

ISTI NEWS

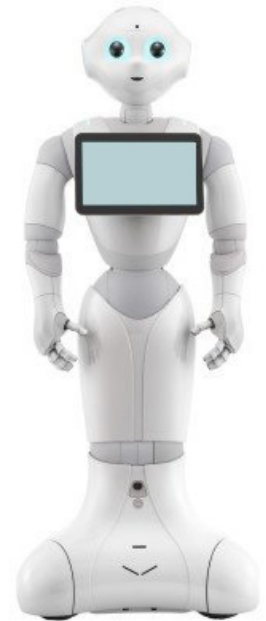
Industry 4.0

Research opportunities and challenges for ISTI

It is generally claimed that the third industrial revolution began in the seventies with the introduction of robots and computers for advanced automation. It now appears that we are witnessing the beginning of a fourth revolution, which is changing the design, operation and maintenance of production plants and systems. This is being called Industry 4.0 and it implies a significant change in design and implementation

perspective: the focus is moving from the search of specific solutions for the automation of individual machines or modules in a production system to a holistic approach, in which machines, production lines, and often the entire value-chain, are fully interconnected and integrated, both in the physical world and in their digital counterpart.

[continues on page 4]



ASTRail

Co-funded by Horizon 2020 under the Shif2Rail program

Compared with other transport sectors, the railway sector is notoriously cautious about adopting technological innovations. This is commonly attributed to the sector's robust safety requirements. An example is smart route planning: while GNSS-based positioning systems have been in use for quite some time now in the avionics and automotive sectors to provide accurate positioning and smart route planning, the current train separation system is still based on fixed blocks - a block being the section of the track between two fixed points.

[continues on page 6]

NESTORE

Co-funded by Horizon 2020

The ageing population is increasing rapidly in the EU; ICT can provide solutions for Active Ageing. However, the success of novel ICT solutions depends on the user perception with respect to their efficacy to support health promotion and global wellness. In this context, NESTORE will develop an innovative, multi-dimensional, personalized coaching system to support healthy ageing by: generating and sustaining motivation to take care of health; suggesting healthy nutrition and personalized physical and mental coaching, as well as social interaction, to prevent decline and preserve wellbeing.

[continues on page 7]

Editorial

Before I go...

May 15th 2018 is an important date for me. It marks the day that my position as Director of the Institute of Information Science and Technologies "A. Faedo" formally ended. Of course, I will continue to serve as acting director of ISTI until the selection procedure for the new director is completed but I feel that now is the right moment to briefly sum up what I feel are the most important achievements of the Institute, which it has been my privilege to direct for the last twelve years.

[continues on page 3]

Contents

Cover Story

Industry 4.0 4

New Projects

ASTRail 6

NESTORE. 7

Data4Impact. 8

E-RIHS 9

HAMLET 10

Selected Papers

Environmental monitoring for smart cities 11

Exploring nominal cellular automata 12

Natural language requirements processing: a 4D vision . . . 12

A tour of secure software engineering solutions for connected vehicles 13

Challenges in certification of autonomous driving systems . 13

Customizable automatic detection of bad usability smells in mobile accessed web applications 14

A web framework for cross-device gestures between personal devices and public displays 14

Clustering individual transactional data for masses of users . 15

Tiles: an online algorithm for community discovery in dynamic social networks 15

On the effects of low-quality training data on information extraction from clinical reports 16

Preservation of database concepts: from data mining to concept theory 16

Energy-efficient query processing in web search engines . . . 17

ANALYTIC: an active learning system for trajectory classification. 17

Bayesian Volterra system identification using reversible jump MCMC algorithm 18

An intelligent cooperative visual sensor network for urban mobility 18

Position-based tensegrity design 19

From paper to web: automatic generation of a web-accessible 3D repository of pottery types 19

On the secular decay of the LARES semi-major axis. 20

On the derivative of the stress–strain relation in a no-tension material 20

Software and Tools

Hexalab: visualizing complex hexahedral meshes on the web 21

Awards and Achievements

Young open science award 22

Makers of merit at the European edition of the maker faire 22

ProgettISTI 2017 23

Grants for Young Mobility 23

Miscellaneous

The giant squid across myth and science 24

Summer school on natural space risks 25

Ph.D. dissertations 26

Welcome aboard! 29

Conferences - Co-organized by ISTI 30

Before I go

Looking Back – with satisfaction; Looking forward – with optimism

[Continued]



Twelve years represent a decidedly long period in Computer Science and, more generally, in Information and Communications Technologies (ICT); sectors in rapid and constant evolution. So much has changed under the pressure of distributed processing, increased computational power, the ever-increasing size of memories, the enormous amount of data made available by the myriad of sensors scattered around us. However, I'm proud to say that, despite the technological revolution in act, over this period the Institute has succeeded in not just maintaining but, in many ways, increasing its already high level of success both nationally and internationally.

Among our strengths, I can list:

(a) A wide spectrum of skills, critical mass and dynamism. The internal organization currently includes 12 research laboratories that can count on approximately 130 researchers, and about 50 units of staff in training and 50 technicians and administrative staff;

(b) A high international visibility. Over 15 researchers and technicians have an h-index of more than 30, many researchers are members of program committees of important international conferences or editorial boards of prestigious journals in the sector;

(c) A strong capacity to acquire research projects. On average, there are 90 ac-

tive projects per year (for a total equivalent value of approximately 20 MEuro); almost 30 of them are international projects (a countervalue of about 12 MEuro and an acceptance rate of over 28%). While the whole of CNR has participated in 36 FET projects in the FP7 program (a 6% acceptance rate program); ISTI has participated in no less than 9 FET projects;

(d) A strong propensity for synergies between groups or research laboratories demonstrated by the numerous (mostly national) projects in which researchers with different backgrounds and skills cooperate;

(e) A large number of young researchers in training (45 with a non permanent contracts, 48 research fellows, grant holders, ...). The Institute has also provided internal awards for young people who have distinguished themselves for their scientific and / or planning activities;

(f) Very effective technical/administrative support. Over the years, software tools have been developed to support the administrative management of laboratories, personnel, project reporting, warehouse management, ... tools that have facilitated daily life in our Institute. In fact, ISTI is too large and too complex to be managed with the tools made available by the Central Administration.

As I said at the beginning, the research activity of the Institute has remained at the

highest levels in these years, even though the research themes have undergone profound changes. To give just one example, In the Visual Computing research laboratory, my lab, 12 years ago I was talking about computer graphics, volume visualization, 3D scanning, today we work on digital fabrication, 3D printing, and so on ...

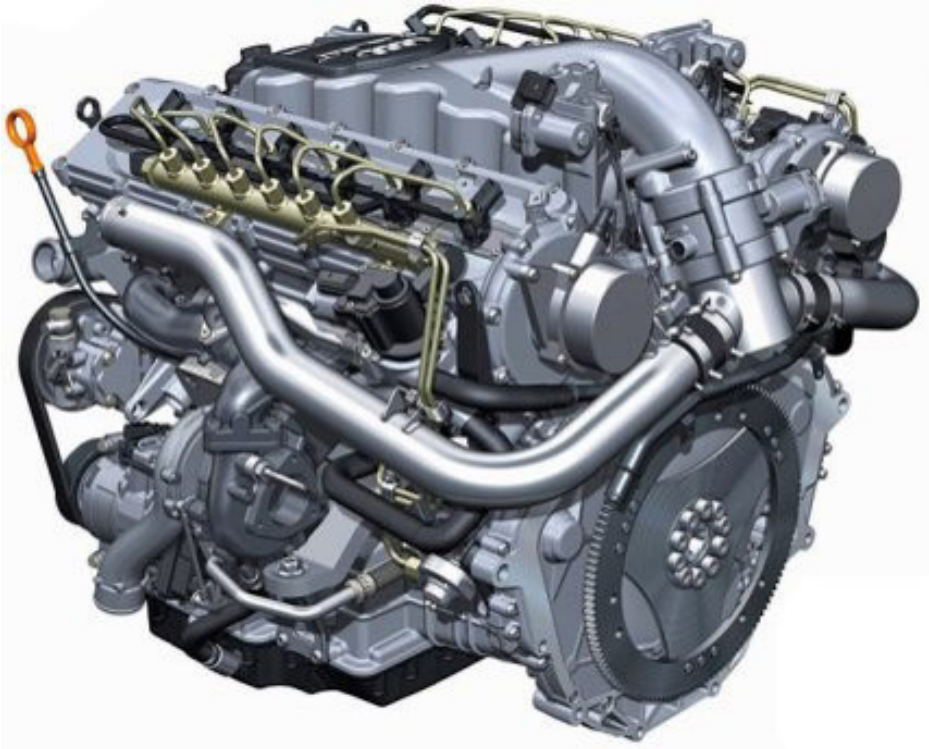
A similar evolutionary process has obviously occurred in all the research areas active in the Institute. Scientific research, by definition, can only evolve by exploring new fields, and facing new challenges. However, I feel that I can take little personal merit for the documented success of the Institute. That accolade goes to the many brilliant and hard-working scientists, technicians and administrators, with whom it has been my great good fortune to share what has been, for me, a truly beautiful and exciting experience.

Finally, let me add that I take great satisfaction in the thought that I am handing over to my successor an active, vibrant and dynamic Institute, which I am certain will continue to achieve important national and international results and recognition in the years to come.

Claudio Montani, ISTI
claudio.montani@isti.cnr.it

Industry 4.0

Research opportunities and challenges for ISTI



[continued]

It is not surprising that the key enabling technologies driving this new wave of transformation are deeply linked to advances in the ICT domain. With minor variations in the various regional, national and international implementation programs, the pillars of the new revolution range from Artificial Intelligence (AI) and Cloud Computing to Augmented Reality (AR), from the Internet of Things (IoT) to Additive Manufacturing.

The research labs of ISTI are playing a leading role in this transformation both by addressing current research challenges in this domain and by fostering the technological transfer of advanced solutions. In the following, we will present a few examples.

Machine learning for anomaly detection

An important theme in Industry 4.0 is the early detection of anomalous events in order to identify potential problems, optimize adaptive maintenance plans and avoid production stops.

In collaboration with the FCA research cen-

ter (CRF, Turin), innovative machine learning techniques are being developed at ISTI to be able to promptly detect events or observations in a data stream that do not conform to an expected pattern. The goal of the research is to implement intelligent data-driven techniques to detect problematic issues in the design of new cars.

A software framework based on deep learning Long Short-Term Memory (LSTM) architectures has been introduced to address the identification of anomalies in the time series data collected by heterogeneous automotive sensors.

Computer vision for monitoring and control

Cameras are another class of sensors in which the use of intelligent systems turns out to be essential if the visual content is to be fully exploited. Methods based on image processing and deep learning are being used to monitor production lines. Processing ranges from real time analysis of the single images, which is performed on smart cameras in an edge computing approach, for monitoring and quality assessment to the integrated analysis of multiple video streams

acquired along the entire production line. On the basis of this global analysis, adaptive policies can be deployed to set optimal parameters at the line and maintenance levels. These ideas are currently being applied in two projects involving the paper industry for tissue production and converting (respectively with A.Celli and Futura Converting), while predictive maintenance strategies are being studied for the optimal management of diagnostic imaging devices in a joint project with Imaginalis.



Augmented reality

Augmented Reality is an emerging technology which can have an important impact in Industry 4.0 with respect to the operation and maintenance of industrial plants, the simulation and design of new solutions and the training of new personnel.

ISTI is developing AR solutions to support operators providing performance dashboard, maintenance manuals and troubleshooting instructions. This is achieved through an Artificial Intelligence (AI) component connected to the Supervisory Control and Data Acquisition (SCADA) system and to a knowledge base of instructions. By understanding the context in which the operators are working, AI can proactively support them during their activity.

Indoor localization

AI should also rely on a precise indoor localization service. In a project in the tissue converting domain, this is achieved through the analysis of the visual content either acquired by RGB or Time-of-Flight



cameras mounted on tablets or see-through glasses. Techniques based on radiofrequency (eg Bluetooth Beacon) and on data fusion are also being studied. In this context, ISTI has led one of the main standardization efforts first with the EvAAL competition (2011-2013), then with the IPIN competition (2014-2017), and now by leading the IPIN-ISC subcommittee dedicated to the standardisation of evaluation procedures, including indoor localization.

Multimedia retrieval for decision support

Providing the right information at the right time means that the available documentation must be correctly indexed. Technologies are being developed in the framework of the ADA project to enable the automatic acquisition, organization, analysis and smart retrieval of information from texts and images contained in technical documentation. In this way, the documents produced by companies during the various stages of the production cycle can be effectively managed and analysed.

