ISTI Young Research Award 2015

Alessia Bardi, Leonardo Candela*, Gianpaolo Coro, Matteo Dellepiane, Andrea Esuli,
Lorenzo Gabrielli, Alberto Gotta, Claudio Lucchese, Diego Marcheggiani,
Franco Maria Nardini, Filippo Palumbo, Nico Pietroni, Giulio Rossetti

Abstract
The ISTI Young Researcher Award is an award for young people of Institute of Information Science and Technologies (ISTI) with high scientific production. In particular, the award is granted to young staff members (less than 35 years old) by assessing the yearly scientific production of the year preceding the award. This report documents procedure and results of the 2015 edition of the award.

Keywords
Young Research Award

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", Consiglio Nazionale delle Ricerche, Via G. Moruzzi 1, 56124, Pisa, Italy
***Corresponding author**: leonardo.candela@isti.cnr.it*

Contents

Introduction

The Institute of Information Science and Technologies (ISTI), an institute of the Italian National Research Council (CNR), promotes the growth of its "young researchers" by means of initiatives aiming at encouraging the scientific production and promoting the visit to major international scientific institutions and research groups. Among these initiatives, the *Young Researcher Award* (YRA) yearly awards the Institute staff of less than 35 years old with the best scientific production. This initiative is funded through self-taxation of all research laboratories of the Institute.

The ISTI YRA is awarded to ISTI members belonging to the following categories:

- *Young*: it is awarded to PhD students and PhD researchers less than 32 years old;

- *Young++*: it is awarded to PhD students and PhD researchers less than 35 years old.

The award is presented each year at the ISTI Day, a yearly meeting where the Director meets the Institute staff. Three persons in each category are awarded with a research funding of 1,500€.

Methods

Submission

Nominations for the YRA Award should be submitted by the candidate by using a dedicated online form. The information collected via the form are very basic. They include name, date of birth, and dates related to PhD activity only. The list of publications will be automatically acquired by the ISTI Institutional Repository[1].

YRA Award Committee

The YRA Award Committee is nominated by the Director of the Institute with the following duties:

- Prepare and develop the call for participation and the related procedures;

- Solicit nominations and assess candidates;

- Provide the ISTI Director with documents underlying the entire process and selecting the award candidates.

The Committee members are:

Claudio Lucchese (Chair) (High Performance Computing Laboratory);

Leonardo Candela (Networked Multimedia Information System Laboratory);

Matteo Dellepiane (Visual Computing Laboratory);

Andrea Esuli (Networked Multimedia Information System Laboratory);

Alberto Gotta (Wireless Networks Laboratory);

---

[1] http://http://pumalab.isti.cnr.it

Franco Maria Nardini (High Performance Computing Laboratory);

Nico Pietroni (Visual Computing Laboratory);

Selection

A total of 13 nominations where received: 7 for the Young category and 6 for the Young++ category.

The following criteria was defined to assess and rank each scientific publication of the candidates:

- diverse ranking systems are going to be used to reduce the effects of any bias;

- for Journal papers it is used (*i*) the Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (ANVUR) Journals ranking[2]; (*ii*) the Computing Research and Education Association of Australasia (CORE) Journals ranking[3]; and (*iii*) the SCImago service[4]. Papers receive a score according to the schema reported in Table 1. In case of multiple scores, the maximum one is used;

- for conference papers it is used the Group of Italian Professors of Computer Engineering (GII) and Group of Italian Professors of Computer Science (GRIN) rating service[5]; papers receive a score according to the schema reported in Table 2;

- "short papers", i.e., papers having less than 6 pages, receive half of the score of the homologous papers;

- papers published in workshops receive a score of 2;

- book chapters not associated to a conference receive a score of 2;

- international conference abstracts receive a score of 1.

