

Paolo Merialdo

Full Professor

Roma Tre University

Department of Engineering - Computer Engineering and Automation

Personal information

First name: Paolo

Family name: Merialdo

Year of birth: 1965

Nationality: Italian

Email: paolo.merialdo@uniroma3.it

Summary

I am full professor at “Università degli Studi Roma Tre” from 2019. I received my Computer Engineering degree from “Università degli Studi di Genova”, in 1990, and my PhD in Computer Science from “La Sapienza Università degli Studi di Roma ”, in 1998.

My research interests are centered on data management, information extraction and data integration. My current research activity is carried on with colleagues from international institutions (AT&T Research Labs, University of Alberta, Notre Dame University, University of Manchester, Amazon Research, University of Toronto).

To exploit the results of my research, I have co-founded an academic spinoff, CHI Technologies srl, and two startups, [Big Profiles](#), and [my Biros](#). Given my technology transfer experiences and my connections with the industry, I chair the Industrial Advisory Board of the Computer Science and Engineering School at the Department of Engineering of my university. Every year the Industrial Advisory Board raises funds for student grants and organizes seminars, meetings, and events to promote a concrete cooperation between university and industry.

I am also a passionate teacher. Students have always appreciated my courses, and I maintain excellent and fruitful relations with our alumni. Involving colleagues of other departments (Business, Law, Humanities, Political Science), I have organized and promoted cross fertilization initiatives to stimulate the collaboration between students of different disciplines and to contaminate them with industry and market needs.

To foster entrepreneurship among university students, I co-founded InnovAction Lab (2010-2017), a successful program for students from any Italian university (<http://www.innovactionlab.org>). Dozens of companies have been funded by the participants of the program. I have strong connections with the main actors (companies, entrepreneurs, investors, universities) of several entrepreneurial ecosystems, such as Israel, NYC, Singapore. I am advisor at the LuissEnlabs startup accelerator in Rome (2014-today). In 2018, I co-founded *dock 3 – The Startup Lab* (<http://www.dock3.it>), a program built on the heritage of InnovAction Lab. *dock 3 – The Startup Lab* includes a training and an incubation phase.

I have been member of AI Task Force of the Agency for Digital Italy (AgID).

Education

- 1998 PhD (Computer Engineering) at Università degli Studi “La Sapienza”. Advisor: prof. Paolo Atzeni. Dissertation title: “*Web Based Information Systems: Design and Maintenance*”. Part of the PhD work was carried out at the University of Toronto, under the supervision of prof. Alberto Mendelzon.
- 1990 Laurea Degree (Electronic Engineering) at Università degli Studi di Genova. Advisor: prof. Gianni Vernazza. Thesis title: *Integration of Territorial Maps in the Vision System of an Autonomous Land Vehicle*.

Academic Career

- 2019 – now: Roma Tre University: Full professor (ING-INF/05)
- 2006 – 2019: Roma Tre University: Associate professor (ING-INF/05)
- 2002 – 2006: Roma Tre University: Assistant professor (ING-INF/05)
- 1999 – 2002: Roma Tre University: PostDoc

Research Interests

My research interests focus on the design, development and experimental evaluation of novel methods, algorithms and techniques for the management and the analysis of large amounts of data from unconventional sources.

Current research projects

- **Kelpie**: the project addressed the study of techniques for explaining the outcome of link prediction systems over knowledge graphs.
- **RAF (Redundancy As Friend)**: the project tackles the big data integration issue: data linkage, schema alignment and data fusion of product specifications at the web scale. In collaboration with AT&T Research Labs and Amazon.
- **In Codice Ratio**: the goal of the project is to build a **knowledge graph** from the **medieval manuscripts** of the Vatican Secret Archive. Convolutional and recurrent neural networks have been developed to automatically transcribe the manuscripts. NLP techniques are applied to populate a knowledge graph. In collaboration with the Vatican Apostolic Archive.
- **Certa**: the project aims at studying techniques to develop **explainable** methods for **Machine Learning** data management systems. The project explores both white-box and black-box solutions, as well as pretraining techniques to improve quality and efficiency of Machine Learning technique to data management problems. In collaboration with AT&T Research Labs and University of Toronto.
- **Noah**: the project has the goal of developing novel solutions for extraction and integration of data from Web sources.

