

Curriculum Vitae

Paolo Manghi

May 2, 2019

1 Personal information

Date of birth: 22nd of December, 1970

Place of birth: Firenze

Status: Married

Children: 2

Nationality: Italian

Contacts: *Address:* Via Don Francesco Soldaini 95 - 56023 Cascina (PI), Italia

Office Tel.: +39 050 6212038

Office Fax: +39 050 6213464

E-mail: paolo.manghi@isti.cnr.it

Web: <http://nemis.isti.cnr.it/person/paolo-manghi>

ORCID identifier: 0000-0001-7291-3210

ResearcherID identifier: D-4727-2017

2 Summary

Paolo Manghi is a permanent research staff member (“Ricercatore a Tempo Indeterminato”) at Istituto di Scienza e Tecnologie dell’Informazione (ISTI) of the Consiglio Nazionale delle Ricerche (CNR), Pisa, Italy. He is a member of the InfraScience research group led by Dr. Donatella Castelli, part of the Networked Multimedia Information Systems (NeMIS) laboratory led by Dr. Fausto Rabitti. He is Technical Director of the OpenAIRE infrastructure (<http://www.openaire.eu>), of which he devised and today supervises the underlying services and technologies. OpenAIRE, soon to become a Legal Entity, is the Open Science infrastructure for Europe, funded by the Commission in support of Open Science fostering and monitoring.

Studies and grants He obtained a four-year doctorate in Informatica (Università di Pisa) and two MSc degrees: Laurea in Scienze dell’Informazione (Università di Pisa, 1996) and Laurea Specialistica in Tecnologie Informatiche (Università di Pisa, 2003). His PhD was partly supported by a TMR Marie-Curie Grant at the Universities of Glasgow and Strathclyde (UK).

Scientific carrier Paolo Manghi started his research experience in 1996, with the Fibonacci research group of Prof. Giorgio Ghelli and Prof. Antonio Albano at the Dipartimento di Informatica, Università di Pisa. He completed a four-year doctorate between the University of Pisa and the Universities of Glasgow and Strathclyde in Scotland (UK), partly funded by a European Marie Curie TMR EC Research Training Fellowship grant. After three years in Scotland, from 1997 to 2000, working with Prof. Richard Connor, he returned to the University of Pisa where he re-joined the Fibonacci research group until May 2006. Since May 2006, he joined the InfraScience research group led by Dr. Donatella Castelli, at the Istituto di Scienza e Tecnologie dell'Informazione, CNR. Currently, he has a permanent position as a researcher at CNR.

Scientific interests His curriculum witnesses a research path that started from *programming and typing in “orthogonally” persistent languages*, to apply the same concepts on the Web/Internet, investigate *type correctness of query languages for XML databases* and *type-safe integration of programming languages and XML languages*, and approach *peer-to-peer service architectures for distributed XML databases and Scholarly Communication Systems*. From these experiences on data management, distributed service architectures, and scholarly communication he approached his today's main interests: **ICT infrastructures in support of Open Science publishing** and **ICT infrastructures in support of Open Science scientific process**. The shift towards Open Science, strongly mandated by researchers, organizations, and funders world-wide, has raised a number of open challenges on how to perform and communicate science so as to enable reproducibility and transparent evaluation of science. In this context, two research quests became prominent: (i) how to transparently enable scientists to perform science using complex ICT solutions; and (ii) how to share and package all products of science (e.g. literature, research data, software, experiments) to make science open. Core topics in the area are policies, data models, workflows, services, and standards in support of scientists in the need of sharing, searching, accessing, evaluating research results and getting scientific reward for them.

Scientific Record His scientific record is characterized by the following main and equally important objectives:

Research networks He is today a well-known researcher in the area of information science and scholarly communication, strongly involved by communities, organizations, and funders in finding solutions to open problems but also in the outlining of common avenues at both European and global level; evidence of this is given by 40 invited talks, around 90 article co-authors from Europe and beyond (source: Scopus), memberships in experts groups and advisory boards, leadership roles in research projects, conference/workshop chairing/organization, and active collaborations with important stakeholders in the area, such as DataCite, CrossRef, Scopus, Elsevier, Confederation of Open Access Repositories, CASRAI, Research Data Alliance, World Data System, Jisc-UK, European Open Science Cloud, etc..

Research funding He is (and has been) main author, project coordinator, project technical director, ISTI technical manager, or WP/Task leader for beyond 20 European Commission projects, National projects, and privately funded research projects; he is also responsible for projects of type “Conto Terzi” at ISTI-CNR.

Research and development He has been designing and supervising the realization of several software products, which successfully passed a process of technology transfer and are today adopted as production-ready and operational systems world-wide. Many of these are services and components of the D-NET Software Toolkit [Pr30], which is today powering around ten production systems in Europe and beyond. He is technically responsible of three of these production systems: the OpenAIRE infrastructure (<http://www.openaire.eu>), the European Film Gateway infrastructure (<http://www.europeanfilmgateway.eu>), and the Data-Literature Interlinking Service, also known as ScholeXplorer (<https://scholexplorer.openaire.eu>).

Research production He has been publishing more than a hundred scientific papers, book chapters, books, and technical reports; since 2006 he has built an R&D team of 11 developers and researchers, of which 3 obtained a PhD under his supervision; he is a promoter of Open Science and Open Access to scientific results, including research data and research software.

3 Studies and grants

- 12.12.2003** Seconda Laurea (Degree MSc) *Specialistica in Tecnologie Informatiche*, University of Pisa, final mark 110/110. Certificate at <https://goo.gl/y39oVp>
- 26.10.2001** Four-year Doctorate in *Informatica*, Pisa Consortium, XIIth Series. PhD Thesis: *Extracting Typed Values from Semistructured Databases*; supervisor: Prof. Giorgio Ghelli; referees: Prof. M. Atkinson (Glasgow, UK), Prof. F. Crestani (Glasgow, UK). Certificate at <https://goo.gl/y39oVp> and PhD Thesis (Manghi [2001])
- 31.10.1997-31.10-2000** European TMR Marie Curie Grant Holder at the Department of Computer Science of Glasgow University and at the Department of Computer and Information Sciences of Strathclyde University (UK). EPSRC Project (Advanced Fellowship GR/K 79222 - High Level Internet Programming Systems): *Hippo*, High-level Internet Programming with Persistent Objects, Supervisor Prof. R. Connor, Confirmation letter at <https://goo.gl/y39oVp>
- 19.07.1996** Laurea (Degree MSc) in *Scienze dell'Informazione*, University of Pisa, final mark 110/110. Master Thesis: *Aspetti Linguistici della Costruzione di Applicazioni Persistenti*; supervisor: Prof. Giorgio Ghelli. Certificate at <https://goo.gl/y39oVp>

4 Scientific Career

4.1 Research Group

Paolo Manghi is the head of the D-NET Research Group on “Services for Open Science scholarly communication”, part of the InfraScience Research Group, in turn affiliated to the NeMIS Laboratory at ISTI, CNR. The D-NET group is currently composed by the following research staff: Alessia Bardi, Claudio Atzori, Sandro La Bruzzo, Andrea Dell’Amico, Miriam Baglioni, Michele Artini, Enrico Ottonello, Andrea Mannocci, Michele De Bonis, Emma Lazzeri. Its focus is on studies, research, development and system operation in the area of services for scientific data infrastructures and scholarly communication infrastructures.

4.2 Experience at Consiglio Nazionale delle Ricerche: 2006 - ongoing

- 16.09.2011-ongoing** Research Staff position as *ricercatore a tempo indeterminato III livello*, at ISTI Consiglio Nazionale delle Ricerche, Pisa. Protocollo ISTI N. 2767 del 12/09/2011, copy at <https://goo.gl/y39oVp>
- 01.07.2011-15.09.2011** Research Fellow position as *ricercatore a tempo determinato* (art. 23), at ISTI Consiglio Nazionale delle Ricerche, Pisa. Funding EC Project: OpenAIRE [Pr19]. Protocollo ISTI N. 59314 del 8/08/2011, copy at <https://goo.gl/y39oVp>
- 01.01.2011-30.06.2011** Research Fellow position as *ricercatore a tempo determinato* (art. 23), at ISTI Consiglio Nazionale delle Ricerche, Pisa. Funding EC Project: OpenAIRE [Pr19]. Protocollo ISTI N. 2262 del 12/01/2011, copy at <https://goo.gl/y39oVp>
- 01.07.2010-31.12.2010** Research Fellow position as *ricercatore a tempo determinato* (art. 23), at ISTI Consiglio Nazionale delle Ricerche, Pisa. Funding EC Project: OpenAIRE [Pr19]. Protocollo ISTI N. 59096 del 10/08/2009, copy at <https://goo.gl/y39oVp>
- 01.07.2009-30.06.2010** Research Fellow position as *ricercatore a tempo determinato* (art. 23), at ISTI Consiglio Nazionale delle Ricerche, Pisa. Funding EC Project: DRIVER-II [Pr22]. Protocollo ISTI N. 59096 del 10/08/2009, copy at <https://goo.gl/y39oVp>
- 01.07.2008-30.06.2009** Research Fellow position as *ricercatore a tempo determinato* (art.23), at ISTI Consiglio Nazionale delle Ricerche, Pisa. Funding EC Project: DRIVER-II [Pr22]. Protocollo ISTI N. 1491 del 26/06/2008, copy at <https://goo.gl/y39oVp>
- 01.05.2006-30.06.2008** Research Fellow position as *assegnista di ricerca*, at ISTI, Consiglio Nazionale delle Ricerche, Pisa. Funding EC project: BELIEF [Pr24]. Protocollo ISTI N. 692 del 28/04/2006, copy at <https://goo.gl/y39oVp>

22.02.2006-22.03.2006 Research contract as *prestazione occasionale* at ISTI, Consiglio Nazionale delle Ricerche, Pisa. Funding EC project: BELIEF [Pr24]. Protocollo ISTI N. 235 del 22/02/2006, copy at <https://goo.gl/y39oVp>

4.3 Experience in Academia: 1995 - 2005

November 2003 - University of Pisa Research Fellow position (01.11.2003-31.10.2005), as *assegnista di ricerca*, at Dipartimento di Informatica, University of Pisa, government funds under INF/01.

October 2003 - Ministry of Education Consulting for CRUI Foundation, Ministry of Education (01.10.2003-01.05.2005). Funding project: IT4PS (*Information Technology for Problem Solving*). Definition of lectures, final tests, and drafting of studying books for ECDL (European Computer Driving License) advanced university courses on MS Excel and MS Access, specific to Pharmacy and Medical Faculties.

July 2003 - University of Pisa Research contract (01.07.2003 - 30.09.2003) for the project *FIRB: Enabling Platforms for High-performance Computational Grids Oriented to Scalable Virtual Organizations (grid.it)* on the activity: *design of a distributed XML database over the GRID*. Project coordinator: Prof. M. Vanneschi, Dipartimento di Informatica, University of Pisa.

June 2001 - University of Pisa Research Fellow position (11.06.2001-11.06.2003), as *assegnista di ricerca*, at Dipartimento di Informatica, University of Pisa. Funding project: *Design and implementation of a query language for semistructured-data*, led by Prof. G. Ghelli.

November 2000 - University of Pisa Research contract (29.11.2000-30.05.2001) for the research project *DataX*, coordinated by Prof. G. Ghelli, Dipartimento di Informatica, University of Pisa.

July 1999 - University of Strathclyde Research Fellow position (01.07.1999-31.10.2000), as European TMR Marie Curie Grant Holder, at the Department of Computer and Information Sciences of Strathclyde University (UK). Project *SNAQue* [Pr34], led by Prof. R. Connor, in cooperation with the Dipartimento di Informatica of Pisa University (reference: Prof. G. Ghelli).

October 1997 - University of Glasgow Research Fellow position (31.10.1997-30.06.1999), as European TMR Marie Curie Grant Holder, at the Department of Computer Science of the University of Glasgow (UK). Funding project: *Hippo* [Pr??], led by Prof. R. Connor, in cooperation with the Dipartimento di Informatica of Pisa University (reference: Prof. G. Ghelli).

September 1996 - University of Pisa PhD Student Grant (01.09.1996-31.10.1997) at the Computer Science Department of the University of Pisa (reference: Prof. G. Ghelli).

5 Scientific Interests

Paolo Manghi research activities gravitated around the areas of database systems, query and programming languages, type systems, service architectures and infrastructures, Open Science infrastructures, with particular focus on:

5.1 e-Infrastructures for Open Science (2009-ongoing)

The shift towards Open Science, strongly mandated by researchers, organizations, and funders worldwide, has raised a number of open challenges on how to perform and communicate science so as to enable reproducibility and transparent evaluation of science. Since he moved to CNR he has supervised a team of researchers and developers on the realization of production-ready service infrastructures in support of Open Science, with respect to two main areas:

Transparently enabling scientists to perform science using complex ICT solutions : most scientists need today to use complex ICT technologies (e.g. parallel execution, big data management) but do not have the skills to understand what is required; in addition, research communities often share a set of common tools, software and services to perform their science, which are not integrated and hardly combinable into scientific workflows; on this respect his research has been on (on-demand) Virtual Research Environments, which are high-level web applications capable of

addressing both challenges and facilitating scientists at benefiting from advanced technologies and at integrating their tools in a meaningful way.

Publications: Candela et al. [2017b], Massimiliano Assante and Pagano [2015], Assante et al. [2008], Castelli et al. [2013], Assante et al. [2015]

Projects: PARTHENOS [Pr7], SoBigData [Pr4], eInfraCentral [Pr5], EOSCpilot [Pr9]

Research software: D-NET software toolkit [Pr30]

Sharing products of science to make science reproducible and transparent Open Science publishing opens up a number of challenges, mainly related with the definition of publishing workflows inclusive of all kinds of scientific products (e.g. literature, research data, software, experiments) and the ability to share an encoding of the scientific process. These broad objectives span from the definition of policies, to the revision of reviewing practices, and the upgrade of standards and tools for describing and sharing publications, data, and software, which may differ from discipline to discipline. His work in this area is related with the definition of models for open science publishing and the design of services facilitating scientists at implementing its overarching requirements.

Publications: Burton et al. [2017a], Burton et al. [2017b], Manghi et al. [2016], Bardi and Manghi [2015b], Candela et al. [2015], Manghi et al. [2014b], Manghi [2014], Bardi and Manghi [2014], D. et al. [2012], Candela et al. [2012], Manghi et al. [2010f], Manghi et al. [2008a], Houssos et al. [2014], Kazai et al. [2010a], Tang et al. [2010], Candela et al. [2009], Mannocci and Manghi [2014], Candela et al. [2007c], Candela et al. [2006]

Projects: BELIEF [Pr24], DRIVER-II [Pr22], DRIVER Research Desktop [Pr21], OpenAIREplus [Pr16], OpenAIRE2020 [Pr6], OpenAIRE-Connect [Pr3], RDA Europe series [Pr13][Pr12][Pr8]

Research software: D-NET software toolkit [Pr30], ScholarLynk [Pr31]

5.2 Aggregative Data Infrastructure (2006-ongoing)

The multidisciplinary character of science and the strong requirements for immediate access to new research results called for systems that facilitate researchers at discovering and accessing content available from several kinds of sources, exposing articles, software, research data, or any other research product of interest to scientists. Such novel systems, hereafter aggregative infrastructures, are capable of collecting information objects (i.e. metadata and the files they describe) from heterogeneous sources in order to populate uniform, cleaned, and enriched cross-source information spaces. Numerous challenges arise in this context, functional and non-functional (e.e. scalability, performance) and his team has been working on the definition of tools that can be configured/customized to different domain and combined into data processing workflows to achieve the specific requirements of a community. The D-NET software toolkit [Pr30] is the result of such activities. It is a workflow-oriented and service-oriented platform whose services target a number of data processing functionalities (e.g. aggregation, cleaning, deduplication, enrichment, consumption according to different query languages) and can be combined and configured to match domain specific needs. D-NET is in fact widely used across several projects and production systems in Europe and beyond.

Publications: Michele Artini and Mannocci [2015], Paolo Manghi [2014], Manghi et al. [2014a], Bardi et al. [2014], Schirrwagen et al. [2013], Hoogerwerf et al. [2013], Bardi A. [2013], Bardi et al. [2012b], Manghi et al. [2012b], Manghi et al. [2012d], Manghi et al. [2010b], Manghi et al. [2010d], Artini et al. [2009], Feijen et al. [2007], Mannocci and Manghi [2016], Mannocci et al. [2016], Mannocci et al. [2014], Kobos et al. [2014], La Bruzzo et al. [2013], Artini et al. [2013], Bardi et al. [2012a], Bardi et al. [2012d], Manghi et al. [2012c], Manghi and Mikulicic [2011a], Manghi et al. [2010c], Artini et al. [2008b], Artini et al. [2008a], Candela et al. [2007b], Candela et al. [2007a], Manghi et al. [2010a]

Projects: DRIVER [Pr23], DRIVER-II [Pr22], EFG [Pr20], OpenAIRE [Pr19], OpenAIREplus [Pr16], OpenAIRE2020 [Pr6], OpenAIRE-Connect [Pr3], HOPE [Pr18], EFG1914 [Pr17], S&TDL [Pr14], EAGLE [Pr15]

Research software: D-NET software toolkit [Pr30], MoniQ [Pr25], GDup [Pr26], OAI-zer [Pr27], TagTick [Pr28], PACE [Pr29]

5.3 Peer-to-peer service architectures for distributed XML databases (2003-2007)

He worked on the design and implementation of *XPeer* [Pr32], an XML peer-to-peer (P2P) database. XPeer differs from standard P2P database systems for two main key ideas: it decouples query compilation, demanded to an *overlay network* of administrative nodes, from query execution, which is directly managed by peers; unlike other P2P systems, XPeer requires minimal human intervention and does not require nodes with special administrative duties.