YRA 2015 Recipients

The recipients of the award for the "*Young*" category are:

Diego Marcheggiani (Networked Multimedia Information System Laboratory)

Lorenzo Gabrielli (Knowledge Discovery and Data Mining Laboratory)

Giulio Rossetti (Knowledge Discovery and Data Mining Laboratory)

The recipients of the award for the "*Young++*" category are:

---

[2] http://www.anvur.org/index.php?option=com_content&view=article&id=254&Itemid=623&lang=it
[3] http://www.core.edu.au/
[4] http://www.scimagojr.com/
[5] http://www.consorzio-cini.it:8080/consultazioneclassificazioni/

Alessia Bardi (Networked Multimedia Information System Laboratory)

Filippo Palumbo (Wireless Networks Laboratory)

Gianpaolo Coro (Networked Multimedia Information System Laboratory)

A per recipient introduction to the research activity as well as to the set of publications leading to the award is reported in the following sections.

Diego Marcheggiani (Young): Publications 2014

The following publications was produced by Marcheggiani during 2014 and evaluated for the YRA 2015: [1, 2, 3].

The work of Marcheggiani is at the crossroads of machine learning, natural language processing, and sentiment analysis. In these last years he focused on sequence learning tasks using Conditional Random Fields (CRFs). CRFs is a supervised learning algorithm specifically devised for sequence classification problems, where the aim of the classifier is not the classification of single elements (i.e., as in text classification) but of the elements as part of an entire sequence one which one has to take into consideration dependencies between elements. These problems are extremely common in natural language processing; part of speech tagging, phrase chunking, named entity recognition and some forms of opinion extraction are few examples of sequence classification tasks.

In [1] Marcheggiani expanded the the most common instantiation of CRFs, the linear chain CRFs, to deal with multiclass, multi-label problems. He tested the multi-label CRFs on a aspect-oriented opinion extraction task. He showed empirically that multi-label CRFs outperform standard linear chain CRFs on a hotel reviews dataset.

In [2] Marcheggiani investigated the problem of active learning for semi-supervised CRFs. Active learning is a procedure that helps to find the most informative examples that have to be annotated in order to get an accurate classification model. Since human annotation is very expensive, active learning helps to reduce the annotation effort, making the human annotator focus to those examples that are more informative for the problem solution. Marcheggiani mixed together active learning and semi-supervised learning to reduce the annotation effort to the minimum possible. He empirically showed (on four different natural language processing tasks) that with the best active learning strategy, paired with semi-supervised CRFs, it is possible to learn an accurate CRFs model with few, but very informative examples.

Finally, in [3] Marcheggiani studied the benefits of having high quality training data versus low quality training data to train CRFs. For this work he trained CRFs to extract medical concepts from medical reports written in Italian.

From these studies he discovered that having low quality training data does not impact in a sensible way the classification performance.

Table 1. Papers in Journal: score

ANVUR	CORE	Scimago	Score
1	A*	Q1	10
2	A	Q2	8
3	B	Q3	6
4	C	Q4	4

Table 2. Conference Papers: score

CORE	Score
A*	8
A	8
B	6
C	4

Lorenzo Gabrielli (Young): Publications 2014

The following publications was produced by Gabrielli during 2014 and evaluated for the YRA 2015: [4, 5, 6, 7, 8, 9, 10].

During the year Lorenzo Gabrielli worked on the analysis of big data to measure mobility habits and levels of wellbeing of society.

Within the study of mobility, it has been realized a methodology to engineer the classification process called Sociometer [10], in order to make portable the entire process. It was also verified the robustness of the methodology with respect to the variation of the time window used. The results in [11] show: (*a*) The proposed improvement allows to automatically label the typical behaviors of call making use of archetypes that describe the typical behavior, avoiding to make use in the phase of classification of a domain expert, obtaining an accuracy of 85%; (*b*) The minimum time window, so that the Sociometer can produce reliable results, it is empirically proved to be of three weeks. Also as part of the study of mobility, was carried out the analytical part, and their interfaces at the base of the RESTful application called Urban Mobility Atlas (UMA), whose purpose is to tell the story of individual mobility in an area, using a set of statistical indicators obtained with Mobility Data Mining technique and Machine Learning.

