advances in AHA, few of the resulting platforms have provided operational services, or found large-scale uptake. In order to maximise the success and impact of previous interventions, research teams from the SI-Lab and the Wireless Network Lab (WN-Lab) at ISTI are actively participating in the PlatformUptake.eu project, funded under the H2020 call "SC1-HCC-02-2019 Support for the Large-scale Uptake of Open Service Platforms in the Active and Healthy Ageing Domain".

The project delivers an inventory of the state-of-the-art by analysing the use of open service platforms in the AHA domain, covering both open platforms and partly-open/proprietary platforms developed by industry, and addressing the interactions between them. A methodology for monitoring open platform development, adoption and spread across Europe has been created by listing the key factors that determine success or hindrance in their uptake by end-user groups, and the evolution of their ecosystems and stakeholder networks [7]. Collecting and processing data from past and currently running European projects and other initiatives built upon such platforms, the project is establishing evaluation guidelines and best practice models for integrating multiple platforms. In addition, the technical, organisational, financial/business and legal aspects relevant to their future evolution and wider uptake by the end-user communities are being considered. The project analyses and depicts the entire ecosystem of open service platforms and their related networks, their achievements and potential, by targeting user groups and by generating synergies between platforms and related projects in the AHA domain.

Figure 2. The TiAssisto multi-platform dashboard

With the current increase in the size of the ageing population in many developed countries, the decline in cognitive functioning in older adults is becoming a key concern. Adequate cognitive training activities can preserve older adults' health conditions by decreasing the rate of intellectual decay and allowing seniors to live independently in their familiar environments longer. Recent studies are increasingly considering the effectiveness of serious games for people with cognitive impairment, thanks to their ability to combine the motivational character of games with the more serious goals of supporting cognitive stimulation. The Human Interfaces in Information Systems Lab (HIIS-Lab) at ISTI-CNR has been working in the area of interactive serious games for cognitively impaired seniors for several years. For example, in a previous study, they developed and assessed Micogito [8], a web game that, by using the book metaphor, supports cognitive stimulation in the elderly through tasks that replicate daily living activities, as a way of introducing them to new technologies. A strong focus has been on understanding the impact of humanoid robots in supporting serious games for seniors affected by Mild Cognitive Impairment (MCI). In particular, the HIIS-Lab has investigated how seniors with MCI relate to and perceive serious games accessed through humanoid robots (see Figure 3), as part of a training programme aimed at improving their cognitive status [9]. For this purpose, two versions of a music-based memory game have been designed by a multi-disciplinary team, one for the hu-

Figure 3. One of the humanoid robots used by the HIIS-Lab

manoid robot and one for tablet platform, and assessed during a between-subject study that involved MCI seniors. The results show that the robot was received with more enthusiasm by the elderly, thus improving their level of engagement. Most contemporary approaches do not attempt to consider emotional features in humanoid robots in an active manner to modulate robot decision-making and dialogue for the user's benefit. The HIIS-Lab is currently investigating the design of robots that show social attitude and the ability to adapt their behaviour by recognizing the user's emotion, also considering the user's specific, associated cognitive disabilities.

Another important scenario in which the use of innovative technologies plays a key role is the assistance of people with neuro-behavioural disorders, in particular children. ADHD (difficulty in focusing attention and controlling behaviour) and dyslexia (problems with language and reading) are two of the most common neuro-behavioural childhood disorders (even though distinct, they can also appear together). The HIIS-Lab has designed, developed and assessed two serious games that aim to promote the rehabilitation of children with dyslexia and children with ADHD, stimulating, in a playful form, the skills required for rehabilitation. By using different techniques (i.e. interviews, empathy maps, personas and customer journey maps), it is possible to understand the problems and needs of this class of users and thus design games that provide positive stimuli. The two games created have been designed for children of 6-8 years, and support several levels of difficulty.

The goal of the first game, "Balloons and letters", is to enhance reading process by focusing on visuospatial dyslexia, which also involves the ability to convert a grapheme into a sound (phoneme). The game design is based on the dyslexic's difficulty in distinguishing letters that are mirrored, or graphically similar when reading. The game includes a user-controlled robot-like character which fluctuates over some balloons; the user has to return the robot safely to the ground, by clicking on the balloons and bursting them, one at a time. Each balloon contains a letter or a syllable, and the user needs to burst the one containing the relevant target letter or syllable, according to the context.

The purpose of the second game, "Robot at school", is to stimulate children's sustained attention and planning skills. In this game, the user interactively controls a robot-like character that needs to collect different types of items, which are related or not, according to the task, to the school domain. A further module has been developed to allow the caregiver to personalise the games according to parameters more appropriate for the child under consideration.

Health and care are already benefitting from innovative technologies and it will increasingly do in the coming years to realise a complete paradigm shift. ISTI-CNR is contributing with viable solutions that address image-based diagnosis as well as remote monitoring and assistance. Several issues are still open and need to be carefully tackled. Our research community is ready to address these challenges to ensure the digital revolution might have a real impact and effective uptake.

Contacts

ARTIFICIAL INTELLIGENCE and MEDICAL IMAGING

Signals & Images Lab

Claudia Caudai (claudia.caudai@isti.cnr.it)
Sara Colantonio (sara.colantonio@isti.cnr.it)
Maria Antonietta Pascali (maria.antonietta.pascali@isti.cnr.it)

Artificial Intelligence for Media and Humanities Lab

Fabio Carrara (fabio.carrara@isti.cnr.it)
Luca Ciampi (luca.ciampi@isti.cnr.it)
Marco Di Benedetto (marco.dibenedetto@isti.cnr.it)

Formal Methods and Tools Lab

Vincenzo Ciancia (vincenzo.ciancia@isti.cnr.it)

INNOVATIVE TECHNOLOGIES for REMOTE MONITORING and ASSISTANCE

Signals & Images Lab

Massimo Martinelli (massimo.martinelli@isti.cnr.it)
Andrea Carboni (andrea.carboni@isti.cnr.it)
Davide Moroni (davide.moroni@isti.cnr.it)

Wireless Networks Lab

Filippo Palumbo (filippo.palumbo@isti.cnr.it)
Dario Russo (dario.russo@isti.cnr.it)

Human Interfaces in Information Systems Lab

Letizia Angileri (letizia.angileri@isti.cnr.it)
Marco Manca (marco.manca@isti.cnr.it)
Fabio Paternò (fabio.paterno@isti.cnr.it)
Carmen Santoro (carmen.santoro@isti.cnr.it)
Eleonora Zedda (eleonora.zedda@isti.cnr.it)

References

1. Communication from the Commission to the European Parliament […] on enabling the digital transformation of health and care in the Digital Single Market; empowering citizens and building a healthier society, 2018

2. Lekadir K., et al. FUTURE-AI: Guiding Principles and Consensus Recommendations for Trustworthy Artificial Intelligence in Future Medical Imaging. arXiv e-prints, arXiv: 2109.09658

3. Bertelli E. et al. Machine and deep learning prediction of prostate cancer aggressiveness using multiparametric MRI, accepted in Frontiers in Oncology, to appear in the Research Topic "Radiomics in Prostate Cancer Imaging"

4. Colantonio S. et al. A Deep Learning Approach for Hepatic Steatosis Estimation from Ultrasound Imaging, in Proc. of the International Conference on Computational Collective Intelligence, ICCCI 2021, pp. 703-714, Springer, Cham., 2021

5. Di Benedetto M. et al., Deep Networks for Behavioral Variant Frontotemporal Dementia Identification from Multiple Acquisition Sources, under revision at Artificial Intelligence In Medicine - ScienceDirect - October 2021

6. Ciancia V., Belmonte G., Latella D., Massink M. (2021) A Hands-On Introduction to Spatial Model Checking Using VoxLogicA. In: Laarman A., Sokolova A. (eds) Model Checking Software. SPIN 2021. Lecture Notes in Computer Science, vol 12864. Springer, Cham. https://doi.org/10.1007/978-3-030-84629-9_2

7. Carboni A. et al. (2021) Success and Hindrance Factors of AHA-Oriented Open Service Platforms. In: Wojtkiewicz K., Treur J., Pimenidis E., Maleszka M. (eds) Advances in Computational Collective Intelligence. ICCCI 2021. Communications in Computer and Information Science, vol 1463. Springer, Cham. https://doi.org/10.1007/978-3-030-88113-9_53

8. Palumbo V., Paternò F., Micogito: a Serious Gamebook Based on Daily Life Scenarios to Cognitively Stimulate Older Adults, GoodIT '21: Proceedings of the Conference on Information Technology for Social Good, September 2021 Pages 163–168, ACM Press

9. Manca M., Paternò F., Santoro C., Zedda E., Braschi C., Franco R., Sale A., The Impact of Serious Games with Humanoid Robots on Mild Cognitive Impairment Older Adults, International Journal of Human-Computer Studies, Volume 145, January 2021, 102509, https://doi.org/10.1016/j.ijhcs.2020.102509

EcoScope

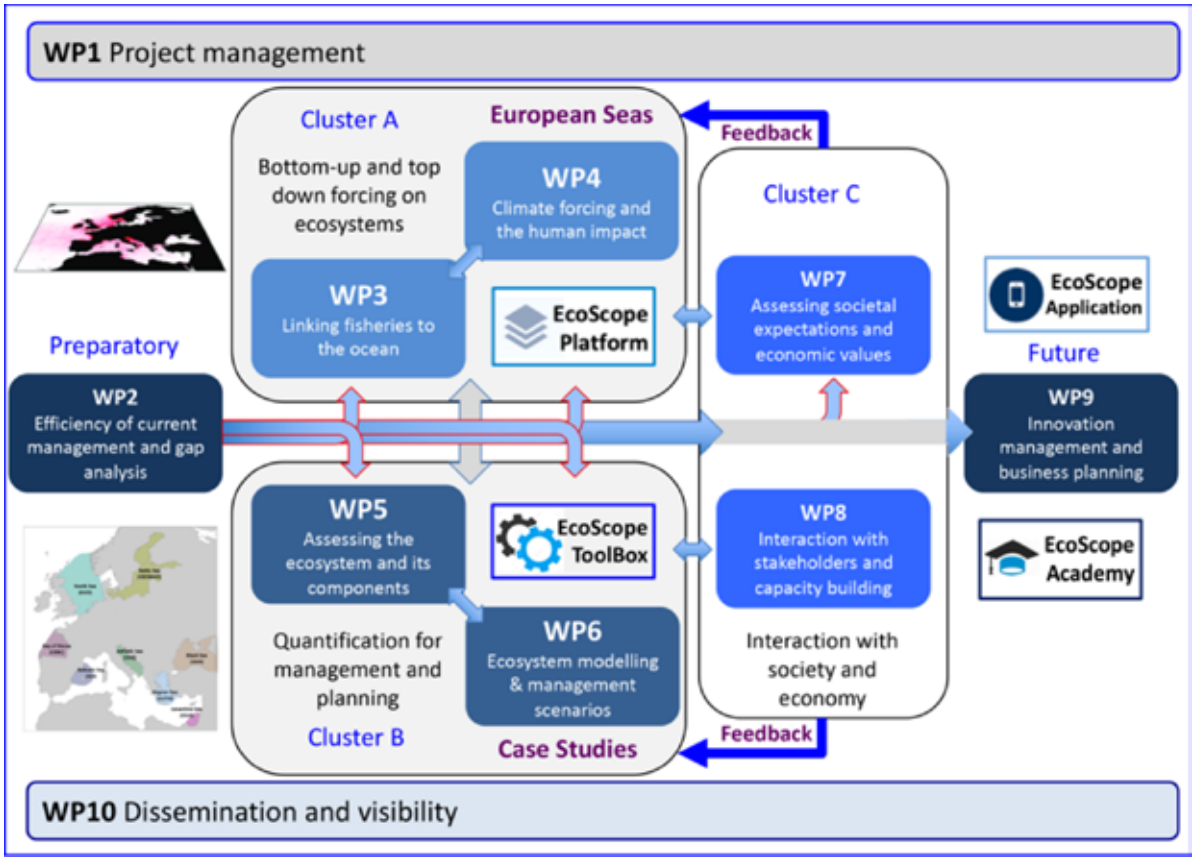
Ecocentric management for sustainable fisheries and healthy marine ecosystems Funded by Horizon 2020

The EcoScope project will develop an interoperable platform and a robust decision-making toolbox, available through a single public portal, to promote an efficient, ecosystem-based approach to the management of fisheries. It will be guided by policy makers and scientific advisory bodies, and address ecosystem degradation and the anthropogenic impact that cause fisheries to be unsustainably exploited across European seas. The EcoScope Platform will organise and homogenise climatic, oceanographic, bio-geochemical, biological and fishery datasets for European seas to a common standard type and format that will be made available through interactive mapping layers. The EcoScope Toolbox, a scoring system linked to the platform, will host ecosystem models, socio-economic indicators, fisheries and ecosystem assessment tools that can be used to examine and develop fisheries management and marine policy scenarios as well as maritime spatial planning simulations. Various groups of end-users and stakeholders will be involved in the design, development and operation of both the platform and the toolbox. Novel assessment methods for data-poor fisheries, including non-commercial species, as well as for biodiversity and the conservation status of protected megafauna, will be used to assess the status of all ecosystem components across European seas and to test new technologies for evaluating the environmental, anthropogenic and climatic impact on ecosystems and fisheries.

A series of sophisticated capacity building tools, such as online courses, documentary films, webinars and games, will be available to stakeholders through the EcoScope Academy. By filling these knowledge gaps and developing new methods and tools, the EcoScope project will provide an effective toolbox to decision makers and end-users that will be adaptive to their capacity, needs and data availability. The toolbox will incorporate methods for dealing with uncertainty and deep uncertainty; thus, it will promote efficient, holistic, ecosystem-based fisheries management that will aid towards restoring fisheries sustainability and ensuring balance between food security and healthy seas.

Contact: Gianpaolo Coro, InfraScience Lab
gianpaolo.coro@isti.cnr.it

https://cordis.europa.eu/project/id/101000302/it



SmaRIERS

Smart Railway Infrastructures: Efficiency, Reliability and Safety
Co-funded by POR FeSR 2014-2020 Ricerca e Sviluppo e innovazione (Research, Development and Innovation, Region of Tuscany)

Smart Railway Infrastructures: Efficiency, Reliability and Safety (SmaRIERS) is a technology transfer project financed by the Tuscany Region for a duration of 24 months. The project partners include ECM s.p.a. (Progress Rail of the multinational Caterpillar Inc. group), a Pistoia-based company and leader in the railway sector that develops solutions and technologies for the safety and control of railway infrastructure. The project is conducted in strong collaboration with the Department of Computer Engineering of the University of Pisa. A railway infrastructure is a highly complex system that includes several interconnected and co-operating devices, components and subsystems. In the context of the project, we will focus on the study of Uninterruptible Power Supply (UPS). UPS is a modular system which ensures the uninterruptible power supply of signaling systems. The main components of Signaling Systems are control logic devices and so-called "yard devices" installed along the railway infrastructure (switches, train detection elements, light signals, etc.) that perform the signaling tasks; vital elements for the safety of people and property that therefore require an uninterruptible power supply. The idea is to equip the UPS with a "SMART" diagnostic system capable of becoming proactive, i.e. capable of preventing critical situations or future problems, so as to allow the operator to plan appropriate actions in advance, through the analysis of plant diagnostic data. The project also includes studies of measurements related to system availability by means of stochastic modelling, as a support to the design of these station power systems.

Contact: Giorgio O. Spagnolo, FMT Lab
spagnolo@isti.cnr.it
https://smariers.isti.cnr.it

BetterWorld

Co-funded by POR FeSR 2014-2020 (Tuscan Project for Research, Development and Innovation)

The growing relevance of sustainability is leading more and more companies (in Italy and around the world) to equip themselves with a sustainability report, a strategic instrument regarding corporate activities from a corporate social responsibility perspective.