Digital manufacturing

In the digital manufacturing domain, computational methods are developed to design and produce physical, tangible objects, by exploiting additive manufacturing technologies or other digitally controlled machines. Some of the focus on production processes has led to the automatic design of 3D printable molds, which allow users to cheaply produce multiple copies of a di-

gital object like the ones shown in the top figure. Similarly, an alternative production method has been developed to generate flat lasercut pieces that can be assembled to form a target 3D object. Research is also studying the reproduction of mechanical behavior: a method has been developed to effectively reproduce a wide range of elastic properties using a set of 3D printed, single-material microstructures.

Health and Safety at work

Industry 4.0 is committed to paying attention to the quality of life, and the health and safety of industrial workers and others involved in organized activities. The noteworthy proliferation of sensing technologies and IoT in sports has attracted the attention of the scientific community. At ISTI, models based on machine learning techniques have been studied to predict the probability of players being injured during their next training sessions, given their recent workloads. These models can also be ported to the factory in order to foster the creation of

safer environments.

Personalizable working environments

Work environments are becoming increasingly richer in sensors and devices, and operators should be able to work safely while still being highly efficient and productive. At ISTI, researchers are investigating effective and usable ways to provide workers and managers with the possibility of intuitively specifying personalization rules in order to customize their smart, IoT-based working environments in a context-dependent manner, by generating actions, alarms and reminders in reaction to dynamic events occurring in working settings. To this aim, at ISTI, intuitive authoring tools are currently being developed, supporting trigger-action personalization rules for IoT applications in the paper industry sector, and for programming humanoid robots. Another key challenge being addressed is how to provide users with effective and high-quality interactive multimodal displays to control the factory of the future, by exploiting various interaction modalities (e.g. graphics, voice, gesture, vibro-tactile feedback, gaze, brain activity).

Industry 4.0 has recently gained significant popularity in both academia and the industrial sector. In this discussion, we analysed the main research challenges associated with Industry 4.0 that are currently being addressed by ISTI, to create smarter systems and services for the factory of the future.

Contact: Davide Moroni, SI Lab

Carmen Santoro, HIIS Lab

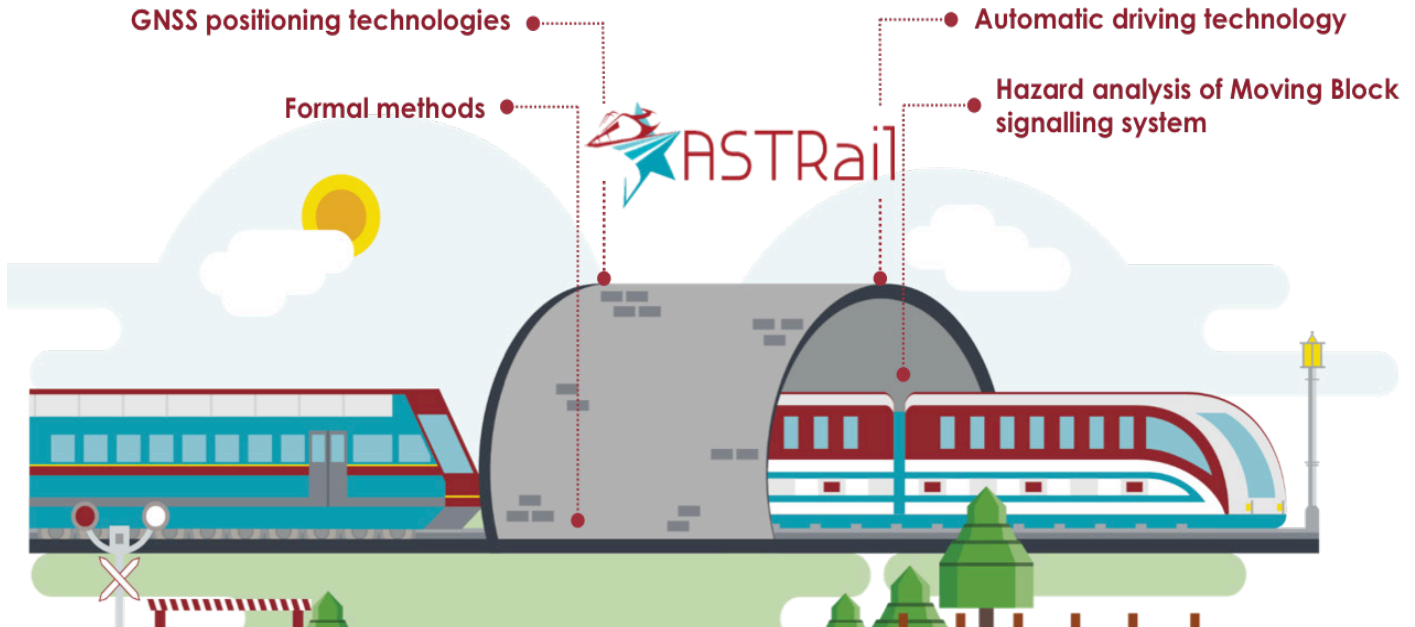
davide.moroni@isti.cnr.it

carmen.santoro@isti.cnr.it



ASTRail

SATellite-based Signalling and Automation Systems on Railways along with Formal Method and Moving Block validation
Co-funded by Horizon 2020 under the Shif2Rail program



Compared with other transport sectors, the railway sector is notoriously cautious about adopting technological innovations. This is commonly attributed to the sector's robust safety requirements. An example is smart route planning: while GNSS-based positioning systems have been in use for quite some time now in the avionics and automotive sectors to provide accurate positioning and smart route planning, the current train separation system is still based on fixed blocks – a block being the section of the track between two fixed points. The block sizes are determined based on parameters such as the line's speed limit, the train's speed, the braking characteristics of trains, sighting and reaction time of drivers, etc. However, the faster trains are allowed

to run, the longer the braking distance and the longer the blocks need to be, thus decreasing the line's capacity. This is because stringent safety requirements impose the length of fixed blocks to be based on the worst-case braking distance, regardless of the actual speed of the train.

With a moving block signalling system, in contrast, a safe zone around the moving train can be computed, thus optimising the line's exploitation. For this solution to work, it requires the precise absolute location, speed and direction of each train to be determined by a combination of sensors: active and passive markers along the track, as well as train borne speedometers. One of the current challenges in the railway sector

is to make moving block signalling systems as effective and precise as possible, including GNSS and leveraging on an integrated solution for signal outages (think, e.g., of tunnels) and the problem of multipaths. This is one of the main topics addressed by the ASTRail project: SATellite-based Signalling and Automation Systems on Railways along with Formal Method and Moving Block Validation.

Contact: Stefania Gnesi and Alessio Ferrari, FMT Lab
stefania.gnesi@isti.cnr.it
alessio.ferrari@isti.cnr.it
<http://www.astrail.eu/>

NESTORE

Novel Empowering Solutions and Technologies for Older people to Retain Everyday life activities

Co-funded by Horizon 2020



The ageing population is increasing rapidly in the EU; ICT can provide solutions for Active Ageing. However, the success of novel ICT solutions depends on the user perception with respect to their efficacy to support health promotion and global wellness. In this context, NESTORE will develop an innovative, multi-dimensional, personalized coaching system to support healthy ageing by: generating and sustaining motivation to take care of health; suggesting healthy nutrition and personalized physical and mental coaching, as well as social interaction, to prevent decline and preserve wellbeing.

NESTORE leverages on novel ICT technologies and will develop a multi-domain unobtrusive monitoring system, including wearable, environmental sensors and tangible objects. It will have an intelligent Decision Support System, to analyse user behaviour and provide personalized targets toward wellbeing. Active coaching, develop-

ped as a conversational agent will be embodied in a physical companion that assumes different forms, able to establish affective communication through multimodal communication channels thus engaging older people with personalized coaching activities in single or multiple domains.

NESTORE will adopt codesign tools and methods that will seek to engage and elicit information related to perception, acceptance and usability of technology to support healthcare. Such methods have been shown to be beneficial in the evaluation and design of health-care interventions.

NESTORE will last 3 years, from September 2017 to August 2020. The NESTORE consortium is composed of 16 partners from 7 European countries (Italy, Belgium, Netherland, Romania, Spain, Switzerland and United Kingdom): Politecnico di Milano – coordinator – supported by Fondazione

Politecnico di Milano, Universitat de Barcelona, University of Applied Sciences and Arts Western Switzerland, Sheffield Hallam University, University of Zurich, Technische Universiteit Delft, Loughborough University, Consiglio Nazionale delle Ricerche (CNR), Eurecat, Flextronics Design, Ropardo, Neosperience, La Meridiana due Cooperativa Sociale, Fundació Salut i Envel·liment, Preventie Collectief and AGE Platform Europe AISBL.

CNR participates in the NESTORE project with the joint collaboration of the Institute for Information Science and Technologies (ISTI), Pisa and the Institute of Bioimaging and Molecular Physiology (IBFM), Milan.

Contact: Filippo Palumbo, WN Lab
filippo.palumbo@isti.cnr.it
<https://nestore-coach.eu/>

Data4Impact

Big DATA approaches FOR improved monitoring of research and innovation performance and assessment of the societal IMPACT in the Health, Demographic Change and Wellbeing Societal Challenge
 Co-funded by Horizon 2020

Recent technological developments in data mining, data treatment and data analysis have been rapid and far reaching, offering new dimensions and opportunities for performance analytics in various domains. The introduction of new technologies and initiatives, including open access mechanisms and social media/online media, has been generating increasing volumes of new data on the research domain. Data4Impact aims to capitalise on these developments and utilise big data approaches to improve the monitoring of research and innovation performance and assessment of the societal impact in the Health, Demographic Change and Wellbeing Societal Challenge (see Fig. Traditional indicators to assess/monitor/analyse traditional innovations (R&D paradigm).

The project has the following main objectives:

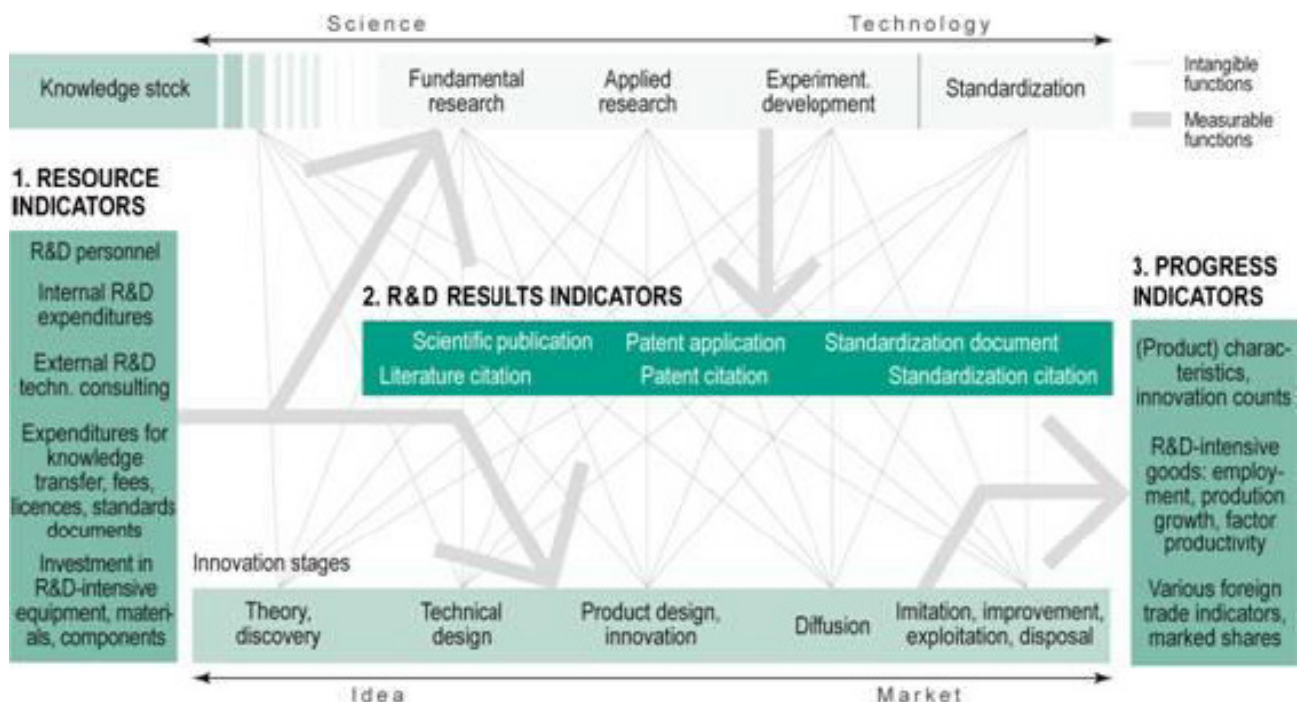
- a) define, develop, analyse and disseminate new indicators for assessing the performance of EU and national research and innovation systems;
- b) explore and collect “big” data on health-related societal challenges at input, throughput, output/result and impact levels;
- c) employ big data approaches to yield more data on the societal impact of national and EU funding on tackling health-related societal challenges;
- d) engage stakeholders in the project activities, validate the project results and develop new indicators and tools using a hands-on approach.