Publications: Conforti et al. [2007b], Sartiani et al. [2005], Conforti et al. [2007a], Conforti et al. [2004]

Projects: GRID-IT (FIRB)

Research software: XPeer [Pr32]

5.4 Type correctness of query languages for XML database (1998-2004)

He worked on the definition of a characterization of query correctness for XML query languages, with the aim of realizing a type system capable to identify at static time all possible structural anomalies of XML queries (Colazzo et al. [2004], Colazzo et al. [2004], Albano et al. [2000]). He experimented such concepts by implementing a compiler over a subset of XQuery, called *microXQuery* and experimenting with new XML query paradigms with the language *TQL* [Pr33] (Albano et al. [2001]).

Publications: Colazzo et al. [2006], Albano et al. [2002], Colazzo et al. [2004], Conforti et al. [2002a], Colazzo et al. [2005], Albano et al. [2000]

Projects: Big Top (Microsoft funding)

Research software: TQL [Pr??]

5.5 Type-safe integration of programming languages and XML languages (1998-2001)

As part of his PhD (Manghi [2001]) he worked on the theoretical design and implementation of a framework (*SNAQue*) for extracting typed programming language (Java) values from XML data and import them type-safely within a run-time computation. Programs can preserve values as XML self-described trees and reload the same values by casting them under the same or compatible types (e.g. subtypes). The resulting framework inspired the definition of novel programming language run-time environments, enabling the realization of “hybrid” applications, where programmers can alternate XQuery-like operations to procedural operations.

Publications: Manghi [2001], Simeoni et al. [2003], Simeoni et al. [2002], Manghi et al. [2002], Connor et al. [2001]

Projects: Marie-Curie TMR grant

Research software: SNAQue [Pr34]

5.6 Programming and typing in “orthogonally” persistent languages (1995-1998)

Orthogonally persistent languages provide run-time support for persistent storage of values of any type, blurring the separation between run-time and persistent values typically stored into files or databases. For his MsC thesis he worked on identifying programming language primitives and methodologies to exploit at best this paradigm (Connor et al. [1996]). This work evolved by developing such results as core concepts of the Web-oriented programming language Hippo (supported by a Marie-Curie grant project), adopting a persistent typed storage distributed across web servers, supporting REST APIs capable of storing and providing access to type values.

Publications: Manghi [1996], Connor et al. [1999], Connor et al. [1996]

Projects: Marie-Curie TMR grant

6 Research Networks

Paolo Manghi has a wide research network, characterized by active collaborations with research organizations and important stakeholders in the ambit of scholarly communication (including scientific publishers, data centres, libraries) and international organizations (including DataCite, CrossRef, DS). Moreover he has acted as referee for conferences and journals, as member of working groups, expert groups, and other

organizations of technical and scientific nature. Finally, he has been invited to present his results, ideas, and vision world-wide.

6.1 Research Collaborations

Among the wide network of collaborations he has been and he is currently involved (as proven by his co-authorships and projects participation) the ones worth mentioning are the following:

- In 2019 he started a collaboration with FREYA project (ref. Martin Fenner), ResearchGraph (ref. Amir Aryani), and ANDS (ref. Adrian Burton) to work on Open Science graphs as co-chairs of an Interest Group in RDA Europe.
- In 2017 he started a collaboration with EUDAT, EGI and IndigoDataCloud to work on a common roadmap towards the European Open Science Cloud.
- In 2014, still ongoing, he started a cooperation with Elsevier (Ref. Hylke Koers and Wouter Haak), DataCite (ref. Martin Fenner), CrossRef (ref. Goeffrey Builder) and PANGAEA data centre to work on standards for facilitating exchange of scholarly links between literature and datasets. As a result in 2016 the Scholix Guidelines have been defined and made public (<http://www.scholix.org/>) and adopted by several stakeholders. Moreover, the ScholeXplorer Service (<http://scholexplorer.openaire.eu>) based on D-NET [Pr30] was launched.
- Since 2013, still ongoing, he started a cooperation with the group of Prof. Carole Goble, University of Manchester, to work on the definition of languages and standards for sharing packages of research products. As a result of this cooperation he is now part of the expert group of ResearchObjects and Prof. Goble is involved in the advisory board of the OpenAIRE-Connect project [Pr3].
- In 2012, still ongoing, he started a cooperation with Jisc UK (ref. Neil Jacobos) and SHARE US (ref. Jeff Spies) on the definition of interoperability standards for subscription and notification services for scholarly communication data sources.
- In 2012 he started a cooperation with the European Grid Infrastructure (EGI, ref. Sergio Andreozzi and Tiziana Ferrari) to work on services to monitor research impact of EGI. The result of this cooperation are visible via the OpenAIRE infrastructure portal.
- In 2009 he started a cooperation with CERN library services (Ref. Tim Smith) on realizing a data model for Zenodo repository (<http://www.zenodo.org>) and integrating its back-ends with OpenAIRE services.
- In 2009 he cooperated with Microsoft Research, departments of Cambridge (ref. Fabrizio Gagliardi, Scarlet Schwiderski-Grosche, Natasa Milic-Frayling and Gabriella Kazai) and Redmond (ref. Alex Wide and Lee Dirks), on integration between Microsoft Desktop and the D-NET-based DRIVER infrastructure [Pr23][Pr22]. The outcome of the cooperation is the project R2D2 [Pr21] and the software tool ScholarLynk (Manghi et al. [2010e]).
- In 2006 he started the DRIVER and DRIVER-II project activities, where he started a flourishing cooperation with the MaDgIK Laboratory of Prof. Yannis Ioannidis, Kapodistrian University of Athens, Dept. of Informatics and Telecommunications, which has led to common work on several projects, including OpenAIRE [Pr19], OpenAIREplus [Pr16], OpenAIRE2020 [Pr6], OpenAIRE-Connect [Pr3].
- In 2005 he cooperated with Redmond Microsoft Research to work on the project *Big Top*. Goal of the cooperation is the integration of *TQL* language within *HighWire*'s programming environment.
- In 2001 he joined the project *Network-Aware Programming and Interoperability* (NAPI), the result of a cooperation between Cambridge Microsoft Research (UK, references: Luca Cardelli and Bruno Quarta) and a selection of Italian Universities (Pisa, Bologna, Firenze, and Milano).

6.2 Conference Organization, Program Committees, and Peer-reviewing

- 2019 Program Committee** OpenRepositories Conference 2019, Co-organizer: Jyrki Ilva. invitation email, 18 December 2018
- 2019 Program Committee** Open Science Fair 2019, Co-organizer: Pedro Principe (Universiti of Minho). Program Committee page at <https://www.opensciencefair.eu/programme/committees-2019>
- 2019 Chair and Organizer** 15th Italian Research Conference on Digital Libraries 2019. Chairs: Paolo Manghi, Leonardo Candela, Gianmaria Silvello. Proceedings (?)
- 2018 Journal Reviewer** Journal of the Association for Information Science and Technology. Contact: Lokman Meho. invitation email, 4 April 2018.
- 2018 Journal Reviewer** EAAI Elsevier. Contact: Salma Begum. invitation email, 8 March 2018.
- 2018 Program Committee** MTSR Conference 2018, Co-Chair: Marios Zervas (Cyprus University of Technology, Cyprus). invitation email, 27 June 2018
- 2018 Program Committee** International Workshop on ResearchObjects, in conjunction with IEEE eScience Conference 2018. Co-Chair: Carole Goble. invitation email, 22 March 2018
- 2018 Program Committee** 7th International Workshop on Mining Scientific Publications 2017. Co-Chair: Dasha Herrmannova. invitation email, 2 January 2018
- 2018 Program Committee** ELPUB Conference 2018, Co-Chair: Pierre Mournier (OpenEdition). invitation email, 8 February 2018
- 2017 Journal Reviewer** Journal of the Association for Information Science and Technology. Contact: Lokman Meho. invitation email, 25 July 2017.
- 2017 Journal Reviewer** Springer Nature's Scientific Data. Contact: Hannah Tippett Simpson. invitation email, 30 August 2017.
- 2017 Program Committee** Italian Research Conference on Digital Libraries 2017. Co-Chair: Costantino Grana. invitation email, 12 September 2016
- 2017 Program Committee** OpenRepositories Conference 2017, Chair: Prof. Sue Hutley (Queensland University of Technology (QUT), Brisbane, Australia). Invitation email, 11 November 2016
- 2017 Program Committee** 6th International Workshop on Mining Scientific Publications 2017. Co-Chair: Petr Knuth. invitation email, 1 May 2017
- 2017 Program Committee** Open Science Fair 2017, Chair: Prof. Yannis Ioannidis (Athena Research and Innovation Centre, Greece). Program Committee page at <https://www.opensciencefair.eu/committees>
- 2016-ongoing Section Editor** Section Editor for the CODATA Data Science Journal since 2016. Editorial team pages at <http://datascience.codata.org/about/editorialteam/>.
- 2016 Organizer and editor** First International Workshop on Reproducible Open Science 2016. Chairs: Paolo Manghi, Oscar Corcho, Amir Aryani, Jochen Shirrwagen. Proceedings (Aryani et al. [2017])
- 2016 Program Committee** Digital Infrastructure for Research 2016, Chair: Sarah Coelho. Program Committee page at <https://www.digitalinfrastructures.eu/content/sponsors-organisers>
- 2016 Program Committee** Open Repositories Conference 2017. Chair: Sue Hutley. invitation email, 11 November 2016
- 2016 Program Committee** IRCDL conference 2017. Chairs: Costantino Grana, Lorenzo Baraldi, Simone Calderara. invitation email 12 September 2016 and Program Committee page at <http://ircdl2017.unimore.it/index.php/people/>
- 2016 Program Committee** 5th International Workshop on Mining Scientific Publications 2016. Chair: Petr Knuth. invitation email, 5 May 2016 and program Committee page at <https://wosp.core.ac.uk/jcdl2016>

- 2016 Program Committee** Workshop on Scholarly Web 2016. Chair: Robert Patton. invitation email, 29 September 2016 and Program Committee at <https://ornlcda.github.io/SWM2017/>
- 2016 Program Committee** Open Repositories Conference 2016. Chair: David Minor. invitation email, 4 February 2016 and Program Committee at <http://or2016.net/about/programme-committee/>
- 2016 Program Committee** Metadata and Semantics Research Conference 2016. Chair: Emmanouel Garoufallou, Imma Subirats Coll, Armando Stellato, Jane Greenberg. Front Matter of Conference Proceedings, ISBN 978-3-319-49156-1
- 2016 Program Committee** IRCDL conference 2016. Chairs: Diego Calvanese, Dario De Nart, Carlo Tasso. Preface of Conference Proceedings, DOI 10.1007/978-3-319-41938-1
- 2015 Journal Reviewer** IEEE-TCDL Bulletin. Contact: Salvatore Silvello. invitation email, 18 August 2015.
- 2015 Program Committee** 4th International Workshop on Mining Scientific Publications 2015. Chair: Petr Knoth. invitation email, 20 March 2015 and program Committee page at <https://wosp.core.ac.uk/jcd12015>
- 2015 Journal Editor** International Journal of Digital Libraries 2015, special issue on Research Data Publishing. mail11September2015,articlespublished,volumestilltobeissued
- 2015 Organizing Committee** Open Research Data Conference 2015: implications for Science and Society. Poland. Chairs: Kevin Ashley, Lukasz Bolikowski, Sarah Callaghan, Paolo Manghi, Lidia Stipiska-Ustasiak, Wojtek Sylwestrzak, Jakub Szprot. Conference web site at <http://ord2015.pon.edu.pl/conference/organisational-committee/>
- 2015 Program Committee** Open Repositories Conference 2015. Chair: Jon Dunn. invitation email, 22 December 2014 and Program Committee at <http://www.or2015.net/program/program-committee/>
- 2015 Program Committee** Metadata and Semantics Research Conference 2015. Chair: Emmanouel Garoufallou, Richard J. Hartley, Panorea Gaitanou. Front Matter of Conference Proceedings, ISBN 978-3-319-24128-9
- 2014 Program Committee** 3rd International Workshop on Mining Scientific Publications. Chair: Petr Knoth. invitation email, 6 May 2014 and program Committee page at <https://wosp.core.ac.uk/d12014/>
- 2014 Program Committee** Scientific and Statistical Database Management Conference 2014. Chair: Christian S. Jensen. Front Matter of Conference Proceedings, ISBN: 978-1-4503-2722-0
- 2014 Program Committee** Open Repositories Conference 2014. Chair: Jon Dunn. invitation email, 11 February 2014
- 2014 Program Committee** Metadata and Semantics Research Conference 2014. Chair: Sissi Closs, Rudi Studer, Emmanouel Garoufallou, Miguel-Angel Sicilia. Front Matter of Conference Proceedings, ISBN 978-3-319-13673-8
- 2014 Program Committee** IRCDL conference 2014. Chair: Maristella Agosti, Tiziana Catarci, Floriana Esposito. Preface of Conference Proceedings, DOI /10.1016/j.procs.2014.10.001
- 2013 Organizer and Editor** Second Workshop on Linking and Contextualizing Publications and Datasets 2013. Proceedings (Bolikowski et al. [2015])
- 2013 Program Committee** Open Repositories Conference 2013. Chair: Jon Dunn. invitation email, 11 March 2013 and program Committee page at http://or2013.net/program_committee/index.html
- 2013 Organization Chair** Metadata and Semantics Research Conference 2013. Editors: Emmanouel Garoufallou and Jane Greenberg. Front Matter of Conference Proceedings, ISBN 978-3-319-03436-2
- 2013 Program Committee** 5th International Conference on Knowledge Engineering and Ontology Development. Program Committee at Conference web site, <http://www.keod.ic3k.org>

- 2013 Program Committee** 2nd International Workshop on Mining Scientific Publications. Chair: Petr Knuth. invitation email, 22 March 2013
- 2013 Organizer and Editor** First Workshop on Linking and Contextualizing Publications and Datasets 2013. Proceedings (Manghi et al. [2014b])
- 2013 Program Committee** World Wide Web Conference 2013, Demo Track. Chair: Antoine Isaac, Daniel Muller, and Bebo White. Lyon, France. Front Matter of ACM Conference Proceedings, ISBN 978-1-4503-2038-2
- 2012 Program Committee** Scientific and Statistical Database Management Conference 2012. Chair: Mike Chatzopoulos. Chania, Crete, Greece. Front Matter of Conference Proceedings, DOI: 0.1007/978-3-642-31235-9
- 2011 Program Committee** Metadata and Semantics Research Conference 2011, OAR Track. Chair: Elena Garcia-Barriocanal (University of Alcalà); co-chairs: Zeynel Cebeci (Aukurova University, Turkey), Mehmet C. Okur, (Yasar University, Turkey). Front Matter of Conference Proceedings, ISBN 978-3-642-24730-9
- 2011 Organizer and Editor** Fourth International Workshop on Very Large Digital Libraries (TPDL 2011), Sept. 2011, Berlin, Germany. Proceedings (Candela et al. [2011a], Candela et al. [2012])
- 2011 Program Committee** BooksOnline2011 Workshop, in conjunction with CIKM Conference 2011. Chair: Gabriella Kazai (Microsoft Research, UK). Front Matter of ACM Workshop Proceedings, ISBN 978-1-4503-0961-5
- 2010 Program Committee** . BooksOnline2010 Workshop, in conjunction with CIKM Conference 2011. Chair: Gabriella Kazai (Microsoft Research, UK). Front Matter of ACM Workshop Proceedings, ISBN 978-1-4503-0377-4
- 2010 Organizer and Editor** Third International Workshop on Very Large Digital Libraries (ECDL 2010), Sept. 2010, Glasgow, UK. Proceedings (Ioannidis et al. [2010])
- 2009 Organizer and Editor** Second International Workshop on Very Large Digital Libraries (ECDL 2009), Sept. 2009, Corfu, Greece Proceedings (Ioannidis et al. [2009b], Ioannidis et al. [2009a], Manghi et al. [2010f])
- 2008 Organizer and Editor** First International Workshop on Very Large Digital Libraries (ECDL 2008), Sept. 2008, Aarhus, Denmark Proceedings (Manghi et al. [2008b], Manghi et al. [2008a])
- 2008 Journal Reviewer** International Journal of Digital Libraries 2008, Editor: Carol Ann Peters. invitation email, 21 October 2008
- 2001 Production Editor** 8th International Workshop on Database Programming Languages (DBPL 2001), Sept. 2001, Rome, Italy. Chair: Prof. Giorgio Ghelli. Front Matter of Springer Proceedings, ISBN 978-3-540-44080-2