This type of reporting is typically performed by specialized consultancy firms in a purely manual form. There are very few IT tools at a global level dedicated to the purpose. Existing instruments are mostly traditional Corporate Performance Management tools (focused on economic-financial aspects), simply enriched with some sustainability indicators, therefore unable to exploit all the data relevant internally and externally to the company (web, IoT). Neither do they use advanced business intelligence functionalities or targeted predictive analysis to identify ways to improve sustainable performance. In addition, these tools are almost exclusively the prerogative of large companies, whereas the issues of sustainability are increasingly also becoming crucial for SMEs.

The BetterWorld project is aimed at creating a product/service which enables all companies interested in moving towards sustainability to collect relevant external and internal data and draw up their own sustainability report, automatically and in line with the best international reporting standards (GRI, IIRC). Timely indications for improving sustainability performance, based on comparison with the market, are also provided.

BetterWorld will collect, harmonize and systematize data from internal sources (management, CRM), from semantic analysis of web content, from IoT devices and from manual inputs. The project will leverage the combination of Business Intelligence, Big Data analysis, artificial intelligence and predictive analysis, in full Industry 4.0 perspective. This data will be exploited to perform sustainable performance analysis, benchmark analysis, and gap analysis. It will enable continuous monitoring and provide forecasts of company performance based on pre-established objectives.

BetterWorld will be based on an ontology of processes expressed in the standards of the semantic web and will incorporate an evolved speech-to-machine interaction system to interpret and receive information in natural language. The usability of the system will be enhanced compared to the majority of existing software. It will also offer intuitive, functional, interactive and state-of-the-art interfaces for input and display of results.

Contact: Carlo Meghini, AIMH Lab
carlo.meghini@isti.cnr.it

SERENI

Funded by Progetti di Ricerca@Cnr (CNR Research Projects - National Research Council)

The rise in the proportion of older people in society makes it of increasing economic and social importance to understand how to maintain the health of the aging mind.

One way is to offer the older adults the opportunity to engage in mentally stimulating activities. Serious interactive games can be useful for this. So far, most of such games have been implemented via tablets, which are a cost-effective solution but offer limited possibilities for full engagement in a multimodal way. Some studies have found that older adults complain about the complexity of both hardware and software, and this can be problematic due to age-related cognitive and physical limitations.

Thanks to their morphological characteristics and multimodal interaction, humanoid robots can open up new possibilities for more effective involvement of the elderly. They appear to have the potential to help formal and informal caregivers monitor and support older adults by providing a mixture of empathy, motivation, encouragement and companionship.

This project aims at developing innovative, serious games that help the elderly to maintain their cognitive and social functional level and thus increasing their prospects of living independently longer.

A combination of innovative technological solutions, including humanoid robots, and knowledge on age-related cognitive fragility will potentially have a positive effect on known key indicators of independent living, such as cognitive performance (e.g. memory, attention and processing speed), walking speed and balance.

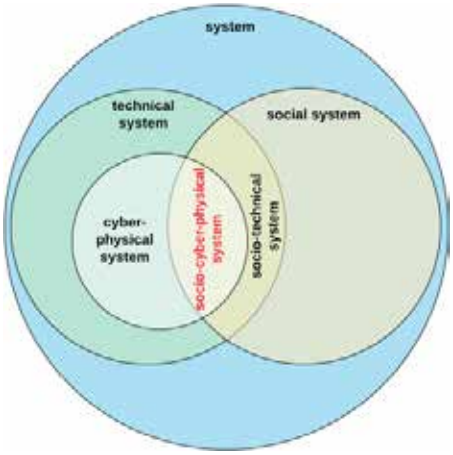
The humanoid robots will act as personal trainers, proposing exercises, communicating through various modalities and stimulating users in cognitive games relevant to their daily life. One aspect that will be carefully considered is that of personality, both of the elderly and that programmed into the robot. In this respect, we will consider those personality traits most examined in the literature. For example, we propose robots that will play with a more or less introverted personality, and will investigate if this can influence the involvement of older adults in games and to what extent. The robot's "personality" will manifest itself both in verbal (e.g. volume, tone, speed, rhythm, length of expressions) and non-verbal (e.g. gestures, sounds, movements) communication.

Contact: Fabio Paternò, HIIS Lab
fabio.paterno@isti.cnr.it
https://hiis.isti.cnr.it/sereni/

Digital transformation of agriculture and rural areas: a socio-cyber-physical system framework to support responsibilisation

K. Rijswijk, L. Klerkx, M. Bacco, F. Bartolini, E. Bulten, L. Debruyne, J. Dessein, I. Scotti, G. Brunori

Journal of Rural Studies, vol. 85. Elsevier, 2021.



Hierarchy of system concepts.

Digital technologies are often seen as an opportunity to enable sustainable futures in agriculture and rural areas. However, this digital transformation process is not inherently good as it impacts on many aspects (e.g. economic, environmental, social, technological, institutional) and their relations. The responsible research and innovation approach calls for a better understanding and anticipation of the often unknown impacts. To meet this aim we have developed a framework that allows to gain insight on the relations between the social, the cyber and the physical, i.e. a socio-cyber-physical system and have described conditions for a successful digital transformation of such a system. These are design of, and creating access to digital technologies, and navigating system complexity. This framework allows for a better problematisation of digital transformation and has been illustrated through an example of digital dairy farming. It supports an enhanced understanding of moral responsibilities regarding digital transformation, fitting within the responsible research and innovation approach, as well as a better understanding who is responsible or accountable for the identified (positive or negative) impacts, i.e. responsibilisation.

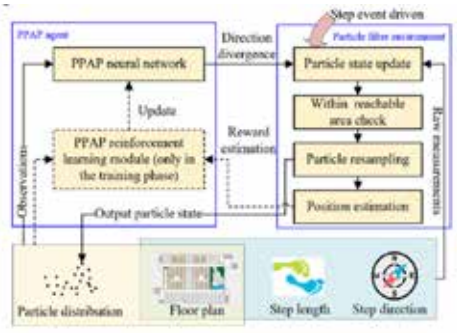
DOI: 10.1016/j.jrurstud.2021.05.003

Particle filter reinforcement via context-sensing for smartphone-based pedestrian dead reckoning

W. Shao, F. Zhao, H. Luo, H. Tian, J. Li, A. Crivello

IEEE communications letters, vol. 25. IEEE, 2021.

Pedestrian dead reckoning based on particle filter is commonly used for enabling seamless smartphone-based indoor positioning. However, compass directions indoor are heavily distorted due to the presence of ferromagnetic materials. Conventional particle filters convert the raw compass direction to a distribution adding a constant variance noise and leveraging a particle swarm to simulate the distribution. Finally, the selection of eligible directions is performed applying external constraints mainly imposed from the indoor map. However, the choice of a constant parameter decreases the positioning performances because the variance of nearby context, including topography, ferromagnetic materials, and particle distribution, is not represented. Therefore, we propose the particle filter reinforcement able to adaptively learn and adjust the variance of the direction observing the context in real-time. Experiments in real-world scenarios show that the proposed method improves the positioning accuracy by more than 20% at the 80% probability compared with state-of-the-art methods..



Overview of the proposed particle filter reinforcement learning.

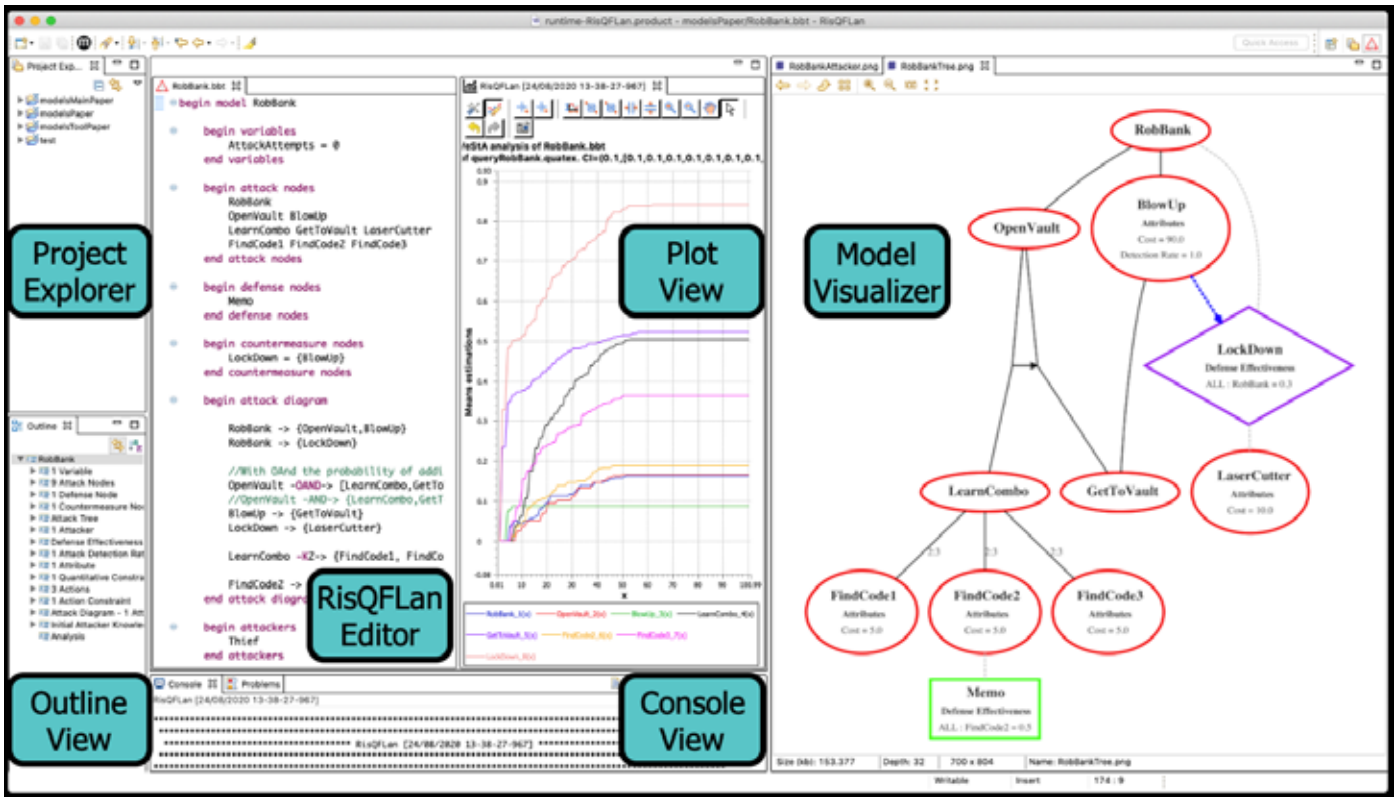
DOI: 10.1109/LCOMM.2021.3090300

Quantitative security risk modeling and analysis with RisQFLan

M.H. ter Beek, A. Legay, A. Lluch Lafuente, A. Vandin
Computers & Security , vol. 109. Elsevier, 2021.

Domain-specific quantitative modeling and analysis approaches are fundamental in scenarios in which qualitative approaches are inappropriate or unfeasible. In this paper, we present a tool-supported approach to quantitative graph-based security risk modeling and analysis based on attack-defense trees. Our approach is based on QFLan, a successful domain-specific approach to support quantitative modeling and analysis of highly configurable systems, whose domain-specific components have been decoupled to facilitate the instantiation of the QFLan approach in the domain of graph-based security risk modeling and analysis. Our approach incorporates distinctive features from three popular kinds of attack trees, namely enhanced attack trees, capabilities-based attack trees and attack countermeasure trees, into the domain-specific modeling language. The result is a new framework, called RisQFLan, to support quantitative security risk modeling and analysis based on attack-defense diagrams. By offering either exact or statistical verification of probabilistic attack scenarios, RisQFLan constitutes a significant novel contribution to the existing toolsets in that domain. We validate our approach by highlighting the additional features offered by RisQFLan in three illustrative case studies from seminal approaches to graph-based security risk modeling analysis based on attack trees.

DOI: 10.1016/j.cose.2021.102381



A screenshot of RisQFLan.

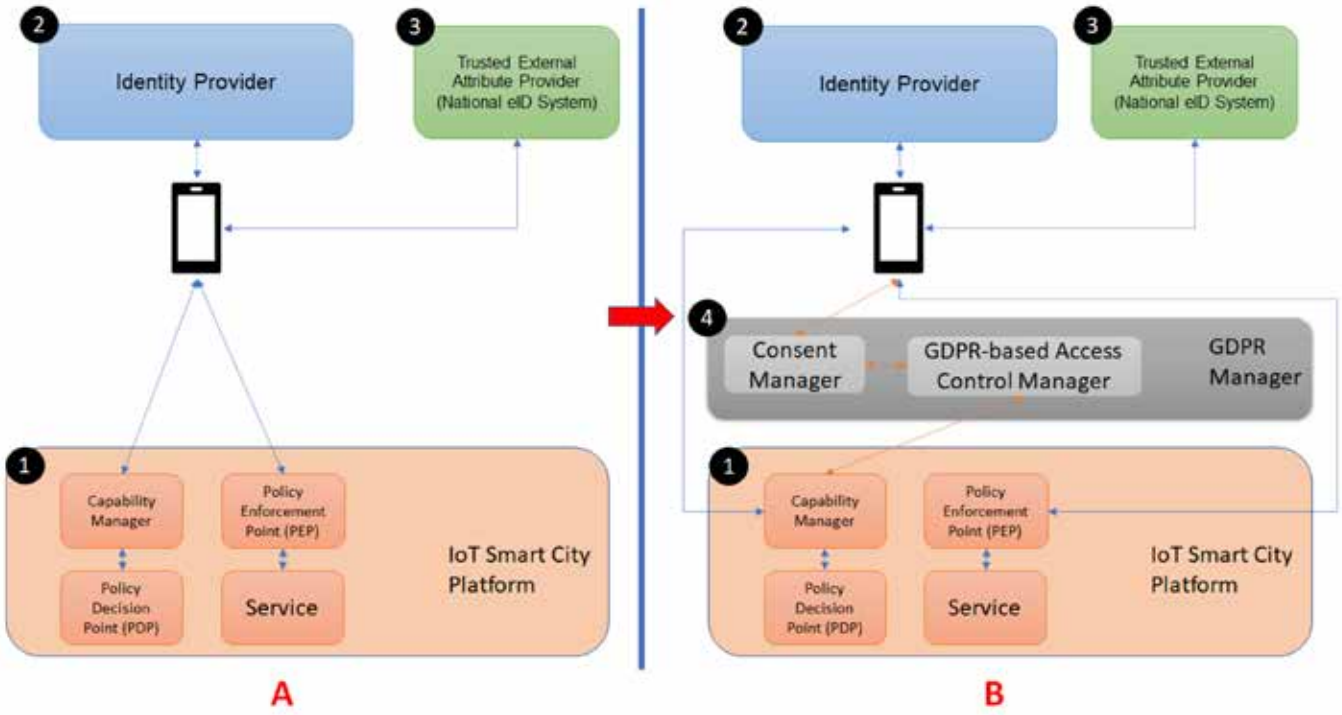
Data protection by design in the context of smart cities: a consent and access control proposal

S. Daoudagh, E. Marchetti, V. Savarino, J.B. Bernabe, J. García-Rodríguez, R. Torres Moreno, J.A. Martinez, A.F. Skarmeta

Sensors, vol. 21. MDPI, 2021.