This will be achieved by developing a robust conceptual model addressing all key

stages of the innovation process; mining large volumes of data on research results and impacts; and analysing of this data using topic modelling, machine learning and other techniques aimed at natural language processing.

The Data4Impact consortium possesses specialist knowledge of the health domain & indicator systems, and is uniquely placed to mine data and apply big data approaches thanks to the partners' long-standing involvement in Open Access e-infrastructures and big data analytics. Through its activities Data4Impact addresses both of the key issues described in the specific CO-CREATION-08-2016-2017 call.

Contact: Paolo Manghi, NeMIS Lab
paolo.manghi@isti.cnr.it
<http://www.data4impact.eu/>



E-RIHS

The European Research Infrastructure for Heritage Science Co-funded by Horizon 2020

E-RIHS PP is a coordination and support action addressing the topic “Preparatory Phase and support to early phase of ESFRI projects” within the call H2020-INFRA-DEV-02-2016. The specific action addressed by this project is the preparation of E-RIHS, the European Research Infrastructure for Heritage Science (HS), one of the six new projects that entered the ESFRI Roadmap in 2016, and the only research infrastructure project in the Social and Cultural Innovation section of the Roadmap. E-RIHS will help the preservation of the World’s Heritage by enabling cutting-edge research in HS, liaising with governments and heritage institutions to promote constant development and, finally, raising the appreciation of the general public for cultural and natural heritage and recognition of its historic, social and economic significance.

E-RIHS PP began in 2017 and will last three years (2017-2020). The first two will be used to address governance, financial aspects, legal documents and logistics. This will lead to a business plan ready for application to ERIC, or to another suitable legal form, by



2019. The final year will be devoted to negotiations with stakeholders, further strategic planning and the start up of activities in order to enter the transition phase. E-RIHS will hopefully be launched as a stand-alone Research Infrastructure in 2021.

CNR coordinates the project with Luca Pezzati, from the Istituto Nazionale di Ottica. ISTI participates in the development of a

Catalogue for Heritage Science resources (Carlo Meghini, NeMIS Lab) and in the provision of visual services (Roberto Scopigno, VC Lab) for Heritage Science.

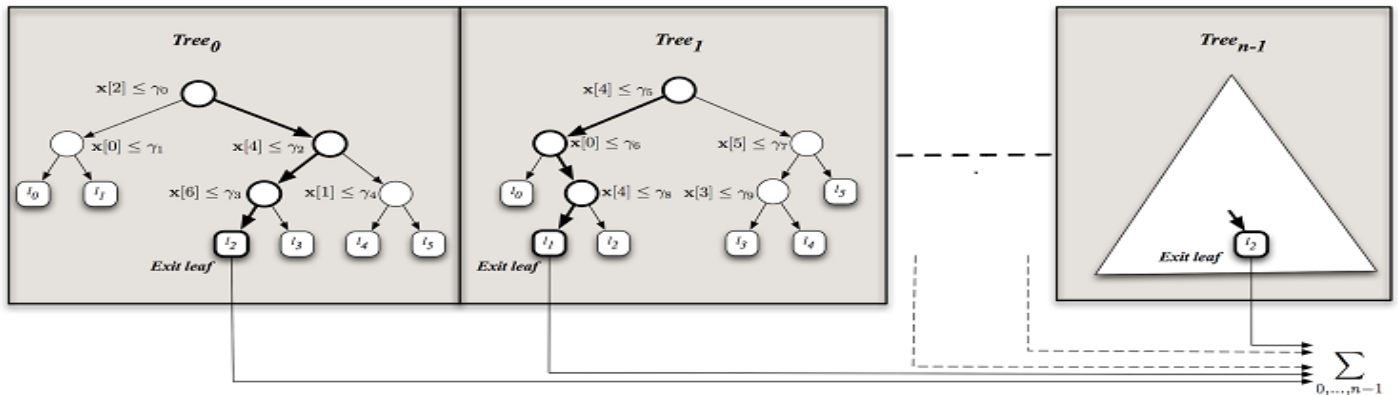
Contact: Carlo Meghini, NeMIS Lab
Roberto Scopigno, VC Lab
carlo.meghini@isti.cnr.it
roberto.scopigno@isti.cnr.it
<http://www.e-rihs.eu/>



HAMLET

Hardware Acceleration of Machine Learning Tasks

Funded by the National Research Council (CNR), Italy, and the National Scientific and Technical Research Council (CONICET), Argentina



Recent advances in Machine Learning (ML) and High Performance Computing (HPC) enabled new solutions for modeling phenomena that were previously too complex for computers to handle, thus opening new horizons for business and science.

However, the complexity of machine-learned models and their widespread use requires novel algorithmic solutions, aimed at rendering fast and scalable both the learning phase and the use of these models in large-scale applications.

The HAMLET project aims at investigating novel approaches to make ML models efficient and scalable. The Argentinean partner (National University of San Luis) has experience in the deployment of HPC models on Systems on Chip (SoC) Field-Programmable

Gate Arrays (FPGA). The Italian partner (HPC Lab @ ISTI-CNR) has experience in algorithmic solutions for large-scale machine learning applications.

The HAMLET project exploits this complementary expertise to investigate how hardware acceleration platforms can offer interesting characteristics diverse from traditional CPU to make ML models fast and scalable.

The preliminary focus of the project is towards the exploitation of FPGA-based hardware acceleration to speed-up ML models based on additive ensembles of decision trees. These models have been shown to be particularly effective in several classification, regression, and ranking tasks. Unfortunately, they are very computationally

demanding because all the decision trees of the ensemble have to be traversed for each item to which the model is applied. This high computational cost becomes a challenging issue in applications where the time budget available to apply the learnt model to a possibly huge number of items is limited and the users' expectations in terms of quality-of-service very high.

The project will generalize the lessons learnt from this first case-study to investigate how FPGAs can be successfully exploited to accelerate other computationally-intensive ML tasks.

Contact: Raffaele Perego, HPC Lab
raffaele.perego@isti.cnr.it

Environmental monitoring for smart cities

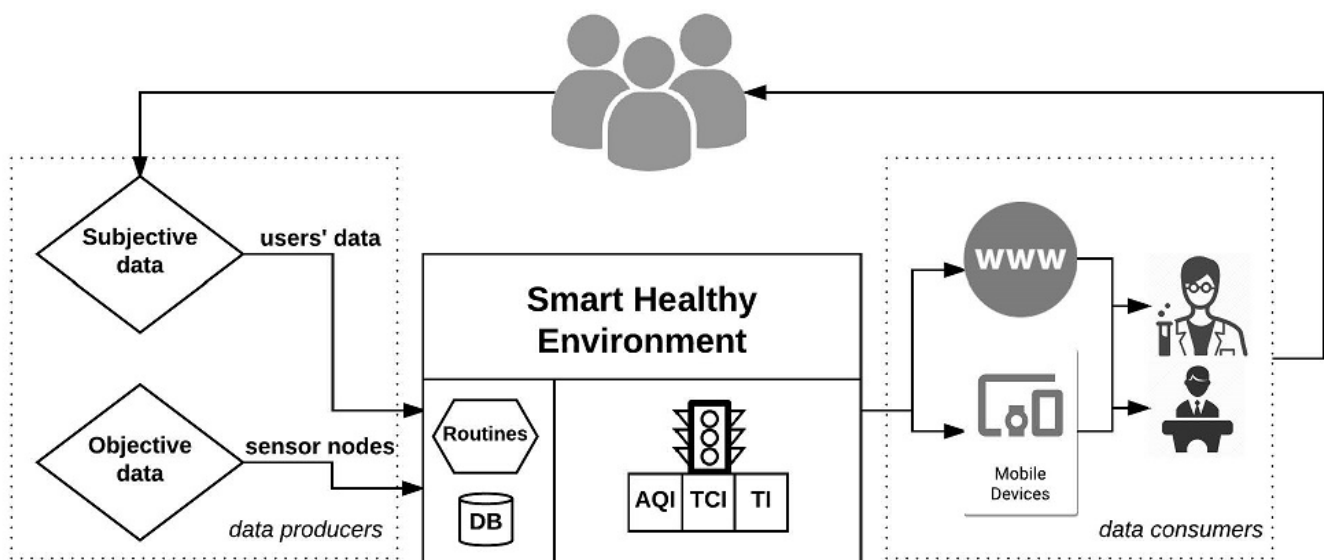
F.M. Bacco, F. Delmastro, E. Ferro, A. Gotta
 IEEE Sensors Journal, vol. 17. IEEE, 2017.

Sleep behavior is a key factor in maintaining good physiological and psychological health. A well-known approach to sleep monitoring is polysomnography. However, this is costly and intrusive, which may disturb sleep. Consequently, polysomnography is not suitable for sleep behavior analysis. Other approaches are based on actigraphy and sleep diaries. Although a good source of information for sleep quality assessment, sleep diaries can be affected by cognitive bias related to the subject's sleep perception, while actigraphy overestimates sleep periods and night-time disturbance compared to sleep diaries. Machine learning techniques can improve the objectivity and reliability of the observations. However, since signal morphology varies widely between people, conventional machine learning is complex to set

up. We have designed an adaptive, reliable, and innovative computational approach to provide per-night assessment of sleep behavior to the end-user. We exploit heartbeat rate and wrist acceleration data, gathered via a smartwatch, in order to identify the subject's sleep behavioral pattern. More specifically, heartbeat rate and wrist motion samples are processed via computational stigmergy, a bio-inspired scalar and temporal aggregation of samples. Stigmergy associates with each sample to a digital pheromone deposit (mark) defined in a mono-dimensional space and characterized by evaporation over time. As a consequence, samples close in terms of time and intensity are aggregated into functional structures called trails. The stigmergic trails allow to compute the similarity between time series

on different temporal scales in order to support classification or clustering processes. The overall computing schema includes a parametric optimization for adapting the structural parameters to individual sleep dynamics. The outcome is a similarity between nights of sleep for the same subject, to generate clusters of nights with different quality levels. Experimental results are shown for three real world subjects. The resulting similarity is also compared with the dynamic time warping, a popular similarity measure for time series.

DOI: [10.1109/JSEN.2017.2722819](https://doi.org/10.1109/JSEN.2017.2722819)



Exploring nominal cellular automata

T. Bolognesi, V. Ciancia

Journal of Logical and Algebraic Methods in Programming, vol. 93. Elsevier, 2017.

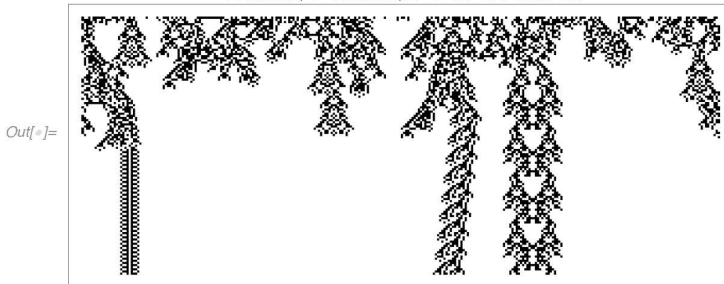
The emerging field of Nominal Computation Theory is concerned with the theory of Nominal Sets and its applications to Computer Science. We investigate here the impact of nominal sets on the definition of Cellular Automata and on their computational capabilities, with a special focus on the emer-

gent behavioural properties of this new model and their significance in the context of computation-oriented interpretations of physical phenomena. An investigation of the relations between Nominal Cellular Automata and Wolfram's Elementary Cellular Automata is carried out, together with an

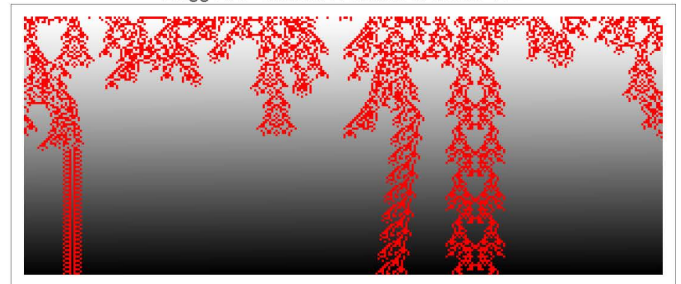
analysis of interesting particles, exhibiting "nominal" behaviour, in a particular kind of rules, reminiscent of the class of *totalistic* Cellular Automata, that we call "bagged".