Technology Transfer and Entrepreneurship

- Co-founder and scientific advisor of myBiros.com (2021-today), a startup that aims at exploiting the results of my research activity in handwritten text recognition

- Co-founder of [Dock3 - The Startup Lab](#), a program to promote entrepreneurship among university students (2018- today)
- Co-founder of the startup [Data Falls – Big Profiles srl](#) (2016 - today), a startup that aims at exploiting the results of my research activity in machine learning for data management
- Co-founder and program curator of [Data Driven Innovation Open Summit](#) (2016-2021), an interdisciplinary event that aims at explaining to a broad public the impact of the big data revolution in the social, economic, industrial fields of our society
- Advisor for KPI6 s.r.l. (2016 – today), an AI startup
- Advisor at the LuissEnlabs Accelerator in Rome (2013 – today)
- Co-founder of InnovAction Lab, a program to foster entrepreneurship among university students (2010- today)
- Chair of the Industrial Advisory Board of the Computer Engineering School, Department of Engineering of Roma Tre University (2008 – today)
- Co-founder of the academic spinoff CHI Technologies srl (2003 - 2008)

Teaching Activities

Courses Taught at the Department of Engineering

- 2021 – today “Digital Entrepreneurship” (6 CFU)
Laurea Magistrale in Ingegneria Informatica
- 2010 – today “Soft Skills for Computer Engineers” (1 CFU)
Laurea Magistrale in Ingegneria Informatica
- 2008 – today “Data Engineering”, formerly “Web data management and analysis” (6 CFU)
Laurea Magistrale in Ingegneria Informatica
- 2002 – today “Web Based Information Systems” (6 CFU)
Laurea in Ingegneria Informatica
- 2002 – 2014 “Object Oriented Programming” (6 CFU)
Laurea in Ingegneria Informatica
- 2006 – 2012 “Fundamentals of Computer Science” (6 CFU)
Laurea in Ingegneria Meccanica

Lectures in University Masters

- 2010-2017 Databases and data mining in the Roma Tre University Master “*Governance, sistema di controllo e auditing negli enti pubblici e private*” (Department of Business Studies)
- 2001-2004 Databases and Information Systems in the Roma Tre University Master “*Economia e Tecnologia della Società dell’Informazione*” (Department of Engineering and Department of Business Studies)

Evaluator/reviewer for funding agencies

- 2008- today Research Grants Council di Honk Kong (external reviewer)
- 2013 UE FP7-SME-2013 (expert evaluator)

Research Funding (as Principal Investigator)

- 2021-2023 MISE. Nome progetto: “*Casa delle Tecnologie Emergenti – Comune di Roma*”. Ruolo: PI.

- 2021-2022 Programma Strategico Regionale per la ricerca, l'innovazione ed il trasferimento tecnologico 2014-2020 Legge Regionale n. 13 del 04/08/2008. Nome progetto: “*Matrices*”. Ruolo: PI.
- 2018-2020 Erasmus+ KA2 - Cooperation for innovation and the exchange of good practices in Entrepreneurship Development “*STARTED - STArtup skills for Researchers and innovaTors*”. Role: P.I. for the Roma Tre research unit.
- 2021-2022 Programma Strategico Regionale per la ricerca, l'innovazione ed il trasferimento tecnologico 2018-2020 Legge Regionale n. 13 del 04/08/2008. Nome progetto: “*In Codice Ratio*”. Ruolo: PI.
- 2014-2015 POR FESR Lazio 2007/2013 “*SocialEyers: Sentiment Analysis Techniques for Real Time Social Media Monitoring*”. Role: P.I. for the Roma Tre research unit.
- 2011-2013 POR FESR Lazio 2007/2013 “*MORNING: MONitoraggio della Reputazione su social Network e Informazione Generata da utenti*”. Role: P.I. for the Roma Tre research unit.
- 2010-2012 Italian Ministry of Education, University and Research, (PRIN2008) “*EASE: Entity Aware Search Engines*”. Role: P.I. for the Roma Tre research unit.
- 2004-2006 Italian Ministry of Education, University and Research, PRIN 2004 “*WISDOM: Intelligent Web Search based on Domain Ontologies*”. Role: P.I. for the Roma Tre research unit.