6.3 Involvement in international expert groups by invitation and project evaluation

- 2019 Inaugural member of the World Data System Technical Advisory Committee** In April 2019 he was appointed as become an inaugural member of the World Data System (WDS) International Technology Office (ITO) Technical Advisory Committee (TAC) Mail from Karen Payne (Associate Director, International Technology International Science Council, World Data Systems), 27th of Apr, 2019
- 2018 Advisory board member of the Research Graph initiative** In 2018 he was appointed as member of the advisory board for the Research Graph initiative Mail from Peter Vats (Research-Graph.com, Australia), 16th of May, 2018
- 2018 Project evaluator** German Federal Ministry of Education and Research (BMBF), Research Initiative “Computational Life Sciences”. Contact: Ren Eulendorf invitation email, 20 April 2018

- 2017 Advisory board member of the H2020 HIRMEOS project** In 2017 he was appointed as member of the external advisory board for the H2020 project HIRMEOS (grant agreement: 731102) Mail from Dasa Radovic (OpenEdition), 27th of July, 2017
- 2017 Advisory board member of the H2020 OpenMinTeD project** In 2013 he was appointed as member of the external advisory board for the H2020 project OpenMinTeD (grant agreement: 654021) Mail from Stelios Piperidis, 20th of January, 2017
- 2014 Member of the Research Object expert group** Since 2014 is part of the Research Object Advisory Board, initiative led by Prof. Carole Goble (University of Manchester). Advisory board page at <http://www.researchobject.org/about>
- 2014 Member of the GreyNet's Resource Policy Committee** Since October 2014 he is member of the GreyNet's Resource Policy Committee, appointed by its Chair Dr. Dobrica Savic, NIS-IAEA. Letter of recognition in <https://goo.gl/y39oVp>
- 2013 Member of the of EuropeanaCloud** In 2013 he was appointed member of the European Research Coordinators Group of EC FP7 EuropeanaCloud project (grant agreement: 325091) Expert list in deliverable <https://goo.gl/AU6Dr2>
- 2012 European Commission evaluator** He was appointed as **expert and evaluator** for project proposals and hearings by the European Commission for the call: "ICT call 9 - objective 8.2 on ICT for access to cultural resources" (2012). Commission confirmation email 2nd of May, 2012
- 2013 Advisory board member of the FP7 EC CHARMe project** In 2013 he was appointed as member of the external advisory board for the FP7 project CHARMe (grant agreement: 312641) Final report at <https://goo.gl/7GzBtG>
- 2011 Membro del Gruppo di Lavoro CNR "Repository dei prodotti della ricerca"** From 11/08/2011 to 31/03/2016 he was appointed member of the Working Group "repository dei prodotti della ricerca" Protocollo CNR num. 60090 del 11/08/2011
- 2009 Member of the core Expert Group of Europeana** From 2009 to 2011 he joined the Core Expert Group of the Europeana FP7 EC Project (grant agreement: 558001), which aimed at refining and identifying the specifications of the data model and architecture of Europeana. Core expert list in deliverable <https://goo.gl/kZYI7z>
- 2008 Member of the DL.org Expertise Group** From 2008 he became a member of the Content Working Group *DL.org Expertise* an activity of the EU project DL.org (www.dlorg.eu). The working group aims at defining best practices, models and standards on interoperability for Digital Libraries with respect to Content. Core expert list available at <https://goo.gl/3x30nh>

6.4 Invited talks and panels

- 19.06.2019** Invited speaker at OAI11 Conference. Presentation title: "Bridging the World where Science is Performed and the World where Science is Published". Ref: Tony Ross-Hellauer (KNOW Center, Austria) invitation email, 7 March 2019
- 27-28.05.2019** Invited speaker at workshop on Open Science Monitoring and Impact in Ghent. Presentation title: "Unlocking OPen Science Monitoring". Ref: Emilie Hermans (University of Ghent, Brussels) invitation email, 28 March 2019
- 19-2.03.2019** Keynote speaker at Open Science Conference 2019. Presentation title: "The OpenAIRE Research Graph". Ref: Guido Scherp (ZBW) invitation email, 10 January 2019
- 18-21.11.2018** Invited expert at Dagstuhl perspective workshop 18472 "Implementing FAIR Data Infrastructures". Ref: Peter Mutschke (GESIS) invitation email, 24 October 2018
- 08.11.2018** Invited speaker at European Commission DG RTD Library talk on "Traceability of research outputs and their authors. New ways of collecting information using persistent identifiers including DOIs" (Brussels, Belgium). Presentation title: "The OpenAIRE Research Graph". Ref: Christian Cuciniello (EC) invitation email and letter, 28 September 2018

- 17.10.2018** Invited speaker at “Future Research Information Forum 2018” in Seoul (South Korea). Presentation title: “Scholarly Communication trends for Research Data in Europe”. Ref: Dr Choi Hee-yoon (KISTI President) invitation email and letter, 28 September 2018
- 05.10.2018** Invited at the Open Data panel, at the GARR Conference on Data Evolution 2018 (Cagliari). Presentation title: OpenAIRE: fostering Open Science publishing: British Library (London, UK). Ref: Federica Tanlongo (GARR) invitation email, 2 August 2018
- 11.09.2018** Invited speaker at TACC Workshop on “Discovering and Connecting Related Research Objects in the Modern Science Landscape”, in conjunction with International Council for Scientific and Technical Information (ICSTI) Conference 2018. Presentation title: OpenAIRE: fostering Open Science publishing: British Library (London, UK). Chair: Brian Hitson (DOE/OSTI) invitation email, 20 February 2018
- 01.02.2018** Invited speaker at AIUCD conference for the session “Nuove Competenze e Nuovi Servizi per la Ricerca nella European Open Science Cloud: Il Futuro degli Open Research Data nelle Digital Humanities”. Presentation title: OpenAIRE and the European Open Science Cloud, Bari (Italy). Ref: Giorgio Maria Di Nunzio (University of Padua) program available at <http://www.aiucd2018.uniba.it/programme.html>
- 15.03.2018** Invited speaker at the OpenAIRE national workshop for the UK. Presentation title: OpenAIRE - Implementing Open Science, Birmingham (UK). Ref: Frank Manista (Jisc) program and slides available at <https://www.jisc.ac.uk/events/openaire-national-workshop-for-the-uk-15-mar-2018>
- 30.11.2017** Invited speaker at the DI4R conference. Presentation title: OpenAIRE services in support of Open Science as-a-Service, Brussels (Belgium). Ref: Franciska de Jong (CLARIN) program available at <https://indico.egi.eu/indico/event/3455/timetable/#20171130>
- 06.09.2017** Session chair at the Open Science FAIR. Session title: Open Science as-a-Service, Athens (Greece). Ref: Natalia Manola (Athena Research and Innovation Center) program available at <http://www.opensciencefair.eu/workshops/parallel-day-2-1/open-science-as-a-service-tools-for-research>
- 09.05.2017** invited speaker at the event COAR and EUDAT Workshop. Presentation title: Notification Brokers for Repositories, Venice (Italy). Ref: Kathleen Sheerer (COAR) program available at <https://www.coar-repositories.org/community/events/archive/coar-annual-meeting-2017/programme/>, presentation available at <https://zenodo.org/record/580032#.Wr-SQNNub0Q>
- 26-30.10.2016** invited speaker at the event Workshop Design your Infrastructure. Presentation title: The OpenAIRE Catalogue of Services, Krakow (Poland). Ref: Tiziana Ferrari (EGI) program available at <https://indico.egi.eu/indico/event/3025/>
- 30.5.2016** invited speaker at the event OpenAIRE Italian Workshop. Presentation title: OpenAIRE Infrastructure Services: technologies for Open Science, Roma (Italy). Ref: Paola Gargiulo (CINECA) program available at https://www.cineca.it/sites/default/files/NationalOpenAIREWorkshop30-31-052015-COM-ITA_1.pdf
- 6-7.4.2016** invited speaker at the event EGI Forum. Presentation title: New services for repeatability and linking of research objects, Amsterdam. Ref: Tiziana Ferrari (EGI) program available at <https://indico.egi.eu/indico/event/2875/session/21/#20160407>
- 4-6.1.2016** invited speaker at the NSF Workshop on Data and Publication. Presentation title: The Data -Literature Interlinking Service On enabling data-literature links interoperability, Washington, Ref: Matthew Mayernik (National Center for Atmospheric Research, US) program available at <https://library.ucar.edu/datapubs-workshop-2016/agenda>
- 24-25.11.2015** invited speaker at the event e-IRG Workshop. Presentation title: *The OpenAIRE Infrastructure: A Vision towards e-infrastructure Commons*, Luxembourg. Ref: Jan Wiebelitz (University of Hannover) program available at <http://e-irg.eu/documents/10920/290578/e-IRG+Luxembourg+Workshop+report+24-25+Nov+2015+final.pdf>
- 14-16.09.2015** invited speaker at the event Workshop on Reproducible and Citable Data and Models. Presentation title: *The OpenAIRE infrastructure and RDA Data Publishing Working Group: results and vision*, Rostock (Germany). Ref: Carole Goble (University of Manchester) program available at <https://sems.uni-rostock.de/reproducible-and-citable-data-and-models/>

- 6.03.2015** invited speaker at the event Scienza Aperta per una Ricerca Migliore. Presentation title: *Science 2.0 Repositories: Time for a Change in Scholarly Communication*, Roma (Italy). Ref: Giovanni Destrobisoli (University of Roma) program available at <https://sites.google.com/site/scienzaapertaricercamigliore/programma>
- 4.11.2014** invited speaker at the event Dichiarazione di Messina 2.0: la vita italiana all'Accesso Aperto. Presentation title: *Scientific and Technological Challenges towards Open Access to any Research Outcome*, Messina (Italy). Ref: Benedetta Alosi (University of Messina) invitation email 21 October 2014 and program at http://decennale.unime.it/?page_id=588
- 25.09.2014** invited speaker at EGI Forum. Presentation title: *Tracking the Scientific Output enabled by EGI with OpenAIRE*, Amsterdam, The Netherlands. Ref: Sergio Andreozzi (EGI) program available at <https://indico.egi.eu/indico/event/2160/session/7/contribution/65>
- 11.06.2014** invited speaker at the OCLC Research Workshop on "Libraries and Research: Supporting Change/Changing Support". Presentation title: *OpenAIRE: the European Scholarly Communication Infrastructure*, Amsterdam, The Netherlands. Ref: Titia van der Werf (OCLC) program available at <http://www.oclc.org/research/events/2014/06-10.html>, select Agenda Day 1
- 08.04.2014** invited speaker at "Open Science 2020 – Harmonizing Current OA practices with H2020 Guidelines" workshop. Presentation title: *Scientific and Technological Challenges to Open Access for any Research Outcome*, Pisa, Italy. Ref: Donatella Castelli (ISTI-CNR) program available at <https://goo.gl/s8U5Rb>
- 07.04.2014** invited speaker at "Grey Literature and Policy Development: The Pisa Declaration" workshop. Presentation title: *OpenAIRE guidelines for Repositories: PUMA's use-case*, Pisa, Italy. Ref: Stefania Biagioni (ISTI-CNR) program at <http://eventi.isti.cnr.it/index.php/en/programme-grey>
- 16.09.2013** invited speaker at EGI Technical Forum. Presentation title: *On Measuring Research Impact*, Madrid, Spain. Ref: Sergio Andreozzi (EGI) list of contributions at <https://indico.egi.eu/indico/event/1417/contributions>
- 03.09.2013** invited speaker at Scientific Data Sharing Workshop. Presentation title: *The OpenAIRE Scholarly Communication Infrastructure*, Anagni, Roma. Ref: Giovanni Destrobisoli (University of Rome) workshop program at <https://sites.google.com/site/openingsciencetosociety/news/anagni-programme> and post-workshop publication (Destro-Bisol et al. [2014])
- 30.04.2013** invited speaker at PREPARDE Cross-linking Workshop. Presentation title: *The OpenAIRE Scholarly Communication Infrastructure, On Interlinking Datasets, Literature, Fundings, and Research Initiatives*, British Atmospheric Data Centre (STFC), Oxfordshire, United Kingdom. Ref: Sarah Callaghan (BADC) workshop program at <http://proj.badc.rl.ac.uk/preparde/wiki/CrosslinkingWorkshop>
- 03-04.12.2013** invited speaker at e-IRG Workshop on Data Issues. Presentation title: *Data and Publication Linking*, Max Planck Institute, Amsterdam, The Netherlands. Ref: Peter Wittenburg (Max Planck) no evidence available due to lost archives
- 09.07.2012** invited speaker at UK RepositoryNET. Presentation title: *On Enabling Data Infrastructures: the D-NET Software Toolkit*, Edinburgh, UK. Ref: Andrew Dorward (University of Edinburgh) no evidence available due to lost archives
- 20-21.03.2012** invited co-speaker with Yannis Ioannidis at the pre-ICRI workshop, Copenhagen, Denmark. Presentation title: *OpenAIRE Data Architecture*. Ref: Peter Wittenburg (Max Planck) no evidence available due to lost archives
- 17-19.01.2012** invited speaker at the EuroMarine Workshop: Linking Data to Scientific Publications, Bremen, Germany. Presentation title: *Towards solutions for publications, projects, and datasets interlinking*. Ref: Michael Diepenbroek (Pangaea) no evidence available due to lost archives
- 12-13.10.2011** invited speaker at the e-IRG workshop, session "Data infrastructure and data Management", Poznan, Poland. Presentation title: *OpenAIREplus project: Towards an European Open Access infrastructure for Research*. Ref: Norbert Meyer. agenda available at <http://e-irg.eu/workshop-2011-10-programme>

- 11-33.09.2011** invited speaker at the CERIF-CRIS and Repositories workshop (EuroCRIS), Brussels, Belgium. Presentation title: *Introducing CRIS Information in Repository Federations – The OpenAIRE Project Experience*. Ref: Keith Jeffery, Anne Asserson. no evidence available due to lost archives
- 22-23.05.2011** invited speaker at the Integrating Research Information CRIS+OAR workshop (EuroCRIS and CNR), CNR offices, Rome, Italy. Presentation title: *Architectures based on Repositories: The OpenAIRE experience*. Ref: Keith Jeffery, Daniela Luzi, Anne Asserson. no evidence available due to lost archives
- 04.05.2011** invited speaker at the INFISO-RTD Open Access coordination workshop, EC Offices, Brussels, Belgium. Presentation title: *Data acquisition in OpenAIRE*. Ref: Carlos Morais Pires (Head of Scientific Data Infrastructure sector, e-Infrastructures Unit, 7th Framework Programme). no evidence available due to lost archives
- 07.07.2010** invited speaker at Birth of Feather session *Interoperability in Digital Libraries* at the Open Repository conference, Madrid, Spain, 6-9 of July, 2010. Ref: Giuseppina Vullo, University of Glasgow, Scotland. no evidence available due to lost archives no evidence available due to lost archives
- 06.07.2010** invited speaker at Birth of Feather session *Handling of compound objects from an e-Infrastructure point of view* at the Open Repository conference, Madrid, Spain, 6-9 of July, 2010. Ref: Jochen Schirrwagen, University of Bielefeld, Germany. no evidence available due to lost archives
- 18.02.2010** invited speaker at European Science Foundation meeting, ILC-CNR, Pisa. Presentation title: *D-NET Software Toolkit and the DRIVER Infrastructure experience*. Ref: Dr. Andrea Bozzi, director of the Istituto di Linguistica Computazionale, CNR. no evidence available due to lost archives
- 03.09.2009** invited speaker at Società Internazionale per lo Studio del Medioevo Latino (SISMEL), Certosa del Galluzzo, Florence. Presentation title: *Aggregating content from heterogeneous metadata data sources*. no evidence available due to lost archives
- 16-18.03.2009** invited speaker at DRIVER/Jisc Digital Repositories Workshop, Amsterdam, The Netherlands. Presentation title: *Typed Compound Object Models for Digital Libraries*. no evidence available due to lost archives
- 13-14.01.2009** invited speaker at EuropeanaLocal Knowledge Sharing Workshop, Den Haag, The Netherlands. Presentation title: *The DRIVER Infrastructure*. no evidence available due to lost archives
- 16.11.2008** invited speaker at SPARC Digital Repositories Meeting, Baltimore, USA. Presentation title: *Building Sustainable Aggregative Digital Library System*. report email, 19 November 2008
- 02.10.2008** invited speaker at TrebleCLEF Workshop, Wolfsburg, Swiss. Presentation title: *The DRIVER Infrastructure architecture: multilinguality issues*. invitation email, 26 September 2008
- 25.06.2007** invited speaker at SURF Foundation - Utrecht, The Netherlands. Presentation title: *Compound Objects*. no evidence available due to lost archives
- 29.05.2007** invited speaker at the DELOS Network of Excellence School on Digital Libraries, Settignano, Florence, Italy. Presentation title: *The DRIVER Repository Infrastructure*. no evidence available due to lost archives
- 27.05.2005** invited speaker at *BigTop* project Workshop (MS Research), Dipartimento di Informatica, University of Bologna. Presentation title: *XPeer: a P2P XML database system* no evidence available due to lost archives
- 11-12.02.2002** invited speaker at Second Workshop on Network-Aware Programming and Interoperability (NAPI) in San Miniato, Pisa. no evidence available due to lost archives
- 07-09.12.2001** invited speaker at First Workshop on Network-Aware Programming and Interoperability (NAPI) at Microsoft Research (Cambridge, England). no evidence available due to lost archives

07-09.12.1998 invited speaker at Workshop on Persistence (Kinloch Rannoch, Scotland). Presentation title: *Interning Neuer Externed Data*. no evidence available due to lost archives

17-19.05.1998: invited speaker at Pastel Workshop (Fort Williams, Scotland). Presentation title: *On the Unification of Persistent Programming and the World Wide Web*. no evidence available due to lost archives

20-22.08.1996: invited speaker at EC-US Workshop on Persistence (Kinloch Rannoch, Scotland). Presentation title: *Programming methodologies for Persistent Programming Languages*. no evidence available due to lost archives

7 Research Funding

Paolo Manghi was and is heavily involved in the authoring, research, partner technical management, project technical direction, and project coordination activities of 22 projects from the European Commission, National, and enterprises. The research projects he has been involved in are mainly concerned with research and development activities with a strong bias on delivering production-level systems for scientific communities and eventually accomplish technology transfer.