DOI: [10.1016/j.jlamp.2017.08.001](https://doi.org/10.1016/j.jlamp.2017.08.001)

Totalistic, 5-context, 2-color CA Rule 20



Bagged 5-context Nominal CA Rule 102

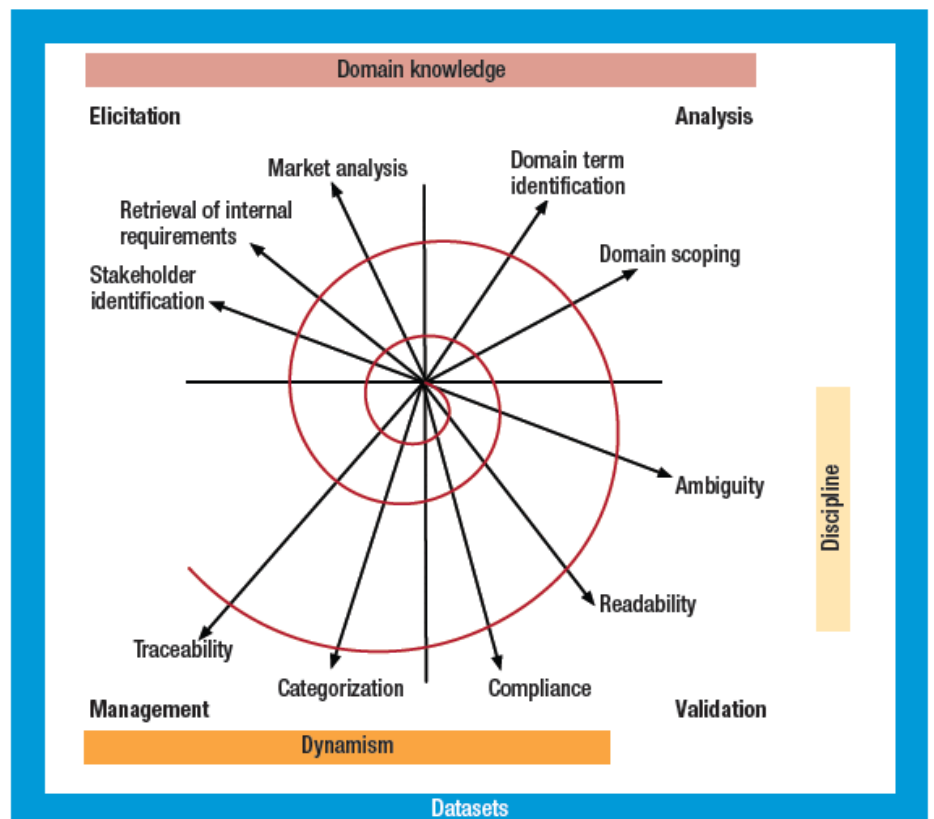


Natural language requirements processing: a 4D vision

A. Ferrari, F. Dell'Orletta, A. Esuli, V. Gervasi, S. Gnesi
IEEE Software, vol. 34 (6). IEEE, 2017.

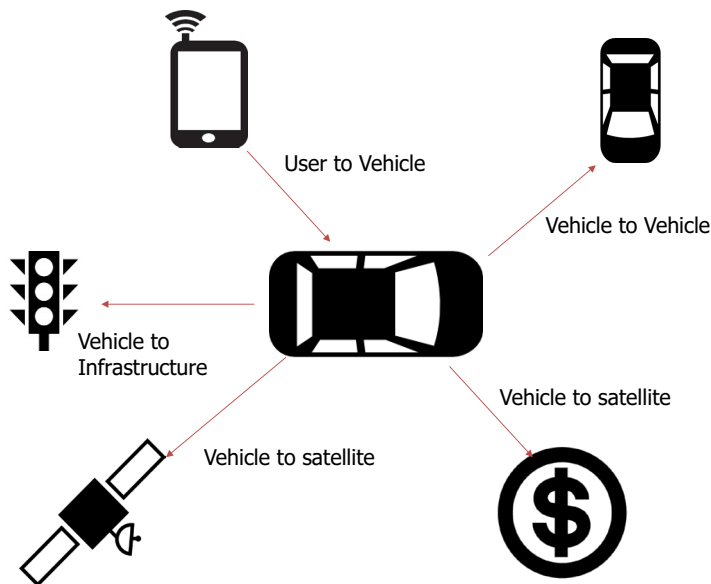
Natural language processing (NLP) and requirements engineering (RE) have had a long relationship, yet their combined use is not well established in industrial practice. This situation should soon change. The future evolution of the application of NLP technologies in RE can be viewed from four dimensions: discipline, dynamism, domain knowledge, and datasets.

DOI: [10.1109/MS.2017.4121207](https://doi.org/10.1109/MS.2017.4121207)



A tour of secure software engineering solutions for connected vehicles

A. Bertolino, A. Calabró, F. Di Giandomenico, G. Lami, F. Lonetti, E. Marchetti, F. Martinelli, I. Matteucci, P. Mori
Software quality journal. Chapman & Hall, in press.



The growing number of vehicles daily moving on roads increases the need to protect the safety and security of passengers, pedestrians, and vehicles themselves. This need is intensified when considering

the pervasive introduction of Information and Communication Technologies (ICT) systems into modern vehicles, because this makes such vehicles potentially vulnerable from the point of view of security. The con-

vergence of safety and security requirements is one of the main outstanding research challenges in software-intensive systems. This work reviews existing methodologies and solutions addressing security issues in the automotive domain with a focus on the integration between safety and security aspects. In particular, we identify the main security issues with vehicular communication technologies and existing gaps between state-of-the-art methodologies and their implementation in the real world. Starting from a literature survey and referring to widely accepted standards of the domain, such as AUTOSAR and ISO 26262, we discuss research challenges and set baselines for a holistic secure-by-design approach targeting safety and security aspects all along the different phases of the development process of automotive software.

DOI: [10.1007/s11219-017-9393-3](https://doi.org/10.1007/s11219-017-9393-3)

Challenges in certification of autonomous driving systems

F. Falcini, G. Lami
ISSREW 2017 –IEEE International Symposium on Software Reliability Engineering Workshops. IEEE, 2017.

Market readiness of on-board automotive software-intensive systems is tightly linked to the availability of appropriate certification schemes aimed at keeping the car makers confident and the consumers safe - especially in the context of Autonomous Driving, which is the next frontier of the automotive industry. Advanced driver assistance systems (ADAS) are going to be pervasively used in modern automobiles. New ADAS are principally based on Arti-

cial Intelligence (AI) technology, and in particular on deep learning. While the automotive community is aware of the important changes such a technology demands in terms of technical skills, development paradigms, and cultural approach, there is still an important gap to be filled in the availability of technical standards and, consequently, in terms of certification capability. Currently, the global automotive industry is subject to a series of standards that are more or less

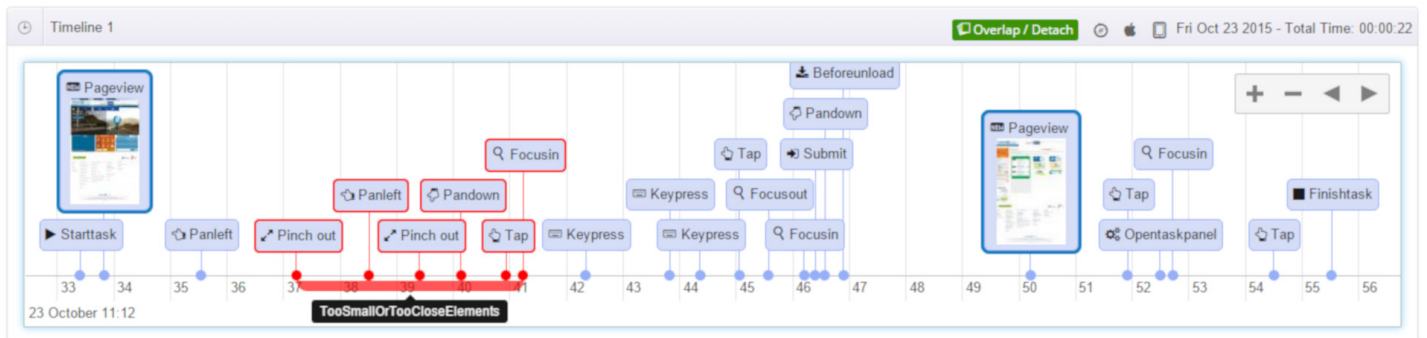
explicitly referring to a traditional way of developing software and systems, so that they are not at all suitable to be applied to ADAS. In this paper the open issues in certification of AI technologies in the automotive industry are addressed by providing an overview of the existing standards and the related applicability issues.

DOI: [10.1109/ISSREW.2017.45](https://doi.org/10.1109/ISSREW.2017.45)

Customizable automatic detection of bad usability smells in mobile accessed web applications

F. Paternò, A.G. Schiavone, A. Conte

MobileHCI '17 - 19th International Conference on Human-Computer Interaction with Mobile Devices and Services. ACM, 2017.



Remote usability evaluation enables the possibility of analysing users' behaviour in their daily settings. We present a method and an associated tool able to identify potential usability issues through the analysis of client-side logs of mobile Web interactions.

Such log analysis is based on the identification of specific usability smells. We describe an example set of bad usability smells, and how they are detected. The tool also allows evaluators to add new usability smells not included in the original set. We also report

on the tool use in analysing the usability of a real, widely used application accessed by forty people through their smartphones whenever and wherever they wanted.

DOI: [10.1145/3098279.3098558](https://doi.org/10.1145/3098279.3098558)

A web framework for cross-device gestures between personal devices and public displays

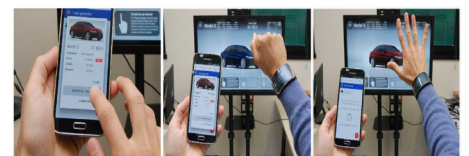
M. Barsotti, F. Paternò, F. Pulina

MUM '17 - 16th International Conference on Mobile and Ubiquitous Multimedia. ACM, 2017.

In order to exploit the wide availability of public displays and personal devices on the mass market at affordable prices it is important to provide developers with frameworks that ease obtaining cross-device user interfaces able to exploit such device ecosystems. We present the design and implementation of a Web framework for the development of cross-device user inter-

faces able to take advantage of both personal devices and public displays, and support various types of gestures and their combinations in such multi-device environments. We introduce the design space addressed, describe the framework functionality, its application interface and run-time support, show some example applications, and report on a first test with developers.

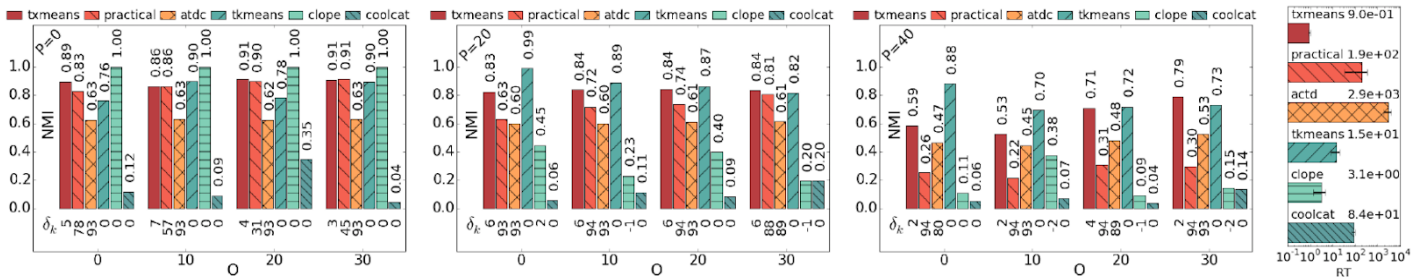
DOI: [10.1145/3152832.3152858](https://doi.org/10.1145/3152832.3152858)



Clustering individual transactional data for masses of users

R. Guidotti, A Monreale, M Nanni, F Giannotti, D Pedreschi

KDD '17 - 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017.



Mining a large number of datasets recording human activities in order to make sense of individual data is the key enabler of a new wave of personalized knowledge-based services. In this paper we focus on the problem of clustering individual transactional data for a large mass of users. Transactional data is a very pervasive kind of information that is collected by several services, often involving huge pools of users. We propose *txmeans*, a parameter-free clustering algorithm able to efficiently partition transactional data in a completely automatic way. *Txmeans* is designed for the case where clustering must



be applied on a massive number of different datasets, for instance when a large set of users need to be analyzed individually and each of them has generated a long history of transactions. A deep experimentation on both real and synthetic datasets shows the practical effectiveness of *txmeans* for

the mass clustering of different personal datasets, and suggests that *txmeans* outperforms existing methods in terms of quality and efficiency. Finally, we present a *personal cart* assistant application based on *txmeans*.