Editorial Boards

- 2017 – today: Associate Editor for the journal of “*Knowledge and Information Systems*” (Springer)

Program Committees (Chair)

- 2019. Vice-Chair 35th IEEE International Conference on Data Engineering (ICDE 2019), April 8-12 2019, Macau SAR, China
- Co-chair of the *Entrepreneurs Track* at the 24th International World Wide Web Conference (WWW 2015), May 18-22, 2015, Florence, Italy. Other chairs: Andrei Broder (Google), Stefano Leonardi (La Sapienza). For the track, I organized panels and keynotes inviting Bill Coughran (partner at Sequoia Capital), David Rose (founder of NY Business Angels), Yossi Matias (Google Eng. Vice President, director of Google Tel Aviv R&D), Uri Levine (founder of Waze), Talia Rafaeli (Vice President on the StageOne Ventures), Tak Lo (director of Techstars London), Luigi Capello (ceo of LVentures)
- Co-chair of the *Fourth Int. Workshop on Data Extraction and Object Search* (DEOS2014), in conjunction with the 23rd International World Wide Web Conference (WWW'14), April 7-11, Seoul, South Korea. Other chairs: Tim Furche (Oxford), Giorgio Orsi (Oxford), Valter Crescenzi (RomaTre)

Program Committees (member)

- ACM SIGMOD/PODS International Conference on Management of Data (SIGMOD2022)
- 31th ACM International Conference on Information and Knowledge Management (CIKM 2021) - PC member
- 47th International Conference on Very Large Databases (VLDB2021)
- 39th IEEE International Conference on Data Engineering (ICDE 2021) - PC member
- 37th IEEE International Conference on Data Engineering (ICDE 2019) - PC member