The growing availability of mobile devices has led to an arising development of smart cities services that share a huge amount of (personal) information and data. Without accurate and verified management, they could become severe back-doors for security and privacy. In this paper, we propose a smart city infrastructure able to integrate a distributed privacy-preserving identity management solution based on attribute-based credentials (p-ABC), a user-centric Consent Manager, and a GDPR-based Access Control mechanism so as to guarantee the enforcement of the GDPR's provisions. Thus, the infrastructure supports the definition of specific purpose, collection of data, regulation of access to personal data, and users' consents, while ensuring selective and minimal disclosure of personal information as well as user's unlinkability across service and identity providers. The proposal has been implemented, integrated, and evaluated in a fully-fledged environment consisting of MiMurcia, the Smart City project for the city of Murcia, CaPe, an industrial consent management system, and GENERAL_D, an academic GDPR-based access control system, showing the feasibility.

DOI: 10.3390/s21217154



Privacy-By-Design Concept.

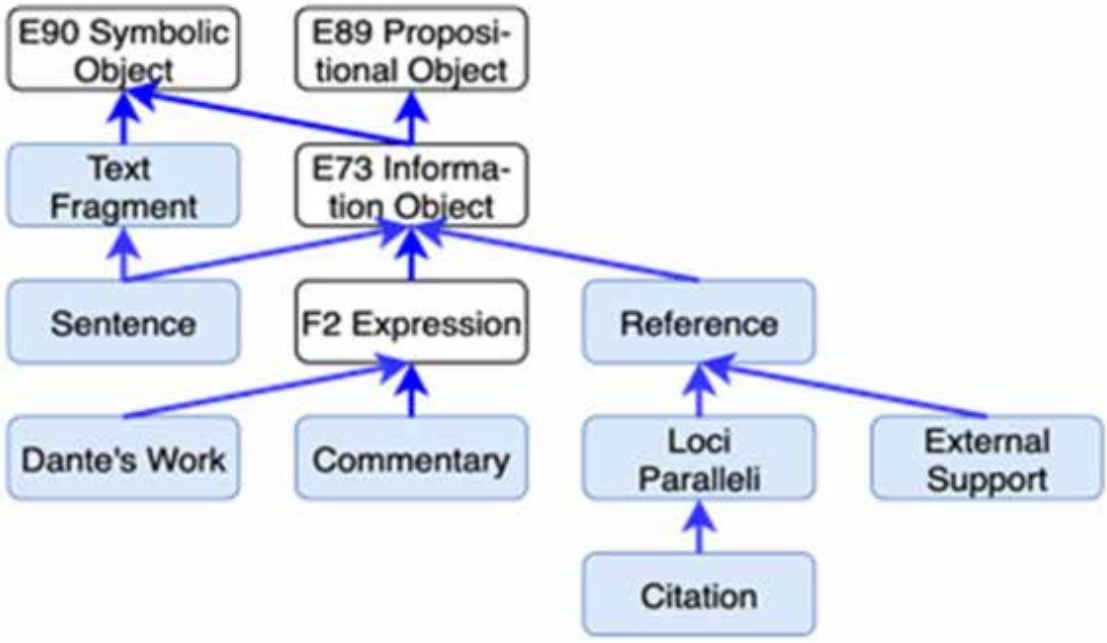
A formal representation of the divine comedy's primary sources: the Hypermedia Dante Network ontology

V. Bartalesi, N. Pratelli, C. Meghini, D. Metilli, G. Tomazzoli, L.M.G. Livraghi, M. Zaccarello

Digital Scholarship in the Humanities, in press. OUP, 2021.

Hypermedia Dante Network (HDN) is a 3-year Italian National Research Project, started in 2020, which aims to enrich the functionalities of the DanteSources Digital Library to efficiently represent knowledge about the primary sources of Dante's Comedy. DanteSources allows users to retrieve and visualize the list and the distribution of Dante's primary sources that have been identified by recent commentaries of five of Dante's minor works (i.e. Vita nova, De vulgari eloquentia, Convivio, De Monarchia, and Rime). The digital library is based on a formal ontology expressed in Resource Description Framework Schema (RDFS) language. Based on the DanteSources experience, the HDN project aims to formally represent the primary sources of the Divine Comedy whose identification is based on several commentaries included in the Dartmouth Dante Project corpus. To reach this goal, we restructured and extended the DanteSources ontology to provide a wider and more complete representation of the knowledge concerning the primary sources of the Comedy. In this article, we present the result of this effort, i.e. theHDN ontology. The ontology is expressed inOWL and has as reference ontologies the CIDOC CRMand its extension FRBRoo, including its in-progress reformulation LRMoo.We also briefly describe the semi-automatic tool that will be used by the scholars to populate the ontology.

DOI: 10.1093/llc/fqab080



The ontology classes representing the references and the main structural components of Divine Comedy andCommentaries. The classes we defined are highlighted in light blue and the subclass relations in blue.

Towards a knowledge base of medieval and renaissance geographical Latin works: The IMAGO ontology

V. Bartalesi, D. Metilli, N. Pratelli, P. Pontari

Digital Scholarship in the Humanities, in press. OUP, 2021.

In this article we present the first achievement of the Index Medii Aevi Geographiae Operum (IMAGO)—Italian National Research Project (2020–23), that is, the ontology we have created in order to formally represent the knowledge about the geographical works written in Middle Ages and Renaissance (6th–15th centuries). The IMAGO ontology is derived from a strict collaboration between the Institute of Information Science and Technologies (ISTI) of the Italian National Research Council (CNR) and the scholars who are involved in the project, who have supported ISTI-CNR in defining a conceptualization of the domain of knowledge. Following the re-use logic, we have selected as reference ontologies the International Committee on Documentation CRM vocabulary and its extension FRBRoo, including its in-progress reformulation, LRMoo. This research is included in a wider project context whose final aim is the creation of a knowledge base (KB) of Latin geographic literature of the Middle Ages and Renaissance Humanism in which the data are formally represented following the Linked Open Data paradigm and using the Semantic Web languages. At the end of the project, this KB will be accessed through a Web application that allows retrieving and consulting the collected data in a user-friendly way for scholars and general users, e.g. tables, maps, CSV files.

DOI: 10.1093/llc/fqab060

Lost in transduction: transductive transfer learning in text classification

A. Moreo, A. Esuli, F. Sebastiani

ACM Transactions on Knowledge Discovery from Data. ACM, 2021.

Obtaining high-quality labelled data for training a classifier in a new application domain is often costly. *Transfer Learning* (a.k.a. "Inductive Transfer") tries to alleviate these costs by transferring, to the "target" domain of interest, knowledge available from a different "source" domain. In transfer learning the lack of labelled information from the target domain is compensated by the availability at training time of a set of unlabelled examples from the target distribution. *Transductive Transfer Learning* denotes the transfer learning setting in which the only set of target documents that we are interested in classifying is known and available at training time. Although this definition is indeed in line with Vapnik's original definition of "transduction" current terminology in the field is confused. In this article we discuss how the term "transduction" has been misused in the transfer learning literature, and propose a clarification consistent with the original characterization of this term given by Vapnik. We go on to observe that the above terminology misuse has brought about misleading experimental comparisons, with inductive transfer learning methods that have been incorrectly compared with transductive transfer learning methods. We then give empirical evidence that the difference in performance between the inductive version and the transductive version of a transfer learning method can indeed be statistically significant (i.e., that knowing at training time the only data one needs to classify indeed gives an advantage). Our clarification allows a reassessment of the field, and of the relative merits of the major, state-of-the-art algorithms for transfer learning in text classification.

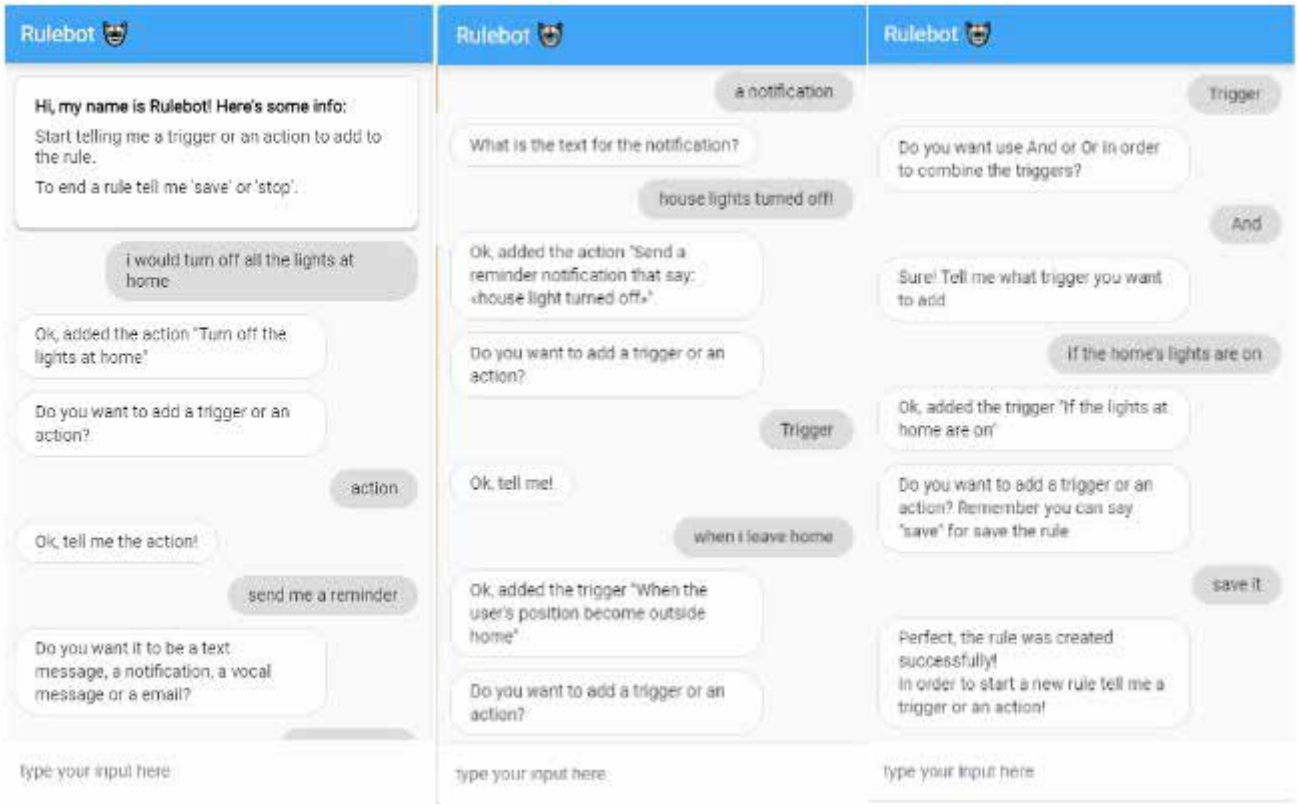
DOI: 10.1145/3453146

Comparative analysis of composition paradigms for personalization rules in IoT settings

S. Gallo, M. Manca, A. Mattioli, F. Paternò, C. Santoro
IS-EDU 2021 - 8th International Symposium on End User Development. Springer, 2021.

The rapid pervasive diffusion of Internet of Things technologies has opened up many opportunities for people to directly personalise the behaviour of surrounding objects and devices based on the dynamic events that can occur. To this end, several tailoring environments have been proposed supporting the end-user creation of trigger-action rules. Such tools can support different composition paradigms. In this paper we present a study that analyses three composition paradigms (graphical wizard, block-based, and conversational) to better understand how well they support rule creation activities. In order to make the analysis consistent we considered three implementations of such composition paradigms supporting the same set of triggers and actions. We have carried out a first user study in order to gather empirical feedback for substantiating our analysis, which provides indications of the pros and cons of each approach.

DOI: 10.1007/978-3-030-79840-6_4



Example of dialogue with the conversational interface.

The role of technology and digital innovation in sustainability and decarbonization of the Blue Economy

E.F. Campana, E. Ciappi, G. Coro
Bulletin of Geophysics and Oceanography, vol. 62 (suppl. 3). Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, 2021.

The development of a sustainable technology for the Blue Economy (a new Blue Technology) sets out three core research objectives, reflecting key challenges to be tackled by the sea industries and scientific and technological communities: The fast development of doable decarbonization processes through development and demonstration of deployable, competitive, and sustainable technological solutions for energy transition (climate neutral blue economy), a sustainable exploitation and exploration of oceans, seas and coastal areas to provide new resources, from raw materials to products, including food (sustainable use and management of marine resources), and the development and exploitation of digital-based knowledge while accumulating data from new observation networks (persistent monitoring and digitalization of seas and oceans). To meet these operational objectives, different topics and related technologies need to be further developed. A possible list of disciplinary objectives is the following.

Data poor approach for the assessment of the main target species of rapido trawl fishery in Adriatic Sea

M. Scanu, E. Armelloni, G. Coro, F. Masnadi, S. Angelini, G. Scarcella
Frontiers in Marine Science, vol. 8. Frontiers, 2021.

Information on stock status is available only for a few of the species forming the catch assemblage of rapido fishery of the North-central Adriatic Sea (Mediterranean Sea). Species that are caught almost exclusively by this gear, either as target (such as Pectinidae) or accessory catches (such as flatfishes apart from the common sole), remain unassessed mainly due to the lack of data and biological information. Based on cluster analysis, the catch assemblage of this fishery was identified and then assessed using the CMSY Bayesian model. The results of this data-poor methodology showed that, among the species analyzed, no one is sustainably exploited. The single-species CMSY results were used as the input to an extension of the same model, to test the effect of four different harvest control rule (HCR) scenarios on the entire catch assemblage, through 15-years forecasts.

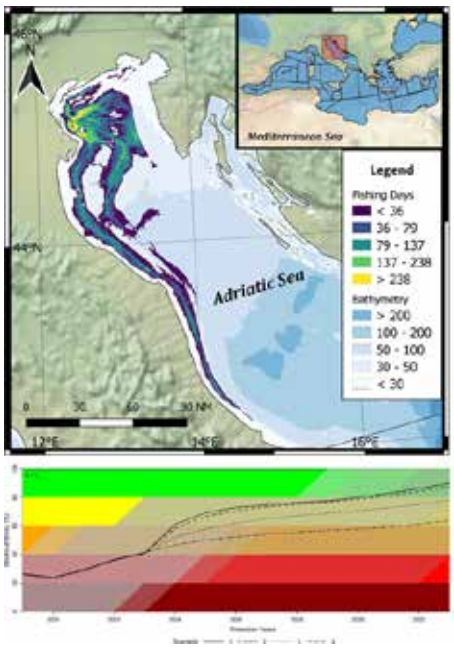


Diagram of a standard ASR with alternative acoustic models: (a) tri-phonetic HMMs using GMM emission probabilities, (b) tri-phonetic HMMs using a DNN to simulate emission probabilities.

The analysis showed that the percentage of the stocks that will reach sustainability at the end of the projections will depend on the HCR applied. Forecasts showed that a reduction of 20% of fishing effort may permit to most of the target and accessory species of the rapido trawl fishery in the Adriatic Sea to recover to sustainable levels within 15 years, also providing a slight increase in the expected catches.