DOI: [10.1145/3097983.3098034](https://doi.org/10.1145/3097983.3098034)

Tiles: an online algorithm for community discovery in dynamic social networks

G. Rossetti, L. Pappalardo, D. Pedreschi, F. Giannotti

Machine Learning, vol. 106 (8). Springer 2017.

Community discovery has emerged during the last decade as one of the most challenging problems in social network analysis. Many algorithms have been proposed to find communities on static networks, i.e. networks which do not change in time. However, social networks are dynamic realities (e.g. call graphs, online social networks): in such scenarios static community discovery fails to identify a partition of the graph that is semantically consistent with the tem-

poral information expressed by the data. In this work we propose Tiles, an algorithm that extracts overlapping communities and tracks their evolution in time following an online iterative procedure. Our algorithm operates following a domino effect strategy, dynamically recomputing nodes community memberships whenever a new interaction takes place. We compare Tiles with state-of-the-art community detection algorithms on both synthetic and real world networks

having annotated community structure: our experiments show that the proposed approach is able to guarantee lower execution times and better correspondence with the ground truth communities than its competitors. Moreover, we illustrate the specifics of the proposed approach by discussing the properties of identified communities that it is able to identify.

DOI: [10.1007/s10994-016-5582-8](https://doi.org/10.1007/s10994-016-5582-8)

On the effects of low-quality training data on information extraction from clinical reports

D. Marcheggiani, F. Sebastiani

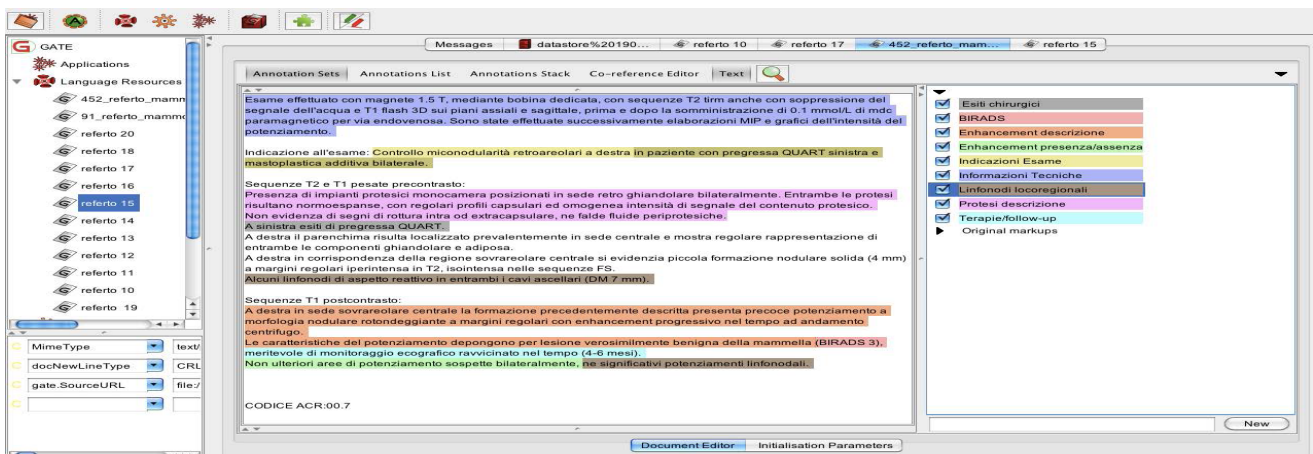
Journal of Data and Information Quality, vol. 9 (1). ACM 2017.

In the last five years there has been a flurry of work on information extraction from clinical documents, i.e., on algorithms capable of extracting, from the informal and unstructured texts that are generated during everyday clinical practice, mentions of concepts relevant to such practice. Many of these research works are about methods based on supervised learning, i.e., methods for training an information extraction system from manually annotated examples. While a lot of work has been devoted to devising learning methods that generate more and more accurate information extractors, no work has been devoted to investigating

the effect of the quality of training data on the learning process for the clinical domain. Low quality in training data often derives from the fact that the person who has annotated the data is different from the one against whose judgment the automatically annotated data must be evaluated. In this paper we test the impact of such data quality issues on the accuracy of information extraction systems as applied to the clinical domain. We do this by comparing the accuracy deriving from training data annotated by the authoritative coder (i.e., the one who has also annotated the test data, and by whose judgment we must abide), with the

accuracy deriving from training data annotated by a different coder, equally expert in the subject matter. The results indicate that, although the disagreement between the two coders (as measured on the training set) is substantial, the difference is (surprisingly enough) not always statistically significant. While the dataset used in the present work originated in a clinical context, the issues we study in this work are of more general interest.

DOI: [10.1145/3106235](https://doi.org/10.1145/3106235)



Preservation of database concepts: from data mining to concept theory

E.I. Locuratolo

Interdisciplinary Digital Preservation Tools and Technologies. IGI Global, 2017.

The algorithm for the integration of classes/concepts which results in an ontology suitable for the preservation of database classes/concepts is described. It has been designed by means of original research at the boundary between mathematics, computer science and concept theory. The input

conceptual graph of this algorithm is determined as a solution of a data mining problem, which has been approached as inverse mapping of conceptual database design. Similarities and differences are discussed between an algorithm formalized in concept theory, which is not suitable for

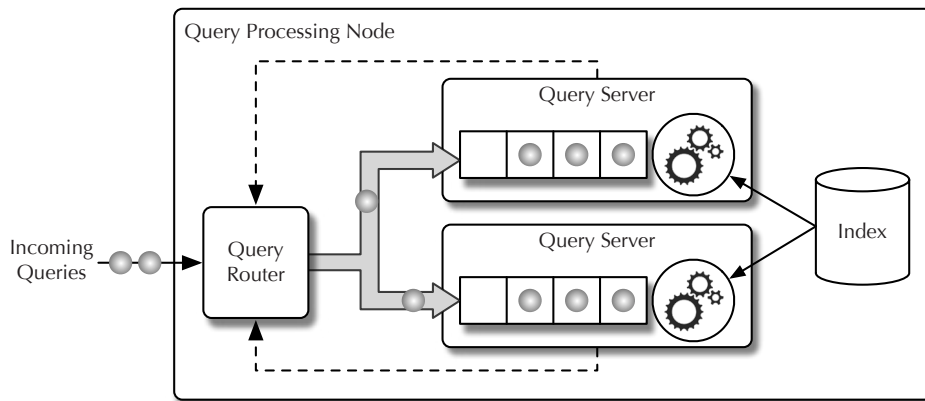
implementation, and the algorithm for the integration of classes/concepts, which exploits the benefits of database notations.

DOI: [10.4018/978-1-5225-1653-8.ch001](https://doi.org/10.4018/978-1-5225-1653-8.ch001)

Energy-efficient query processing in web search engines

M. Catena, N. Tonellotto

IEEE Transactions on Knowledge and Data Engineering, vol. 29 (7). IEEE, 2017.



Web search engines are composed by thousands of query processing nodes, i.e., servers dedicated to process user queries. So many servers consume a significant amount of energy, mostly accountable to their CPUs, but they are necessary to ensure low latencies, since users expect sub-second response times (e.g., 500 ms).

However, users can hardly notice response times that are faster than their expectations. Hence, we propose the Predictive Energy Saving Online Scheduling Algorithm (**PESOS**) to select the most appropriate CPU frequency to process a query on a per-core basis. **PESOS** aims at processing queries by their deadlines, and leverage high-le-

vel scheduling information to reduce the CPU energy consumption of a query processing node. **PESOS** bases its decision on query efficiency predictors, estimating the processing volume and processing time of a query. We experimentally evaluate **PESOS** on the TREC ClueWeb09B collection and the MSN2006 query log. Results show that **PESOS** can reduce the CPU energy consumption of a query processing node up to ~48 percent compared to a system running at maximum CPU core frequency. **PESOS** outperforms also the best state-of-the-art competitor with a ~20 percent energy saving, while the competitor requires a fine parameter tuning and it may incur in uncontrollable latency violations.

DOI: [10.1109/TKDE.2017.2681279](https://doi.org/10.1109/TKDE.2017.2681279)

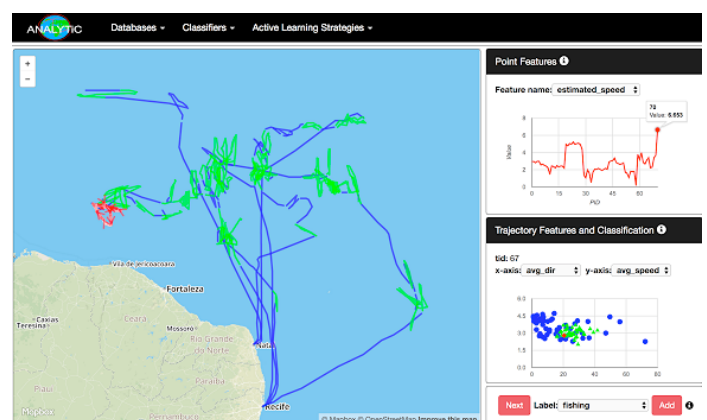
ANALYTIC: an active learning system for trajectory classification

A.S. Junior, C. Renso, S. Matwin

IEEE Computer Graphics and Applications, vol. 37 (5). IEEE, 2017.

An increasing volume of trajectory data is becoming available through the tracking of various moving objects, like animals, vessels, vehicles and humans. However, these large collections of movement data lack semantic annotations, since they are typically done by domain experts in a time-consuming activity. A promising approach is the use of machine learning algorithms to try to infer semantic annotations from the trajectories by learning from sets of labeled data. We experiment active learning, a machine learning approach minimizing the set of trajectories to be annotated while pre-serving good performance measures. We test some active learning strategies with three different

trajectory datasets in order to evaluate how this technique may limit the human effort required for the learning task. We support the annotation task by providing the ANALYTIC platform, a web-based interactive tool to visually assist the user in the active learning process over trajectory data.

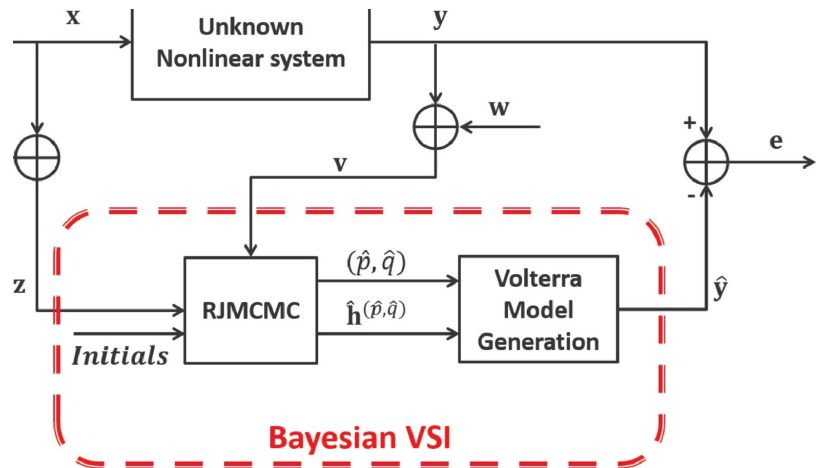


DOI: [10.1109/MCG.2017.3621221](https://doi.org/10.1109/MCG.2017.3621221)

Bayesian Volterra system identification using reversible jump MCMC algorithm

O. Karakuş, E.E. Kuruoğlu, M.A. Altınkaya
Signal Processing, vol. 141. Elsevier, 2017.

Volterra systems have had significant success in modelling nonlinear systems in various real-world applications. However, it is generally assumed that the nonlinearity degree of the system is known beforehand. In this paper, we contribute to the literature on Volterra system identification (VSI) with a numerical Bayesian approach which identifies model coefficients and the nonlinearity degree concurrently. Although this numerical Bayesian method, namely reversible jump Markov chain Monte Carlo (RJMCMC) algorithm has been used with success in various model selection problems, our use is in a novel context in the sense that both memory size and nonlinearity degree are estimated. The aforementioned study ensures an anomalous approach to RJMCMC and provides a new understanding on its



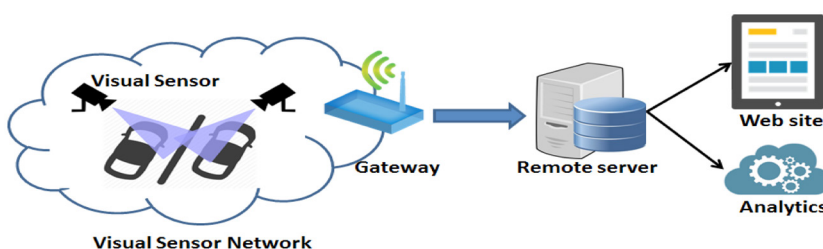
flexible use which enables trans-structural transitions between different classes of models in addition to transdimensional transitions for which it is classically used. We study the performance of the method on synthetically generated data including

OFDM communications over a nonlinear channel.