- 29th ACM International Conference on Information and Knowledge Management (CIKM 2019) - PC member
- 35th IEEE International Conference on Data Engineering (ICDE 2019) - senior PC member
- 27th ACM International Conference on Information and Knowledge Management (CIKM 2018) - senior PC member
- 43rd International Conference on Very Large Databases (VLDB2017)
- 34th IEEE International Conference on Data Engineering (ICDE 2018)
- 43rd Int. Conference on Very Large Databases (VLDB2017)
- 17th 2017 International Conference on Big Data and Smart Computing (BigComp 2017)
- 16th International Conference on Web Engineering (ICWE 2017)
- 2016 International Conference on Big Data and Smart Computing (BigComp 2016)
- 2015 International Conference on Big Data and Smart Computing (BigComp 2015)
- 1st Int. Workshop on Novel Web Search Interfaces and Systems (NWSearch 2015)
- International Workshop on Enterprise Web Application Dependability (EWAD 2015)
- Fourth International Workshop on Linked Web Data Management (LWDM2015)
- 15th International Conference on Web Engineering (ICWE 2015)
- 7th Balkan Conference in Informatics (BCI 2015)
- 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD 2015)
- 8th Balkan Conference in Informatics (BCI 2015) PC member
- Fourth International Workshop on Semantic Search Over the Web (SSW 2014)
- Third International Workshop on Linked Web Data Management (LWDM 2013)
- Third International Workshop on Semantic Search Over the Web (SSW 2013)
- Second International Workshop on semantic search over the Web (SSW 2012)
- Int. Workshop on Web Information and Data Management (ACM WIDM 2012)
- Second International Workshop on Linked Web Data Management (LWDM 2012)
- 38th International Conference on Very Large Databases (VLDB 2012)
- 1st International Workshop on Linked Web Data Management (LWDM 2011)
- 1st International Workshop on Databases and Social Networks (DBSocial 2011)
- 12th International Workshop on the Web and Databases (WebDB 2009)
- 2009 IEEE/WIC/ACM Int. Joint Conference on Web-Intelligence (WI-2009)
- Tenth International Conference on Web Information Systems Engineering (WISE 2009)
- 2008 IEEE/WIC/ACM Int. Joint Conference on Web-Intelligence (WI-2008)
- 3rd International ICST Conference on Scalable Information Systems (Infoscale 2008)
- First International Workshop on Modeling Mobile Applications and Services (M2AS08)
- Ninth International Conference on Web Information Systems Engineering (WISE 2008)
- Sixteenth ACM Conference on Information and Knowledge Management (CIKM 2007)
- 2007 IEEE/WIC/ACM Int. Joint Conference on Web-Intelligence (WI-2007)
- 22nd IEEE International Conference on Data Engineering (ICDE 2006)
- 2005 IEEE/WIC/ACM Int. Joint Conference on Web-Intelligence (WI-2005)
- Seventh Int. Workshop on Web Information and Data Management (ACM WIDM 2005)
- 2004 IEEE/WIC/ACM Int. Joint Conference on Web- Intelligence (WI-2004)
- Sixth Int. Workshop on Web Information and Data Management (ACM WIDM 2004)
- Fifth Int. Workshop on Web Information and Data Management (ACM WIDM 2003)
- Eighteenth Annual ACM Symposium on Applied Computing (ACM SAC 2003)
- Fourth Int. Workshop on Web Information and Data Management (ACM WIDM 2002)
- Fourth Int. Workshop on the Web and Databases (WebDB 2001)

Invited Lectures and Tutorials

- Conférence Universitaire de Suisse Occidentale - Université de Neuchâtel: “*Crowdsourcing Data Management*” - Dec 2017
- Extending Database Technology (EDBT) Summer School: “*Intelligent Web Data Extraction*” – with Rim Furche (University of Oxford, UK) - Sept 2017
- AT&T Research Labs at New York City: *Fact harvesting from natural language corpora* - July 2016
- University of Manchester: *Crowdsourcing wrapper inference* - Nov 2014
- University of Alberta: *Wrapper inference needs feedback* - July 2013
- Twenty-third Italian Symposium on Advanced Database Systems: *The startup ecosystem: a quick tour* - June 2014
- University of Manchester: *Web Data Extraction* - May 2007

PhD Students

Graduated

- 2017 Elena Nieddu. “*In Codice Ratio: a crowd-enabled, low resource handwritten text recognition approach for the Vatican Registers*”. Now at Fondazione Ugo Bordoni (Italy)
- 2020 Federico Piai: Instance level attribute alignment for product specifications. Now at Progesoftware spa (Italy)
- 2018 Matteo Cannaviccio: “*Knowledge Graphs Augmentation with Natural Language Evidence*”. Now at Diffbot (San Francisco, California)
- 2014 Disheng Qiu: “*Crowdsourcing Large scale Data Extraction from the Web: Bridging Unsupervised and Supervised Approaches*”. Now co-founder and CTO of Wanderio (Italy).
- 2012 Mirko Bronzi: “*Exploiting Information Redundancy to Collect Data from the Web*”. Now at Nuances (Montreal, Canada)
- 2011 Lorenzo Blanco: “*Extraction, integration and probabilistic characterization of web data*”. Now at Google (London, UK)

