


INFORMAZIONI GENERALI

Merialdo Paolo

 Università degli Studi Roma Tre
Dipartimento di Ingegneria
Via Vito Volterra 60, 00146 Roma (Italy)

 <http://merialdo.inf.uniroma3.it>

ESPERIENZE
PROFESSIONALI

1 giu 2019–oggi **Professore Ordinario (settore scientifico disciplinare ING-INF/05)**
Dipartimento di Ingegneria- Università degli studi Roma Tre

1 nov 2006–31 mag 2019 **Professore Associato (settore scientifico disciplinare ING-INF/05)**
Dipartimento di Ingegneria- Università degli studi Roma Tre

Feb. 1, 2002–Oct. 31, 2006 **Ricercatore Universitario (settore scientifico disciplinare ING-INF/05)**
Facoltà di Ingegneria- Università degli studi Roma Tre

STUDI

1998 **Dottorato in Ingegneria Informatica– Università di Roma "La Sapienza"**

1990 **Laurea in Ingegneria Elettronica - Università degli Studi di Genova**

ATTIVITA' DIDATTICA

Insegnamenti

"Analisi e gestione dell'informazione su Web" – 6 CFU – Laurea Magistrale in Ingegneria Informatica Roma Tre. 2009/10, 2010/11, 2011/12, 2012/13, 2013/14, 2014/15, 2015/16, 2016/17, 2017/18, 2018/19. Circa 50 studenti/anno

"Advanced Topics in Computer Science" – 6 CFU – Laurea Magistrale in Ingegneria Informatica Roma Tre. 2018/19. Circa 12 studenti/anno

"Conoscenze utili per l'inserimento nel mondo del lavoro" – 1 CFU – Laurea Magistrale in Ingegneria Informatica Roma Tre 2008/09, 2009/10, 2010/11, 2011/12, 2012/13, 2013/14, 2014/15, 2015/16, 2016/17, 2017/18, 2018/19. Circa 100 studenti/anno

"Sistemi informativi su Web" – 6 CFU (5 CFU fino al 2007/2008) –Laurea in Ingegneria Informatica Roma Tre 2002/03, 2003/04, 2004/05, 2005/06, 2006/07, 2007/08, 2008/09, 2009/10, 2010/11, 2011/12, 2012/13, 2013/14, 2014/15, 2015/16, 2016/17, 2017/18, 2018/19. Circa 120 studenti/anno

"Programmazione Orientata agli Oggetti" – 6 CFU (5 CFU fino al 2007/2008) – Laurea in Ingegneria Informatica Roma Tre 2002/03, 2003/04, 2004/05, 2005/06, 2006/07, 2007/08, 2008/09, 2009/10, 2010/11, 2011/12, 2012/13. Circa 200 studenti/anno

"Computer Programming Basics" – 6 CFU (4 CFU until 2007/2008) –Laurea in Ingegneria Meccanica Roma Tre 2006/07, 2007/08, 2008/09, 2009/10, 2010/11, 2011/12. Circa 200 studenti/anno