The following sections present the projects where he was directly involved, each accompanied by the project details: name, call, aim, research topics, scientific roles he covered, contract at the time of the project, and funding granted. Projects are categorized in *ongoing* and *closed* and a summary can be found in Table 1. From the Table, in summary, he was:

- **Project Coordinator** for 1 project
- **Technical Manager** for ISTI for 6 projects
- **Technical Director** for 9 projects
- **Work Package or Task Leader** for 5 projects
- **Proposal's main author** for 6 projects
- **Researcher** (no other role) for 6 projects

7.1 Ongoing projects

[Pr1] *OpenAIRE-Advance* (OpenAIRE Advancing Open Scholarship)

Topic: EINFRA-12-2017 - Data and Distributed Computing e-infrastructures for Open Science (grant agreement: 777541, Call: H2020-EINFRA-2017).

Aim: OpenAIRE-Advance continues the mission of OpenAIRE to support the Open Access/Open Data mandates in Europe. By sustaining the current successful infrastructure, comprised of a human network and robust technical services, it consolidates its achievements while working to shift the momentum among its communities to Open Science, aiming to be a trusted e-Infrastructure within the realms of the European Open Science Cloud.

Roles in project: Technical Director and researcher Role specified in project Grant Agreement

Project topics: Open Science publishing, Open Science as a Service, global alignment

Position at ISTI: Ricercatore a Tempo Indeterminato

Funding: 1,100,750 (project total: 10,000,000 EUR)

Duration: January 2018 - December 2020

[Pr2] *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)

Call: CO-CREATION-08-2016-2017 - Better integration of evidence on the impact of research and innovation in policy making (grant agreement: 770531, Call: H2020-SC6-CO-CREATION-2017).

Start Date	End Date	Project	Roles	ISTI-CNR grant	Project grant	Funder
01/2018	12/2020	OpenAIRE-Advance	R, TD, TM	€ 1,100,750	€ 10,000,000	EC-H2020
11/2017	10/2019	Data4Impact	R, TM	€ 199,250	€ 1,499,978	EC-H2020
01/2017	06/2019	eInfraCentral	R	€ 57,935	€ 1,499,037	EC-H2020
09/2015	08/2019	SoBigData	R, WPL, TM	€ 204,000	€ 5,917,500	EC-H2020
01/2017	06/2019	OpenAIRE-Connect	R, PC	€ 496,250	€ 1,997,837	EC-H2020
05/2015	04/2019	PARTHENOS	R, TL, TM	€ 437,219	€ 11,999,711	EC-H2020
01/2017	12/2018	EOSCPilot	R	€ 275,500	€ 9,953,067	EC-H2020
01/2015	06/2018	OpenAIRE2020	R, WPL, TD	€ 947,750	€ 13,132,500	EC-H2020
01/2017	02/2018	RDA Europe 3	R	€ 149,395	€ 3,999,187	EC-H2020
11/2016	11/2017	EFG oper. 2016/17	TD	€ 38,000	€ 38,000	DIF Institute
11/2015	11/2016	EFG oper. 2015/16	TD	€ 47,000	€ 47,000	DIF Institute
10/2014	03/2016	RDA Europe 2	R	€ 64,200	€ 1,169,853	EC-FP7
09/2012	12/2014	RDA Europe	R	€ 103,683	€ 3,317,577	EC-FP7
09/2013	10/2015	S&TDL	R, TM	€ 170,000	-	USID-CNR
04/2013	03/2016	EAGLE	R, TD	€ 222,245	€ 3,749,765	EC-FP7
02/2012	01/2014	EFG1914	R, WPL, TD	€ 586,000	€ 2,299,200	EC-FP7
12/2011	11/2014	OpenAIREplus	R, WPL, TD	€ 586,000	€ 2,299,200	EC-FP7
04/2010	03/2014	HOPE	R, TC	€ 300,000	€ 2,700,000	EC-FP7
12/2009	11/2012	OpenAIRE	R, TD	€ 660,000	€ 4,170,000	EC-FP7
10/2009	09/2010	R2D2	R, TM	€ 100,000	€ 300,000	Microsoft
09/2008	08/2011	EFG	R, TM	€ 500,000	€ 4,500,000	EC-FP7
12/2007	11/2009	DRIVER II	R, WPL, TM	€ 580,000	€ 2,700,000	EC-FP7
06/2006	11/2007	DRIVER	R, WPL, TM	€ 430,000	€ 2,700,000	EC-FP6
11/2005	12/2007	BELIEF	R	€ 220,000	€ 950,000	EC-FP6

Table 1: Research projects: a summary (Legenda: R = Researcher; TD = Technical Director for Project, WPL, TL = Work Package or Task Leader, TM = Technical Manager for ISTI, TC = Technical Coordinator, PC = Project Coordinator); projects in grey are those for which Paolo Manghi was conceiver and main author of the proposal.

Aim: Data4Impact 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. It will do so 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 analysis of these data using topic modelling, machine learning and other techniques aimed at natural language processing.

Roles in project: Researcher WP leader to be found

Project topics: impact of research in society, data aggregation and provision

Position at ISTI: Ricercatore a Tempo Indeterminato

Funding: 199,250 EUR (project total: 1,500,000 EUR)

Duration: November 2017 - October 2019

[Pr3] *OpenAIRE-Connect* (CONNECTing scientific results in support of Open Science)

Call: H2020-EU.1.4.1.3. - Development, deployment and operation of ICT-based e-infrastructures (grant agreement: 731011, Call: H2020-EINFRA-2016-1).

Aim: OpenAIRE-Connect aims to provide technological and social bridges, and deliver services enabling uniform exchange of research artefacts (literature, data, and methods), with semantic links between them, across research communities and content providers in scientific communication. It will introduce and implement the concept of Open Science as a Service (OSaaS) on top of the existing OpenAIRE infrastructure, delivering out-of-the-box, on-demand deployable tools.

Roles in project: Project Coordinator and proposal main author Role specified in project Grant Agreement

Project topics: reproducibility, transparent evaluation, linking, mining

Position at ISTI: Ricercatore a Tempo Indeterminato

Funding: 496,250 EUR (project total: 1,997,837 EUR)

Duration: January 2017 - June 2019

[Pr4] *SoBigData* (SoBigData Research Infrastructure)

Call: H2020-EU.1.4.1.2. - Integrating and opening existing national and regional research infrastructures of European interest (grant agreement: 654024, Call: H2020-INFRAIA-2014-2015).

Aim: SoBigData proposes to create the Social Mining & Big Data Ecosystem: a research infrastructure (RI) providing an integrated ecosystem for ethic-sensitive scientific discoveries and advanced applications of social data mining on the various dimensions of social life, as recorded by “big data”.

Topics: virtual research environments, service integration, workflows

Roles: Researcher and Technical manager for ISTI, WP10 Leader e lettera d’incarico ISTI, del 1/9/2015 available at <https://goo.gl/y39oVp>

Funding: 1,085,000 EUR of which 204,000 EUR to InfraScience group (Total project: 5,917,500 EUR)

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Duration: September 2015 - August 2019

[Pr5] *eInfraCentral* (European E-Infrastructure Services Gateway)

Call: INFRASUPP-03-2016 - Support to policies and international cooperation for e-infrastructures (grant agreement: 731049, Call: H2020-INFRASUPP-2016-1).

Aim: eInfraCentral’s mission is to ensure that by 2020 a broader/ more varied set of users (including industry) benefits from European infrastructures and relative services via a central registry. The challenge is to shift towards an e-Infrastructure ”market place” to initiate new service offerings and to engage with a broader set of users and needs.

Topics: registry, services, research infrastructures

Roles: Researcher role specified in Project Grant Agreement

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 57,935 (project total: 1,499,037 EUR)

Duration: January 2017 - June 2019

7.2 Closed projects

In the following only the major projects he was involved while being at CNR are reported. However, he participated to several others, such as GRID.it (FIRB IT), DataX (IT), EPSRC (UK), Training Movement Grant (EC), NAPI project (Microsoft Research Cambridge).

[Pr6] *OpenAIRE2020* (Open Access Infrastructure for Research in Europe towards 2020)

Call: H2020 e-infrastructure program (grant agreement: 643410, Call: H2020-EINFRA-2014-1).

Aim: OpenAIRE2020 continues and extends OpenAIREs scholarly communication infrastructure to manage and monitor the outcomes of EC-funded research, including datasets and software. It enriches the OpenAIRE platform with services for brokering information to repositories, monitoring usage statistics, data anonymization, de-duplication, and research impact monitoring.

Topics: deduplication, research data, data flow monitoring.

Roles: Proposal main co-author (3 authors), Researcher, WP8 leader, and Technical Officer ; Technical Officer role also specified in Project Grant Agreement; lettera d’incarico ISTI, 19/12/2014 available at <https://goo.gl/y39oVp>

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 947,750 EUR (project total: 13,132,500 EUR)

Duration: January 2015 - June 2018

[Pr7] *PARTHENOS* (Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies)

Call: H2020-EU.1.4.1.1. - Developing new world-class research infrastructures (grant agreement: 654119, Call: H2020-INFRADEV-1-2014-1).

Aim: PARTHENOS aims at strengthening the cohesion of research in the broad sector of Linguistic Studies, Humanities, Cultural Heritage, History, Archaeology and related fields through a thematic cluster of European Research Infrastructures, integrating initiatives, e-infrastructures and other world-class infrastructures, and building bridges between different, although tightly, interrelated fields.

Topics: metadata aggregation, registries, scientific process support

Roles: Researcher, Technical manager for ISTI, and Task leader e lettera d'incarico ISTI, 27/04/2015 available at <https://goo.gl/y39oVp>

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 1,551,625 EUR of which 437,219 EUR to InfraScience group (Project total:11,999,711 EUR)

Duration: May 2015 - April 2019

[Pr8] *RDA Europe* (Research Data Alliance - Europe 3)

Call: H2020 CSA - Coordination and support action (grant agreement: 653194 call: H2020-EINFRA-2014-2)

Aim: This RDA/Europe proposal will mainly build on the successful consortium of the previous years but strengthen its anchoring at policy, scientific and industrial uptake to drive forward RDA Europe objectives that are all grounded in technological excellence and innovation and have clear international ties

Topics: data processing, data management, data infrastructures, global alignment, data-literature interlinking, data citation

Roles: Researcher lettera d'incarico ISTI del 1/9/2015 available at <https://goo.gl/y39oVp>

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 149,395 EUR (project total: 3,999,187 EUR)

Duration: September 2015 - February 2018

[Pr9] *EOSCPilot* (The European Open Science Cloud for Research Pilot Project) **Call:** H2020 RIA - Research and Innovation action (grant agreement: 739563, Call: H2020-INFRADEV-2016-2).

Aim: The EOSCpilot project will support the first phase in the development of the European Open Science Cloud (EOSC) as described in the EC Communication on European Cloud Initiatives [2016, <https://goo.gl/Fp07QD>].

Topics: open science, cloud services, storage, computation

Roles: Researcher;

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato role specified in Project Grant Agreement

Funding: 275,500 (project total: 9,953,067 EUR)

Duration: January 2017 - December 2018

[Pr10] *EFG infrastructure operation 2016-2017*

Call: Contratto ISTI-CNR conto terzi per servizio, contract available at <https://goo.gl/y39oVp>

Aim: The contract covers the effort required to operate and maintain the EFG infrastructure resulting from EFG [Pr20] and EFG1914 [Pr17] EC projects between 2016 and 2017

Topics: data aggregation, data management

Roles: Researcher e Technical Director

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 38,000 EUR

Duration: November 2016 - November 2017

[Pr11] *EFG infrastructure operation 2015-2016*

Call: Contratto ISTI-CNR conto terzi per servizio, contract available at <https://goo.gl/y39oVp>

Aim: The contract covers the effort required to operate and maintain the EFG infrastructure resulting from EFG [Pr20] and EFG1914 [Pr17] EC projects between 2015 and 2016

Topics: data aggregation, data management

Roles: Researcher e Technical Director

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 47,000 EUR

Duration: November 2015 - November 2016

[Pr12] *RDA Europe* (RDA Europe the European plug-in to the global Research Data Alliance (RDA))

Call: FP7 CSA - Coordination and support action (grant agreement: 632756 call: FP7-INFRASTRUCTURES-2013-2)

Aim: RDA Europe - the European plug-in to the global Research Data Alliance ensures that European political, research, industrial and e-infrastructure stakeholders are aware of, engaged with and actively involved in the global RDA activities.

Topics: data processing, data management, data infrastructures, global alignment, data-literature interlinking, data citation

Roles: Researcher

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 64,200 EUR (project total: 1,169,853 EUR)

Duration: October 2014 - March 2016

[Pr13] *RDA Europe* (Research Data Alliance Europe)

Call: FP7 CSA - Coordination and support action (grant agreement: 312424 call: FP7-INFRASTRUCTURES-2012-1)

Aim: RDA Europe aims to be the premier global forum driving convergence between emerging global data infrastructures, with a particular focus on EU-US links. Its prime objective is to establish an EU-US coordination platform to improve the interoperability of today's and tomorrow's scientific data infrastructures on both continents.

Topics: data processing, data management, data infrastructures, global alignment, data-literature interlinking, data citation

Roles: Researcher

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 103,683 EUR (project total: 3,317,577 EUR)

Duration: September 2012 - December 2014

[Pr14] *S&TDL* (Science & Technology Digital Library)

Call: Fondi USID-CNR prot. ISTI 51029 del 3/9/3013

Aim: S&TDL realized the first Italian digital library for science and technology, making available the Italian heritage of literature, datasets, and digitized material of historical and cultural interest, crossed with research funding, people, their skills, and disciplines

Topics: digital library, deduplication, research data and literature interlinking

Roles: Researcher and Technical Manger

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 170,000 EUR (project total:)

Duration: September 2013 - October 2015

[Pr15] *EAGLE* (Electronic Archive of Greek and Latin Epigraphy)

Call: FP7 Best Practice Network (grant agreement: 325122, call: FP7-INFRASTRUCTURES-2012-1).

Aim: The project brings together some of the most prominent European institutions and archives in the field of ancient epigraphy to provide a significant quantity and quality of content for Europeana.

Topics: service infrastructures for the aggregation of Digital Library systems.

Roles: Researcher and Technical director

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 344,480 EUR of which 222,245 EUR to InfraScience group (project total: 3,749,765 EUR)

Duration: April 2013 - March 2016

[Pr16] *OpenAIREplus* (2nd-Generation Open Access Infrastructure for Research in Europe)

Call: FP7 Combination of Collaborative projects and Coordination and Support Actions - Integrated Infrastructures Initiative project (I3) proposal (grant agreement: 283595, call: FP7-INFRASTRUCTURES-2011-2).

Aim: The project continues and extends the OpenAIRE System in the direction of openly share research data and link it to the relative publications.

Topics: service infrastructures for the aggregation of Digital Library systems and Digital Library Management Systems, automatic deposition in remote repositories, collaborative data curation, authority file management, research data.

Roles: Proposal Author, Researcher, and Technical director, and Work Package 6 Leader

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 586,000 EUR (project total: 2,299,200 EUR)

Duration: December 2011 - December 2014.

[Pr17] *EFG1914* (The European Film Gateway 1914)

Call: ICT Policy Support Programme - Pilot Type B (Grant agreement: 292276, Call: CIP-ICT-PSP-2011-5, 2.2. "Digitising content for Europeana").

Aim: EFG1914 will digitise and make available 710 hours of film and 6.800 film-related items on the theme of World War I held by 21 European archives in 15 countries. The content will be made available through the EFG Portal and Europeana and, addition to that, in a special virtual exhibition dedicated to the content digitised in the project and to themes around WWI, film, the European film industries and their audiences in a decade of conflict and cataclysm. The project is supported by ACE and the Europeana Foundation itself.

Topics: service infrastructures for the aggregation of Digital Library systems, data curation, authority file management.

Roles: Researcher, Technical director for the project and WP6 leader

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 137,100 EUR (project total: 4,200,000 EUR)

Duration: February 2012 - January 2014

[Pr18] *HOPE* (Heritage of the People's Europe) (2010-2013)

Call: Best Practice Networks project (grant agreement: 250549 - FP7 EU eContentplus 2009).

Aim: HOPE is a Best Practice Network of archives, libraries and museums of social and labor history institutions across Europe. It aims to improve access to the vast amount of highly significant but scattered digital collections on social history. It proposes to achieve this by promoting the adoption of standards and best practices for digital libraries amongst its partners, by ensuring that the metadata and the content become available through Europeana and by implementing a full scale discovery-to-delivery model. Manghi et al. [2010d], ? Bardi et al. [2012c] Bardi et al. [2012d]

Topics: service infrastructures for the aggregation of Digital Library systems;

Roles: Project Technical Coordinator, Proposal Author, Researcher

Position at ISTI-CNR: Ricercatore a Tempo Indeterminato

Funding: 300,000 EUR (project total: 2,700,000 EUR)

Duration: April 2010 - March 2014.