Regarding to the other area of research, measuring well-being of an area, he participated in the creation of an analytic process designed to explore how mobility and sociability of individuals are related to the economic well-being. Thanks to the measures of mobility and sociability extracted from phone data it is possible to create a model to relate these measures and measures of well-being (European Deprivation Index). The results, which are going to be published, highlight how the model is able to explain more than 40% of the variance of the phenomenon and in particular that the mobility measures are much more predictive of those of sociability. This result is of great impact since the availability of mobility measures is significantly higher than those of sociability, especially for reasons of privacy.

Finally, a new research topic under study concerns the use of open government data for measuring the performance of public administrations

All the researches were performed within the Cimplex[6] and SoBigData[7] EU projects.

Giulio Rossetti (Young): Publications 2014

The following publications was produced by Rossetti during 2014 and evaluated for the YRA 2015: [12, 13].

During the year, Rossetti worked on the analysis of network dynamics and science of success. In his research on complex networks he primarily addressed the *dynamic community discovery* problem proposing an algorithm, TILES, able to dynamically track communities in online streamed graphs. Even though network science has received huge attention in the last decade, only recently time-aware approaches started to be proposed in order to cope with the intrinsic dynamism of the phenomena modeled by complex networks: social contacts, business exchanges, online communications are only few examples for which a static analysis is not able to produce up-to-date and meaningful results the often under- (or over-)estimate the real network connectivity.

Starting from this observation, Rossetti focuses his attentions on the analysis of how the temporal dimension shapes network topology. In such rapidly changing context, communities are used as the basilar bricks able to bound the analytical spectrum to well defined subsets of nodes: communities can be tracked trough time, used as boundaries to study network homophily, seen as preferential playgrounds when trying to predict future interactions among nodes. All this themes were addressed in Rossetti's PhD thesis *"Network Dynamics"* [14] and constituted the backbone of the subjects of DyNo, the 1st workshop on *Dynamics In Networks* (held in conjunction with the 2015 IEEE/ACM International Conference ASONAM, Paris, France, August 2015) for which he acted as co-chair.

The other main theme he pursued during this year is related to the science of success. Several studies have tried to

---

[6]`www.cimplex-project.eu`
[7]`www.sobigdata.eu`

capture the connections amongst success and innovation: so far economic theories proposed strategies to organize adopters in different categories in order to model how both the level of risk and the social pressure affect their decision to acquire a novel good. One of the major question that remains open, once used such models to classify the adopters, is: can we forecast the "*success*" of a good by analyzing its early adopters? To answer such question, a data driven analysis was performed on different contexts (music listenings, retail market purchases...) and it was observed how there exist a peculiar category of *innovators* able to sense successful goods and to selectively adopt them before the others.

All the researches on network dynamics and science of success were performed within the Cimplex[8] and SoBigData[9] EU projects.

Alessia Bardi (Young++): Publications 2014

The following publications was produced by Bardi during 2014 and evaluated for the YRA 2015: [15, 16, 17, 18, 19, 20].

The research activities of Bardi are conducted in the field of data infrastructures for scholarly communication. In 2014 she focused on two main research topics: *enhanced publications* and *aggregative data infrastructures*.

The notion of enhanced publication is an emerging paradigm for the representation and dissemination of scientific research results that responds to the request of research communities of including in the scholarly communication the whole range of research outcomes used and produced during a scientific investigation, together with the traditional scientific article. An enhanced publication enriches a digital publication with additional research products (e.g., research data, workflows, scientific processes, software) to complete the description of the research and support research assessment, reproduction and re-use. In [15] Bardi and Manghi analysed the state of the art on enhanced publications and proposed a common terminology and classification schemes for the comparison and discussion of existing data models for enhanced publications and the information systems that adopt them. The study highlighted the existence of a plethora of solutions for enhanced publications, each realised to serve a specific research community, and suggests the need of a more systemic approach at the realization of information systems for enhanced publications in order to support their sustainability, decrease their realisation and maintenance cost and, therefore, foster their wider adoption.