DOI: 10.3389/fmars.2021.552076

An Open Science approach to infer fishing activity pressure on stocks and biodiversity from vessel tracking data

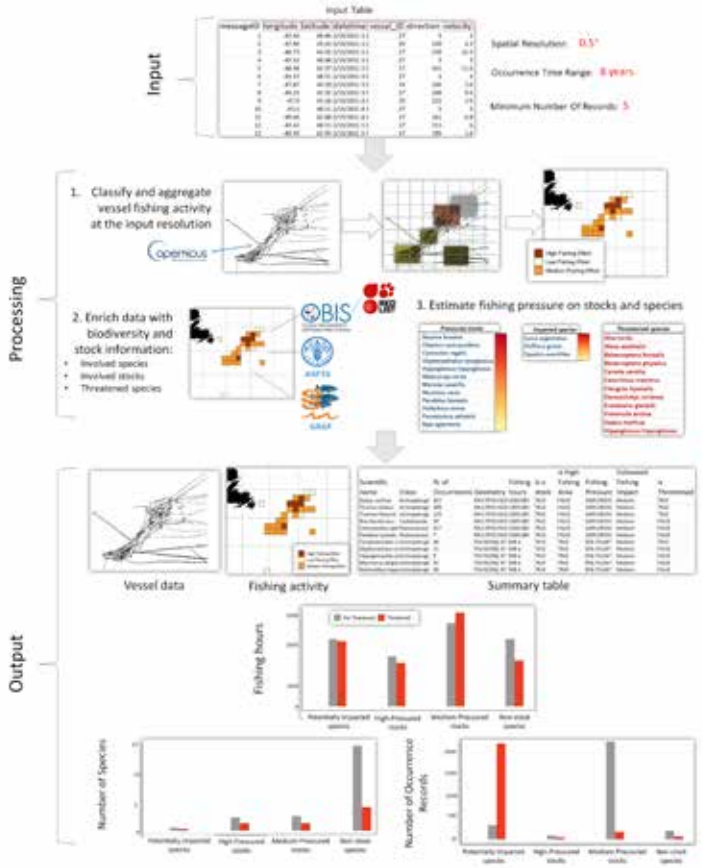
G. Coro, A. Ellenbroek, P. Pagano
Ecological Informatics, vol. 64. Elsevier, 2021.

Vessel tracking data help study the potential impact of fisheries on biodiversity and produce risk assessments. Existing workflows process vessel tracks to identify fishing activity and integrate information on species vulnerability. However, there are significant data integration challenges across the data sources needed for an integrated impact assessment due to heterogeneous nomenclatures, data accessibility issues, geographical and computational scalability of the processes, and confidentiality and transparency towards decision making authorities.

This paper presents an Open Science data integration approach to use vessel tracking data in integrated impact assessments. Our approach combines heterogeneous knowledge sources from fisheries, biodiversity, and environmental observations to infer fishing activity and risks to potentially impacted species. An Open Science e-Infrastructure facilitates access to data sources and maximises the reproducibility of the results and the method's reusability across several application domains.

Our method's quality is assessed through three case studies: The first demonstrates cross-dataset consistency by comparing the results obtained from two different vessel data sources. The second performs a temporal pattern analysis of fishing activity and potentially impacted species over time. The third assesses the potential impact of reduced fishing pressure on marine biodiversity and threatened species due to the 2020 COVID-19 lockdown in Italy. The method is meant to be integrated with other systems through its Open Science-oriented features and can rapidly use new sources of findable, accessible, interoperable, and reusable (FAIR) data. Other systems can use it to (i) classify vessel activity in data-limited scenarios, (ii) identify bycatch species (when catchability data are available), and (iii) study the effects of fisheries on habitats and populations' growth.

DOI: 10.1016/j.ecoinf.2021.101384



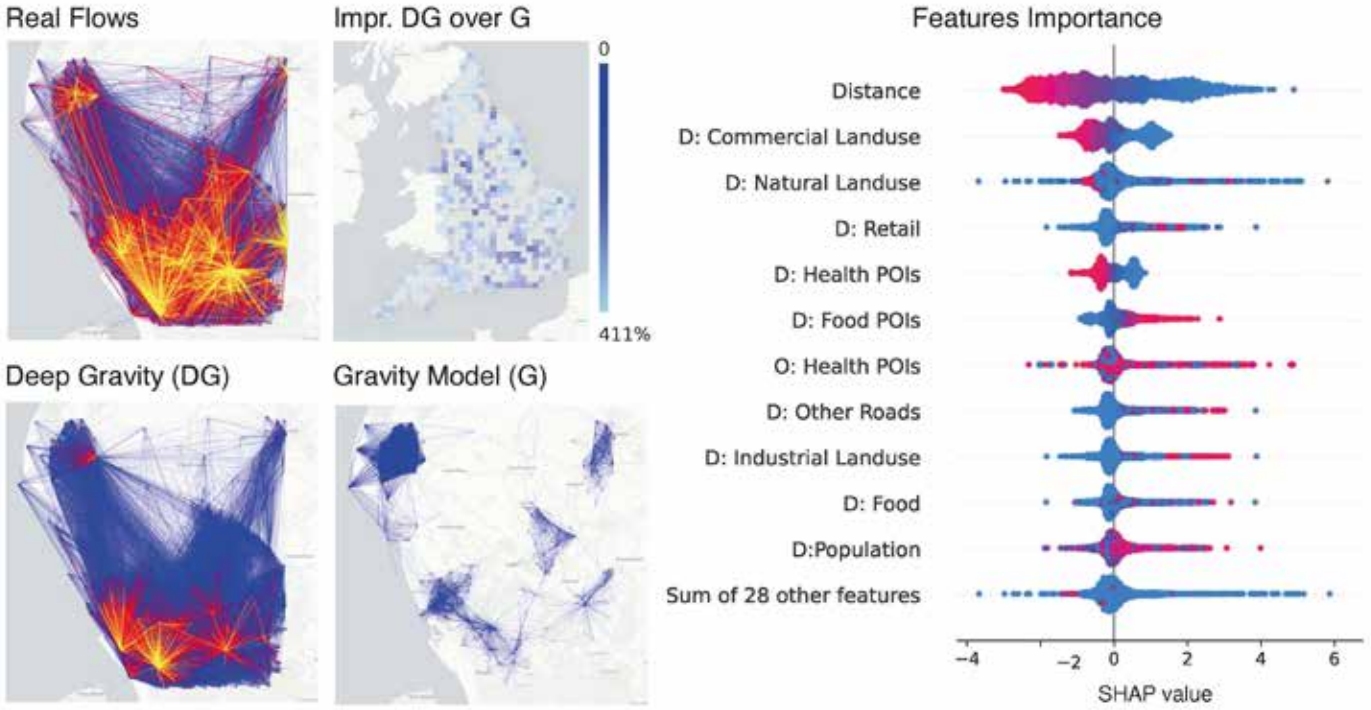
Schema of our methodological workflow, with processing divided into three separate steps. Red boxes in the input vessel tracking data table highlight mandatory fields of our process.

A deep gravity model for mobility flows generation

F. Simini, G. Barlacchi, M. Luca, L. Pappalardo
Nature Communications, vol. 12. Springer, 2021.

The movements of individuals within and among cities influence critical aspects of our society, such as well-being, the spreading of epidemics, and the quality of the environment. When information about mobility flows is not available for a particular region of interest, we must rely on mathematical models to generate them. In this work, we propose Deep Gravity, an effective model to generate flow probabilities that exploits many features (e.g., land use, road network, transport, food, health facilities) extracted from voluntary geographic data, and uses deep neural networks to discover non-linear relationships between those features and mobility flows. Our experiments, conducted on mobility flows in England, Italy, and New York State, show that Deep Gravity has good geographic generalization capability, achieving a significant increase in performance, especially in densely populated regions of interest, with respect to the classic gravity model and models that do not use deep neural networks or geographic data. Deep Gravity has good generalization capability, generating realistic flows also for geographic areas for which there is no data availability for training. Finally, we show how flows generated by Deep Gravity may be explained in terms of the geographic features and highlight crucial differences among the three considered countries interpreting the model's prediction with explainable AI techniques.

DOI: 10.1038/s41467-021-26752-4



(LEFT) Visualization of the mobility network describing real flows, the flows generated by Deep Gravity (DG), and those generated by the gravity model (G) on a region of interest with 1001 locations (OAs) in the north of Liverpool, England, UK. Coloured edges denote observed or average flows: blue edges indicate flows with a number of commuters between 0 and 3, red edges between 3 and 5, and yellow edges above 5 commuters. On the map of England, we visualize the relative improvement of DG over G in terms of CPC (Common Part of Commuters). While both DG and G underestimate the flows, DG captures the overall structure of the flow network more accurately than G.

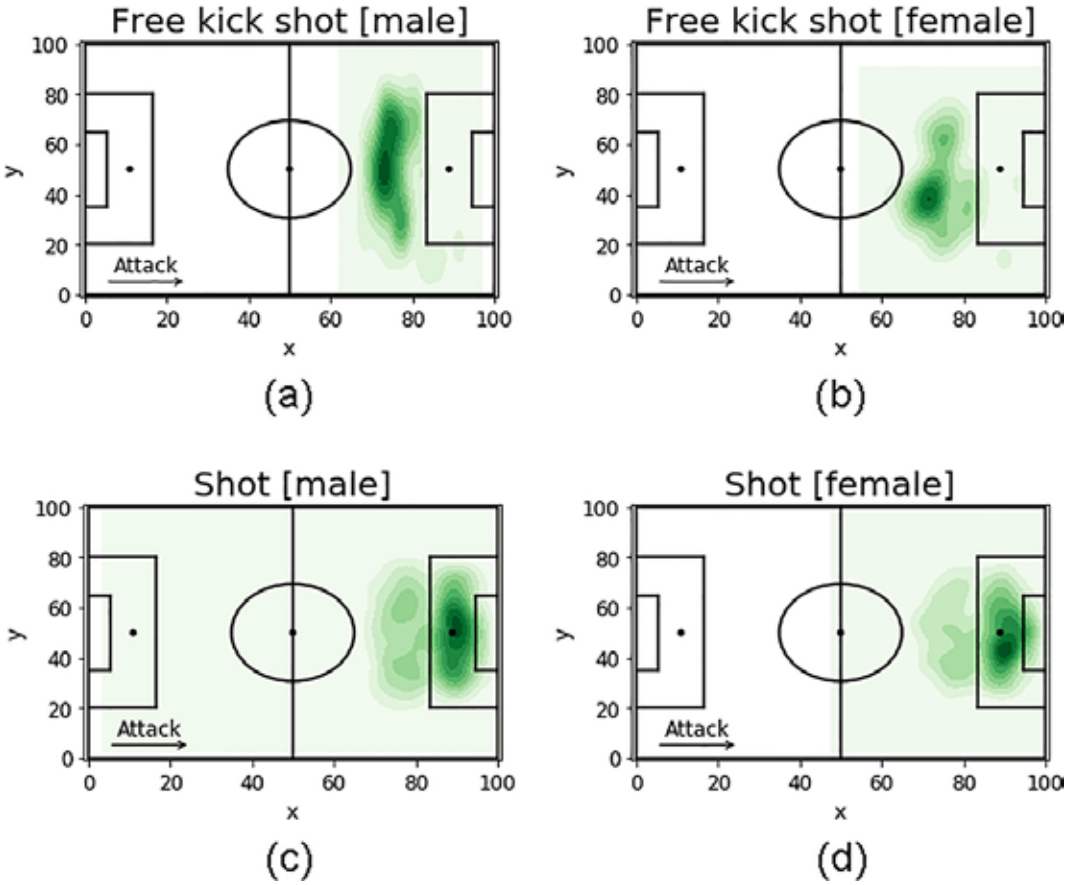
(RIGHT) Distribution of Shapely values for all features in Deep Gravity for England. Features are reported on the vertical axis and are sorted from the most relevant on top to the least relevant on the bottom. Feature names starting with "D:" and "O:" indicate features of the destination and origin, respectively. Each point denotes an origin-destination pair, where blue points represent pairs where the feature has a low value and red points pairs with high values. For each variable, these points are randomly jittered along the vertical axis to make overlapping ones visible. The point's position on the horizontal axis represents the feature's Shapely value for that origin-destination pair, that is, whether the feature contributes to increasing or decreasing the flow probability for that pair. For example, high distances are a travel deterrence while short distances are associated with an increment of commuters.

Explaining the difference between men's and women's football

L. Pappalardo, A. Rossi, M. Natilli, P.Cintia
PLOS ONE, vol. 16. Public Library of Science, 2021.

Women's football is gaining supporters and practitioners worldwide, raising questions about what the differences are with men's football. While the two sports are often compared based on the players' physical attributes, we analyze the spatio-temporal events during matches in the last World Cups to compare male and female teams based on their technical performance. We train an artificial intelligence model to recognize if a team is male or female based on variables that describe a match's playing intensity, accuracy, and performance quality. Our model accurately distinguishes between men's and women's football, revealing crucial technical differences, which we investigate through the extraction of explanations from the classifier's decisions. The differences between men's and women's football are rooted in play accuracy, the recovery time of ball possession, and the players' performance quality. Our methodology may help journalists and fans understand what makes women's football a distinct sport and coaches design tactics tailored to female teams.

DOI: 10.1371/journal.pone.0255407



Heatmaps describing the pitch zones from where free-kickshots and shots in motion are more likely to be made by male and female players , computed as the kernel estimate of the first grade intensity function, where the event points are the free-kickshots and the shots in motion, and the football field is the region of interest. The darker is the green, and the higher is the number of free-kickshots and shots in motion in that field zone. The pitch length (x) and width (y) are in the range [0,100], which indicates the percentage of the field starting from the left corner of the attack in gteam. The darker is the green, and the higher is the number of free-kick shots and shots in motion in that field zone. The pitch length (x) and width (y) are in the range [0, 100], which indicates the percentage of the field starting from the left corner of the attacking team.

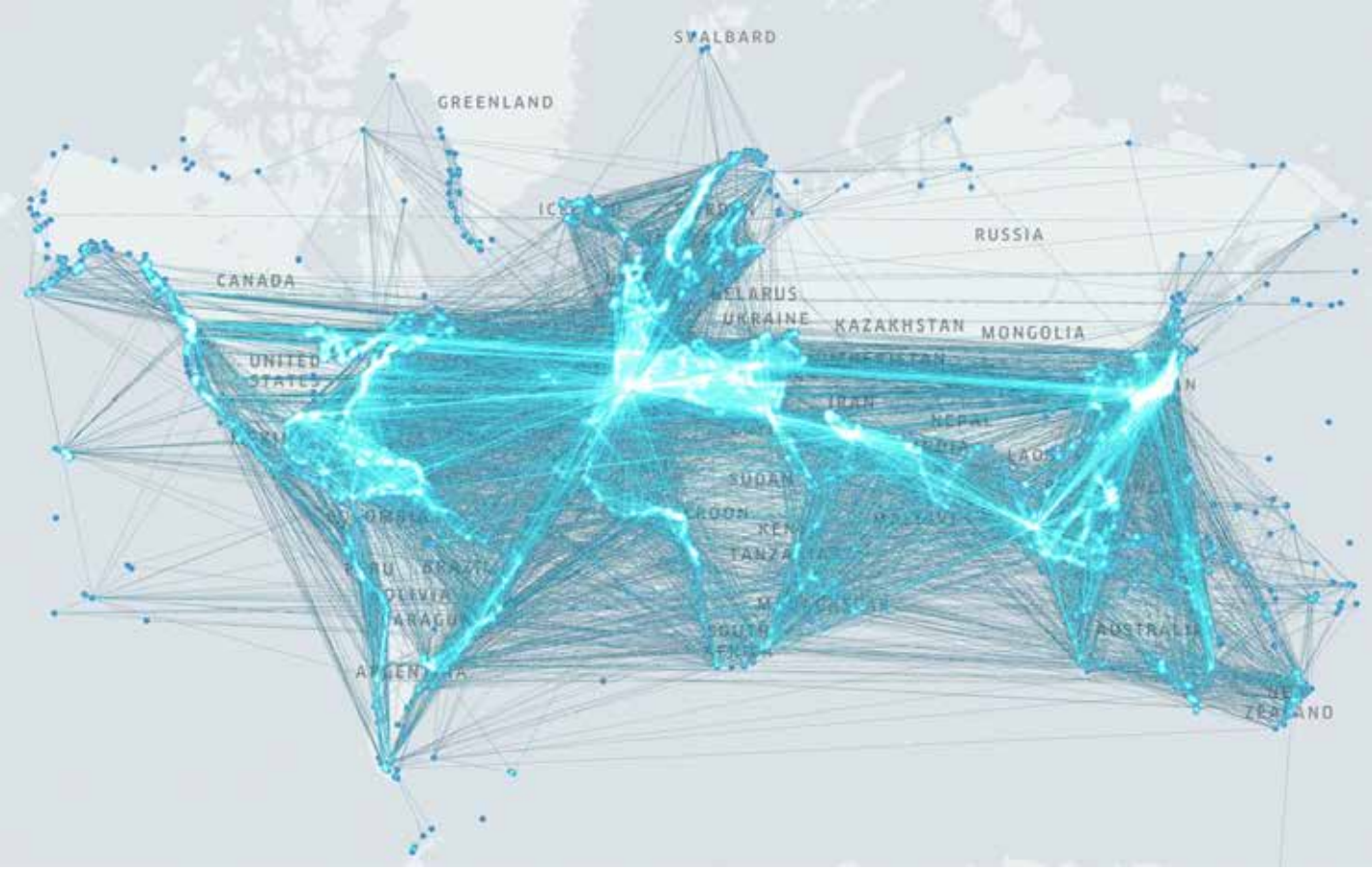
Understanding evolution of maritime networks from automatic identification system data

E. Carlini, V. Monteiro de Lira, A. Soares, M. Etemad, B. Brandoli, S. Matwin
GeoInformatica, in press. Springer, 2021.

The study of vessel movements is a well-established source of information to understand the role of maritime routes and ports in economic, social, and environmental contexts. Such a role cannot be adequately unraveled by looking at ports and routes in isolation; instead, they must be put in relation to one another. Therefore, recent studies on maritime traffic model the interplay between vessels and ports as a graph, which is often built using Automatic Identification System (AIS) data. However, only a few works explicitly study the evolution of such graphs and, when they do, generally consider coarse-grained time intervals.