DOI:[10.1016/j.sigpro.2017.05.031](https://doi.org/10.1016/j.sigpro.2017.05.031)

An intelligent cooperative visual sensor network for urban mobility

G.R. Leone, D. Moroni, G. Pieri, M. Petracca, O. Salvetti, A. Azzarà, F. Marino
Sensors, vol. 17 (11). MDPI, 2017.



Smart cities are demanding solutions for improved traffic efficiency, in order to guarantee optimal access to mobility resources available in urban areas. Intelligent video analytics deployed directly on board embedded sensors offers great opportunities to gather highly informative data about traffic and transport, allowing reconstruction of a real-time neat picture of urban mobility patterns. In this paper, we present a visual

sensor network in which each node embeds computer vision logics for analyzing in real time urban traffic. The nodes in the network share their perceptions and build a global and comprehensive interpretation of the analyzed scenes in a cooperative and adaptive fashion. This is possible thanks to an especially designed Internet of Things (IoT) compliant middleware which encompasses in-network event composition as well as

full support of Machine-2-Machine (M2M) communication mechanism. The potential of the proposed cooperative visual sensor network is shown with two sample applications in urban mobility connected to the estimation of vehicular flows and parking management. Besides providing detailed results of each key component of the proposed solution, the validity of the approach is demonstrated by extensive field tests that proved the suitability of the system in providing a scalable, adaptable and extensible data collection layer for managing and understanding mobility in smart cities.

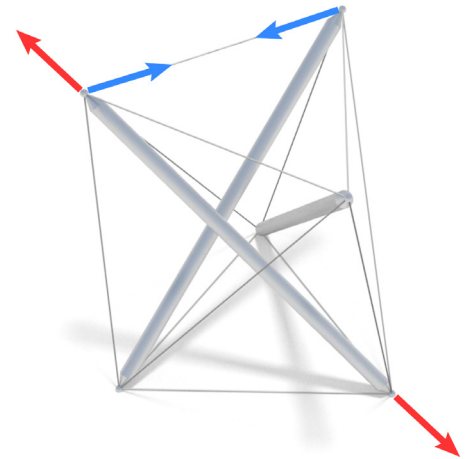
DOI: [10.3390/s17112588](https://doi.org/10.3390/s17112588)

Position-based tensegrity design

N. Pietroni, M. Tarini, A. Vaxman, D. Panozzo, P. Cignoni
ACM Transactions on Graphics, vol. 36 (6). ACM, 2017.

We propose a novel framework for the computational design of tensegrity structures, which are constructions made of struts and cables, held rigid by continuous tension between the elements. Tensegrities are known to be difficult to design---existing design methods are often restricted to using symmetric or templated configurations, limiting the design space to simple constructions. We introduce an algorithm to automatically create free-form stable tensegrity designs that satisfy both fabrication

and geometric constraints, and faithfully approximate input geometric shapes. Our approach sidesteps the usual force-based approach in favor of a geometric optimization on the positions of the elements. Equipped with this formulation, we provide a design framework to explore the highly constrained space of tensegrity structures. We validate our method with simulations and real-world constructions.



DOI: [10.1145/3130800.3130809](https://doi.org/10.1145/3130800.3130809)

From paper to web: automatic generation of a web-accessible 3D repository of pottery types

M. Dellepiane, M. Callieri, F. Banterle, D. Arenga, M. Zalocco, R. Scopigno
EG GCH 2017 - 15th Eurographics Workshops on Graphics and Cultural Heritage, Eurographics Association, 2017.

3D web repositories are a hot topic for the research community in general. In the Cultural Heritage (CH) context, 3D repositories pose a difficult challenge due to the complexity and variability of models and to the need for structured and coherent metadata for browsing and searching.

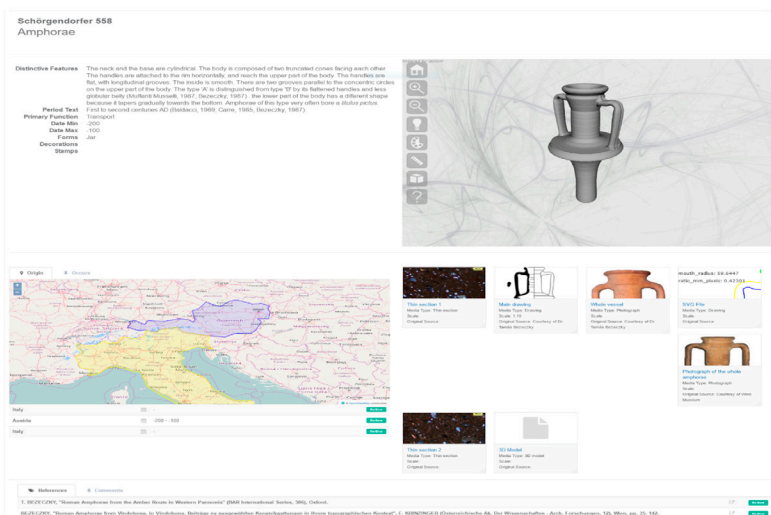
This paper presents one of the efforts of the ArchAIDE project: to create a struc-

tured and semantically-rich 3D database of pottery types, usable by archaeologists and other communities. For example, researchers working on shape-based analysis and automatic classification.

The automated workflow described here starts from pages of a printed catalog, extracts the textual and graphical description of a pottery type, and processes those data

to produce structured metadata information and a 3D representation. This information is then ingested in the database, where it becomes accessible by the community using dynamically-created web presentation pages, showing in a common context: 3D, 2D and metadata information.

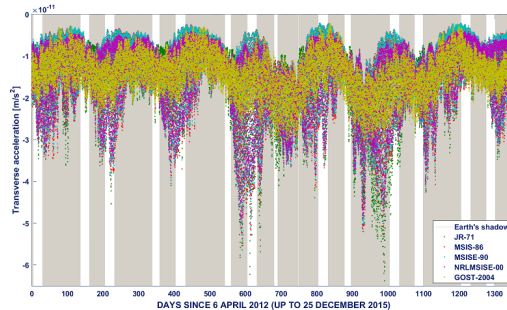
DOI: [10.2312/gch.20171293](https://doi.org/10.2312/gch.20171293)



On the secular decay of the LARES semi-major axis

C. Pardini, L. Anselmo, D.M. Lucchesi, R. Peron
Acta Astronautica, vol. 140. Elsevier, 2017.

The laser-ranged satellite LARES is expected to provide new refined measurements of relativistic physics, as well as significant contributions to space geodesy and geophysics. The very low area-to-mass ratio of this passive and extremely dense satellite was chosen to reduce as much as possible the disturbing effects of non-gravitational perturbations. However, because of its height, about 1450 km compared with about 5800-5900 km for the two LAGEOS satellites, LARES is exposed to a much stronger drag due to neutral atmosphere. From a precise orbit determination, analyzing the laser ranging normal points of LARES over a time span of about 3.7 years with the GEODYN II (NASA/GSFC) code, an average semi-major axis decay rate of -0.999 m per year was found, corresponding to a non-conservative net force acting nearly opposite to the velocity vector of the satellite and with a mean along-track acceleration of -1.444×10^{-11} m/s². By means of a modified



version of the SATRAP (ISTI/CNR) code, the neutral drag perturbation acting on LARES was evaluated over the same time span, taking into account the real evolution of solar and geomagnetic activities, with five thermospheric density models (JR-71, MSIS-86, MSISE-90, NRLMSISE-00 and GOST-2004). All of them provided consistent results, well within their acknowledged uncertainties. Moreover, when the same models (JR-71 and MSIS-86) were used within GEODYN II in a least-squares fit of the tracking data, the differences between the average drag

coefficients estimated with SATRAP and GEODYN were of the order of 1% or less. Unlike what happened for the two LAGEOS, where Yarkovsky thermal drag and charged particle drag were the leading causes, it was found that neutral atmosphere drag alone was able to explain most (~98.6%) of the observed semi-major axis decay of LARES. The remaining ~1.4%, corresponding to an average along-track acceleration of about -2×10^{-13} m/s² (i.e. ~1/70 of neutral drag), was probably linked to thermal thrust effects. It was 50%, or less, of the value previously reported in the literature, but further and more detailed investigations, including the detection of the signature of the periodic terms, will be needed in order to characterize such smaller non-gravitational perturbation.

DOI: [10.1016/j.actaastro.2017.09.012](https://doi.org/10.1016/j.actaastro.2017.09.012)

On the derivative of the stress–strain relation in a no-tension material

C. Padovani, M. Silhavý
Mathematics and Mechanics of Solids, vol. 22 (7). Sage, 2017.

The stress–strain relation of a no-tension material, used to model masonry structures, is determined by the nonlinear projection of the strain tensor onto the image of the convex cone of negative-semidefinite stresses under the fourth-order tensor of elastic compliances. We prove that the stress–strain relation is indefinitely differentiable on an open dense subset Ω of the set of all strains.

The set Ω consists of four open connected regions determined by the rank $k=0, 1, 2, 3$ of the resulting stress. Further, an equation for the derivative of the stress–strain relation is derived. This equation cannot be solved explicitly in the case of a material of general symmetry, but it is shown that for an isotropic material this leads to the derivative established earlier in [1] by different means.

For a material of general symmetry, when the tensor of elasticities does not have the representation known in the isotropic case, only general steps leading to the evaluation of the derivative are described.

DOI: [10.1177/1081286515571786](https://doi.org/10.1177/1081286515571786)

References

[1] Lucchesi M., Padovani C., Zani N., (1996). Masonry-like solids with bounded compressive strength. International Journal of Solids and Structures, vol.33 (14) p. 1961-1994 and Masonry constructions: Mechanical models and numerical applications. Berlin: Springer, 2008.

Hexalab: visualizing complex hexahedral meshes on the web

The Visual Computing Lab has now released another free open source tool to the research community: HexaLab. HexaLab is an online WebGL application for interactive visualization, exploration and assessment of hexahedral meshes; it runs directly in a web browser, so you can just go to www.hexalab.net and use it.

Hex-meshes, i.e. volumetric meshes composed of hexahedral cells, are one of the most used 3D representations for numerical simulation, most notably by Finite Element Analysis (FEA). However the shape and the configuration of the cubic cells filling the volume, affects the quality of the simulation results and the analysis of these structures. In particular an intelligible visualization of their interior, is a non trivial task.

HexaLab is an advanced visualization tool

that fill this niche. It targets both users and scholars who employ hexahedral meshes for Finite Element Analysis, and it facilitates inspection and checking of mesh quality, assessing its usability for simulations.