The second research topic of Bardi is that of Aggregative Data Infrastructures (ADIs). ADIs are digital infrastructures supporting the aggregation of data and metadata from a set of heterogeneous data sources. ADIs play an important role in scholarly communication as they bring together, integrate and combine digital objects deposited at different locations and offer a single entry point to support researchers in the

discovery of and access to relevant research material. Interoperability issues of data and technologies, scarcity of computational resources, and highly evolving requirements typically represent an obstacle to practitioners constructing ADIs capable of performing such an integration. [17] presented the architectural principles and the services of the D-NET Software Toolkit, a framework where developers find the tools for the construction and operation of customized, robust, scalable, and autonomic ADIs in a cost-effective way as instances of service-oriented data infrastructures. The D-NET Software Toolkit is a general-purpose service-oriented framework adopted by several national consortia and EC projects, such as OpenAIRE[10] and HOPE (Heritage of the People's Europe) [15, 19], to create aggregative infrastructures under diverse application domains.

Filippo Palumbo (Young++): Publications 2014

The following publications was produced by Palumbo during 2014 and evaluated for the YRA 2015: [21, 22, 23, 24, 25, 26, 27].

During the year, Palumbo worked on the development of a middleware infrastructure that can easily abstract and export sensors pervasively deployed in smart environments. Having such a communication infrastructure, running on desktop and mobile platforms, let him collect data from running test-sites coming from the EU FP7 GiraffPlus [11] project (in which he led the Work Package 2 - System Design and Development). These data were used to create context-aware applications useful for the target scenario: Ambient Assisted Living (AAL). AAL aims at the creation of services oriented to the assistance of elderly people. This research area is becoming more and more popular due to the increasing age of population in developed countries. In the vision of smart cities, elderly people are supported by pervasive and smart environments in their daily mobility and activities, with which they interact by means of their personal devices (like smartphones). Such environments are composed by a large number of objects that offer a variety of services which, in turn, are requested, negotiated, and consumed by the personal devices of the users. For this reason, the presence of a lightweight distributed communication infrastructure, like the developed middleware, is an important requirement. AAL services are build exploiting contextual information coming from the sensorized home, like the mobility of the user, his activities and behavioural patterns. In this regard, one of the most important source of context information is the position of the user in the house. For this reason, part of the year's activities were focused on the development of a device-free indoor localization system that opportunistically exploits the capabilities offered by the smart environment. AAL applications generally have low accuracy requirements for indoor localization; this opens the opportunity for parasitizing the existing smart envi-

---

[8]www.cimplex-project.eu
[9]www.sobigdata.eu

[10]The OpenAIRE initiative, Open Access Infrastructure for Research in Europe: www.openaire.eu
[11]http://www.giraffplus.eu/

ronment infrastructure without adding dedicated positioning sensors. In this scenario, one can exploit simple binary sensors that are usually present in the smart environment, such as light and appliance switches or intrusion detection sensors, to obtain a rough estimate of the position of the user. The application developed, called CEO, is device-free, meaning that the user is not required to carry any device in order to be localized. Future works will further analyse the possibilities offered by smart environments equipped with distributed sensor network and the middleware infrastructure developed coming from the EU FP7 DOREMI [12] project.

Gianpaolo Coro (Young++): Publications 2014

The following publications was produced by Coro during 2014 and evaluated for the YRA 2015: [28, 29, 30, 31, 32].