Our goal is to fill this gap by providing a conceptual framework for the fine-grained systematic study of maritime graphs evolution. To this end, this paper presents the month-by-month analysis of world-wide graphs built using a 3-years AIS dataset. The main goal is to provide a systematic study of the evolutionary aspects of ports networks, with the purpose of identifying recurrent patterns in their evolution. The analysis focuses on the evolution of several topological graph features and their stationarity and statistical correlation. Results have revealed the difference in the graphs between long- and short-range vessels, with the former generating well-connected giant strongly connected components that are relatively stable over time, while the latter generating more variable and fragmented networks. Further, we revealed seasonal and trending patterns (e.g., the average distance of passenger vessels routes has a seasonal period of one year) that can provide insights in the worldwide maritime context and be used as building blocks toward the prediction of graphs topology.

DOI: 10.1007/s10707-021-00451-0

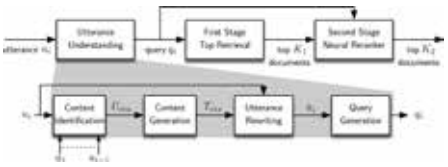


World wide cargo routes in 2019 extrapolated from the dataset used in the paper. The nodes represent ports and the edges are voyages between two ports.

Adaptive utterance rewriting for conversational search

I. Mele, C.I. Muntean, F.M. Nardini, R. Perego, N. Tonellotto, O. Frieder
Information Processing & Management, vol. 58. Elsevier, 2021.

Conversational Information Retrieval (IR) has recently gained interest due to the widespread popularity of conversational assistant systems. Thanks to the recent advances in automatic speech recognition and understanding, conversational assistant are widely used in chatbots and smart home devices, as well as in wearable devices and smartphones. In a conversational context, a user converses with a system through a sequence of natural-language questions, i.e., utterances. Starting from a given subject, the conversation evolves through sequences of user utterances and system replies. The retrieval of documents relevant to an utterance is difficult due to informal use of natural language in speech and the complexity of understanding the semantic context coming from previous utterances. Our objective is to prove that adding missing context chosen in a proper way



Conversation understanding architecture.

improves the retrieval task in conversational systems. We also prove that utterance classification can be used to find the best rewriting technique for transforming a raw utterance into an enriched and self-explanatory. We adopt the 2019 TREC Conversational Assistant Track (CAsT) framework to experiment with a modular architecture performing in order: (i) automatic utterance understanding and rewriting, (ii) first-stage retrieval of candidate passages for the rewritten utterances, and (iii) neural re-ranking of candidate passages. By understanding the conversational context, we propose adaptive utterance rewriting strategies based on the current utterance and the dialogue evolution of the user with the system. The classifier identifies those utterances lacking context information as well as the dependencies on the previous utterances. The best performance was achieved by a BERT-BERTctx binary classification cascade with a weighted F1 score of 0.76. Experimentally, we evaluate the proposed architecture in terms of traditional information retrieval metrics at small cutoffs. Results demonstrate the effectiveness of our techniques, achieving an improvement up to 0.6512 (+201%) for P@1 and 0.4484 (+214%) for nDCG@3 w.r.t. the CAsT baseline.

DOI: 10.1016/j.ipm.2021.102682

Fast filtering of search results sorted by attribute

F.M. Nardini, R. Trani, R. Venturini
ACM Transactions on Information Systems (TOIS). ACM,

Modern search services often provide multiple options to rank the search results, e.g., sort “by relevance”, “by price” or “by discount” in e commerce. While the traditional rank by relevance effectively places the relevant results in the top positions of the results list, the rank by attribute could place many marginally relevant results in the head of the results list leading to poor user experience. In the past, this issue has been addressed by investigating the relevance aware filtering problem, which asks to select the subset of results maximizing the relevance of the attribute sorted list. Recently, an exact algorithm has been proposed to solve this problem optimally. However, the high computational cost of the algorithm makes it impractical for the Web search scenario, which is characterized by huge lists of results and strict time constraints. For this reason, the problem is often solved using efficient yet inaccurate heuristic algorithms. In this paper, we first prove the performance bounds of the existing heuristics. We then propose two efficient and effective algorithms to solve the relevance aware filtering problem. First, we propose OPT Filtering, a novel exact algorithm that is faster than the existing state of the art optimal algorithm. Second, we propose an approximate and even more efficient algorithm, ϵ-Filtering, which, given an allowed approximation error ϵ, finds a $(1-\epsilon)$–optimal filtering, i.e., the relevance of its solution is at least $(1-\epsilon)$ times the optimum. We conduct a comprehensive evaluation of the two proposed algorithms against state of the art competitors on two real world public datasets. Experimental results show that OPT Filtering achieves a significant speedup of up to two orders of magnitude with respect to the existing optimal solution, while ϵ-Filtering further improves this result by trading effectiveness for efficiency. In particular, experiments show that ϵ-Filtering can achieve quasi optimal solutions while being faster than all state of the art competitors in most of the tested configurations.

DOI: 10.1145/3477982

Online communication and body language

P. Paradisi, M. Raglianti, L. Sebastiani
Frontiers in Behavioral Neuroscience, vol. 15. Frontiers, 2021.

The COVID-19 emergency brought out the role of online digital technologies. The increase in online social interactivity was accelerated by social distancing, which has been recognized to have adverse effects due to physical and emotional isolation (Canet-Juric et al., 2020). Body language is central to social interactions, and its role is clearly diminished when going online, but the relevance of this change is still not clear. This transition toward online could affect the wellness of the people, especially the population with specific fragilities, e.g., young people and seniors (Beam and Kim, 2020; Canet-Juric et al., 2020; Fernández Cruz et al., 2020). We here briefly present our viewpoint on some issues concerning changes in body interactions in online interpersonal communication. Our aim is to encourage constructive discussion and raise awareness about these very topical issues.

DOI: 10.3389/fnbeh.2021.709365

Patient perceptions and knowledge of ionizing radiation from medical imaging

L. Bastiani, F. Paolicchi, L. Faggioni, M. Martinelli, R. Gerasia, C. Martini, P. Cornacchione, M. Ceccarelli, D. Chiappino, D. Della Latta, K. Negri, D. Pertoldi, D. Negro, G. Nuzzi, V. Rizzo, P. Tamburrino, C. Pozzessere, G. Aringhieri, D. Caramella
JAMA Network Open. American Medical Association, 2021.

Importance Although imaging has become a standard tool of modern medicine, its widespread use has been paralleled by an increasing cumulative radiation dose to patients despite technological advancements and campaigns calling for better awareness and minimization of unnecessary exposures.

Objective To assess patients' knowledge about medical radiation and related risks.

Design, Setting, and Participants A survey study of hospitals in Italy was conducted; all patients in waiting rooms for medical imaging procedures before undergoing imaging examinations at 16 teaching and nonteaching hospitals were approached to take the survey. The survey was performed from June 1, 2019, to May 31, 2020.

Main Outcomes and Measures Survey respondents' basic knowledge of ionizing radiation levels and health risks, earlier imaging tests performed, and information and communication about radiation protection issues.

Results Among 3039 patients invited to participate, the response rate was 94.3% (n = 2866). Participants included 1531 women (53.4%); mean (SD) age was 44.9 (17.3) years. Of the 2866 participants, 1529 (53.3%) were aware of the existence of natural sources of ionizing radiation. Mammography (1101 [38.4%]) and magnetic resonance imaging (1231 [43.0%]) were categorized as radiation-based imaging modalities. More than half of the 2866 patients (1579 [55.1%]; P = .03) did not know that chest computed tomography delivers a larger dose of radiation than chest radiography, and only 1499 (52.3%) knew that radiation can be emitted after nuclear medicine examinations (P = .004). A total of 667 patients (23.3%) believed that radiation risks were unrelated to age, 1273 (44.4%) deemed their knowledge about radiation risks inadequate, and 2305 (80.4%) preferred to be informed about radiation risks by medical staff. A better knowledge of radiation issues was associated with receiving information from health care professionals (odds ratio [OR], 1.71; 95% CI, 1.43-2.03; P < .001) and having a higher educational level (intermediate vs low: OR, 1.48; 95% CI, 1.17-1.88; P < .001; high vs low: OR, 2.68; 95% CI, 2.09-3.43; P < .001).

Conclusions and Relevance The results of this survey suggest that patients undergoing medical imaging procedures have overall limited knowledge about medical radiation. Intervention to achieve better patient awareness of radiation risks related to medical exposures may be beneficial.

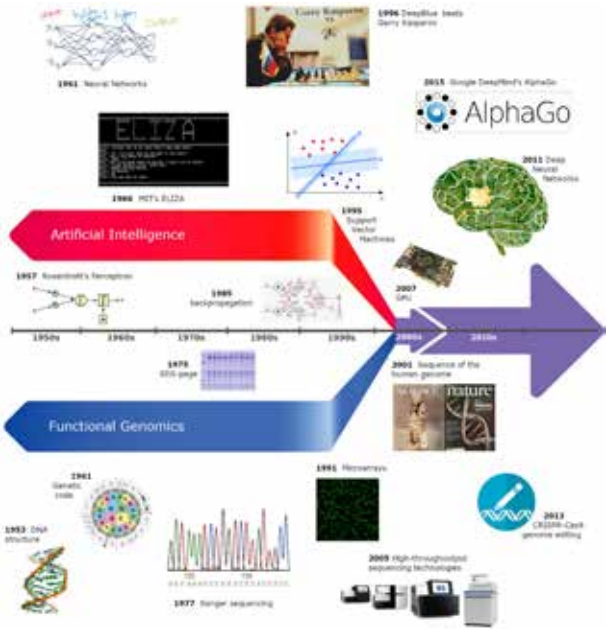
DOI: 10.1001/jamanetorkopen.2021.28561

AI applications in functional genomics

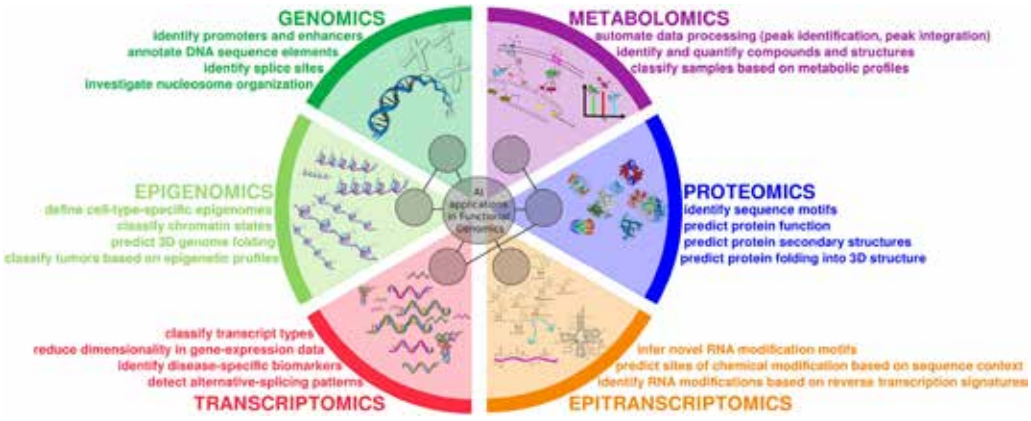
C. Caudai, A. Galizia, F. Geraci, L. Le Pera, V. Morea, E. Salerno, A. Via, T. Colombo
Computational and Structural Biotechnology Journal, vol. 19. Elsevier, 2021.

We review the current applications of artificial intelligence (AI) in functional genomics. The recent explosion of AI follows the remarkable achievements made possible by ''deep learning", along with a burst of ''big data" that can meet its hunger. Biology is about to overthrow astronomy as the paradigmatic representative of big data producer. This has been made possible by huge advancements in the field of high throughput technologies, applied to determine how the individual components of a biological system work together to accomplish different processes. The disciplines contributing to this bulk of data are collectively known as functional genomics. They consist in studies of: i) the information contained in the DNA (genomics); ii) the modifications that DNA can reversibly undergo (epigenomics); iii) the RNA transcripts originated by a genome (transcriptomics); iv) the ensemble of chemical modifications decorating different types of RNA transcripts (epitranscriptomics); v) the products of protein-coding transcripts (proteomics); and vi) the small molecules produced from cell metabolism (metabolomics) present in an organism or system at a given time, in physiological or pathological conditions. After reviewing main applications of AI in functional genomics, we discuss important accompanying issues, including ethical, legal and economic issues and the importance of explainability.

DOI: 10.1016/j.csbj.2021.10.009



A timeline of momentous events in functional genomics and artificial intelligence from their foundation until the time they crossed their paths.



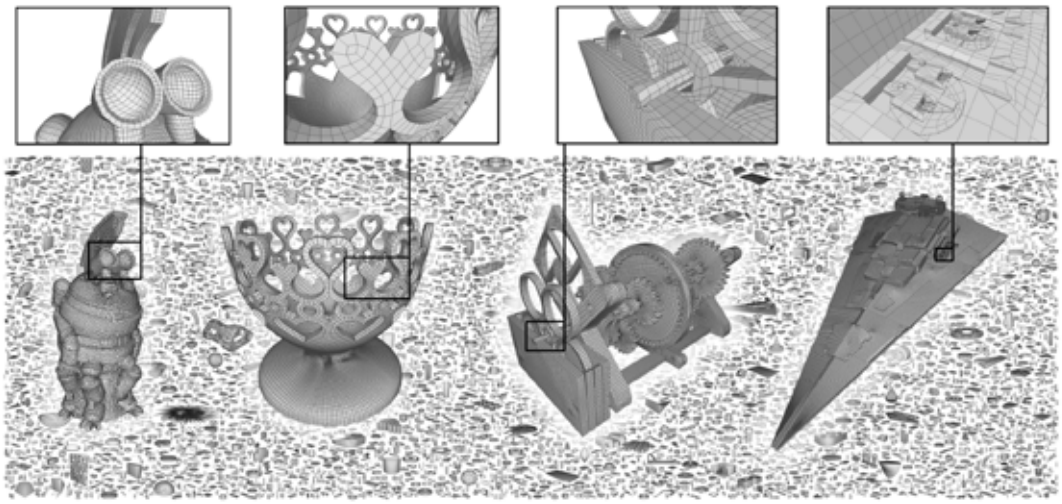
AI applications in functional genomics.