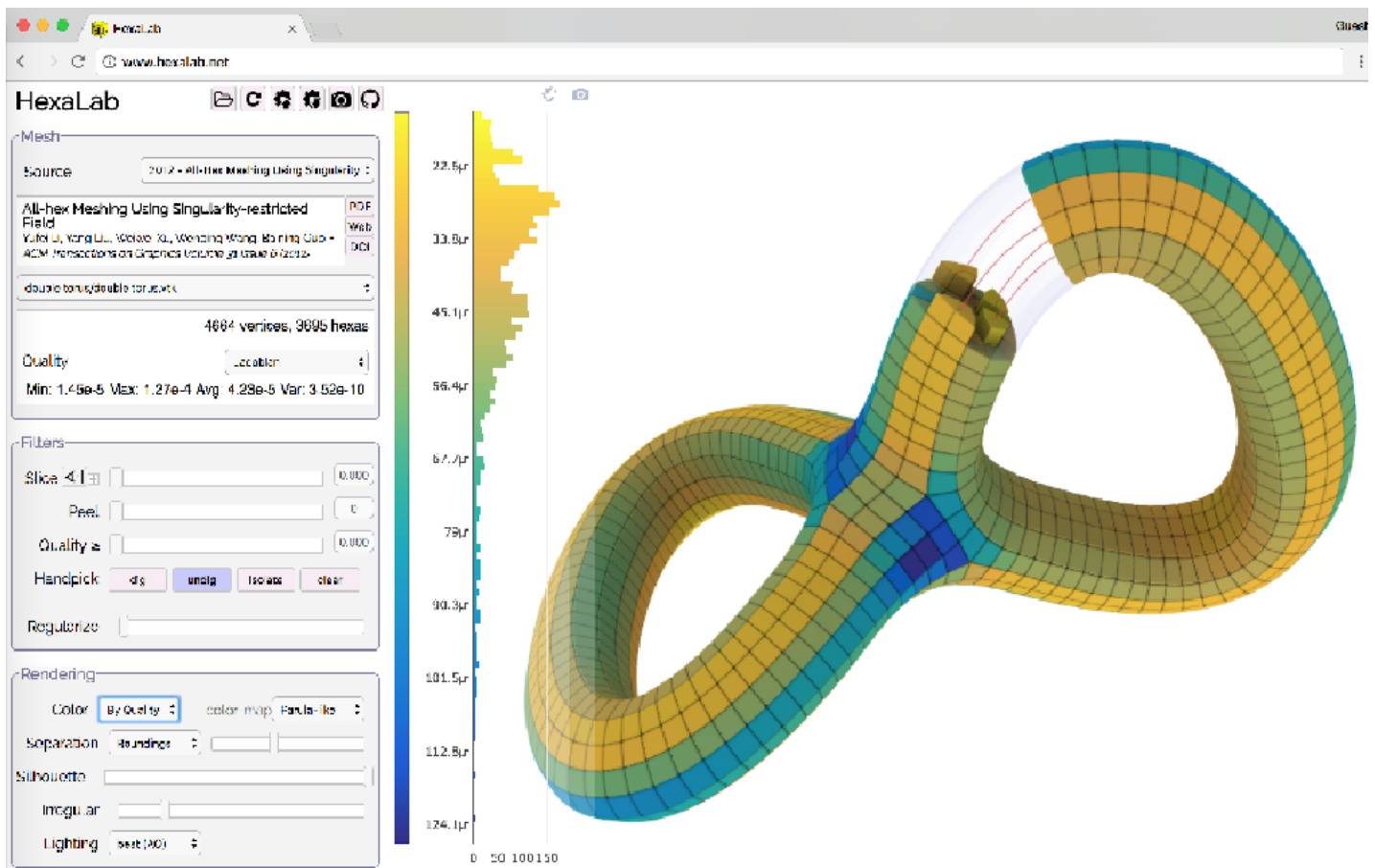
HexaLab can be used to perform a detailed analysis of the mesh structure, isolating weak points using all the quantitative and quality measures proposed in the literature, providing statistical information and color mapping visualization.

To allow effective analysis of the meshing, it supports a wide variety of interactive visualization and volume inspection tools such as slicing by planes, peeling from the outside, filtering out well shaped cells to reveal the bad ones and even a 'minecraft' style removal and reveal picking of individual cubic elements.

To generate high quality images and to help the immediate spatial perception of shape relations, HexaLab uses advanced global illumination effects, like ambient occlusion and sophisticated shading of the elements.

Finally, to promote reproducible research and simple access to the state of the art in this field, HexaLab also offers integrated access to a repository containing all the publicly available meshes produced together with the most recent papers for hex mesh generation. For more information see the companion paper available on arxiv <https://arxiv.org/abs/1806.06639>.

Contact: Paolo Cignoni, VC Lab
paolo.cignoni@isti.cnr.it



Young open science award

Edition 2017

The ISTI Young Open Science Award aims at promoting the sharing and use of software and dataset produced during the research activities of the ISTI personnel. The award selects a publication connected to a software or dataset that has been made available to the community, and generated an impact in terms of uses, citations, visibility.

Recipient 2017:

“A multisource and multivariate dataset for indoor localization methods based on WLAN and geo-magnetic field fingerprinting” by Davide La Rosa.

Contact: Matteo Dellepiane, VC Lab
matteo.dellepiane@isti.cnr.it

Davide La Rosa



Makers of merit at the European edition of the maker faire

Rome 1-3 Dicembre 2017

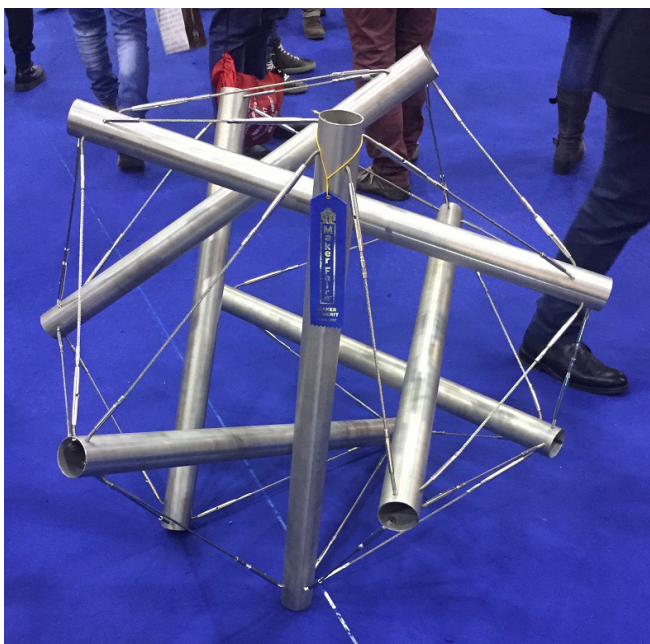
Maker Faire is the Greatest Show (and Tell) on Earth – a family-friendly showcase of invention, creativity and resourcefulness, and a celebration of the Maker Movement. It's a place where people of all ages and backgrounds gather together to show what they are making, and share what they are learning.

Maker Faire originated in 2006 in the San Francisco Bay Area as a project of the editors of “Make: magazine” It has since grown into a significant worldwide network of both flagship and independently-produced events.

The VC Lab new technology for designing tensegrity structures, has been featured at CNR stand of the Maker Faire in Rome – the European Edition 4.0.

Maker Faire Rome is the world's largest exhibition after the “Area Bay” and “New York” in the US, with more than 100.000 visitors attending.

The new technology introduced by the VC Lab has been awarded the “Makers of Merit” blue ribbon as one of the projects that really stand out, that always have a crowd around it.



Contact: Paolo Cignoni, VC Lab
paolo.cignoni@isti.cnr.it
<http://vcg.isti.cnr.it/Publications/2017/PT-VPC17/>

ProgettISTI 2017

The ProgettISTI initiative calls for the presentation of short project proposals by young researchers from different laboratories. The aim is to improve the mutual knowledge of the research activities in the Institute, to encourage young researchers to experiment the creation and writing of

a project proposal, and hopefully to help in starting new collaborations.

The proposal selected for 2017 was:

- TAPAS: Tensor Algorithms for Performance Analysis of Large Systems

Proponents

Giulio Masetti, SEDC Lab and Leonardo Robol, MMS Lab.

Contact: Matteo Dellepiane, VC Lab
matteo.dellepiane@isti.cnr.it

Grants for Young Mobility

The ISTI Grants for Young Mobility (GYM) program enables young researchers (under 34) to carry out research in cooperation with foreign Universities and Research In-

stitutions of clear international standing. It complements similar CNR programs.

The winners in 2017 were:

Contact: Matteo Dellepiane, VC Lab
matteo.dellepiane@isti.cnr.it



Fabio Carrara



Matteo Catena

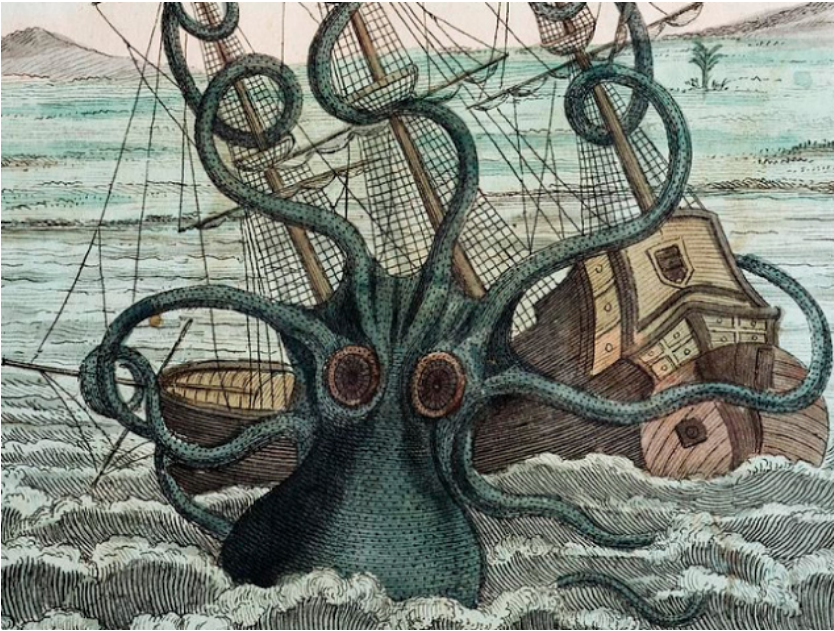


Ioanna Miliou



Manuele Sabbadin

The giant squid across myth and science



Stories of sea monsters from different era have strong resemblance to each other. In medieval times, these have particularly referred to the mythological Kraken, sometimes depicted as an island with tentacles or as a lurking monstrous sea serpent capable of sinking an entire fleet. At the end of the 18th century, these representations have converged into the form of a giant squid. In the 19th century, myth faded more and more away into literature (e.g. J. Verne's "Twenty Thousand Leagues Under the Sea"), allowing the giant squid to officially enter scientific taxonomies as *Architeuthis dux*. Despite this long history, this species was scientifically known only from stranded or dead specimens, thus myth has continued to prosper and inspire narrative literature, music, poetry, movies, and even video-games during the 20th and the first decade of the 21st century. At the same time, indirect scientific investigations of the giant squid's life, behaviour, and habitat have begun to collect data in a systematic way and to organise them in databases, building the overall profile of an abyssal 20 m-long

cephalopod. It was only in 2012 that, based on this collected knowledge, an expedition of biologists was able to find and film a giant squid alive in its native habitat for the first time.

In 2015, the NeMIS InfraScience Lab used the D4Science e-Infrastructure to combine giant squids sightings reported all over the world, with environmental information from satellites and buoys, in order to estimate the species' preferences to biotic and abiotic parameters and consequently produce a map of habitat suitability. Generating this map required using a cloud computing platform (the gCube DataMiner) in combination with other geographically spread data publication and access services of D4Science. These included providers of species occurrence data (e.g. GBIF, OBIS etc.) and environmental data (e.g. Copernicus, NASA, NOAA etc.). D4Science facilitated the combination of these data through the use of standard representations, autonomous management of access policies, security, and heterogeneous representations, which would have

required months of work otherwise. Cloud computing was used to train ~ 100,000 machine learning models parametrised (based on Maximum Entropy and Artificial Neural Networks) and to select the model with the highest prediction performance of certified sea level sightings. This work, published on the Ecological Modelling Journal, produced the first time ever habitat suitability map of the giant squid *Architeuthis dux*. This map revealed the ubiquity of the giant squid in the abysses of the global oceans and shared several patterns with an expert hand-drawn map from 1987. Further, it identified suitable areas where two other *Architeuthis dux* subspecies have been recently identified.

In 2017, the story of the myth and science around the giant squid has been represented as an interactive timeline by means of the Narrative Building Tool (NBT) developed by the NeMIS Digital Libraries group. NBT is a semi-automatic tool to construct and visualise narratives, intended as semantic networks of events related to each other through semantic relations, based on an "ontology for narratives" developed by the same DL group. The tool facilitates the construction of events to form narratives by using Wikidata/Wikimedia Commons as external knowledge base of images and entities. The produced narratives can be exported as Linked Open Data and are automatically represented using the Web Ontology Language (OWL). The produced timelines can be also visualised in other formats like tables and network graphs.

Contact: Gianpaolo Coro, NeMIS Lab

gianpaolo.coro@isti.cnr.it

<https://dlnarratives.eu/timeline/squid.html>

References

- ¹ Candela, L., Castelli, D., & Pagano, P. (2009, July). D4Science: an e-Infrastructure for Supporting Virtual Research Environments. In IRCDL (pp. 166-169).
- ² Coro, G., Panichi, G., Scarponi, P., & Pagano, P. (2017). Cloud computing in a distributed e-infrastructure using the web processing service standard. *Concurrency and Computation: Practice and Experience*, 29(18).
- ³ Coro, G., Magliozzi, C., Ellenbroek, A., & Pagano, P. (2015). Improving data quality to build a robust distribution model for *Architeuthis dux*. *Ecological Modelling*, 305, 29-39.
- ⁴ <https://dlnarratives.eu/timeline/squid.html>.
- ⁵ <https://dlnarratives.eu/>.