Research e-Infrastructures are the main focus of the work of Coro. These are computer systems enabling researchers in different locations across the world to collaborate and work on the same scientific initiatives. E-Infrastructures usually provide either sharing and storage facilities (Data e-Infrastructures) or computational facilities (Computational e-Infrastructures). Generally, these platforms allow researchers to work together by accessing shared and distributed scientific facilities, including data, instruments, computing and communication. Applications of e-Infrastructure span several domains, e.g., biology, environmental science, etc. E-Infrastructures usually embed the concept of Virtual Research Environments (VREs, [33]), web-based working environments, where groups of scientists have transparent and seamless access to a shared set of remote resources (e.g., data, tools and computing capabilities). VREs are e-Infrastructures' main support to collaboration across institutions, disciplines, and nations. Hybrid Data e-Infrastructures (HDIs) are a particular type of e-Infrastructure assuming that computational, storage and sharing technologies can be integrated to provide elastic access and usage of facilities [34]. HDIs rely on mediator services orchestrated by scalable Information Systems, which register and monitor the available facilities. Additionally, HDIs deliver computing, storage, data and software as-a-Service and organise resources using VREs.

The work of Coro focussed on Cloud computing in the context of HDIs. Since HDIs address user communities, issues in this context are slightly different from those of standard Cloud computing systems. In particular, a Cloud computing platform in an HDI is demanded to (*i*) integrate community-developed algorithms written using different programming languages, (*ii*) integrate external computational platforms providing algorithms as-a-Service, (*iii*) use computational resources from other e-Infrastructures, (*iv*) dynamically use the resources available in the e-Infrastructure (e.g., data, GIS maps, etc.). Furthermore, the Cloud computing platform should provide cross-domain algorithms that may enrich methods commonly used in a certain scientific domain, as well as state-of-the-art algorithms for specific supported communities.

The work by Coro focussed on solving the above issues [31]. To this aim, he contributed to create a system named Statistical Manager [35] for the D4Science e-Infrastructure [36]. This Cloud computing system parallelises data mining procedures provided by a scientific community in the form of compiled software or R scripts. This system is able to correctly manage the dependency of an R script on the version of the R interpreter and to parallelise its execution on several machines. Furthermore, the system requires minimal (or no) alteration of the original code and allows for fast updating of the algorithms when a new version is ready. The Statistical Manager can also integrate external computational facilities that provide algorithms via the Web Processing Service standard. Furthermore, it can use computational resources hosted also by external e-Infrastructures (e.g., the European Grid Infrastructure) as well as by D4Science.

The Statistical Manager was endowed with appealing algorithms, especially for the computational biology communities of practice. In particular, an intelligent taxonomic search workflow [37] was developed to allow biologists to identify the official transcription of a species' scientific name, according to authoritative taxonomic names collections (e.g., GBIF, OBIS, Catalog of Life). Furthermore, algorithms to compare GIS data were developed to allow biologists as well as environmental scientists to numerically evaluate the similarity between two data sets [29]. Finally, the Statistical Manager was endowed with algorithms that were used (and are still used) to assess the health status of fisheries stocks [30]. As a result of this work, biologists communities have expressed their interest in using this platform and in having their algorithms published through the Statistical Manager. Furthermore, this system is also used to teach courses in computational biology and computer engineering. More than 24,000 scientists and students per month currently use the Statistical Manager. Furthermore, this system is being used to produce life history traits information on marine species, which is used by the European Commission to produce regulations and recommendations for fisheries in Europe.

Conclusion

This brief report document the 2015 edition of the ISTI Young Research Award, one of the initiatives promoted by the Istituto di Scienza e Tecnologie dell'Informazione to support the young members of its staff. This is the third edition of the award that started in 2013. In the reality, it continues similar initiatives promoted in the previous years.

YRA goes in tandem with the Grants for Young Mobility (GYM), a program enabling the ISTI staff of less than 34 years old to carry out research in cooperation with foreign Universities and Research Institutions of clear international standing. It complements CNR similar programs.

Both the initiatives are funded through self-taxation of all research laboratories of the Institute thus demonstrating the willingness to incentivise the activity and growth of young re-

---

[12]http://www.doremi-fp7.eu/

searchers. In fact the initiatives will be in place in 2016 also and they are going to be further reinforced by a third initiative aiming at supporting project proposals having principal investigators that are both young and belonging to diverse laboratories.