Current

- 2019 Tommaso Teofili. Explaining ML solutions to data management issues. Expected 2023
- 2018 Andrea Rossi. Machine learning techniques for link prediction. Expected 2022

Publications

Citation Indices (as of May 14, 2021)

- Elsevier Scopus (<https://www.scopus.com/authid/detail.uri?authorId=6602539867>) documents: 82, h-index: **16**, citations: **1686**
- Google Scholar (<https://scholar.google.com/citations?user=njKmu5cAAAAJ>) documents: >100, h-index: **26**, citations: **4288**

Journals

1. Elena Nieddu, Donatella Firmani, Paolo Merialdo, Marco Maiorino. *In Codice Ratio: A crowd-enabled solution for low resource machine transcription of the Vatican Registers*. Information Processing & Management, 58:5 (2021)

2. Andrea Rossi, Denilson Barbosa, Donatella Firmani, Antonio Martinata, Paolo Merialdo: *Knowledge Graph Embedding for Link Prediction: A Comparative Analysis*. ACM Transactions on Knowledge Discovery from Data (TKDD) 15:2, 1-49 (2021)
3. Disheng Qiu, Valter Crescenzi, Paolo Merialdo: *Hybrid Crowd-Machine Wrapper Inference*. ACM Transactions on Knowledge Discovery from Data (TKDD) 13:5 (2019)
4. Luciano Barbosa, Valter Crescenzi, Xin Luna Dong, Paolo Merialdo, Federico Piai, Disheng Qiu, Yanyan Shen, Divesh Srivastava: *Big Data Integration for Product Specifications*. IEEE Data Engineering Bulletin. 41(2):74-84 (2018)
5. Valter Crescenzi, Alvaro A.A. Fernandez, Paolo Merialdo, Norman Paton: *Crowdsourcing for Data Management*. Knowledge and Information Systems 53(1): 1-41 (2017)
6. Valter Crescenzi, Paolo Merialdo, Disheng Qiu: *Crowdsourcing large scale wrapper inference*. Distributed and Parallel Databases 33(1): 95-122 (2015)
7. Tim Weninger, Rodrigo Palácios, Valter Crescenzi, Thomas Gottron, Paolo Merialdo: *Web Content Extraction: a MetaAnalysis of its Past and Thoughts on its Future*. SIGKDD Explorations 17(2): 17-23 (2015)
8. Mirko Bronzi, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Extraction and Integration of Partially Overlapping Web Sources*. PVLDB 6(10): 805-816 (2013)
9. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Web Data Reconciliation: Models and Experiences*. SeCO Book: 1-15 (2012)
10. Valter Crescenzi, Paolo Merialdo: *Wrapper Inference for Ambiguous Web Pages*. Journal of Applied Artificial Intelligence 22(1&2): 21-52 (2008)
11. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo: *Structure and Semantics of Data-Intensive Web Pages: An Experimental Study on their Relationships*. J. UCS 14(11): 1877-1892 (2008)
12. Valter Crescenzi, Paolo Merialdo, Paolo Missier: *Clustering Web pages based on their structure*. Data Knowledge Engineering 54(3): 279-299 (2005)
13. Paolo Merialdo, Paolo Atzeni, Giansalvatore Mecca: *Design and development of data-intensive web sites: The araneus approach*. ACM Trans. On Internet Technology 3(1): 49-92 (2003)
14. Paolo Atzeni, Giansalvatore Mecca, Paolo Merialdo: *Managing Web-Based Data: Database Models and Transformations*. IEEE Internet Computing 6(4): 33-37 (2002)
15. Giansalvatore Mecca, Alberto O. Mendelzon, Paolo Merialdo: *Efficient Queries over Web Views*. IEEE Trans. Knowledge Data Engineering 14(6): 1280-1298 (2002)
16. Paolo Atzeni, Paolo Merialdo, Giansalvatore Mecca: *Data-Intensive Web Sites: Design and Maintenance*. World Wide Web 4(1-2): 21-47 (2001)
17. Giansalvatore Mecca, Paolo Merialdo, Paolo Atzeni: *Araneus in the Era of XML*. IEEE Data Engineering Bulletin 22(3): 19-26 (1999)
18. Paolo Atzeni, Giansalvatore Mecca, Paolo Merialdo: *Semistructured und Structured Data in the Web: Going Back and Forth*. ACM SIGMOD Record 26(4): 16-23 (1997)