[Pr19] *OpenAIRE* (Open Access Infrastructure for Research in Europe)

Call: FP7 Combination of Collaborative projects and Coordination and Support Actions - Integrated Infrastructures Initiative project (I3) proposal (grant agreement: 246686, call: FP7-INFRASTRUCTURES-2009-1).

Aim: The project is a key actor in the process of dissemination and uptake of the EU Open Access mandate. It aims at deploying and maintaining the OpenAIRE System, which enables the European Infrastructure for Open Access articles published under fundings from projects in the FP7. OpenAIRE offers a portal (www.openaire.eu) from which authors can deposit their articles, compliant Open Access repositories can register to be pro-actively harvested by the infrastructure, and users can search and access the articles or statistics on articles per project and per research area

Topics: service infrastructures for the aggregation of Digital Library systems and Digital Library Management Systems, automatic deposition in remote repositories, collaborative data curation, authority file management.

Roles: Proposal Author, Researcher, and Technical director

Position at ISTI-CNR: Ricercatore a Tempo Determinato until September 2011, then Ricercatore a Tempo Indeterminato

Funding: 660,000 EUR (project total: 4,170,000 EUR)

Duration: December 2009 - November 2012.

[Pr20] *EFG* (European Film Gateway)

Call: Best Practice Networks project (grant agreement: ECP 517006-EFG, call: FP7 EU eContent^{plus} 2007).

Aim: The project aims at deploying and maintaining a pan-European Film Gateway infrastructure, capable of gathering, aggregating and exposing film information and content available from Movie archives across European countries. The EFG online portal, will provide direct access to about 790,000 digital objects including films, photos, posters, drawings, sound material and text documents. Users will have the possibility to search (in multiple languages) and to browse through the digital objects. EFG aims at supporting technical and semantic interoperability between cinematographic archives and it will support the export to Europeana. EFG will also evaluate measures to be taken to deal with IPR issues.

Research topics: service infrastructures and data models, multi-lingual search, authority file management;

Roles: Proposal Author, Researcher and Technical manager for ISTI

Position at ISTI-CNR: Ricercatore a Tempo Determinato

Funding: 500,000 EUR (project total: 4,500,000 EUR);

Duration: September 2008 - August 2011.

[Pr21] *DRIVER Research Desktop (R2D2)*

Call: research project between Microsoft Corporation, Research Departments of Cambridge and Redmond (ref. Scarlet Schwiderski-Grosche), the InfraScience research group of ISTI-CNR (ref. Donatella Castelli) and the MaDgIK Laboratory (ref. Prof. Yannis Ioannidis) of the Dept. of Informatics and Telecommunications, National and Kapodistrian University of Athens.

Aim: The project aims at the realization of a Microsoft desktop tool called ScholarLynk, capable of enabling collaboration of researchers by exploiting support of service oriented infrastructures (D-NET software [Pr30]) and exploring new avenues in collaborative tagging systems.

Topics: tagging and collaborative systems for very large digital libraries, service infrastructures for the aggregation of Digital Library systems, Digital Library Management Systems.

Roles: Proposal Author, Researcher and Technical manager for ISTI

Position at ISTI-CNR: Ricercatore a Tempo Determinato

Funding: 100,000 EUR (project total 300,000 EUR).

Duration: October 2009 - September 2010.

[Pr22] *DRIVER-II* (the Digital Repository Infrastructure Vision for European Research - Phase Two)

Call: FP7 EU Combination of Collaborative projects and Coordination and support actions project (grant agreement: 212147, call: INFRA-2007-1.2.1 - Scientific Digital Repositories).

Aim: The project goal is to continue the operation of the DRIVER Infrastructure [Pr23] and to extend it with a Digital Library System for the management of “enhanced publications”, special compound objects grouping together articles and research data related with them.

Topics: service infrastructures for the aggregation of Digital Library Systems, Compound Object data models and DLMS.

Roles: Proposal Author, Researcher and Technical director for the project

Position at ISTI-CNR: Assegnista di Ricerca until June 2008, then Ricercatore a Tempo Determinato

Funding: 580,000 EUR (project total: 2,700,000 EUR)

Duration: December 2007 - November 2009

[Pr23] *DRIVER* (the Digital Repository Infrastructure Vision for European Research)

Call: FP6 EC Specific Targeted Research Project (grant agreement: IST-034047, Call: FP6-2005-IST-5 - 2.5.6.3 “Research Networking Testbeds”).

Aim: The project’s applicative goal is that of encouraging Open Access business models among researchers and publishers by giving centralized access and visibility to this critical mass of publications already available from repositories in Europe. To this aim, the project funded the realization and deployment of the D-NET software toolkit and today operates the first European Open Access Repository Infrastructure. D-NET services are combined to form multiple distributed applications over the aggregation of content from a federation of institutional repositories (today 240, for about 2,500,000 digital object surrogates), i.e. archives containing any form of scientific output, including scientific/technical reports, working papers, pre-prints, articles and original research data.

Topics: service infrastructures for the aggregation of Digital Library systems;

Roles: Researcher, Technical manager for ISTI and Work Package 3 Leader

Position at ISTI-CNR: Assegnista di Ricerca

Funding: 430,000 EUR (project total: 2,700,000 EUR)

Duration: June 2006 - November 2007

[Pr24] *BELIEF* (Bringing Europe’s eElectronic Infrastructures to Expanding Frontiers)

Call: FP6 EC Specific Supporting Action project (grant agreement: 026500, Call: FP6-2004-IST-6 - 3.2.3 “Communication Network Development - eInfrastructure Consolidating Initiatives”).

Aim: The aim of the project is to facilitate knowledge-exchange on eInfrastructures a one-stop home for public eInfrastructure documentation, e.g. project DoWs, project deliverables. This information will be readily accessible to BELIEF Community Members through the BELIEF Digital Library especially developed to provide a central repository for eInfrastructure Information.

Topics: Digital Library Systems and methodologies for document collection.

Roles: Researcher;

Position at ISTI-CNR: Contratto per prestazione occasionale until March 2006, then Assegnista di Ricerca

Duration: November 2005 - December 2007

Fundings: 220,000 EUR (project total: 950,000 EUR)

8 Research and Development

Paolo Manghi activities resulted in research software which in several cases reached production-ready quality and is currently being used to operate production systems. He is also technically responsible for some of these production systems, listed below.

8.1 Research software

- [Pr25] *MoniQ* (2015-ongoing). **Roles:** Researcher and Technical director. **Affiliated Institution:** Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa. software available at <https://doi.org/10.5281/zenodo.343903> (Mannocci and Manghi [2017]). *MoniQ* is a software for controlling data flows within complex data infrastructures. The underlying framework captures features that are typical of workflows and others that are typical of data quality and provides the elements to control an expected quality and behaviour over time.
- [Pr26] *GDup* (2015-ongoing). **Roles:** Researcher and Technical director. **Affiliated Institution:** Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa. software available at <https://doi.org/10.5281/zenodo.292980> (Atzori and Manghi [2017]). *GDup* is a powerful suite of tools allowing admin users to manage deduplication within a very large graph-like information space. The tool is developed using Hadoop MapReduce and allows configuration of deduplication algorithms, relationship propagation strategies, and ground truth management and exploitation.
- [Pr27] *OAI-zer* (2013-2014). **Roles:** Researcher and Technical director. **Affiliated Institution:** Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa. Bardi et al. [2012a] software available at <https://doi.org/10.5281/zenodo.292992> (La Bruzzo and Manghi [2017a]). This software offers a suite of tools for relational database administrators to make their databases compatible with OAI-ORE protocol (<https://www.openarchives.org/ore/> with RDF exports). Admins can define relational-RDF mappings with an intuitive user interface and deploy on demand multiple OAI-ORE APIs (one for each ORE view of the database).
- [Pr28] *TagTick* (2012-2014). **Roles:** Researcher and Technical director. **Affiliated Institution:** Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa. software described in Bardi et al. [2012c] and in operation at <http://www.socialhistoryportal.org/collections>. *TagTick* is a wrapper of Solr Full-Text index capable of supporting real-time tagging of the index collection without refreshing its content. It allows admin users to tag and test large portions of the index without going through cumbersome and potentially useless re-indexes of the collection.
- [Pr29] *PACE (Programmable Authority Control Engine)* (2010-ongoing). **Roles:** Researcher and Technical director. **Affiliated Institution:** Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa. software described in Manghi and Mikulicic [2011a]Manghi et al. [2012d]. *PACE* is a stand-alone software devised to maintain aggregative authority files, which are disambiguated lists of records resulting from aggregation of multiple and heterogenous sources.
- [Pr30] *D-NET Software Toolkit* (2007-ongoing). **Roles:** Researcher and Technical director. **Affiliated Institution:** Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa. software available at <https://doi.org/10.5281/zenodo.51385> (Artini et al. [2016]) *D-NET* proposes a service-oriented framework specically designed to support developers at constructing custom aggregative infrastructures in a cost-effective way. *D-NET* offers data management services capable of providing access to different kinds of external data sources, storing and processing (cleaning, deduplication, enrichment) information objects of any data models, converting them into common formats, and exposing information objects to third-party applications through a number of standard access APIs. It is today used as powering platform for several production systems world-wide, including: CEON national repository aggregator of Poland (ref. ICM), La Referencia national repository aggregator of Argentina (ref. MINCYT Minstery), and Recolecta the national repository aggregator of Spain (ref. FECYT Minstery), Turkey national aggregator (ref. Tubitak Minstery), EAGLE project infrastructure [Pr15], HOPE project infrastructure [Pr18], and the services described in Section 8.2. Among *D-NET* services the most relevant are *MoniQ* [Pr25] for data flow monitoring, *GDup* [Pr26] for graph deduplication (Manghi and Mikulicic [2011a], Manghi

and Mikulicic [2011b], Manghi et al. [2012d]), enhanced publication publishing (Bardi and Manghi [2015a], Bardi and Manghi [2014], Bardi and Manghi [2015b]), scholarly communication notification broker services (Michele Artini and Mannocci [2015]), ORE services for relational databases [Pr27] (La Bruzzo et al. [2013]), real-time annotation systems for full-text indices [Pr28] Artini et al. [2014a] Paolo Manghi [2014], and recently *HyWare* (Candela et al. [2017b]), a hybrid workflow language for research infrastructures.

[Pr31] *ScholarLynk* (2009-2011). **Roles:** Researcher and Technical director. **Affiliated Institution:** Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa, National and Kapodistrian University of Athens, Microsoft Research Cambridge. Software described in Manghi et al. [2010e] and available at <https://goo.gl/ApyWfQ>

[Pr32] XPeer: P2P XML Database (2003-2006). **Roles:** Researcher and Technical director. **Affiliated Institution:** Dipartimento di Informatica, Università di Pisa. software described in ?

[Pr33] TQL (2000-2001). **Roles:** Researcher. **Affiliated Institution:** Dipartimento di Informatica, Università di Pisa. software described in Albano et al. [2001] and available at <http://pages.di.unipi.it/ghelli/tql/>

[Pr34] SNAQue (1998-2000). **Roles:** Researcher and Technical director. **Affiliated Institution:** Department of Computer Science, Strathclyde University (UK). software described in Manghi [2001]

8.2 Production systems

He and his teams are technically responsible of the operation and QoS of the following production services:

OpenAIRE infrastructure The OpenAIRE infrastructure (<http://www.openaire.eu>) is both a networking and technological infrastructure whose mission is to advocate and monitor the adoption of the European Commission Open Access mandates, and to evaluate the impact of EC funding and National funders. Its technological infrastructure is powered by D-NET [Pr30] provides services to monitor funders, project research impact, and track Open Access trends in terms of related publications and datasets. To this aim, the services offer functionalities to populate a European (and beyond) graph-like information space that aggregates metadata information about 18 million publications, datasets, organizations, persons, projects, and funders collected, harmonized (semantic and structural cleaning), and deduplicated from around 800 (and growing) online data sources (e.g. publication repositories, dataset repositories, and CRIS systems, journals, publishers). The service takes advantage of cutting edge big data technology to mine a growing knowledge base of PDFs (5 millions today) and infer metadata and scholarly links between objects in the graph.

Adoption: The European Commission, several funders worldwide (e.g. European Commission, Wellcome Trust, Fundacao para a Ciencia e a Tecnologia, Australian Research Council, National Science Foundation), and project coordinators are using its services to measure the impact of their investments and to report in the results of projects.

Funding: The OpenAIRE infrastructure is the result of a number of funding projects: OpenAIRE, OpenAIREplus, OpenAIRE2020 and OpenAIRE-Connect. It will become a Legal Entity in 2017.

EFG infrastructure Film archives, containing collections of cinema-related digital material, have been created in many European countries. Today, the European Film Gateway infrastructure (powered by D-NET [Pr30]) provides a single access point to 59 collections from 19 archives and across 14 European countries, for a total of 640,000 digital objects (<http://www.europeanfilmgateway.eu>).

Adoption: The Service is used today by scientists in the field to share their products with the broader European audience and to easily find the products they need via a unique entry point to the heritage of film research.

Funding: The initiative is and has been funded by the EC projects EFG [Pr20] and EFG1914 [Pr17] and today is being sustained by the DFI institute via “conto terzi” projects [Pr11][Pr10].

Data-Literature interlinking service The Data Literature Interlinking service (<http://dliservice.research-infrastructures.eu>) is powered by D-NET [Pr30] populates and provides access to a graph of 8 million dataset-publication and dataset-dataset links (assertions) aggregated from 20 data sources (and growing). The links are collected from data sources managed by publishers (e.g.

Elsevier, Springer Nature, IEEE), data centers (e.g. PANGAEA, CCDC, ICPSR), or other organizations providing services to manage links between datasets and publications (e.g. DataCite, OpenAIRE, Thomson Reuters), to be then harmonized, resolved, and deduplicated.

Adoption: Today the Service's API are used by Scopus (Thomson Reuters), some libraries, the RD-Switchboard service (<http://www.rd-switchboard.org/>, ANDS), and OpenAIRE itself to provide extra information on datasets related with articles and vice versa.

Funding: The initiative is and has been funded by OpenAIRE project series [Pr16][Pr6] and by Research Data Alliance projects [Pr13][Pr12][Pr8].

9 Other activities

Paolo Manghi had a regular teaching activity, which since 2006 he had to reduce due to growing research and project coordination responsibilities. His major experience is in teaching database design principles, but also other topics were touched, such as architectures and programming languages and methodologies. Teaching has naturally evolved in the years into supervision of MSc students and PhD students.

Moreover, he has been member or president of 16 personnel selection boards (commissioni per la selezione del personale).

9.1 Post-doc supervision (relatore)

PhD supervision He has supervised three PhD students:

Alessia Bardi, 2016 *topic:* enabling infrastructures for enhanced publications PhD thesis at <https://etd.adm.unipi.it/theses/available/etd-05062016-113902/>

Claudio Atzori, 2016 *topic:* de-duplication and relationship propagation of objects in graphs of entities PhD thesis at <https://etd.adm.unipi.it/theses/available/etd-05092016-090250/>

Andrea Mannocci, 2017 *topic:* monitoring data flows over existing data infrastructures PhD thesis at <http://etd.adm.unipi.it/theses/available/etd-12232016-151401/>

9.2 MsC supervision (relatore)

Laurea specialistica 2009-2010 - MSc thesis Sandro La Bruzzo, Dipartimento di Informatica, Università di Pisa: *Implementazione di interfacce OAI per l'esportazione di Compound Objects nel sistema DOROTY*; work on design and implementation of standard data-export interfaces for the typed data model of the DOROTY DLMS (for more information see research activities section of this CV). MsC Thesis at <https://etd.adm.unipi.it/theses/available/etd-06252010-124050/>

Laurea specialistica 2008-2009 - MSc thesis Alessia Bardi, Dipartimento di Informatica, Università di Pisa: *DOROTY: Digital Object Repository with Types* MsC Thesis at <https://etd.adm.unipi.it/theses/available/etd-03092009-110518/>

MsC external evaluator MsC dissertation examination, candidate student: Lebeko Bernard Nkoebele Poulo, supervisor Hussain Suleman, University of Cape Twon invitation email 29 April 2013

9.3 University, Master, and Post-Doc lecturer

October 2010 Lecture: *Content Interoperability in Digital Libraries* at the *DL.org Autumn School*, National and Kapodistrian University of Athens, Athens, Greece, 3-8 of October, 2010. Duration: 2 hours.

December 2001 Course: *Applicazioni Web per Basi di Dati*, Web & Net Master. Organizers: Pisa Province Council, University of Pisa (reference. Dr. Laura Ricci), and Qualital Consortium. Duration: 2 weeks.

June 2009 Module *D-NET Technology* of the DRIVER Tutorial at the CERN workshop on Innovations in Scholarly Communication, OAI6 Conference, University of Geneva, Geneva, 16-19 of June, 2009. Duration: 1 day (17.06.2009).

April 2009 Course: *On Teaching Databases*, SSIS in Informatica (Teachers School of Specialization), Dipartimento di Informatica di Pisa. Duration: 2 months (09.04-06.06.2009).