Reliable feature-line driven quad remeshing

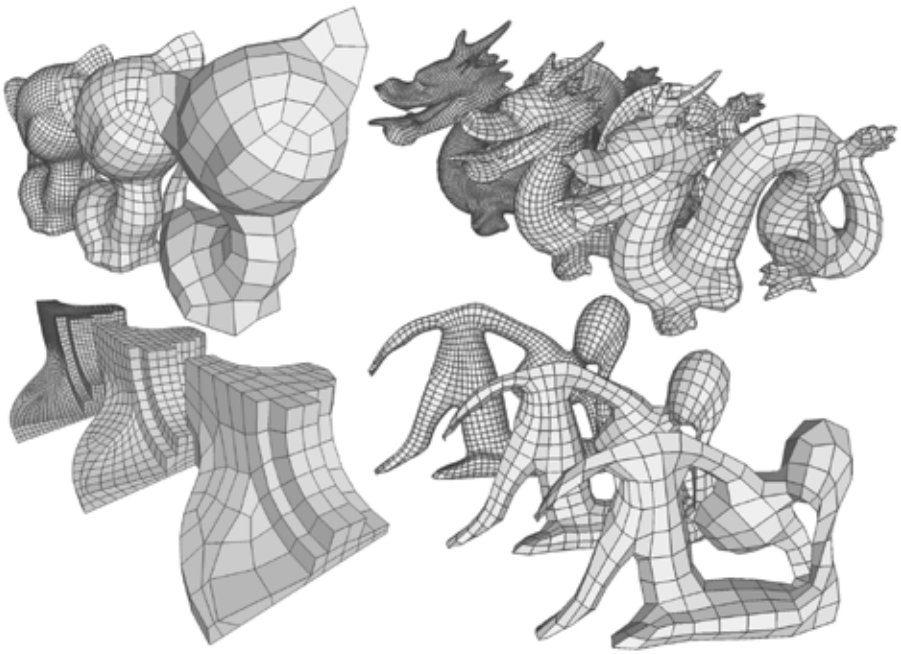
N. Pietroni, S. Nuvoli, T. Alderighi, P. Cignoni, M. Tarini
ACM Transactions on Graphics, vol. 40. ACM, 2021.

In this paper we present a new algorithm for the semi-regular quadrangulation of an input surface, driven by its line features, such as sharp creases. We define a perfectly feature-aligned cross-field and a coarse layout of polygonal-shaped patches where we strictly ensure that all the feature-lines are represented as patch boundaries. To be able to consistently do so, we allow non-quadrilateral patches and T-junctions in the layout; the key is the ability to constrain the layout so that it still admits a globally consistent, T-junction-free, and pure-quad internal tessellation of its patches. This requires the insertion of additional irregular-vertices inside patches, but the regularity of the final-mesh is safeguarded by optimizing for both their number and for their reciprocal alignment. In total, our method guarantees the reproduction of feature-lines by construction, while still producing good quality, isometric, pure-quad, conforming meshes, making it an ideal candidate for CAD models. Moreover, the method is fully automatic, requiring no user intervention, and remarkably reliable, requiring little assumptions on the input mesh, as we demonstrate by batch processing the entire Thingi10K repository, with less than 0.5% of the attempted cases failing to produce a usable mesh.

DOI: 10.1145/3450626.3459941



A mosaic of the processed Thingi10K dataset. The entire dataset is available at https://www.quadmesh.cloud.

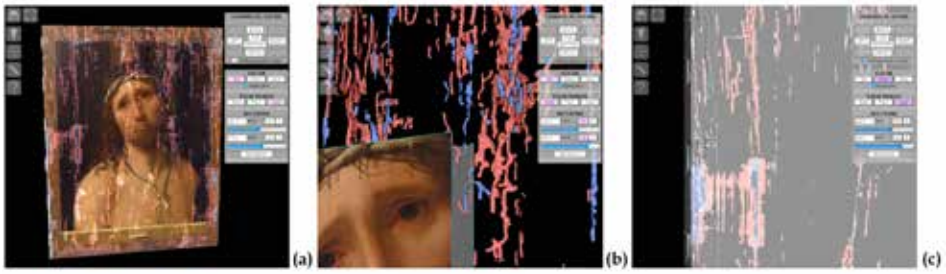


Examples of quadrangulations created at different resolution.

“Ecce Homo” by Antonello da Messina, from non-invasive investigations to data fusion and dissemination

F. Albertin, C. Ruberto, C. Cucci, M. Callieri, M. Potenziani, E. Siotto, P. Pingi, R. Scopigno, M. Bettuzzi, R. Brancaccio, M.P. Morigi, L. Castelli, F. Taccetti, M. Picollo, L. Stefani, F. de Vita
Scientific Reports, vol. 11. Springer, 2021.

In this paper we show how scientific investigations of artworks are nowadays crucial in terms of preservation since they are able to provide a measurable evaluation of the materials and their state of conservation. We report the case of Antonello da Messina’s painting “Ecce Homo”: its delicate state of conservation, with the need for constant monitoring, required a broad and in-depth diagnostic campaign to support the restorers. The project was carried out entirely in situ using non-invasive cutting-edge techniques and proposes a multimodal and data-centric approach, integrating 3D and 2D methodologies. The surface irregularities and the support were analysed with a structured-light 3D scanner and X-ray tomography. The painting materials were investigated with X-ray fluorescence scanning (MA-XRF) and reflectance hyperspectral imaging (HSI). Primarily, the data were jointly used for a scientific scope and provided new knowledge of the painting in terms of materials and painting techniques. In addition, two web-based interactive platforms were developed: one to provide restorers and experts with a new perspective of the hidden geometries of the painting, and the other targeted at the general public for dissemination purposes. The results of the Ecce Homo scientific analysis were exhibited, using a touch-screen interface, and developed for different user levels, from adults to kids.



The combined inside and out dataset. (a) All the tunnels and anomalies, shown in transparency below the surface. Empty and filled tunnels are shown in red and blue, respectively. (b) The networks of different sized tunnels. (c) Loss of cohesion between the wood fibres on the left of the painting.

DOI: 10.1038/s41598-021-95212-2

Volume decomposition for two-piece rigid casting

T. Alderighi, L. Malomo, B. Bickel, P. Cignoni, N. Pietroni
ACM Transactions on Graphics, vol. 40. ACM, 2021.

We introduce a novel technique to automatically decompose an input object’s volume into a set of parts that can be represented by two opposite height fields. Such decomposition enables the manufacturing of individual parts using two-piece reusable rigid molds. Our decomposition strategy relies on a new energy formulation that utilizes a precomputed signal on the mesh volume representing the accessibility for a predefined set of extraction directions. Thanks to this novel formulation, our method allows for efficient optimization of a fabrication-aware partitioning of volumes in a completely automatic way. We demonstrate the efficacy of our approach by generating valid volume partitionings for a wide range of complex objects and physically reproducing several of them.

DOI: 10.1145/3478513.3480555

Given a closed 2-manifold mesh, we split it automatically into double height field (DHF) components that can be fabricated using two-piece rigid casting and assembled afterwards.

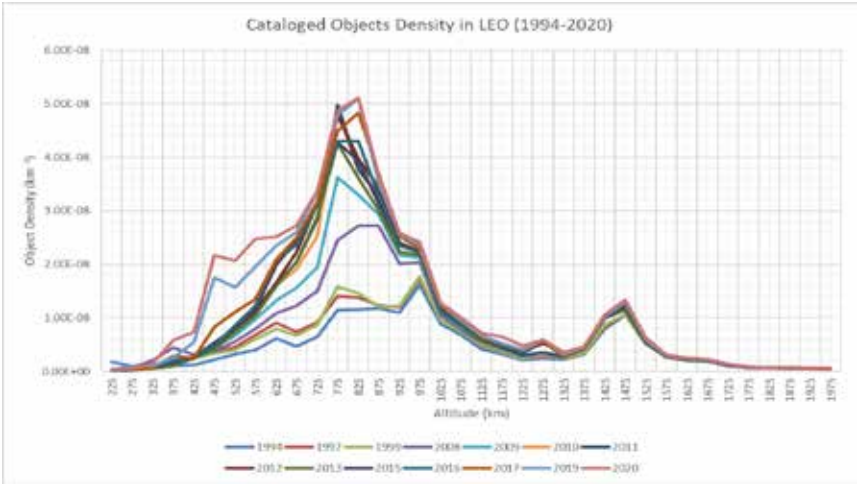
Evaluating the impact of space activities in low Earth orbit

C. Pardini, L. Anselmo

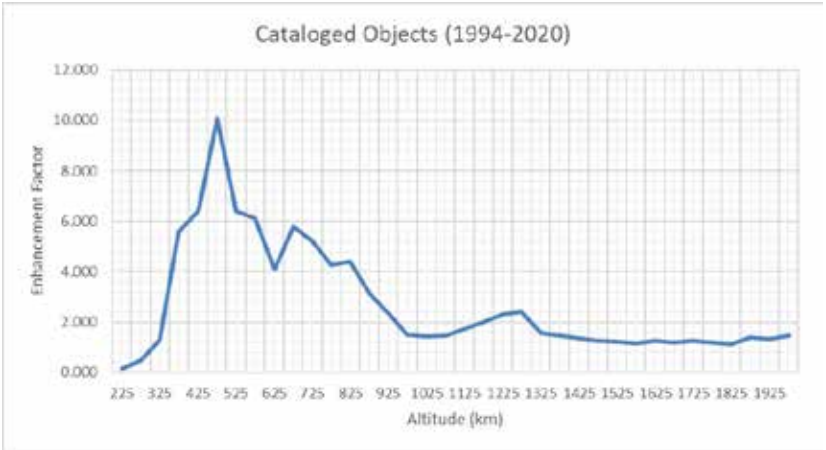
Acta Astronautica, vol. 184. Elsevier, 2021.

The evolution of cataloged orbital debris in low Earth orbit (LEO) over the last quarter of century was analyzed in detail, to gather insights on the development of space activities, on the effectiveness of the debris mitigation measures recommended in the meantime, and on the environmental impact of fragmentations, in particular collisions, both intentional and accidental. The main conclusion was that the observed evolution matched on the whole the predictions of the unmitigated business-as-usual scenarios simulated twenty years ago, and that the benefits caused by the progressive worldwide adoption of mitigation measures were unfortunately offset by a couple of catastrophic collisions and prolonged weak solar activity. Concerning the recorded growth of cataloged fragmentation debris, nowhere have the signs of an exponential increase been revealed so far. Nevertheless, the overall picture has worsened during the last quarter of a century and extreme care is required in planning and conducting new space activities from now on, especially in a phase of increased and ever more rapid exploitation. In order to assess the sustainability of space activities, especially over the next 10–30 years, several environmental criticality indexes have been introduced and discussed, estimating their current values in LEO, as well as their magnitudes associated with specific scenarios of debris growth. They could provide simple tools for evaluating the relative and absolute impact on the debris environment, either in LEO as a whole or in specific altitude shells, of new spacecraft deployments and operations, as in the case of mega-constellations of satellites. The main result of this preliminary analysis was that all indexes were consistent in indicating that from one third to one half of the LEO capacity to sustain long-term space activities – as they are currently conceived – has already been saturated. The 2020s, with their many planned launches, will therefore be crucial years for enforcing more effective debris mitigation and remediation measures.

DOI: 10.1016/j.actaastro.2021.03.030



Evolution of the density of cataloged objects in LEO, averaged over 50 km altitude bins, from 1994 to 2020.



The enhancement factor, plotted as a function of the altitude, shows how many times the cataloged objects in LEO multiplied from 1994 to 2020.

QuaPy: a Python-based framework for quantification

Alejandro Moreo, Andrea Esuli, Fabrizio Sebastiani, ISTI-CNR. CIKM '2 1 - 30th ACM International Conference on Information & Knowledge Management, 1-5 November 2021 (DOI: 10.1145/3459637.3482015).

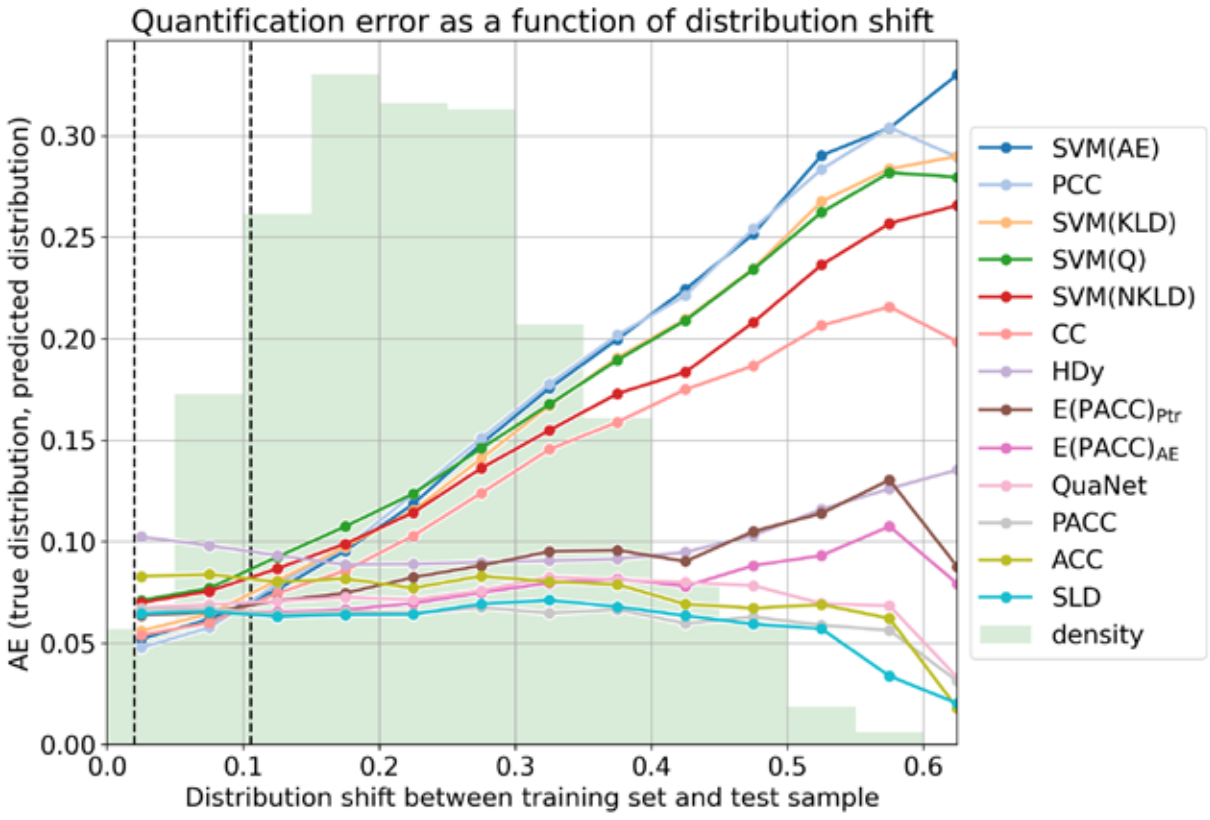
Quantification (variously called "learning to quantify", or "supervised prevalence estimation", or "class prior estimation") is the task of training models ("quantifiers") that estimate the relative frequencies (a.k.a. prevalence values) of the classes of interest in a sample of unlabelled data items. For example, a quantifier might be tasked to estimate the percentage of tweets, from a sample of 100,000 tweets about Donald Trump, which display a Positive, Neutral, or Negative stance towards Trump. Quantification is of special interest in fields inherently concerned with aggregate data, such as the social sciences, epidemiology, market research, and ecological modelling.