Conferences (with peer review)

1. Valerio Cetrelli, Valter Crescenzi, Paolo Merialdo, Roger Voyat: *NOAH: Creating Data Integration Pipelines over Continuously Extracted Web Data*. EDBT/ICDT Workshops 2021
2. Simone Scardapane, Elena Nieddu, Donatella Firmani, Paolo Merialdo: *Multikernel Activation Functions: Formulation and a Case Study*. Proceedings of the INNS Big Data and Deep Learning Conference INNSBDDL 2019.
3. Vincenzo Di Cicco, Donatella Firmani, Nick Koudas, Paolo Merialdo, Divesh Srivastava. *Interpreting deep learning models for entity resolution: an experience report using LIME*.

Proceedings of the Second International Workshop on Exploiting Artificial Intelligence Techniques for Data Management. aiDM@SIGMOD 2019.

4. Donatella Firmani, Marco Maiorino, Paolo Merialdo, Elena Nieddu: *Towards Knowledge Discovery from the Vatican Secret Archives. In Codice Ratio - Episode 1: Machine Transcription of the Manuscripts*. ACM KDD 2018.
5. Matteo Cannaviccio, Denilson Barbosa, Paolo Merialdo: *Towards Annotating Relational Data on the Web with Language Models*. The Web Conference 2018" (WWW 2018)
6. Disheng Qiu, Luciano Barbosa, Valter Crescenzi, Paolo Merialdo, Divesh Srivastava: *Big Data Linkage for Product Specification Pages*. ACM SIGMOD 2018.
7. Matteo Cannaviccio, Lorenzo Ariemma, Denilson Barbosa, Paolo Merialdo: *Leveraging Wikipedia Table Schemas for Knowledge Graph Augmentation*. WebDB 2018
8. Serena Ammirati, Donatella Firmani, Marco Maiorino, Paolo Merialdo, Elena Nieddu, Andrea Rossi: *In Codice Ratio: Scalable Transcription of Historical Handwritten Documents*. SEBD 2017
9. Donatella Firmani, Paolo Merialdo, Elena Nieddu, Simone Scardapane: *In Codice Ratio: OCR of Handwritten Latin Documents using Deep Convolutional Networks*. The 11th Italian Workshop on Artificial Intelligence for Cultural Heritage. AI*CH 2017.
10. Matteo Cannaviccio, Denilson Barbosa, Paolo Merialdo: *Accurate fact harvesting from natural language text in wikipedia with Lector*. WebDB 2016
11. Paolo Merialdo: *The Startup Ecosystem: a Quick Tour*. SEBD 2015
12. Lorenz Bühmann, Ricardo Usbeck, Axel-Cyrille Ngonga Ngomo, Muhammad Saleem, Andreas Both, Valter Crescenzi, Paolo Merialdo, Disheng Qiu: *Web-Scale Extension of RDF Knowledge Bases from Templated Websites*. Semantic Web Conference (1) 2014
13. Yoones A. Sekhavat, Francesco Di Paolo, Denilson Barbosa, Paolo Merialdo: *Knowledge Base Augmentation using Tabular Data*. LDOW 2014
14. Valter Crescenzi, Paolo Merialdo, Disheng Qiu: *Wrapper Generation Supervised by a Noisy Crowd*. DBCrowd 2013
15. Valter Crescenzi, Paolo Merialdo, Disheng Qiu: *A framework for learning web wrappers from the crowd*. WWW 2013
16. Valter Crescenzi, Paolo Merialdo, Disheng Qiu: *ALFRED: crowd assisted data extraction*. WWW (Companion Volume) 2013
17. Rolando Creo, Valter Crescenzi, Disheng Qiu, Paolo Merialdo: *Minimizing the Costs of the Training Data for Learning Web Wrappers*. VLDS 2012
18. Paolo Merialdo, Paolo Papotti: *Flint: From Web Pages to Probabilistic Semantic Data*. Semantic Search over the Web 2012
19. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Contextual Data Extraction and Instance-Based Integration*. VLDS 2011
20. Mirko Bronzi, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Wrapper Generation for Overlapping Web Sources*. Web Intelligence 2011
21. Lorenzo Blanco, Mirko Bronzi, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Automatically building probabilistic databases from the web*. WWW (Companion Volume) 2011
22. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Probabilistic Models to Reconcile Complex Data from Inaccurate Data Sources*. CAiSE 2010
23. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Probabilistic Reconciliation of Records from Inaccurate Web Sources (Extended Abstract)*. SEBD 2010
24. Paolo Papotti, Valter Crescenzi, Paolo Merialdo, Mirko Bronzi, Lorenzo Blanco: *Redundancy-Driven Web Data Extraction and Integration*. WebDB 2010