August 2008 Module *Put Yourself in the DRIVER's Seat - Practical Training for Building a European Repository Network* at the Summer School *Digital Libraries la Carte 2008*, Tilburg University, The Netherlands. Duration: 1 day (28.08.2008).

October 2007 Course: *On Teaching Databases*, SSIS in Informatica (Teachers School of Specialization), Dipartimento di Informatica di Pisa. Duration: 1 month.

September 2007 Course: Master "Open Source". Seminars on *SQL Data Definition Language*. Dipartimento di Informatica, University of Pisa. Duration: 3 days (14.09, 19.10, 11.11.2007).

September 2006 Course: Master "Open Source". Seminars: *Il ruolo della progettazione nello sviluppo delle Basi di Dati e Il Linguaggio SQL*. Dipartimento di Informatica, University of Pisa. Duration: 2 days (08.09.2006, 15.09.2006).

March 2005 Course: Master "Open Source". Seminars: *Le Interrogazioni SQL e Ambienti per la definizione delle interrogazioni*. Dipartimento di Informatica, University of Pisa. Duration: 2 days (11.03.2005, 16.03.2005).

December 2004 Course: *IT4PS - Access*. Organizers: Dipartimento di patologia sperimentale biotecnologie mediche, infettivologia e epidemiologia, University of Pisa, CRUI (Conferenza dei Rettori delle Università Italiane), and AICA (Associazione Italiana per l'Informatica e il Calcolo Automatico). Duration: 2 weeks.

May 2004 Course: *Database, MASTER SIT*. Polo Scientifico e Tecnologico dell'Area Livornese Srl, Livorno. Duration: 2 weeks.

January 2002 Course: *Basi di Dati*, IFTS course *Tecnico Informatico per il supporto al commercio elettronico e ai servizi informativi territoriali*. Organizers: Regione Toscana and University of Pisa (reference: Prof. Dino Pedreschi). Duration: semester.

December 2000 Course: *Basi di Dati*, Master Web & Wireless. Organizers: University of Pisa and Vodafone (ex Omnitel). Duration: 1 day.

9.4 University teaching assistant (contratti di supporto alla didattica)

April 2006 Course: *General concepts of informatics (Informatica II)*, Corso di Laurea in Comunicazione Pubblica Sociale e Impresa, Facoltà di Lettere, University of Pisa. Duration: semester (from 05.04.2006).

September 2005 Course: *Introduction to programming - Java laboratory*, Laurea in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 29.09.2005).

March 2004 Course: *Databases - Oracle platform laboratory*, Laurea in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 01.03.2004).

October 2003 Course: *Introduction to programming - Java laboratory*, Laurea in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 20.10.2003).

March 2003 Course: corso *Databases - Oracle platform laboratory*, Laurea in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 01.03.2003).

October 2002 Course: *Introduction to programming - Java laboratory*, Laurea in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 16.10.2002).

March 2002 Course: *Databases - Oracle platform laboratory*, Laurea in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 18.03.2002).

March 2001 Course: *Databases - Oracle platform laboratory*, Laurea in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 12.03.2001).

October 2000 Course: *Architetture degli Elaboratori I - Operative Systems*, Diploma in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 23.10.2000).

March 2000 Course: *Databases - Oracle platform laboratory*, Laurea in Informatica, Dipartimento di Informatica, University of Pisa. Duration: semester (from 20.03.2000).

9.5 Member of CNR commissions for public selection of personnel

2019.04 Membro di commissione esaminatrice di pubblica selezione' Selection for 1 senior research grant ("assegno di ricerca professionalizzante"). bandi ISTI 05/2019, protocollo n. 1062 del 05/04/2019

2019.03 Presidente di commissione esaminatrice di pubblica selezione' Selection for 2 senior research grant ("assegno post-dottorale"). bandi ISTI 03/2019, protocollo n. 936 del 26/03/2019

2018.11 Presidente di commissione esaminatrice di pubblica selezione' Selection for 2 senior research grant ("assegno di ricerca professionalizzante"). bandi ISTI 018/2018, prot. n. 4192 del 19/11/2018

2018.09 Presidente di commissione esaminatrice di pubblica selezione Selection for 1 position as temporary research fellow art. 23 ("contratto a tempo determinato"). selezione 380.1 ISTI RIC, Protocollo n. 3434 del 26/09/2018

2018.04 Presidente di commissione esaminatrice di pubblica selezione Selection for 1 position as temporary research fellow art. 23 ("contratto a tempo determinato"). selezione 02/2018 ISTI, Protocollo n. 1713 del 27/04/2018

2017.11 Presidente di commissione esaminatrice di pubblica selezione Selection for 1 position as temporary research fellow art. 23 ("contratto a tempo determinato") to work on the project OpenAIRE2020. selezione 12/2017 ISTI, 14.11.2017

2017.07 Membro di commissione esaminatrice di pubblica selezione' Selection for 1 position as Art. 15 contract bandi ISTI 006/2017, 31.07.2017

2017.06 Membro di commissione esaminatrice di pubblica selezione' Selection for 1 senior research grant ("assegno di ricerca professionalizzante") to work on the project CURIOSITY DRIVEN - NeMIS. bandi ISTI 008/2017, 30.06.2017

2017.05 Membro di commissione esaminatrice di pubblica selezione' Selection for 1 position as temporary research fellow art. 15 ("CTER"). bandi ISTI 006/2017, Protocollo n. 3565 del 31/07/2017

2017.05 Presidente di commissione esaminatrice di pubblica selezione' Selection for 2 senior research grant ("assegno di ricerca professionalizzante") to work on the project CURIOSITY DRIVEN - NeMIS. bandi ISTI 005/2017, 16.05.2017

2016.10 Presidente di commissione esaminatrice di pubblica selezione' Selection for 1 position as temporary research fellow ("assegnista di ricerca") to work on the project CURIOSITY-DRIVEN. prot. CNR-ISTI 3998, 14.10.2016

2016.09 Presidente di commissione esaminatrice di pubblica selezione' Selection for 1 position as Art. 15 contract. prot. CNR-ISTI 3239, 01.09.2016

2016.08 Membro di commissione esaminatrice di pubblica selezione Selection for 1 position as temporary research fellow ("assegnista di ricerca") to work on the projects EGI-ENGAGE and ENVRI-Plus. prot. CNR-ISTI 3216, 28.08.2015

2015.04 Presidente di commissione esaminatrice di pubblica selezione Selection for 1 position as Art. 15 contract to work on the project OpenAIRE2020. prot. CNR-ISTI1742, 22.04.2015

2015.04 Presidente di commissione esaminatrice di pubblica selezione Selection for 1 position as temporary research fellow ("contratto a tempo determinato") to work on the project OpenAIRE2020. prot. CNR-ISTI 1604, 13.04.2015

2015.03 Membro di commissione esaminatrice di pubblica selezione Selection for 1 research grant ("assegnista di ricerca") to work on the project OpenAIRE2020. prot. CNR-ISTI 1047, 4.03.2015

- 2014.08 Membro di commissione esaminatrice di pubblica selezione** Selection for 1 research grant (“assegnista di ricerca”). prot. CNR-IFC 0006843, 19.08.2014
- 2014.08 Membro di commissione esaminatrice di pubblica selezione** Selection for 1 research grant (“assegnista di ricerca”). prot. CNR-IFC 0006842, 19.08.2014
- 2014.01 Membro di commissione esaminatrice di pubblica selezione** Selection for 1 position as Art. 15 contract to work on the project Science and Technology Digital Library. prot. CNR-ISTI 0093, 14.01.2014
- 2013.10 Membro di commissione esaminatrice di pubblica selezione** Selection for 2 research grants (“assegnista di ricerca”) to work on the project Science and Technology Digital Library. prot. CNR-ISTI 3003, 29.10.2013
- 2012.11 Membro di commissione esaminatrice di pubblica selezione** Selection for 1 position as Art. 15 contract to work on the project OpenAIREplus. prot. CNR-ISTI 3383, 30.11.2012
- 2012.11 Membro di commissione esaminatrice di pubblica selezione** Selection for 1 research grant (“assegnista di ricerca”) to work on the project iMarine, ENVRI, EAGLE, EUBrasil-OpenBio. prot. CNR-ISTI 3098, 9.11.2012
- 2012.10 Membro di commissione esaminatrice di pubblica selezione** Selection for 1 research grant (“assegnista di ricerca”) to work on the project OpenAIREplus, HOPE, and EFG1914. prot. CNR-ISTI 2717, 9.10.2012
- 2012.07 Membro di commissione esaminatrice di pubblica selezione** Selection for 2 research grants (“assegnista di ricerca”) to work on the project OpenAIREplus and EFG1914. prot. CNR-ISTI 2135, 19.07.2012
- 2010.09 Membro di commissione esaminatrice di pubblica selezione** Selection for 1 research grant (“assegnista di ricerca”) to work on the project D4Science-II. prot. CNR-ISTI 2210, 28.09.2010
- 2009.04 Membro di commissione esaminatrice di pubblica selezione** Selection for 1 research grant (“assegnista di ricerca”) to work on the project DRIVER-II. prot. CNR-ISTI 953, 23.04.2009

10 Scientific Publications

Paolo Manghi has published around 140 products (including articles, books, chapters, demos, short papers, software, etc.) and counts around 85 co-authors (source: Scopus).

10.1 Thesis and dissertations

- Manghi [2001] Paolo Manghi. *Extracting Typed Values from Semistructured Databases*. PhD thesis, Dipartimento di Informatica, University of Pisa, 2001. Supervisors: Giorgio Ghelli and Richard Connor
- Manghi [1996] Paolo Manghi. *Aspetti Linguistici della Costruzione di Applicazioni Persistenti*. Master’s thesis, Dipartimento di Informatica Università di Pisa, July 1996

10.2 Journals

- Culina et al. [2018] Antica Culina, Miriam Baglioni, Tom W. Crowther, Marcel E. Visser, Saskia Woutersen-Windhouwer, and Paolo Manghi. Navigating the unfolding open data landscape in ecology and evolution. *Nature Ecology & Evolution*, 2(3):420–426, 2018. ISSN 2397-334X. doi:10.1038/s41559-017-0458-2. URL <https://doi.org/10.1038/s41559-017-0458-2>
- Burton et al. [2017a] Adrian Burton, Hylke Koers, Paolo Manghi, Sandro La Bruzzo, Amir Aryani, Michael Diepenbroek, and Uwe Schindler. The data-literature interlinking service: towards a common infrastructure for sharing data-article links. *Program*, 51(1):null, 2017a. doi:10.1108/PROG-06-2016-0048. URL <http://www.emeraldinsight.com/doi/abs/10.1108/PROG-06-2016-0048>

- Burton et al. [2017b] Adrian Burton, Hylke Koers, Paolo Manghi, Markus Stocker, Martin Fenner, Amir Aryani, Sandro La Bruzzo, Michael Diepenbroek, and Uwe Schindler. The scholix framework for interoperability in data-literature information exchange. *D-Lib Magazine*, 23(1/2), 2017b. doi:doi.org/10.1045/january2017-burton
- Candela et al. [2017b] Leonardo Candela, Fosca Giannotti, Valerio Grossi, Paolo Manghi, and Roberto Trasarti. Hyware: a hybrid workflow language for research e-infrastructures. *D-Lib Magazine*, 23(1/2), 2017b. doi:doi.org/10.1045/january2017-candela
- Manghi et al. [2016] Paolo Manghi, Jochen Schirrwagen, Oscar Corcho, and Amir Aryani. Report on the First International Workshop on Reproducible Open Science. *SIGMOD Record*, 45(4), December 2016
- Michele Artini and Mannocci [2015] Alessia Bardi Sandro La Bruzzo Paolo Manghi Michele Artini, Claudio Atzori and Andrea Mannocci. The openaire literature broker service for institutional repositories. *D-Lib Magazine*, 21(11/12), November/December 2015. doi:DOI: 10.1045/november2015-artini
- Bardi and Manghi [2015b] Alessia Bardi and Paolo Manghi. A framework supporting the shift from traditional digital publications to enhanced publications. *D-Lib Magazine*, 21(1/2), January/February 2015b. doi:DOI: 10.1045/january2015-bardi
- Massimiliano Assante and Pagano [2015] Donatella Castelli Paolo Manghi Massimiliano Assante, Leonardo Candela and Pasquale Pagano. Science 2.0 repositories: Time for a change in scholarly communication. *D-Lib Magazine*, 21(1/2), January/February 2015. doi:DOI: 10.1045/january2015-assante
- Candela et al. [2015] Leonardo Candela, Donatella Castelli, Paolo Manghi, and Alice Tani. Data journals: A survey. *Journal of the Association for Information Science and Technology*, 66(9): 1747–1762, 2015. ISSN 2330-1643. doi:10.1002/asi.23358. URL <http://dx.doi.org/10.1002/asi.23358>
- Paolo Manghi [2014] Alessia Bardi Claudio Atzori Sandro La Bruzzo Marko Mikulicic Paolo Manghi, Michele Artini. High-performance annotation tagging over solr full-text indexes. *Information Technology and Libraries*, 33(3):22–44, 2014. doi:10.6017/ital.v33i3.4633
- Manghi et al. [2014b] Paolo Manghi, Lukasz Bolikowski, Nikos Houssos, and Jochen Schirrwagen. Report on the First Workshop on Linking and Contextualizing Publications and Datasets. *SIGMOD Record*, 43(2), June 2014b
- Manghi [2014] Paolo Manghi. Sfide tecnologiche per l’accesso aperto a tutti i prodotti della ricerca. *Bibliotime*, XVII(3), November 2014
- Destro-Bisol et al. [2014] Giovanni Destro-Bisol, Paolo Anagnostou, Marco Capocasa, Silvia Benicivelli, Andrea Cerroni, Jorge Contreras, Neela Enke, Bernardino Fantini, Pietro Greco, Catherine Heeney, Daniela Luzi, Paolo Manghi, Deborah Mascalzoni, Jennifer C. Molloy, Fabio Parenti, Jelte M. Wicherts, and Geoffrey Boulton. Perspectives on open science and scientific data sharing: an interdisciplinary workshop. *Journal of Anthropological Sciences*, 92:179–200, 2014. doi:10.4436/JASS.92006
- Manghi et al. [2014a] 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, 2014a. doi:10.1108/PROG-08-2013-0045. URL <http://www.emeraldinsight.com/doi/abs/10.1108/PROG-08-2013-0045>
- Bardi et al. [2014] Alessia Bardi, Paolo Manghi, and Franco Zoppi. Coping with interoperability and sustainability in cultural heritage aggregative data infrastructures. *Int. J. of Metadata, Semantics and Ontologies*, 9(2):138 – 154, 2014
- Bardi and Manghi [2014] Alessia Bardi and Paolo Manghi. Enhanced publications: Data models and information systems. *LIBER Quarterly*, 23(4), 2014. ISSN 2213-056X. URL <http://doi.org/10.18352/lq.8445>

- Schirrwagen et al. [2013] Jochen Schirrwagen, Paolo Manghi, Natalia Manola, Lukasz Bolikowski, Najla Rettberg, and Birgit Schmidt. Data curation in the openaire scholarly communication infrastructure. *Information Standards Quarterly*, 25(3):13–19, Fall 2013
- Hoogerwerf et al. [2013] Maarten Hoogerwerf, Mathias Lösch, Jochen Schirrwagen, Sarah Callaghan, Paolo Manghi, Katerina Iatropoulou, Dimitra Keramida, and Najla Rettberg. Linking data and publications: Towards a cross-disciplinary approach. *International Journal of Digital Curation*, 8(1):244–254, June 2013. doi:doi:10.2218/ijdc.v8i1.257
- Castelli et al. [2013] Donatella Castelli, Paolo Manghi, and Costantino Thanos. A vision towards scientific communication infrastructures. *International Journal on Digital Libraries*, 13(3-4):155–169, 2013. ISSN 1432-5012. doi:10.1007/s00799-013-0106-7. URL <http://dx.doi.org/10.1007/s00799-013-0106-7>
- Bardi A. [2013] Manghi P. Bardi A., La Bruzzo S. Oaizer : customized oai-ore and oai-pmh exports of compound objects for the fedora repository. *The Grey Journal (TGJ): An international journal on grey literature*, 9(1):40 – 47, 2013
- D. et al. [2012] Castelli D., Manghi P., and Thanos C. Comunicazione scientifica: la nuova frontiera. *DigItalia WEB. Rivista digitale dei beni culturali*, 2:29 – 39, 2012
- Bardi et al. [2012b] A. Bardi, P. Manghi, and F. Zoppi. Interoperabilità e armonizzazione nelle biblioteche digitali: l’esperienza del progetto European Film Gateway. *DigItalia*, 2:132 – 142, 2012b
- Manghi et al. [2012b] Paolo Manghi, Lukasz Bolikowski, Natalia Manola, Jochen Schirrwagen, and Tim Smith. Openaireplus: the european scholarly communication data infrastructure. *D-Lib Magazine*, 18(9-10), September October 2012b. doi:10.1045/september2012-manghi
- Manghi et al. [2012d] Paolo Manghi, Marko Mikulicic, and Claudio Atzori. De-duplication of aggregation authority files. *Int. J. of Metadata, Semantics and Ontologies*, 7(2):114 – 130, 2012d. doi:10.1504/IJMSO.2012.050014
- Candela et al. [2012] Leonardo Candela, Paolo Manghi, and Yannis Ioannidis. Fourth workshop on very large digital libraries: On the marriage between very large digital libraries and very large data archives. *SIGMOD Record*, 40(4):61–64, January 2012. ISSN 0163-5808. doi:10.1145/2094114.2094130. URL <http://doi.acm.org/10.1145/2094114.2094130>
- Manghi et al. [2010b] Paolo Manghi, Natalia Manola, Wolfram Horstmann, and Dale Peters. An Infrastructure for Managing EC Funded Research Output – The OpenAIRE Project. *International Journal on Grey Literature (TGJ)*, 6(1), Spring 2010b. ISSN 1574-1796
- Manghi et al. [2010d] Paolo Manghi, Marko Mikulicic, Leonardo Candela, Donatella Castelli, and Pasquale Pagano. Realizing and Maintaining Aggregative Digital Library Systems: D-NET Software Toolkit and OAIster System. *D-Lib Magazine*, 16(3/4), March/April 2010d. ISSN 1082-9873. doi:doi:10.1045/march2010-manghi
- Artini et al. [2009] Michele Artini, Leonardo Candela, Donatella Castelli, Paolo Manghi, Marko Mikulicic, and Pasquale Pagano. Aggregative Digital Library Systems in the DRIVER Infrastructure. *World Digital Libraries Journal*, 2(2):113–130, December 2009. ISSN ISSN 0974-567X
- Manghi et al. [2010f] Paolo Manghi, Pasquale Pagano, and Yannis Ioannidis. Second workshop on very large digital libraries: In conjunction with the european conference on digital libraries corfu, greece, 2 october 2009. *SIGMOD Rec.*, 38(4):46–48, June 2010f. ISSN 0163-5808. URL <http://doi.acm.org/10.1145/1815948.1815959>
- Manghi et al. [2008a] Paolo Manghi, Pasquale Pagano, and Pavel Zezula. First workshop on Very Large Digital Libraries – VLDL 2008. *SIGMOD Record*, 37(4):115–117, 2008a
- Feijen et al. [2007] Martin Feijen, Wolfram Horstmann, Paolo Manghi, Mary Robinson, and Rosemary Russell. DRIVER: Building the Network for Accessing Digital Repositories across Europe. *Ariadne*, 53, 2007. ISSN 1361-3200