Aknowledgements

Author contributions

Contributions to the paper are described using the taxonomy described in [38]. Writing the initial draft: LC, AB, GC, LG, DM, FP, GR; Critical review, commentary or revision: LC, AG, CL.

References

[1] Diego Marcheggiani, Oscar Tackstrom, Andrea Esuli, and Fabrizio Sebastiani. Hierarchical multi-label conditional random fields for aspect-oriented opinion mining. In Maarten de Rijke, Tom Kenter, Arjen P. de Vries, ChengXiang Zhai, Franciska de Jong, Kira Radinsky, and Katja Hofmann, editors, *Advances in Information Retrieval, 36th European Conference on IR Research, ECIR 2014, Amsterdam, The Netherlands, April 13-16, 2014. Proceedings*, pages 273–285, 2014.

[2] Diego Marcheggiani and Artières Thierry. An experimental comparison of active learning strategies for partially labeled sequences. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 898–906, 2014.

[3] Diego Marcheggiani and Fabrizio Sebastiani. On the effects of low-quality training data on information extraction from clinical reports. In *Proc. 6th Italian Information Retrieval Workshop (IIR) 2015*, 2014.

[4] Mirco Nanni, Roberto Trasarti, Barbara Furletti, Lorenzo Gabrielli, Peter Van Der Mede andJoost De Bruijn, Erik De Romph, and Gerard Bruil. Transportation planning based on gsm traces: a case study on ivory coast. In Daniel Villatoro Jordi Nin, editor, *Citizen in Sensor Networks, Second International Workshop, CitiSens 2013, Barcelona, Spain, September 19, 2013, Revised Selected Papers*, pages 15–25. Springer International Publishing, 2014.

[5] Mirco Nanni, Roberto Trasarti, Paolo Cintia, Barbara Furletti, Lorenzo Gabrielli, Salvatore Rinzivillo, and Fosca Giannotti. Mobility profiling. In Davy Janssens, Ansar-Ul-Haque Yasar, and Luk Knapen, editors, *Data Science and Simulation in Transportation Research*. IGI Global, 2014.

[6] Lorenzo Gabrielli, Salvatore Rinzivillo, Francesco Ronzano, and Daniel Villatoro. From tweets to semantic trajectories: mining anomalous urban mobility patterns. In *Citizen in Sensor Networks, Second International Workshop, CitiSens 2013, Barcelona, Spain, September 19, 2013, Revised Selected Papers*, pages 26–35. Springer International Publishing, 2014.

[7] Roberto Trasarti, Barbara Furletti, Lorenzo Gabrielli, Mirco Nanni, and Dino Pedreschi. Big data analytics for smart mobility: a case study. In *Workshop Proceedings of the EDBT/ICDT 2014 Joint Conference (March 28, 2014, Athens, Greece)*, pages 363–364. CEUR-WS.org, 2014.

[8] Lorenzo Gabrielli, Barbara Furletti, Fosca Giannotti, Mirco Nanni, and Salvatore Rinzivillo. Use of mobile phone data to estimate visitors mobility flows. In *MoK-MaSD 2014 Modelling and Knowledge Management applications: Systems and Domains*. 2014.

[9] Salvatore Rinzivillo, Lorenzo Gabrielli, Mirco Nanni, Luca Pappalardo, Dino Pedreschi, and Fosca Giannotti. The purpose of motion: learning activities from individual mobility networks. In *The 2014 International Conference on Data Science and Advanced Analytics*. 2014.

[10] Barbara Furletti, Lorenzo Gabrielli, Fosca Giannotti, Letizia Milli, Mirco Nanni, Dino Pedreschi, Roberta Vivio, and Giuseppe Garofalo. Use of mobile phone data to estimate mobility flows. Measuring urban population and inter-city mobility using big data in an integrated approach. In *47th Scientific Meeting of the Italian Statistical Society*. 2014.