Summer school on natural space risks

28 August 28 – 1 September, 2017
Paris Observatory - France

Carmen Pardini, a researcher of the Space Flight Dynamics Laboratory active in the field of reentry predictions of uncontrolled space objects since 1991, was an invited lecturer on this subject at the international Summer School on Natural Space Risks, held at the Astronomical Observatory of Paris, the oldest and largest infrastructure in astrophysics in France, from August 28 to September 1, 2017.

This summer school aimed at providing master and first year PhD students with an intensive training on all natural space risks. The program addressed space weather, space debris, uncontrolled reentries, near-Earth objects, and their societal impacts.

The invited lecture of Carmen Pardini was titled: "Uncontrolled reentries of manmade space objects: how to get reliable products to manage and mitigate the potential risk in the airspace and on the ground".

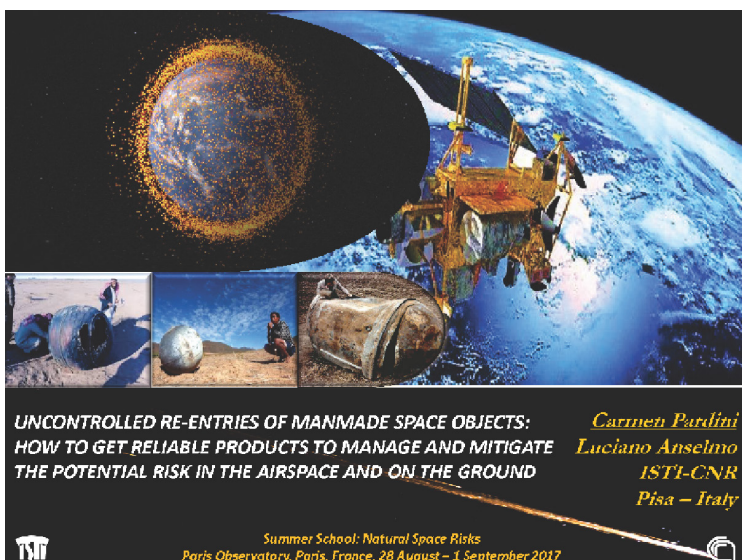
In spite of decades of efforts, predicting the reentry time and location of an uncontrolled satellite remains a very problematic activity as reentry predictions are affected by diverse sources of unavoidable uncertainties. The experience accumulated worldwide suggests that a relative prediction error of

$\pm 20\%$ should be adopted to compute the uncertainty windows associated with nominal reentry epoch predictions, in order to reasonably cover all possible error sources. However, in specific cases, more conservative prediction errors, up to $\pm 30\%$, should be considered, in particular during the last 2-3 days of residual lifetime. Therefore, even predictions issued a few hours before reentry may be affected by a quite large along-track uncertainty, often corresponding to more than one full orbital path.

In consequence, the typical reentry prediction standard products, such as the nominal decay forecasts with the associated reentry uncertainty windows and corresponding sub-satellite ground tracks, are of no, or very limited, use for civil protection applications. In other words, the locations possibly at risk in a given area of the planet cannot be identified reasonably ahead of reentry using such knowledge. For this reason, specific approaches and procedures have been devised

and applied in Italy, since the orbital decay of the BeppoSAX spacecraft in 2003, to provide reasonable and unambiguous information useful for civil protection planning and applications.

Contact: Carmen Pardini, SFD Lab
carmen.pardini@isti.cnr.it
<https://nsr-2017.sciencesconf.org/>



Ph.D. dissertations

Energy efficiency in large scale information retrieval systems

Author: Matteo Catena, Gran Sasso Science Institute, L'Aquila
 Supervisor: Nicola Tonellotto

Web search engines perform query processing on clusters composed by thousands of computers and hosted in large data centers. Such facilities can require megawatts of electricity to operate, raising economic and environmental concerns. Therefore, an important question is how to reduce the energy expenditure of these data centers. Another problem is how to reduce the carbon dioxide emissions and the consequent negative impact of the data centers on the environment.

One solution consists in using green energy to partially power the data centers. We propose a new query forwarding algorithm that exploits both the green energy sources

available at different data centers and variations in market energy prices. The proposed solution maintains a high query throughput, while reducing the energy operational costs of multi-center search engines by up to 25%.

Another option is to reduce the energy consumption of CPUs to mitigate the energy expenditure and carbon footprint of a search company. Currently, CPU core frequencies are typically managed by the operating system. We propose to delegate the CPU power management from the OS to the query processing application. This solution reduces server power consumption by up to 24%, with only limited (but uncontrollable)

drawbacks in the quality of search results with respect to a system running at maximum CPU frequency. Since users hardly notice response times that are faster than their expectations, we reason that Web search engines should not process queries faster than user expectations. Consequently, we propose the Predictive Energy Saving Online Scheduling (PESOS) algorithm to select the most appropriate CPU frequency to process a query by its deadline, on a per-core basis. PESOS can reduce the CPU energy consumption of a query processing server from 24% up to 48% when compared to a high performance system running at maximum CPU core frequency.

Tangible interaction in museums and cultural heritage sites: towards a conceptual and design framework

Author: Daniele Duranti, IMT School for Advanced Studies, Lucca
 Supervisors: Emanuele Pellegrini; Roberto Scopigno; Raffaella Trocchianesi

Drawing on a design perspective, the research explores the application of tangible interaction in museums and cultural heritage sites. Tangible interaction is a consolidated research area inside HCI (Human Computer Interaction) and Interaction Design.

Since the early 2000s, tangible interaction has also been applied to the cultural heritage field for the creation of onsite interactive installations that better integrate digital technologies, the materiality of the objects and the physicality of the experience during the visit.

So far, research in this field within the cultural heritage domain has mainly focused on developing new systems and evaluating

them while a move towards more theoretical and conceptual studies is still missing. As a consequence, there is no common language in the field, neither is there a clear understanding of what has been done and what is missing, nor a formalization of the aspects that make up the design of tangible interaction systems.

This research represents a first attempt at overcoming these problems at least partially by replying to three fundamental questions:

1. How has tangible interaction been applied to onsite interactive installations in the cultural sector?
2. What kind of experience of cultural heritage does tangible interaction allow?

3. What are the aspects that make up the design of a tangible interaction system?

In order to answer these questions a theoretical framework for tangible interaction in museums and cultural heritage sites is proposed. It can be intended as both a conceptual and a design framework. Indeed, not only does it show what tangible interaction is by providing a categorisation of past tangible interaction systems, but it also identifies a set of aspects that make up the design of such interactive systems. These aspects represent themes around which choices have to be made during the design process, and the knowledge of which can facilitate or inspire the design.

Content-aware quad meshing

Author: Giorgio Marcias, Dipartimento di Informatica, Università degli Studi di Pisa

Supervisors: Roberto Scopigno, Nico Pietroni

Quadrilateral meshes are preferred over triangular meshes in the animation and CAD industry, mostly because they fit well the role of control grids for higher-level digital representations of smooth surfaces, such as subdivision surfaces and Non-Uniform Rational B-Splines (NURBS). Nonetheless, most software and hardware designed for 3D visualization and also most acquisition technology is optimized to generate highly accurate triangular meshes.

Due to their different properties, triangular and quadrilateral meshes have each led to different types of applications, even in different scientific disciplines. Research has put effort into the design of methods to convert from one to the other.

Currently, however, no automatic method is able to produce quadrilateral meshes that can be directly employed in the animation pipeline.

In this thesis we investigate the problem of converting a dense, triangular mesh into a coarse and highly structured quadrilateral mesh, suitable to flow into the standard pipeline of production in the animation industry. We discuss what properties are required and what are the most commonly used methods that can be employed for this goal. On the basis of the properties identified, we present three different methods. The first one generates meshes with structurally sound patch layouts. The second in-

troduces a way to drive the quadrangulation with the object deformations expected in an animation sequence.

Finally, the last method overcomes the shortcomings of the others by allowing the user to sketch the high-level patch layout connectivity in an intuitive and interactive manner.

We employed a data-driven approach, which provides precise control over every single aspect of the mesh, without the drawbacks of the classical, manual methods, but leveraging the expertise contained implicitly in other, well-designed, quadrilateral meshes.

Improving the efficiency and effectiveness of document understanding in web search

Author: Salvatore Trani, Dipartimento di Informatica, Università degli Studi di Pisa

Supervisor: Raffaele Perego

Thinking Web Search Engines are probably nowadays the most complex information systems since they need to match billions of web pages with short and often ambiguous queries submitted by a multitude of users. This task is particularly challenging due to the requirement to provide high quality results with sub-second response times.

In recent years, natural language understanding techniques have been profitably exploited to overcome ambiguities in queries and documents. Enriching raw

text with semantic concepts - where concepts can precisely identify persons, companies, locations, etc. - makes it possible to achieve a remarkable improvement in retrieval effectiveness.

In this thesis, we argue that it is possible to improve both efficiency and effectiveness of document understanding in Web search by exploiting learning-to-rank solutions. On one hand, enriching documents with machine-learned semantic annotations leads to an improvement of the effectiveness,

since the retrieval of relevant documents can exploit a finer comprehension of the documents.

On the other hand, by enhancing the efficiency of learning to rank techniques, we can improve both efficiency and effectiveness, since a faster ranking technique can reduce query processing time or, alternatively, allow a more complex and accurate ranking model to be deployed.

Privacy risk assessment in big data analytics and user-centric data ecosystems

Author: Francesca Pratesi, Dipartimento di Informatica, Università degli Studi di Pisa

Supervisors: Fosca Giannotti, Dino Pedreschi

Nowadays, our life is centered on data. We leave traces relating to our movements via our mobile phones, to our relationships within social networks, to our habits from query-logs. The resulting data enables a new class of services that can improve our society. The drawback is the question of privacy: since the data describes our lives at a very detailed level, privacy breaches can occur.

The objective of this thesis is two-fold. Firstly, we propose a framework that aims to enable a privacy-aware data-sharing ecosystem, based on Privacy-by-Design. This framework, called PRISQUIT (Privacy RiSk versus QUallTy), can support a Data

Provider in sharing personal data with an external entity. PRISQUIT helps to decide the right level of aggregation of the data and the most suitable strategies for enforcing privacy, by quantifying the empirical privacy risk of individuals. It then analyzes the data quality guaranteed using only data from users having a privacy risk below a specific threshold. The framework is modular, so it is possible to enrich it with new kinds of data, new privacy risks and utility functions, new end services and new mitigation strategies.

Secondly, we investigate the privacy perspective within a user-centric model, where each individual has full control of the life cycle of his personal data. To do this,

PRISQUIT is employed to study the correlation between some individual features, such as the mobility entropy, and the actual privacy risk, permitting each user to obtain an estimated level of his own privacy risk. After that, we propose three transformations based on the differential privacy paradigm, which offers very strong privacy guarantees regardless of any external knowledge of a malicious agent.

We also conduct a wide range of experiments to prove the flexibility and the utility of PRISQUIT and the usefulness of the two approaches related to the user-centric ecosystem.

Welcome aboard!



Claudio Atzori
Research Staff (Tecnologo)
NeMIS Lab



Miriam Baglioni
Research Staff (Researcher)
NeMIS Lab



Valentina Bartalesi Lenzi
Research Staff (Researcher)
NeMIS Lab



Andrea Carboni
Research Staff (Tecnologo)
SI Lab



Alejandro Moreo Fernandez
Research Staff (Researcher)
NeMIS Lab



Luca Pappalardo
Research Staff (Researcher)
KDD Lab

Conferences - Co-organized by ISTI

Congress Visual Heritage 2018 - 23rd CHNT and EG GCH 2018
Vienna, Austria - November 12 - 15, 2018

<https://www.chnt.at/topic/>



Digitize: Research - Record - Reactivate - Reconstruct
hosted by CHNT 23 2018 | Vienna, Austria - November 12 - 15, 2018



RefRefresh 2018 - 1st Workshop on Reframing Research (@ EuroCSS symposium)
Cologne, Germany - December 5, 2018

<http://refresh.kmi.open.ac.uk>



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>

Editorial Secretariat

segreteria scientifica@isti.cnr.it

Editorial Board

Fabrizio Falchi
Silvia Giannini
Carol Peters
Claudia Raviolo
Roberto Scopigno

Layout and Design

Salvatore Rinzivillo
Daniele Fadda

Copyright notice

All authors, as identified in each article, retain copyright of their work
The authors are responsible for the technical and scientific contents of their work

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 amendment

Legal representative of the Institute of Information Science and Technologies "A. Faedo": Dott. Claudio Montani

ISTI News is not for sale but is distributed for purposes of study and research and published online at
<http://www.isti.cnr.it/ISTINews>

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