- Colazzo et al. [2006] D. Colazzo, G. Ghelli, P. Manghi, and C. Sartiani. Static analysis for path correctness of xml queries. *Journal of Functional Programming*, 16(4-5):621–661, 2006. URL [dx.doi.org/10.1017/S0956796806005983](https://doi.org/10.1017/S0956796806005983). cited By 18
- Simeoni et al. [2003] F. Simeoni, D. Lievens, R. Conn, and P. Mangh. Language bindings to xml. *IEEE Internet Computing*, 7(1):19–27, Jan 2003. ISSN 1089-7801. doi:10.1109/MIC.2003.1167335
- Simeoni et al. [2002] Fabio Simeoni, Paolo Manghi, David Lievens, Richard C.H. Connor, and Steve Neely. An Approach to High-Level Language Bindings for XML. *Elsevier Journal on Information and Software Technology, Special Issue on Objects, <XML> and Databases*, 44(4):217–228, 2002. ISSN 0950-5849. URL [http://dx.doi.org/10.1016/S0950-5849\(02\)00011-3](http://dx.doi.org/10.1016/S0950-5849(02)00011-3)
- Albano et al. [2002] A. Albano, D. Colazzo, G. Ghelli, P. Manghi, C. Sartiani, L. Lini, and M. Paoli. A Typed Text Retrieval Query Language for XML Documents. *JASIST Special Topic Issue on XML and Information Retrieval*, 53(6):467 – 488, February 2002. ISSN 1532-2882. doi:10.1002/asi.10059

10.3 Conferences and Major Workshops

- Bardi et al. [2019] Alessia Bardi, Vittore Casarosa, and Paolo Manghi. Foundations of a framework for peer-reviewing the research flow. In Paolo Manghi, Leonardo Candela, and Gianmaria Silvello, editors, *Digital Libraries: Supporting Open Science*, pages 195–208, Cham, 2019. Springer International Publishing. ISBN 978-3-030-11226-4
- La Bruzzo et al. [2019] Sandro La Bruzzo, Paolo Manghi, and Andrea Mannocci. Openaire’s doiboost - boosting crossref for research. In Paolo Manghi, Leonardo Candela, and Gianmaria Silvello, editors, *Digital Libraries: Supporting Open Science*, pages 133–143, Cham, 2019. Springer International Publishing. ISBN 978-3-030-11226-4
- Atzori et al. [2018a] C. Atzori, P. Manghi, and A. Bardi. Gdup: De-duplication of scholarly communication big graphs. In *2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT)*, pages 142–151, Dec 2018a. doi:10.1109/BDCAT.2018.00025
- Toli et al. [2018] Eleni Toli, Electra Sifacaki, Natalia Manola, Yannis Ioannidis, Tony Ross-Hellauer, Edit Görögh, Michela Vignoli, Vilté Banelytè, Paolo Manghi, and Saskia Woutersen-Windhouwer. Developing sustainable open science solutions in the frame of eu funded research: The openup case. In *Proceedings of the 14th International Symposium on Open Collaboration, OpenSym ’18*, pages 4:1–4:10, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5936-8. doi:10.1145/3233391.3233528. URL <http://doi.acm.org/10.1145/3233391.3233528>
- Poblet et al. [2018] Marta Poblet, Amir Aryani, Paolo Manghi, Kathryn Unsworth, Jingbo Wang, Brigitte Hausstein, Sunje Dallmeier-Tiessen, Claus-Peter Klas, Pompeu Casanovas, and Victor Rodriguez-Doncel. Assigning creative commons licenses to research metadata: Issues and cases. In Ugo Pagallo, Monica Palmirani, Pompeu Casanovas, Giovanni Sartor, and Serena Villata, editors, *AI Approaches to the Complexity of Legal Systems*, pages 245–256, Cham, 2018. Springer International Publishing. ISBN 978-3-030-00178-0
- Atzori et al. [2017] Claudio Atzori, Alessia Bardi, Paolo Manghi, and Andrea Mannocci. The openaire workflows for data management. In Costantino Grana and Lorenzo Baraldi, editors, *Digital Libraries and Archives*, pages 95–107, Cham, 2017. Springer International Publishing. ISBN 978-3-319-68130-6
- de Castro et al. [2017] Pablo de Castro, Jochen Schirrwagen, Dimitris Karaiskos, Jan Dvorak, Andrea Bollini, Vasilis Bonis, Nikon Gasparis, Victoria Tsoukala, Paolo Manghi, and Pedro Principe. Progress in the implementation of the openaire guidelines for cris managers. *Procedia Computer Science*, 106:104 – 111, 2017. ISSN 1877-0509. doi:<https://doi.org/10.1016/j.procs.2017.03.042>. URL <http://www.sciencedirect.com/science/article/pii/S1877050917303101>. 13th International Conference on Current Research Information Systems, CRIS2016, Communicating and Measuring Research Responsibly: Profiling, Metrics, Impact, Interoperability
- Mannocci and Manghi [2016] Andrea Mannocci and Paolo Manghi. Dataq: A data flow quality monitoring system for aggregative data infrastructures. In Norbert Fuhr, László Kovács, Thomas Risse, and Wolfgang Nejdl, editors, *Research and Advanced Technology for Digital Libraries: 20th*

International Conference on Theory and Practice of Digital Libraries, TPDL 2016, Hannover, Germany, September 5–9, 2016, Proceedings, pages 357–369, Cham, 2016. Springer International Publishing. ISBN 978-3-319-43997-6. URL http://dx.doi.org/10.1007/978-3-319-43997-6_28

- Mannocci et al. [2016] Andrea Mannocci, Vittore Casarosa, Paolo Manghi, and Franco Zoppi. The eagle europeana network of ancient greek and latin epigraphy: A technical perspective. In Diego Calvanese, Dario De Nart, and Carlo Tasso, editors, *Digital Libraries on the Move: 11th Italian Research Conference on Digital Libraries, IRCDL 2015, Bolzano, Italy, January 29-30, 2015, Revised Selected Papers*, pages 75–78, Cham, 2016. Springer International Publishing. ISBN 978-3-319-41938-1. URL http://dx.doi.org/10.1007/978-3-319-41938-1_8
- Burton et al. [2015] Adrian Burton, Hylke Koers, Paolo Manghi, Sandro La Bruzzo, Amir Aryani, Michael Diepenbroek, and Uwe Schindler. On bridging data centers and publishers: The data-literature interlinking service. In Emmanouel Garoufallou, RichardJ. Hartley, and Panorea Gaitanou, editors, *Metadata and Semantics Research*, volume 544 of *Communications in Computer and Information Science*, pages 324–335. Springer International Publishing, 2015. ISBN 978-3-319-24128-9. URL http://dx.doi.org/10.1007/978-3-319-24129-6_28
- Assante et al. [2015] Massimiliano Assante, Leonardo Candela, Donatella Castelli, Paolo Manghi, and Pasquale Pagano. Repositories for open science: The scirepo reference model. In Emmanouel Garoufallou, RichardJ. Hartley, and Panorea Gaitanou, editors, *Metadata and Semantics Research*, volume 544 of *Communications in Computer and Information Science*, pages 298–311. Springer International Publishing, 2015. ISBN 978-3-319-24128-9. URL http://dx.doi.org/10.1007/978-3-319-24129-6_26
- Mannocci et al. [2014] Andrea Mannocci, Vittore Casarosa, Paolo Manghi, and Franco Zoppi. The europeana network of ancient greek and latin epigraphy data infrastructure. In Sissi Closs, Rudi Studer, Emmanouel Garoufallou, and Miguel-Angel Sicilia, editors, *Metadata and Semantics Research*, volume 478 of *Communications in Computer and Information Science*, pages 286–300. Springer International Publishing, 2014. ISBN 978-3-319-13673-8. URL http://dx.doi.org/10.1007/978-3-319-13674-5_27
- Kobos et al. [2014] Mateusz Kobos, Łukasz Bolikowski, Marek Horst, Paolo Manghi, Natalia Manola, and Jochen Schirrwagen. Information inference in scholarly communication infrastructures: The openaireplus project experience. *Procedia Computer Science*, 38(0):92 – 99, 2014. ISSN 1877-0509. doi:<http://dx.doi.org/10.1016/j.procs.2014.10.016>. URL <http://www.sciencedirect.com/science/article/pii/S1877050914013763>. 10th Italian Research Conference on Digital Libraries, IRCDL 2014
- Houssos et al. [2014] Nikos Houssos, Brigitte Jörg, Jan Dvořák, Pedro Príncipe, Eloy Rodrigues, Paolo Manghi, and Mikael Karstensen and Elbæk. Openaire guidelines for crisis managers : supporting interoperability of open research information through established standards. In Pablo De Castro Keith Jeffery, Anna Clements and Daniela Luzi, editors, *12th International Conference on Current Research Information Systems, CRIS 2014 — Managing data intensive science – The role of Research Information Systems in realising the digital agenda*, volume 33 of *Procedia Computer Science*, pages 33–38. Elsevier, 2014. doi:10.1016/j.procs.2014.06.006
- La Bruzzo et al. [2013] Sandro La Bruzzo, Paolo Manghi, and Alessia Bardi. Oaizer: Configurable oai exports over relational databases. In Emmanouel Garoufallou and Jane Greenberg, editors, *Metadata and Semantics Research*, volume 390 of *Communications in Computer and Information Science*, pages 35–47. Springer International Publishing, 2013. ISBN 978-3-319-03436-2. URL http://dx.doi.org/10.1007/978-3-319-03437-9_5
- Artini et al. [2013] Michele Artini, Alessia Bardi, Federico Biagini, Franca Debole, Sandro Bruzzo, Paolo Manghi, Marko Mikulicic, Pasquale Savino, and Franco Zoppi. Data interoperability and curation: The european film gateway experience. In Maristella Agosti, Floriana Esposito, Stefano Ferilli, and Nicola Ferro, editors, *Digital Libraries and Archives*, volume 354 of *Communications in Computer and Information Science*, pages 33–44. Springer Berlin Heidelberg, 2013. ISBN 978-3-642-35833-3. URL http://dx.doi.org/10.1007/978-3-642-35834-0_6

- Bardi et al. [2012a] A. Bardi, S. La Bruzzo, and P. Manghi. Oaizer : customized oai-ore and oai-pmh exports of compound objects for the fedora repository. In *Tracking Innovation through Grey Literature*, GL14 - Fourteenth International Conference on Grey Literature, Roma, November 2012a
- Bardi et al. [2012d] Alessia Bardi, Paolo Manghi, and Franco Zoppi. Aggregative data infrastructures for the cultural heritage. In JuanManuel Dodero, Manuel Palomo-Duarte, and Pythagoras Karampiperis, editors, *Metadata and Semantics Research*, volume 343 of *Communications in Computer and Information Science*, pages 239–251. Springer Berlin Heidelberg, 2012d. ISBN 978-3-642-35232-4. URL http://dx.doi.org/10.1007/978-3-642-35233-1_24
- Manghi et al. [2012c] Paolo Manghi, Nikos Houssos, Marko Mikulicic, and Brigitte Jörg. The data model of the openaire scientific communication e-infrastructure. In JuanManuel Dodero, Manuel Palomo-Duarte, and Pythagoras Karampiperis, editors, *Metadata and Semantics Research*, Communications in Computer and Information Science, pages 168–180. Springer Berlin Heidelberg, 2012c. ISBN 978-3-642-35232-4. URL http://dx.doi.org/10.1007/978-3-642-35233-1_18
- Manghi and Mikulicic [2011a] Paolo Manghi and Marko Mikulicic. Pace: A general-purpose tool for authority control. In Elena Garcia-Barriocanal, Zeynel Cebeci, Mehmet C. Okur, and Aydin Ozturk, editors, *Metadata and Semantic Research*, volume 240 of *Communications in Computer and Information Science*, pages 80–92. Springer Berlin Heidelberg, 2011a. ISBN 978-3-642-24731-6. URL http://dx.doi.org/10.1007/978-3-642-24731-6_8
- Kazai et al. [2010a] Gabriella Kazai, Paolo Manghi, Katerina Iatropoulou, Tim Haughton, Marko Mikulicic, Antonis Lempesis, Natasa Milic-Frayling, and Natalia Manola. Architecture for a Collaborative Research Environment based on the notion of Reading List Sharing. In *Proceedings of the 14th European Conference on Digital Libraries*, Glasgow, UK, September 2010a
- Manghi et al. [2010c] Paolo Manghi, Marko Mikulicic, Leonardo Candela, Michele Artini, and Alessia Bardi. *General-Purpose Digital Library Content Laboratory Systems*, pages 14–21. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010c. ISBN 978-3-642-15464-5. URL http://dx.doi.org/10.1007/978-3-642-15464-5_3
- Tang et al. [2010] Cristina Tang, Donatella Castelli, Leonardo Candela, Paolo Manghi, Pasquale Pagano, and Costantino Thanos. An Event-centric Provenance Model for Digital Libraries. In *Sixth Italian Research Conference on Digital Library Management Systems (IRCDL 2010)*, 2010
- Candela et al. [2009] Leonardo Candela, Donatella Castelli, Paolo Manghi, Marko Mikulicic, and Pasquale Pagano. On Foundations of Typed Data Models for Digital Libraries. In *Fifth Italian Research Conference on Digital Libraries (IRCDL)*, Padua, Italy, 2009
- Assante et al. [2008] Massimiliano Assante, Leonardo Candela, Donatella Castelli, Luca Frosini, Lucio Lelii, Paolo Manghi, Andrea Manzi, Pasquale Pagano, and Manuele Simi. An Extensible Virtual Digital Libraries Generator. In *Research and Advanced Technology for Digital Libraries, 12th European Conference (ECDL)*, volume 5173 of *Lecture Notes in Computer Science*, pages 122–134, Aarhus, Denmark, 2008. Springer. ISBN 978-3-540-87598-7
- Artini et al. [2008b] Michele Artini, Leonardo Candela, Donatella Castelli, Paolo Manghi, Marko Mikulicic, and Pasquale Pagano. Sustainable Digital Library Systems over the DRIVER Repository Infrastructure. In *Research and Advanced Technology for Digital Libraries, 12th European Conference (ECDL)*, volume 5173 of *Lecture Notes in Computer Science*, pages 227–231, Aarhus, Denmark, 2008b. Springer. ISBN 978-3-540-87598-7
- Artini et al. [2008a] Michele Artini, Federico Biagini, Paolo Manghi, and Marko Mikulicic. OAI-Publishers in Repository Infrastructures. In *Post-proceedings of the Forth Italian Research Conference on Digital Library Systems (IRCDL)*, pages 93–98, Padua, Italy, January 2008a. DELOS: an Association for Digital Libraries
- Conforti et al. [2007b] Giovanni Conforti, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. Scalable Query Dissemination in XPeer. In *Eleventh International Database Engineering and Applications Symposium (IDEAS 2007)*, pages 199–207, Banff, Alberta, Canada, September 2007b. IEEE Computer Society. ISBN 0-7695-2947-X