[11] Lorenzo Gabrielli, Barbara Furletti, Roberto Trasarti, Fosca Giannotti, and Dino Pedreschi. City users' classification with mobile phone data. In *Big Data (Big Data), 2015 IEEE International Conference on*, pages 1007–1012, Santa Clara (CA) - USA, 11/2015 2015.

[12] Michele Coscia, Giulio Rossetti, Dino Pedreschi, and Fosca Giannotti. Uncovering hierarchical and overlapping communities with a local-first approach. *ACM Transactions on Knowledge Discovery from Data*, 9(1):6–27, 2014.

[13] Diego Pennacchioli, Giulio Rossetti, Luca Pappalardo, Dino Pedreschi, Fosca Giannotti, and Michele Coscia. The patterns of musical influence on the last.fm social network. In *22nd Italian Symposium on Advanced Database Systems*, pages 284–291. 2014.

[14] Giulio Rossetti. *Social Network Dynamics*. PhD thesis, University of Pisa, Italy, 2015.

[15] Alessia Bardi, Paolo Manghi, and Franco Zoppi. Coping with interoperability and sustainability in cultural heritage aggregative data infrastructures. *International Journal of Metadata, Semantics and Ontologies*, 9(2):138–154, 2014.

[16] Alessia Bardi and Paolo Manghi. Enhanced publications: data models and information systems. *Liber Quarterly*, 23(4):240–273, 2014.

[17] Paolo Manghi, Michele Artini, Claudio Atzori, Alessia Bardi, Andrea Mannocci, Sandro La Bruzzo, Leonardo Candela, Donatella Castelli, and Pasquale Pagano. The D-NET software toolkit: a framework for the realization, maintenance, and operation of aggregative infrastructures. *Program*, 48(4):322–354, 2014.

[18] Paolo Manghi, Michele Artini, Alessia Bardi, Claudio Atzori, Sandro La Bruzzo, and Marko Mikulicic. High-performance annotation tagging over solr full-text indexes. *Information Technologies and Libraries*, 33(3):22–44, 2014.

[19] Michele Artini, Claudio Atzori, Alessia Bardi, Sandro La Bruzzo, Paolo Manghi, Marko Mikulicic, and Franco Zoppi. The heritage of the people's europe project: an aggregative data infrastructure for cultural heritage. In *Bridging Between Cultural Heritage Institutions. 9th Italian Research Conference, IRCDL 2013 (Rome, Italy, January 31–February 1, 2013). Revised Selected Papers*. Springer International Publishing, 2014.

[20] Michele Artini, Claudio Atzori, Alessia Bardi, Sandro La Bruzzo, and Paolo Manghi. Tagtick: a tool for annotation tagging over solr indexes. In *2014 IEEE/ACM Joint Conference on Digital Libraries*. IEEE, 2014.

[21] Filippo Palumbo, Jonas Ullberg, Ales Štimec, Francesco Furfari, Lars Karlsson, and Silvia Coradeschi. Sensor network infrastructure for a home care monitoring system. *Sensors*, 14(3):3833–3860, 2014.

[22] Amedeo Cesta, Luca Coraci, Gabriella Cortellessa, Riccardo De Benedictis, Francesco Furfari, Andrea Orlandini, Filippo Palumbo, and Aleš Štimec. From senasor data to user services in giraffplus. In *Ambient Assisted Living*, pages 269–280. Springer, 2014.

[23] Silvia Coradeschi, Amadeo Cesta, Gabriella Cortellessa, Luca Coraci, C Galindo, Javier Gonzalez, Lars Karlsson, Anette Forsberg, Susanne Frennert, Francesco Furfari, et al. Giraffplus: A system for monitoring activities and physiological parameters and promoting social interaction for elderly. In *Human-Computer Systems Interaction: Backgrounds and Applications 3*, pages 261–271. Springer, 2014.