Quantification can trivially be solved via classification, i.e., by training a classifier, applying it to the unlabelled data items, and counting how many data items have been assigned to each class. However, this approach (called "classify and count") is known to deliver strongly biased estimates of class prevalence values, and especially so in presence of distribution shift, i.e., when the class prevalence values in the unlabelled set may substantially differ from those observed during training. For this reason, quantification is no longer considered a by-product of classification, but a task in its own right. As such, quantification has its own learning methods, model selection protocols, evaluation measures, and evaluation protocols.

QuaPy is the first dedicated framework written in Python that supports research, development, and experimentation in quantification. QuaPy provides implementations of many baseline and advanced quantification methods, of quantification-oriented routines of model selection, of specific evaluation measures, and of robust evaluation protocols. QuaPy also makes available datasets commonly used in the field, and offers visualization tools for facilitating the analysis and interpretation of the results.

The software is open-source and publicly available under a BSD-3 licence via GitHub, and can be easily installed via pip.

Contact: Alejandro Moreo, AIMH Lab
alejandro.moreo@isti.cnr.it
https://github.com/HLT-ISTI/QuaPy



SoBigData for the social good

A research infrastructure where Artificial Intelligence enables individuals and society

SoBigData is a European Research Infrastructure for Social Mining and Big Data Analytics.

The SoBigData reseqrch infrastructure (RI) is an ecosystem of human and digital resources comprising data scientists, analytics, and processes. It is designed to enable multidisciplinary scientists and innovators to carry out experiments and make them reusable by the community.

SoBigData is based on the most advanced analysis techniques of Artificial Intelligence and Big Data. It is the only RI that combines issues related to AI and Society, a highly strategic area in the current European and world economy. For this reason, SobigData has been selected by ESFRI, the European Strategy Forum on Research Infrastructures, to be part of its Roadmap 2021.

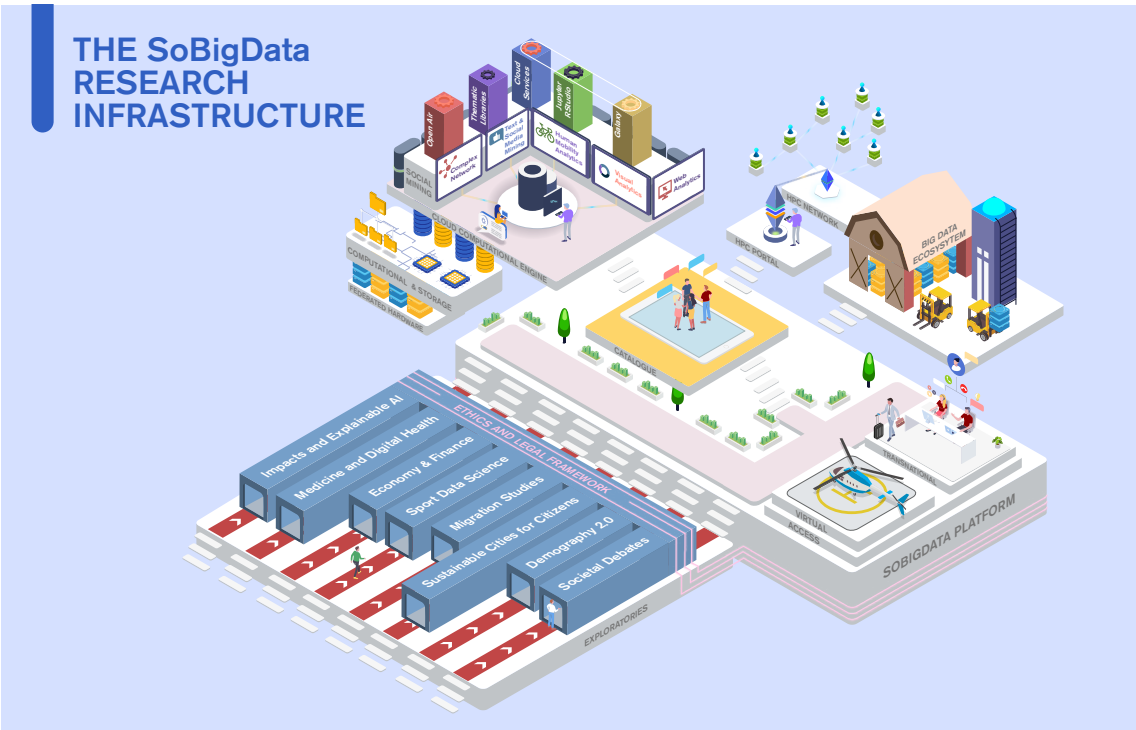
SoBigData promotes open science. All the RI activities are based on the principles of FAIR (Findable, Accessible, Interoperable, Reusable) and FACT (Fair, Accurate, Confidential and Transparent) and linked to the ethical use of data and research results in compliance with concepts such as privacy and non-disclosure of personal information.

SoBigdata offers free innovative services, such as the SoBigData Lab that integrated JupyterHub to become a place where users can develop algorithms, written in any programming language, or create new algorithms in an interactive Jupyter notebook equipped with all the specialized libraries already integrated and ready to use. Users can also deploy methods, written in almost any programming language, in the SoBigdata method engine to use the full power of the cloud computing provided by SoBigData or to execute experiments on the cloud computing center.

All actions in SoBigData are intended to positively impact society and increase the concept of conscience and well-being. This commitment is also evidenced by the link between the research infrastructure and the 17 Sustainable Development Goals (SDGs) promoted by the United Nations.

Contact: Roberto Trasarti, KDD Lab
roberto.trasarti@isti.cnr.it
www.sobigdata.eu

This infrastructure is part of a project that has received funding from the **European Union's Horizon 2020 research and innovation programme under grant agreement N°871042**



TAPAS – Tensor Algorithms for Performability Analysis of large Systems

A tool for the stochastic evaluation of large interdependent composed models with absorbing states in the context of dependability assessment

Stochastic state-space model-based approaches are widely used to perform dependability, performance or performability analysis of complex systems. Generally, such systems are composed of many components interacting with each other in an intricate manner. This makes model-based analysis very challenging, both with respect to the definition of the overall model, which must represent all the individual components and their specific interdependencies, and the state-space explosion.

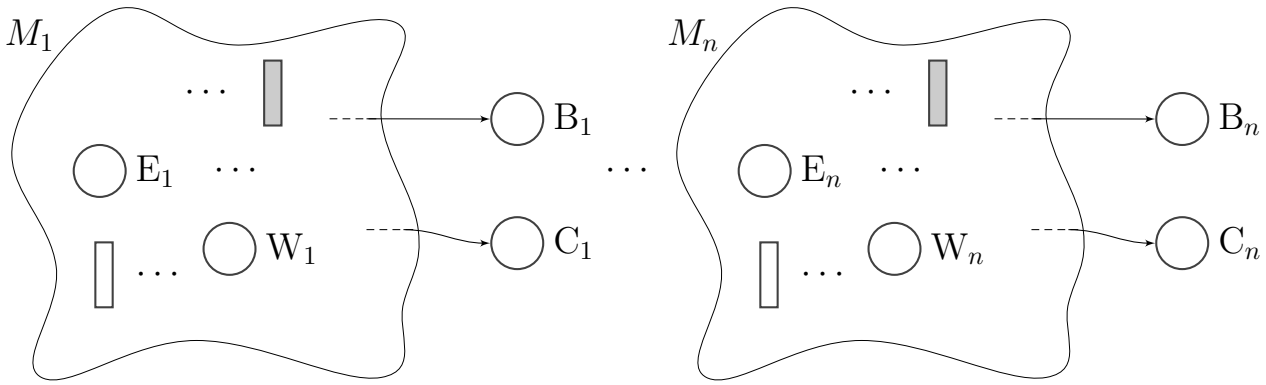
In this context, TAPAS provides solutions to large homogeneous Continuous Time Markov Chains (CTMC), described by Stochastic Automata Networks (SAN) models, structured in submodels with absorbing states. A high level representation of a SAN is shown in the figure, where n submodels, each being a CTMC with absorbing states Bi and Ci, evolve through local actions (white rectangles) and synchronization actions (gray rectangles).

Submodels with multiple absorbing states are needed to distinctly represent components having multiple failure causes (e.g. the benign and catastrophic failures modelled through states Bi and Ci in the figure), or fault tolerant systems suffering failure due to the exhaustion of redundancy.

TAPAS advances current tools for conducting similar analyses, based on implicit representation and manipulation of the state space, by extending the implicit representation to the solution vector in addition to the descriptor matrix. To the best of authors' knowledge, TAPAS is the first tool to tackle complete implicit model representation. Moreover, it addresses the analysis of limiting properties of CTMC with absorbing states, another aspect that distinguishes TAPAS with respect to existing alternatives.

TAPAS is an experimental and academic open-source tool released under BSD license. The tool and related documentation are available at the GitHub page https://github.com/numpi/tapas.

Contact: Giulio Masetti, Silvano Chiaradonna, SEDC Lab
giulio.masetti@isti.cnr.it;
silvano.chiaradonna@isti.cnr.it



Monitoring environment for system of systems

Support SoS monitoring through ontology derived rules

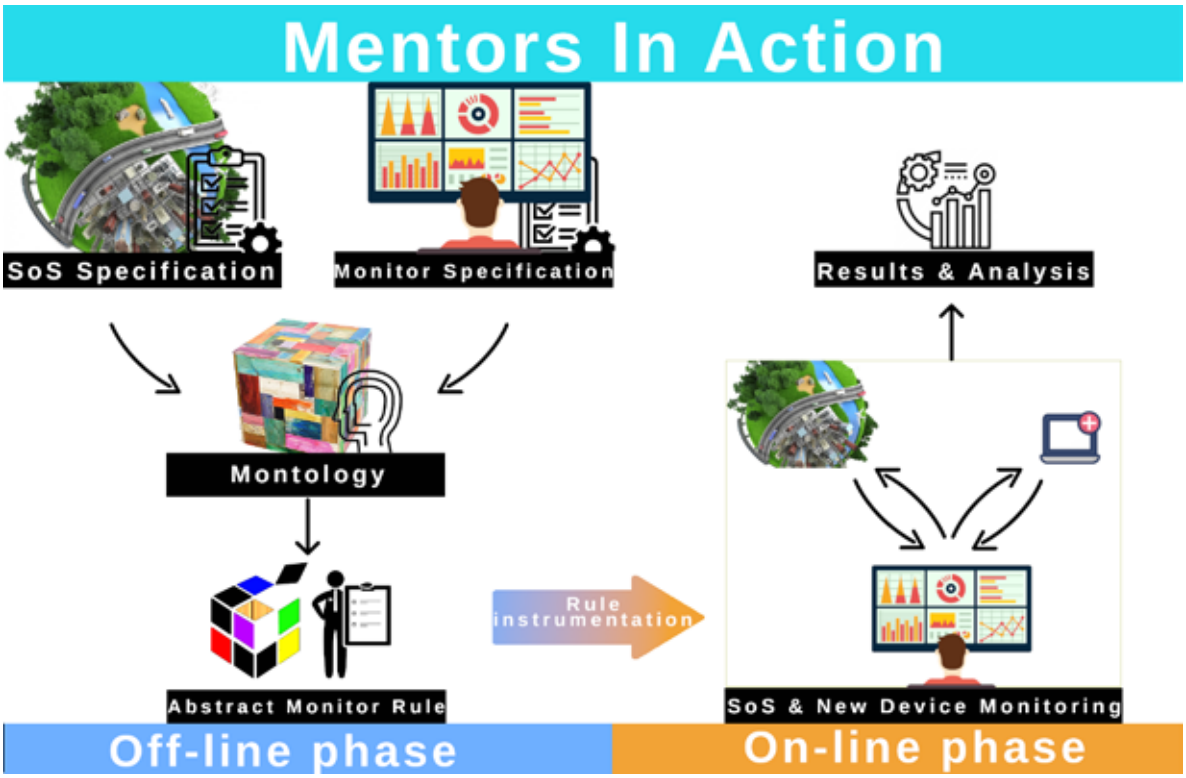
Nowadays, the development of System of Systems (SoSs) requires the integration and collaboration of different ICT components or devices typically developed by several third parties. On the one hand, this process ensures high productivity and competitiveness; on the other, it exposes the SoS to crucial vulnerabilities and risks. A typical consequence is that each of the integrated components could be affected by, and involuntarily propagate, the vulnerabilities of the others included in the SoS, with potential drastically consequences. In this situation, when a new device or component is integrated into an existing SoS, facilities able to efficaciously assess and prevent anomalous and dangerous situations are necessary. One method commonly adopted is the use of a monitoring system. A monitoring engine collects events from different levels and the different system components, inferring more complex patterns. The derived patterns represent the observed normal (or abnormal) behavior of the monitored system.

In this context, we propose MENTORS (Monitoring ENviromenT fOR Sos): a toolchain including a dedicated monitoring engine for SoS. MENTORS is based on semantic web technologies to formally represent the SoS, and monitoring knowledge through a core ontology called MONTOLOGY (MONitoring onTOLOGY)

MENTORS is composed of two phases: the off-line in which MONTOLOGY is defined and used for the specification of monitoring rules, and the on-line where the SoS and the new device are monitored, vulnerabilities are detected and countermeasures applied.

MENTORS is supported by a reference architecture that allows the continuous and automatic enhancement of MONTOLOGY through an inference system. MENTORS is currently under development and validation within the H2020 BIECO Project, agreement No 952702.