25. Lorenzo Blanco, Mirko Bronzi, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Exploiting information redundancy to wring out structured data from the web*. WWW 2010
26. Blanco, Mirko Bronzi, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Data Extraction and Integration from Imprecise Web Sources*. SEBD 2009
27. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Flint: Google-basing the Web*. EDBT 2008
28. Daniele Braga, Diego Calvanese, Alessandro Campi, Stefano Ceri, Florian Daniel, Davide Martinenghi, Paolo Merialdo, Riccardo Torlone: *NGS: a framework for multi-domain query answering*. IEEE ICDE Workshops 2008
29. Claudio Bertoli, Valter Crescenzi, Paolo Merialdo: *Crawling programs for wrapper-based applications*. IEEE IRI 2008
30. Daniele Braga, Diego Calvanese, Alessandro Campi, Stefano Ceri, Florian Daniel, Davide Martinenghi, Paolo Merialdo, Riccardo Torlone: *A New Generation Search Engine Supporting Cross Domain Queries*. SEBD 2008
31. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Searching Entities on the Web by Sample*. SEBD 2008
32. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Supporting the automatic construction of entity aware search engines*. WIDM 2008
33. Valter Crescenzi, Paolo Merialdo: *Efficient Techniques for Effective Wrapper Induction*. IEEE ICDE Workshops 2006
34. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo: *Harvesting Structurally Similar Pages*. SEBD 2005
35. Sonia Bergamaschi, Paolo Bouquet, Paolo Ciaccia, Paolo Merialdo: *Speaking Words of WISDOM: Web Intelligent Search based on DOMain ontologies*. SWAP 2005
36. Lorenzo Blanco, Valter Crescenzi, Paolo Merialdo: *Efficiently Locating Collections of Web Pages to Wrap*. WEBIST 2005
37. Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *Improving the expressiveness of ROADRUNNER*. SEBD 2004
38. Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo, Paolo Missier: *An Automatic Data Grabber for Large Web Sites*. VLDB 2004
39. Luigi Arlotta, Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *Automatic annotation of data extracted from large web sites*. SEBD 2003
40. Luigi Arlotta, Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *Automatic annotation of data extracted from large Web sites*. WebDB 2003
41. Valter Crescenzi, Paolo Merialdo, Paolo Missier: *Fine-grain web site structure discovery*. WIDM 2003
42. Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *Wrapping-oriented classification of web pages*. SAC 2002
43. Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *Back to Gold's Age: Bridging the Gap Between Traditional Grammar Inference and Web Information Extraction*. SEBD 2002
44. Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *RoadRunner: automatic data extraction from data-intensive web sites*. ACM SIGMOD Conference 2002
45. Paolo Atzeni, Paolo Merialdo, Giuseppe Sindoni: *Web Site Evaluation: Methodology and Case Study*. ER (Workshops) 2001
46. Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *Automatic Web Information Extraction in the ROADRUNNER System*. ER (Workshops) 2001
47. Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *The RoadRunner Web Data Extraction System*. SEBD 2001