[24] Filippo Palumbo and Paolo Barsocchi. Salt: Source-agnostic localization technique based on context data from binary sensor networks. In *Ambient Intelligence*, pages 17–32. Springer, 2014.

[25] Paolo Barsocchi, Erina Ferro, Luigi Fortunati, Fabio Mavilia, and Filippo Palumbo. Ems@cnr: An energy monitoring sensor network infrastructure for in-building location-based services. In *High Performance Computing & Simulation (HPCS), 2014 International Conference on*, pages 857–862. IEEE, 2014.

[26] Filippo Palumbo, Davide La Rosa, and Stefano Chessa. Gp-m: Mobile middleware infrastructure for ambient assisted living. In *Computers and Communication (ISCC), 2014 IEEE Symposium on*, pages 1–6. IEEE, 2014.

[27] Paolo Barsocchi, Erina Ferro, Filippo Palumbo, and Francesco Potortì. Smart meter led probe for real-time appliance load monitoring. In *SENSORS, 2014 IEEE*, pages 1451–1454. IEEE, 2014.

[28] Leonardo Candela, Donatella Castelli, Gianpaolo Coro, Lucio Lelii, Francesco Mangiacrapa, Valentina Marioli, and Pasquale Pagano. An infrastructure-oriented approach for supporting biodiversity research. *Ecological Informatics*, 2014.

[29] Gianpaolo Coro, Pasquale Pagano, and Anton Ellenbroek. Comparing heterogeneous distribution maps for marine species. *GIScience & Remote Sensing*, 51(4):13–20, 2014.

[30] Rainer Froese, Gianpaolo Coro, Kristin Kleisner, and Nazli Demirel. Revisiting safe biological limits in fisheries. *Fish and Fisheries*, 2014.

[31] Gianpaolo Coro, Leonardo Candela, Pasquale Pagano, Angela Italiano, and Loredana Liccardo. Parallelizing the execution of native data mining algorithms for computational biology. *Concurrency and Computation-Practice & Experience*, 2014.

[32] Leonardo Candela, Gianpaolo Coro, Roberto Cossu, and Pasquale Pagano. Realizing spatial data infrastructure solutions in ENVRI. GEPW-8 - GEO European Projects' Workshop, Athens, Greece, June2014.

[33] Leonardo Candela, Donatella Castelli, and Pasquale Pagano. Virtual research environments: an overview and a research agenda. *Data Science Journal*, 12:GRDI75–GRDI81, 2013.

[34] Leonardo Candela, Donatella Castelli, Andrea Manzi, and Pasquale Pagano. Realising Virtual Research Environments by Hybrid Data Infrastructures: the D4Science Experience. In *International Symposium on Grids and Clouds (ISGC) 2014 23-28 March 2014, Academia Sinica, Taipei, Taiwan, PoS(ISGC2014)022*, Proceedings of Science, 2014.

[35] Leonardo Candela, Donatella Castelli, Gianpaolo Coro, Pasquale Pagano, and Fabio Sinibaldi. Species distribution modeling in the cloud. *Concurrency and Computation: Practice and Experience*, 2013.

[36] Leonardo Candela, Donatella Castelli, and Pasquale Pagano. D4Science: an e-Infrastructure for Supporting Virtual Research Environments. In Maristella Agosti, Floriana Esposito, and Costantino Thanos, editors, *Post-proceedings of the 5th Italian Research Conference on Digital Libraries - IRCDL 2009*, pages 166–169. DELOS: an Association for Digital Libraries, 2009.

[37] Edward Vanden Berghe, Gianpaolo Coro, Nicolas Bailly, Fabio Fiorellato, Caselyn Aldemita, Anton Ellenbroek, and Pasquale Pagano. Retrieving taxa names from large biodiversity data collections using a flexible matching workflow. *Ecological Informatics*, 28:29–41, 2015.

[38] Liz Allen, Jo Scott, Amy Brand, Marjorie Hlava, and Micah Altman. Publishing: Credit where credit is due. *Nature*, pages 312–313, April 2014.