Contacts: Antonello Calabrò, Said Daoudagh and Eda Marchetti
SEDC Lab
antonello.calabro@isti.cnr.it
said.daoudagh@isti.cnr.it
eda.marchetti@isti.cnr.it



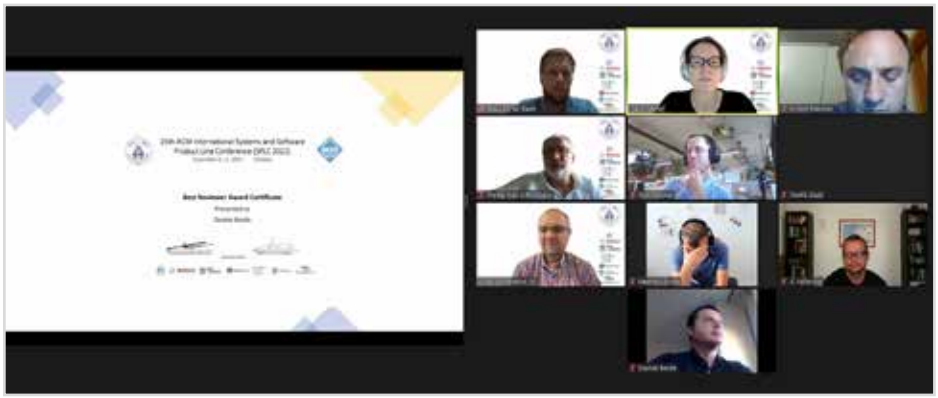
Best Short Paper award at the 11th Italian Information Retrieval Workshop (IIR 2021)

"Garbled-Word Embeddings for Jumbled Text" by Gianluca Sperduti, Alejandro Moreo, and Fabrizio Sebastiani

Aoccdrnig to a reasrech at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a word are, the olny itmopnrat tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a total mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe". We investigate the extent to which this phenomenon applies to computers as well. Our hypothesis is that computers are able to learn distributed word representations that are resilient to character reshuffling, without incurring a significant loss in performance in tasks that use these representations. If our hypothesis is confirmed, this may form the basis for a new and more efficient way of encoding character-based representations of text in deep learning, and one that may prove especially robust to misspellings, or to corruption of text due to OCR. This paper discusses some fundamental psycho-linguistic aspects that lie at the basis of the phenomenon we investigate, and reports on a preliminary proof of concept of the above idea. The paper "Garbled Word Embeddings for Jumbled Text" by Gianluca Sperduti, Alejandro Moreo, and Fabrizio Sebastiani, has received the Best Short Paper Award at the 11th Italian Information Retrieval Workshop (IIR 2021)

Best Reviewer award SPLC2021

Best reviewer award for the 25th ACM International Systems and Software Product Line Conference



Davide Basile, Formal Methods and Tools lab, has served as member of the Program Committee of the 25th ACM International Systems and Software Product Line Conference and has received the Best Reviewer Award. The Systems and Software Product Line Conference (SPLC) is a premier forum where researchers, practitioners, and educators present and discuss the most recent ideas, trends, experiences, and challenges in the area of software and system product lines engineering.

From the official Twitter account of the conference it is possible to see some pictures of the award ceremony:

https://twitter.com/splcconf/status/1435643315794464768.

Contact: Davide Basile, FMT Lab
davide.basile@isti.cnr.it

Facebook Research award 2021

The team of Antonia Bertolino, Breno Miranda, Emilio Cruciani and Roberto Verdecchia wins again!

Facebook is currently working on building a "digital twin" that reproduces the interaction dynamics of the social network. Using this virtual simulation, it should be possible to study the behaviour of the network and identify possible integrity concerns without violating the privacy and security of users. In May 2021, the company launched a Call for Proposals asking for the help of academic researchers in overcoming the many challenges raised by integrity and privacy issues in the area of high-scale agent-based user interaction simulation.

Five projects were selected out of the 29 proposals submitted by 24 academic institutions from 11 countries. One of these five was the project "Testing non-testable programs using association rules" submitted by Antonia Bertolino (SEDC Lab, ISTI-CNR) together with Breno Miranda (Federal University of Pernambuco, Brazil), Emilio Cruciani (University of Salzburg, Austria), and Roberto Verdecchia (Software and Sustainability Group, Vrije Universiteit Amsterdam, NL). This project addresses the problem of how to assess the correctness of simulated system behaviour, when there is no mechanism to determine whether an observation is acceptable or is instead the symptom of a potential problem. The key intuition of the project is to leverage association rules, hidden relationships between data that can be mined using specific learning algorithms. Anomalous or suspicious behaviour can thus be recognized if it departs from relevant extracted rules.

This team is not new to the FB podium: the same group already won a Facebook Research Award in 2019, competing against over 100 proposals, with a project that aimed to identify so-called flaky (or non-deterministic) test cases.

Antonia Bertolino is also among the proponents of another project submitted to the 2021 call, together with Patrizio Pelliccione and Michele Flammini, both from the Gran Sasso Science Institute. This proposal, entitled "SOCIETY: Social testing by gamified behaviour trees", will not receive funding,

but was included among 9 finalists judged as potentially impactful.

For more information, see: https://research.fb.com/blog/2021/09/announcing-the-winners-of-the-2021-rfp-on-agent-based-user-interaction-simulation-to-find-and-fix-integrity-and-privacy-issues.

Contact: Antonia Bertolino, SEDC Lab
antonia.bertolino@isti.cnr.it

IEEE VIS Test of Time award

Marco Tarini from the university of Milan, Paolo Cignoni and Claudio Montani from ISTI-CNR have won the IEEE VIS Test of Time Award for their paper: "Ambient Occlusion and Edge Cueing for Enhancing Real Time Molecular Visualization", IEEE Transactions on Visualization and Computer Graphics, Volume 12, Issue 5, September 2006.

The IEEE VIS Test of Time Award is a prestigious accolade given to recognize articles published 15 years earlier, whose contents are still vibrant and useful today and have had a major impact and influence within and beyond the visualization community. From the award text: "this highly cited paper was awarded to recognize its continued prominence both within the scientific visualization community as well as outside it." The committee recognized a work that is "not tied to a level of technology or research domain". The paper considers the problems of illustrating three dimensional structure in real time and considers the problem at differing levels of complexity, scale, and detail. "It is an elegant treatment of a difficult problem and a model of effective presentation of lasting results in visualization research."

The decisions by the award committee are based on objective measures such as the numbers of citations, and more subjective ones such as the quality and longevity and influence of ideas, outreach, uptake and effect not only in the research community, but also within application domains and visualization practice.

The awarding ceremony was held during the opening session of the prestigious IEEE 2021 Visualization Conference on Tuesday 26th October 2021.

http://ieeevis.org/year/2021/info/awards/test-of-time-awards

Young Researcher award "Matteo Dellepiane" 2021

The ISTI Young Researcher Award (YRA) "Matteo Dellepiane" is an annual award that honors its staff of less than 35 years old for a distinct contribute to the Institute activity with their scientific production.

There are two categories:

- Young Researcher - Beginner - awarded to researchers less than 32 years old;
- Young Researcher - Advanced - awarded to Ph.D. students and Ph.D. researchers less than 35 years old.

Contact: Franco Maria Nardini, HPC Lab
francomaria.nardini@isti.cnr.it

Beginner

Salvatore Citraro (ex-aequo)

Nicola Messina (ex-aequo)

Giulio Ermanno Pibiri

Voukelatou Vasiliki

Advanced

Antonino Crivello

Riccardo Guidotti

Cristina Muntean

1961: The year that changed Italian computing

Sixty years ago, the first Italian scientific computer

On 13 November, 1961, the Calcolatrice Elettronica Pisana (Pisan Electronic Computer - known as CEP) was inaugurated. To celebrate the 60th anniversary of this momentous day, the University of Pisa, with the contribution of three institutes of the National Research Council (IIT-CNR, ISTI-CNR and IAC-CNR) organised a day dedicated to "Pioneers of Information Technology", in which early initiatives in digital computing in Pisa and elsewhere in Italy were illustrated.

Many of the presentations recalled the success of Olivetti's Elea series of computers, some of which were designed in Pisa in the Barbaricina laboratories, and also remembered the unfortunate vicissitudes of the company after the deaths of Mario Tchou and Adriano Olivetti. Other talks dwelt on the completion of hardware, software and computing languages for the CEP after 1961; the evolution of information technology in Pisa and Italy with the crucial contribution of the CSCE (Pisa University Computing Centre), which in 1962 became a CNR Research Center in collaboration with the University of Pisa, to be later reorganized as the Institute for Information Technology (IEI-CNR); the creation of CNUCE (CNR research institute and university computing centre); the establishment of AICA (Italian Computing Association); the launching of a Degree in Information Science in Pisa, the first in Italy. These important milestones were presented in detail in the talk by Domenico Laforenza.

During the day, clips of interviews with several pioneers were screened. In particular, Luciano Azzarelli and Graziano Bertini touched on a number of ground-breaking early activities including the development of the system for Magnetic Tapes Control (1963) and the complex maintenance of the CEP, guaranteed during the sixties by the technicians who had collaborated in its construction.

After 1962, the expertise acquired by the researchers and technicians during the implementation of the CEP led to considerable research activity both in the field of information processing and that of digital electronic design. Valuable collaborations were initiated with Italian universities and a number of departments at the Santa Chiara Hospital in Pisa. Interfaces were created for multichannel acquisition, original analog-to-digital converters with "discrete" components, and signal pre-treatment circuits in the sectors of physics and biomedicine. The data collected and recorded were then processed using the CEP. More details of the day are in the Proceedings, currently in preparation. (Contact: fabio.gadducci@unipi.it)

Early Italian Scientific Computing Systems

As part of the day dedicated to "Pioneers of Information Technology" organized by the University of Pisa in November 2021, Graziano Bertini described some of the most important early systems in whose construction he had participated:

- - a system for the acquisition and analysis of information detected by micro-electrodes from retinal nerve fibers, developed in collaboration with the Institute of Physiology of the University of Pisa (Figure 1, 1962 -63).
- - the creation of a self-synchronizing cardiac pacemaker (February-June 1966) to resolve conflicts between the physiological stimulus of patients and the artificial stimulus of fixed-rate pacemakers in response to a request from the Institute of Physiology (see Figuree 2, ISTI News 2, 2017).
- - the special-purpose computer TO-PI (Turin-Pisa) for the analysis and synthesis of complex Boolean functions, installed at the Istituto Nazionale Galileo Ferraris of Turin (part of the rack is shown in Figure 3).

This computer had the first integrated circuits and involved the commitment of the entire CSCE HW group in 1966-68.

Further details on these systems can be found in publications at http://csce.isti.cnr.it/ and http://pumax.isti.cnr.it/.

Contact: Graziano Bertini, SI Lab
graziano.bertini@isti.cnr.it

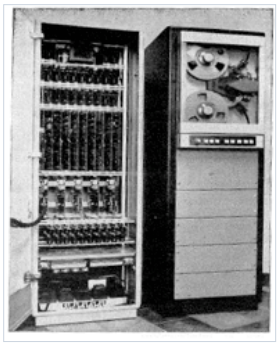


Fig. 1

Fig. 2

Fig. 3

Ph.D. dissertations

The GDPR compliance through access control systems

Author: Said Daoudagh, Department of Computer Science, University of Pisa, Italy
Supervisors: Eda Marchetti, ISTI-CNR, Anna Monreale (University of Pisa)

The GDPR is changing how Personal Data should be processed. It states, in Art.5.1(f), that "[data] should be processed in a manner that ensures appropriate security of the personal data [. . .], using appropriate technical or organizational measures (integrity and confidentiality)". We identify in the Access Control (AC) systems such a measure.

Indeed, AC is the mechanism used to restrict access to data or systems according to Access Control Policies (ACPs), i.e., a set of rules that specify who has access to which resources and under which circumstances.

In our view the ACPs, when suitably enriched with attributes, elements and rules extracted from the GDPR provisions, can suitably specify the regulations and the AC systems can assure a by-design lawfully compliance with the privacy preserving rules. Vulnerabilities, threats, inaccuracies and misinterpretations that occur during the process of ACPs specification and AC systems implementation may have serious consequences for the security of personal data (security perspective) and for the lawfulness of the data processing (legal perspective). For mitigating these risks, this thesis provides a systematic process for automatically deriving, testing and enforcing ACPs and AC systems in compliance with the GDPR. Its data protection by-design solution promotes the adoption of AC systems ruled by policies systematically designed for expressing the GDPR's provisions.

Specifically, the main contributions of this thesis are: (1) the definition of an Access Control Development Life Cycle for analyzing, designing, implementing and testing AC mechanisms (systems and policies) able to guarantee the compliance with the GDPR; (2) the realization of a reference architecture allowing the automatic application of the proposed Life Cycle; and (3) the use of the thesis proposal within five application examples highlighting the flexibility and feasibility of the proposal.

Enhancing the computational representation of narrative and its extraction from text

Author: Daniele Metilli, University of Pisa
Supervisors: Carlo Meghini, ISTI-CNR, Maria Simi, University of Pisa

Narratives are a fundamental part of human life. Every human being encounters countless stories during their life, and these stories contribute to form a common understanding of reality. This is reflected in the current digital landscape, and especially on the Web, where narratives are published and shared everyday. However, the current digital representation of narratives is limited by the fact that each narrative is generally expressed as natural language text or other media, in an unstructured way that is neither standardized nor machine-readable. These limitations hinder the manageability of narratives by automated systems. One way to solve this problem would be to create an ontology of narrative, i.e., a formal model of what a narrative is, then develop semi-automated methods to extract narratives from natural language text, and use the extracted data to populate the ontology. However, the feasibility of this approach remains an open question. This thesis attempts to investigate this research question, starting from the state of the art in the fields of Computational Narratology, Semantic Web, and Natural Language Processing. Based on this analysis, we have identified a set of requirements, and we have developed a methodology for our research work. Then, we have developed an informal conceptualization of narrative, and we have expressed it in a formal way using First-Order Logic. The result of this work is the Narrative Ontology (NOnt), a formal model of narrative that also includes a representation of its textual structure and textual semantics. To ensure interoperability, the ontology is based on the CIDOC CRM and FRBRoo standards, and it has been expressed using the OWL and SWRL languages of the Semantic Web. Based on the ontology, we have developed NarraNext, a semi-automatic tool that is able to extract the main elements of narrative from natural language text. The tool allows the user to create a complete narrative based on a text, using the extracted knowledge to populate the ontology. NarraNext is based on recent advancements in the Natural Language Processing field, including deep neural networks, and is integrated with the Wikidata knowledge base. The validation of our work is being carried out in three different scenarios: (i) a case study on biographies of historical figures found in Wikipedia; (ii) the Mingei project, which applies NOnt to the representation and preservation of Heritage Crafts; (iii) the Hypermedia Dante Network project, where NOnt has been integrated with a citation ontology to represent the content of Dante's Comedy. All three applications have served to validate the representational adequacy of the NOnt and the satisfaction of the requirements we defined. The case study on biographies has also evaluated the effectiveness of the NarraNext tool.

Conferences - Co-organized by ISTI

Third workshop on Human-Centered Computational Sensing (HCCS'22)
in conjunction with PerCom 2022)
March 21-25, 2022 in Pisa, Italy
https://sites.google.com/view/hccs2022/

20th International Conference on Pervasive Computing and Communications (PerCom 2022)
Demo Papers Session
March 21-25, 2022 in Pisa, Italy
https://www.percom.org/call-for-demo-papers/

ETAPS
EUROPEAN JOINT CONFERENCES ON
THEORY & PRACTICE OF SOFTWARE

3rd Workshop on Quantitative Aspects of Variant-rich Systems (QAVS 2022) at ETAPS 2022, Munich, Germany, 3 April, 2022
https://qavs.edgecloud.de/

17th International Conference on integrated Formal Methods (iFM 2022),
Lugano, Switzerland, 7-10 June 2022
https://ifm22.si.usi.ch/

24th IFIP WG 6.1 International Conference on Coordination Models and Languages (COORDINATION 2022), Lucca, Italy, 13-17 June 2022
https://www.discotec.org/2022/coordination

1st CLEF Lab on Learning to Quantify (LeQua 2022)
Bologna, Italy - September 5-8, 2022
https://lequa2022.github.io/

13th Conference and Labs of the Evaluation Forum (CLEF 2022)
Bologna, Italy - September 5-9, 2022
https://clef2022.clef-initiative.eu/

ISTI News is published by the Institute of Information Science and Technologies “A. Faedo”
National Research Council
Via G. Moruzzi 1—56124 Pisa Italy
http://www.isti.cnr.it

ISSN: 2724-6566
DOI: 10.32079/ISTINews

Privacy statement
The personal data (names, email addresses...) and the other information entered in ISTI News will be treated according with the provision set out in Legislative Degree 196/2003 (known as Privacy Code) and subsequently integration and amendement

Legal representative of the Institute of Information Science and Technologies “A. Faedo”: Dott. Roberto Scopigno

ISTI News is not for sale but is distributed for purposes of study and research and published online at
https://www.isti.cnr.it/en/isti-news

To subscribe: send your email to istinews@isti.cnr.it