48. Valter Crescenzi, Giansalvatore Mecca, Paolo Merialdo: *RoadRunner: Towards Automatic Data Extraction from Large Web Sites*. VLDB 2001
49. Giansalvatore Mecca, Paolo Merialdo, Paolo Atzeni, Valter Crescenzi: *Experiences in XML data management*. SEBD 2000
50. Paolo Merialdo, Paolo Atzeni, Marco Magnante, Giansalvatore Mecca, Marco Pecorone: *Homer: a Model-Based CASE Tool for Data-Intensive Web Sites*. ACM SIGMOD Conference 2000
51. Giansalvatore Mecca, Paolo Merialdo, Paolo Atzeni, Valter Crescenzi: *The ARANEUS Guide to Web-Site Development*. SEBD 1999
52. Giansalvatore Mecca, Paolo Merialdo, Paolo Atzeni, Valter Crescenzi: *The (Short) Araneus Guide to Web-Site Development*. WebDB 1999
53. Giansalvatore Mecca, Alberto O. Mendelzon, Paolo Merialdo: *Efficient Queries over Web Views*. EDBT 1998
54. Paolo Atzeni, Giansalvatore Mecca, Paolo Merialdo: *Design and Maintenance of Data-Intensive Web Sites*. EDBT 1998
55. Giansalvatore Mecca, Paolo Merialdo, Alessandro Masci, Giuseppe Sindoni: *The Araneus Project: Extending Database Techniques to the World Wide Web*. SEBD 1998
56. Giansalvatore Mecca, Paolo Atzeni, Alessandro Masci, Paolo Merialdo, Giuseppe Sindoni: *The Araneus Web-Base Management System*. ACM SIGMOD Conference 1998
57. Giansalvatore Mecca, Paolo Merialdo, Paolo Atzeni: *Do we really need a new query language for XML? XML-QL* 1998
58. Paolo Merialdo, Federico Bilotta, Laura Fiorani, Giovanni Ettore Gigante: *Automated Recognition of Cardiac Structures in Echocontrast Perfusion Studies*. CBMS 1997
59. Patrizia Grifoni, Daniela Luzi, Paolo Merialdo, Fabrizio L. Ricci: *ATREUS: A Model for the Conceptual Representation of a Workflow*. DEXA Workshop 1997
60. Paolo Atzeni, Alessandro Masci, Giansalvatore Mecca, Paolo Merialdo, Elena Tabet: *ULIXES: Building Relational Views over the Web*. IEEE ICDE 1997
61. Paolo Merialdo, Giuseppe Sindoni: *MrBrAQue: A Multimedia Medical Report Management System*. ICMCS 1997
62. Paolo Atzeni, Alessandro Masci, Giansalvatore Mecca, Paolo Merialdo, Elena Tabet: *Structures in the Web*. SEBD 1997
63. Paolo Atzeni, Giansalvatore Mecca, Paolo Merialdo: *To Weave the Web*. VLDB 1997
64. Paolo Merialdo, Pier Carlo Pecollo, Carlo S. Regazzoni, Gianni Vernazza, Rodolfo Zunino: *Integration of Territorial Maps in the Vision System of an Autonomous Land Vehicle*. IAS 1989

In preparation:

1. Federico Piai, Paolo Atzeni, Paolo Merialdo, Divesh Srivastava: *Fine-Grained Semantic Typing for Heterogeneous Sources*.

I authorize the use of my personal data in compliance with GDPR.

Roma
May 14, 2021

Paolo Merialdo