- Candela et al. [2007b] Leonardo Candela, Donatella Castelli, Paolo Manghi, and Pasquale Pagano. Enabling Services in Knowledge Infrastructures: The DRIVER Experience. In *Post-proceedings of the Third Italian Research Conference on Digital Library Systems (IRCDL)*, pages 71–77, Padua, Italy, January 2007b. DELOS: a Network of Excellence on Digital Libraries
- Candela et al. [2007a] Leonardo Candela, Donatella Castelli, Paolo Manghi, and Pasquale Pagano. OAI-Aggregator Services for Enhanced Cross-Repository Services. In *Post-proceedings of the Third Italian Research Conference on Digital Library Systems (IRCDL)*, pages 61–66, Padua, Italy, January 2007a. DELOS: a Network of Excellence on Digital Libraries
- Sartiani et al. [2005] Carlo Sartiani, Paolo Manghi, Giorgio Ghelli, and Giovanni Conforti. Xpeer : A self-organizing xml p2p database system. In Wolfgang Lindner, Marco Mesiti, Can Türker, Yannis Tzitzikas, and Athena I. Vakali, editors, *Current Trends in Database Technology - EDBT 2004 Workshops: EDBT 2004 Workshops PhD, DataX, PIM, P2P&DB, and ClustWeb, Heraklion, Crete, Greece, March 14-18, 2004. Revised Selected Papers*, pages 456–465, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg. ISBN 978-3-540-30192-9. URL http://dx.doi.org/10.1007/978-3-540-30192-9_45
- Colazzo et al. [2004] Dario Colazzo, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. Types for Path Correctness of XML Queries. In *Proceedings of the Ninth ACM SIGPLAN International Conference on Functional Programming (ICFP)*, pages 126–137, Snow Bird, UT, USA, September 2004. ACM. ISBN 1-58113-905-5. URL <http://doi.acm.org/10.1145/1016850.1016869>
- Manghi et al. [2002] Paolo Manghi, Fabio Simeoni, David Lievens, and Richard C. H. Connor. Hybrid Applications over XML: Integrating the Procedural and Declarative Approaches. In *Fourth ACM CIKM International Workshop on Web Information and Data Management (WIDM)*, pages 9–14, SAIC Headquarters, LcLean, Virginia, USA, 2002. ACM. ISBN 1-58113-593-9
- Conforti et al. [2002a] Giovanni Conforti, Giorgio Ghelli, Antonio Albano, Dario Colazzo, Paolo Manghi, and Carlo Sartiani. The Query Language TQL. In *5th International Workshop on Web and Data Bases (WebDB02) in conjunction with ACM SIGMOD 2002*, pages 13–18, 2002a
- Connor et al. [1999] Richard Connor, Keith Sibson, and Paolo Manghi. *On the Unification of Persistent Programming and the World-Wide Web*, pages 34–51. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999. ISBN 978-3-540-48909-2. URL http://dx.doi.org/10.1007/10704656_3

10.4 Workshops

- Mannocci and Manghi [2014] Andrea Mannocci and Paolo Manghi. Preliminary analysis of data sources interlinking. In Bolikowski Lukasz, Casarosa Vittore, Goodale Paula, Houssos Nikos, Manghi Paolo, and Schirrwagen Jochen, editors, *Theory and Practice of Digital Libraries – TPDL 2013 Selected Workshops*, volume 416 of *Communications in Computer and Information Science*, pages 53–64. Springer International Publishing, 2014. ISBN 978-3-319-08424-4. URL http://dx.doi.org/10.1007/978-3-319-08425-1_6
- Manghi et al. [2010a] Paolo Manghi, Leonardo Candela, and Pasquale Pagano. Interoperability Patterns in Digital Library Systems Federations. In *Proceedings of the Second DL.org Workshop on Making Digital Libraries Interoperable: Challenges and Approaches, in conjunction with ECDL 2010*, Glasgow, Scotland (UK), September 2010a. ISTI-CNR
- Candela et al. [2007c] Leonardo Candela, Paolo Manghi, and Pasquale Pagano. An Architecture for Type-based Repository Systems. In *Proceedings of the Second Workshop on Foundations of Digital Libraries, in conjunction with ECDL2007*, Budapest, Hungary, September 2007c. DELOS
- Conforti et al. [2007a] Giovanni Conforti, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. Scalable Query Dissemination in XPeer. In *Proceedings of the 1st Workshop on Emerging Research Opportunities for Web Data Management (EROW 2007) Collocated with the 11th International Conference on Database Theory (ICDT 2007)*, volume 229 of *CEUR Workshop Proceedings*. CEUR-WS.org, January 2007a. URL <http://ceur-ws.org/Vol-229/paper2.pdf>

- Candela et al. [2006] Leonardo Candela, Donatella Castelli, Paolo Manghi, and Pasquale Pagano. Typing OpenDLib Repository Service: Strengths of an Information Object Type Language. In *1st European Workshop on the use of Digital Object Repository Systems in Digital Libraries (DORS DL), in conjunction with ECDL2006*, Alicante, Spain, 2006
- Colazzo et al. [2005] Dario Colazzo, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. Types for Path Correctness of XML Queries. In *Proceedings of the Thirteenth Italian Symposium on Advanced Database Systems (SEBD)*, pages 264–271, Brixen-Bressanone, Italy, 2005. ISBN 88-548-0122-4
- Conforti et al. [2004] Giovanni Conforti, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. A Self-organizing XML P2P Database System. In *Proceedings of the Twelfth Italian Symposium on Advanced Database Systems (SEBD)*, pages 394–401, S. Margherita di Pula, Cagliari, Italy, June 2004. ISBN 88-901409-1-7
- Albano et al. [2000] Antonio Albano, Dario Colazzo, Giorgio Ghelli, Paolo Manghi, and Carlo Sartiani. A Type System for Querying XML Documents. In *SIGIR00-XML ACM Workshop held in conjunction with 23rd International SIGIR Conference on Research and Development in Information Retrieval*, Athens, Greece, 2000
- Connor et al. [2001] Richard C.H. Connor, David Lievens, Paolo Manghi, Steve Neely, and Fabio Simeoni. Extracting Typed Values from XML Databases. In *OOPSLA Workshop on Objects, <XML> and Databases*, Tampa Bay, Florida, USA, 2001. URL <http://strathprints.strath.ac.uk/2456/>
- Connor et al. [1996] Richard C.H. Connor, Giorgio Ghelli, and Paolo Manghi. Modules and Type Abstraction in Persistent Systems. In *7th International Workshop in Persistent Object Systems, Persistent Object Systems (POS) - Principles and Practice*, Cape May, New Jersey, May 1996

10.5 Books, chapters, and editorials

- P. et al. [2019] Manghi P., Candela L., and Silvello G. Digital libraries: Supporting open science. In *Digital Libraries: Supporting Open Science*, page 2. Springer International Publishing, 2019. ISBN 978-3-030-11226-4. doi:10.1007/978-3-030-11226-4
- Candela et al. [2017a] Leonardo Candela, Donatella Castelli, Paolo Manghi, and Sarah Callaghan. On research data publishing. *International Journal on Digital Libraries*, 18(2):73–75, Jun 2017a. ISSN 1432-1300. doi:10.1007/s00799-017-0213-y. URL <https://doi.org/10.1007/s00799-017-0213-y>
- Aryani et al. [2017] Amir Aryani, Oscar Corcho, Paolo Manghi, and Jochen Schirrwagen. Guest editorial “repscience2016”. volume 23, 2017. doi:10.1045/january2017-guest-editorial
- Bolikowski et al. [2015] Łukasz Bolikowski, Paolo Manghi, Nikos Houssos, and Jochen Schirrwagen. Guest editorial “data as first-class citizens”. volume 21, 2015. doi:10.1045/january2015-guest-editorial
- Bolikowski et al. [2014] L. Bolikowski, V. Casarosa, P. Goodale, N. Houssos, P. Manghi, and J. Schirrwagen, editors. *Theory and Practice of Digital Libraries – TPD L 2013 Selected Workshops*, volume 416 of *Communications in Computer and Information Science*. Springer Verlag, September 22-26, 2014. doi:10.1007/978-3-319-08425-1
- Candela et al. [2013] L. Candela, P. Castelli D., Manghi, and P. Pagano. *Recent Developments in the Design, Construction, and Evaluation of Digital Libraries: Case Studies*, chapter Infrastructure-Based Research Digital Libraries, pages 1–17. IGI Global, January 2013. doi:10.4018/978-1-4666-2991-2.ch001
- Candela et al. [2011b] Leonardo Candela, Akrivi Katifori, and Paolo Manghi. *Studies on Subject-Specific Requirements for Open Access Infrastructure*, chapter e-Infrastructures, pages 125–164. Universitätsbibliothek Bielefeld, 2011b. ISBN 978-3-943363-00-5. doi:10.2390/PUB-2011-1
- Candela et al. [2011a] Leonardo Candela, Yannis Ioannidis, and Paolo Manghi, editors. *Proceedings of the Fourth Workshop on Very Large Digital Libraries (VL DL 2011), in conjunction with TPD L 2011*, Berlin, Germany, September 2011a. Istituto di Scienza e Tecnologie dell’Informazione, Consiglio Nazionale delle Ricerche, Italy, ISTI-CNR

- Ioannidis et al. [2010] Yannis Ioannidis, Paolo Manghi, and Pasquale Pagano, editors. *Proceedings of the Third Workshop on Very Large Digital Libraries (VLDL 2010), in conjunction with ECDL2010*, Glasgow, Scotland, UK, September 2010. Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Italy, ISTI-CNR
- Ioannidis et al. [2009b] Yannis Ioannidis, Paolo Manghi, and Pasquale Pagano, editors. *Proceedings of the Second Workshop on Very Large Digital Libraries (VLDL 2009), in conjunction with ECDL2008*, Corfu, Greece, October 2009b. Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Italy, DELOS Association for Digital Libraries
- Manghi et al. [2008b] Paolo Manghi, Pasquale Pagano, and Pavel Zezula, editors. *Proceedings of the First Workshop on Very Large Digital Libraries (VLDL 2008), in conjunction with ECDL2008*, Aarhus, Denmark, September 2008b. Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Italy, DELOS Association for Digital Libraries
- P. et al. [2005] Manghi P., Brogi A., Martinelli A., Gervasi V., Fabrizio A., and Pacini G. *Basi di Dati per Medicina e Farmacia*. McGraw Hill, November 2005
- A. et al. [2005] Brogi A., Martinelli A., Gervasi V., Manghi P., and Pacini G. Fabrizio A. *Il Foglio Elettronico per Medicina e Farmacia*. McGraw Hill, May 2005

10.6 Short papers, white papers, communications, demos, and posters

- K. et al. [2018] Shearer K., Rodrigues E., Bollini A., Cabezas A., Castelli D., Carr L., Chan L., Humphrey C., Johnson R., Knoth P., Manghi P., Matizirofa L., Perakakis P., Schirrwagen J., Smith T., Van De Sompel H., Walk P., and Wilcox D. Yamaji K. Next generation repositories: scaling up repositories to a global knowledge commons. Next Generation Repositories WG, Zenodo.org, 2018. doi:<https://doi.org/10.5281/zenodo.1215014>
- Atzori et al. [2018b] C. Atzori, P. Manghi, and A. Bardi. De-duplicating the openaire scholarly communication big graph. In *2018 IEEE 14th International Conference on e-Science (e-Science)*, pages 372–373, Oct 2018b. doi:10.1109/eScience.2018.00104
- Bardi and Manghi [2015a] Alessia Bardi and Paolo Manghi. Enhanced publication management systems: A systemic approach towards modern scientific communication. In *Proceedings of the 24th International Conference on World Wide Web, WWW '15 Companion*, pages 1051–1052, New York, NY, USA, 2015a. ACM. ISBN 978-1-4503-3473-0. doi:10.1145/2740908.2742026. URL <http://doi.acm.org/10.1145/2740908.2742026>
- Massimiliano Assante and Castelli [2015] Paolo Manghi Pasquale Pagano Massimiliano Assante, Leonardo Candela and Donatella Castelli. Providing research infrastructures with data publishing. *Ercim News Magazine - Special Theme: Scientific Data Sharing and Re-use*, January(100):20–22, 2015. ISSN 0926-4981
- Artini et al. [2014b] M. Artini, C. Atzori, and P. Manghi. Keeping your aggregative infrastructure under control. In *Digital Libraries (JCDL), 2014 IEEE/ACM Joint Conference on*, pages 409–410, Sept 2014b. doi:10.1109/JCDL.2014.6970199
- Artini et al. [2014a] M. Artini, C. Atzori, A. Bardi, S. La Bruzzo, and P. Manghi. Tagtick: A tool for annotation tagging over solr indexes. In *Digital Libraries (JCDL), 2014 IEEE/ACM Joint Conference on*, pages 407–408, Sept 2014a. doi:10.1109/JCDL.2014.6970198
- Artini et al. [2014c] 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 Tiziana Catarci, Nicola Ferro, and Antonella Poggi, editors, *Bridging Between Cultural Heritage Institutions*, volume 385 of *Communications in Computer and Information Science*, pages 77–80. Springer Berlin Heidelberg, 2014c. ISBN 978-3-642-54346-3. URL http://dx.doi.org/10.1007/978-3-642-54347-0_9
- Manghi and Mannocci [2013] Paolo Manghi and Andrea Mannocci. Data searchery. In Trond Aalberg, Christos Papatheodorou, Milena Dobrova, Giannis Tsakonas, and Charles J. Farrugia, editors, *Research and Advanced Technology for Digital Libraries*, volume 8092 of *Lecture Notes in*

Computer Science, pages 458–461. Springer Berlin Heidelberg, 2013. URL http://dx.doi.org/10.1007/978-3-642-40501-3_60

- Manghi et al. [2012a] Paolo Manghi, Lukasz Bolikowski, Natalia Manola, Jochen Schirrwagen, Tim Smith, and Najla Rettberg. The openaire data infrastructure services: On interlinking european institutional repositories, dataset archives, and cris systems. Poster, 2012a
- Kazai et al. [2010c] Gabriella Kazai, Natasa Milic-Frayling, Tim Haughton, Natalia Manola, Katerina Iatropoulou, Antonis Lempesis, Paolo Manghi, and Marko Mikulicic. Connecting the local and the online in information management. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management, CIKM '10*, pages 1941–1942, New York, NY, USA, 2010c. ACM. ISBN 978-1-4503-0099-5. doi:10.1145/1871437.1871774. URL <http://doi.acm.org/10.1145/1871437.1871774>
- Kazai et al. [2010b] Gabriella Kazai, Paolo Manghi, Katerina Iatropoulou, Tim Haughton, Marko Mikulicic, and Natasa Milic-Frayling. ScholarLynk: A Reading List Metaphor for Collaborative Research, Poster on proceedings of the Fifth International Conference of Open Repositories, Madrid, Spain, July 2010b
- Castelli et al. [2009] Donatella Castelli, Paolo Manghi, and Stefania Biagioni. OpenAIRE Project - Open Access Infrastructure for Research in Europe, Poster on proceedings of the Eleventh International Conference on Grey Literature (GL11), December 2009
- Jones and Manghi [2009] Sophia Jones and Paolo Manghi. Digital repository infrastructure vision for european research. In *IRCDL*, pages 150–153, 2009
- Ioannidis et al. [2009a] Yannis Ioannidis, Paolo Manghi, and Pasquale Pagano. Second Workshop on Very Large Digital Libraries – VLDL 2009. *D-Lib Magazine*, 15(11/12), November 2009a
- Castelli et al. [2008] Donatella Castelli, Leonardo Candela, Paolo Manghi, Marko Mikulicic, and Pasquale Pagano. Typed Compound Objects Models for Digital Library Repository Systems. Technical report, Istituto di Scienze e Tecnologie dell’Informazione, Consiglio Nazionale delle Ricerche, October 2008
- Conforti et al. [2002b] Giovanni Conforti, Giorgio Ghelli, Antonio Albano, Dario Colazzo, Paolo Manghi, and Carlo Sartiani. The Query Language TQL - Demo Presentation. In *X Convegno nazionale su Sistemi Evoluti per Basi di Dati (SEBD)*, pages 427–431, Portoferraio, Italy, June 2002b

10.7 Research Data and Research Software

- La Bruzzo et al. [2018] Sandro La Bruzzo, Paolo Manghi, and Andrea Mannocci. Doiboost dataset dump. Sep 2018. doi:10.5281/zenodo.1438356
- Bruzzo and Manghi [2018] Sandro La Bruzzo and Paolo Manghi. Openaire scholexplorer service: Scholix json dump. Mar 2018. doi:10.5281/zenodo.1200253
- La Bruzzo and Manghi [2017b] Sandro La Bruzzo and Paolo Manghi. The scholix metadata schema (json and xml). Nov 2017b. doi:10.5281/zenodo.1120275
- Mannocci and Manghi [2017] Andrea Mannocci and Paolo Manghi. MoniQ, February 2017. URL <https://doi.org/10.5281/zenodo.343903>
- Atzori and Manghi [2017] Claudio Atzori and Paolo Manghi. Gdup: a big graph entity deduplication system - release 1.0, February 2017
- La Bruzzo and Manghi [2017a] S. La Bruzzo and P. Manghi. OaiZer, February 2017a. URL <https://doi.org/10.5281/zenodo.292992>
- Artini et al. [2016] M. Artini, A. Bardi, C. Atzori, S. La Bruzzo, P. Manghi, and A. Mannocci. dnet-basic-aggregator: release 1.1.1, May 2016. URL <https://doi.org/10.5281/zenodo.51385>