Descrizione	<p>Paolo Merialdo è professore ordinario ssd ING-INF/05. La sua attività di ricerca abbraccia circa venti anni (1996-2020) e si concentra prevalentemente nello studio di tecniche e metodologie per la gestione, l'estrazione e l'integrazione di dati da sorgenti non strutturate. Inizialmente ha proposto modelli e linguaggi per la gestione di siti web data-intensive. Successivamente, ha focalizzato la ricerca sullo studio di soluzioni per l'estrazione di dati da Web; in questo contesto ha approfondito la possibilità di adottare tecniche di inferenza grammaticale e soluzioni basate su tecniche di <i>crowdsourcing</i> e <i>active learning</i>. Più recentemente, ha spostato l'indagine sui problemi di <i>data integration</i> proponendo soluzioni volte a sfruttare la ridondanza dell'informazione presente su Web. Negli ultimi quattro anni allo scopo di applicare tecniche di <i>information extraction</i> su documenti storici, ha iniziato a studiare soluzioni innovative basate su tecniche di <i>computer vision</i> e <i>deep learning</i> per la trascrizione automatica di manoscritti medioevali. In particolare, ha avviato il progetto In Codice Ratio, in collaborazione con l'Archivio Apostolico Vaticano, finalizzato a produrre trascrizioni e ad estrarre informazioni dai Registri Papali. Google ha invitato a presentare il progetto al Google Tensorflow Summit (la conferenza sulle tecnologie di Deep Learning di Google), indicandolo come miglior progetto in ambito Digital Humanities. Dal progetto sono nati quattro contratti di trasferimento tecnologico con aziende del Lazio ed è stata avviata una collaborazione con la Camera dei Deputati per la trascrizione dei manoscritti dell'Archivio della Camera dei Deputati. Ha co-fondato uno spin-off universitario ed una startup. Si occupa di formazione imprenditoriale: ha co-fondato "<i>InnovAction Lab</i>" e "<i>dock3 - The Startup Lab</i>", programmi di formazione imprenditoriale dai quali sono nate numerose startup laziali. È advisor dell'acceleratore di startup LuissEnlabs.</p>
Indici bibliometrici (28 lug 2020)	<p>Elsevier Scopus. h-index: 16, citazioni: 1653, documenti: 77 Google Scholar. h-index: 25, citazioni: 4089</p>
Comitati editoriali	<p>2017 – oggi: Associate Editor Springer "Knowledge and Information Systems"</p>
Comitati di programma (Chair, vice Chair)	<p>2019. Vice-Chair 35th IEEE International Conference on Data Engineering (ICDE 2019), April 8-12 2019, Macau SAR, China</p> <p>2015. Co-chair Entrepreneurs Track at the 24th International World Wide Web Conference (WWW 2015), May 18-22, 2015, Florence, Italy. Co-chairs: Andrei Broder (Google), Stefano Leonardi (La Sapienza)</p> <p>2014. Co-chair DEOS 2014: Fourth Workshop on Data Extraction and Object Search in conjunction with the 23rd International World Wide Web Conference (WWW'14), April 7-11, Seoul, South Korea. Co-chairs: Tim Furche (Oxford), Giorgio Orsi (Oxford), Valter Crescenzi (Roma Tre)</p>
Pubblicazioni (recenti)	<p>Riviste:</p> <p>Disheng Qiu, Valter Crescenzi, Paolo Merialdo: <i>Hybrid Crowd-Machine Wrapper Inference</i>. ACM Transactions on Knowledge Discovery from Data. Volume: 13. Issue: 5. 2019</p> <p>Luciano Barbosa, Valter Crescenzi, Xin Luna Dong, Paolo Merialdo, Federico Piai, Disheng Qiu, Yanyan Shen, Divesh Srivastava: <i>Big Data Integration for Product Specifications</i>. IEEE Data Engineering Bulletin. 41(2):74-84 (2018)</p> <p>Valter Crescenzi, Alvaro A.A. Fernandez, Paolo Merialdo, Norman Paton: <i>Crowdsourcing for Data Management</i>. Knowledge and Information Systems 53(1): 1-41 (2017)</p> <p>Valter Crescenzi, Paolo Merialdo, Disheng Qiu: <i>Crowdsourcing large scale wrapper inference</i>. Distributed and Parallel Databases 33(1): 95-122 (2015)</p> <p>Tim Wening, Rodrigo Palacios, Valter Crescenzi, Thomas Gottron, Paolo Merialdo: <i>Web Content Extraction: a MetaAnalysis of its Past and Thoughts on its Future</i>. SIGKDD Explorations 17(2): 17-23 (2015)</p>

Mirko Bronzi, Valter Crescenzi, Paolo Merialdo, Paolo Papotti: *Extraction and Integration of Partially Overlapping Web Sources*. PVLDB 6(10): 805-816 (2013)

Atti di Conferenze:

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

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

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*. KDD 2018: 263-272

Disheng Qiu, Luciano Barbosa, Valter Crescenzi, Paolo Merialdo, Divesh Srivastava: *Big Data Linkage for Product Specification Pages*. SIGMOD Conference 2018: 67-81

Matteo Cannaviccio, Denilson Barbosa, Paolo Merialdo: *Towards Annotating Relational Data on the Web with Language Models*. WWW 2018: 1307-1316

Donatella Firmani, Paolo Merialdo, Elena Nieddu, Simone Scardapane: *In Codice Ratio: OCR of Handwritten Latin Documents using Deep Convolutional Networks*. AI*CH@AI*IA 2017: 9-16

Serena Ammirati, Donatella Firmani, Marco Maiorino, Paolo Merialdo, Elena Nieddu, Andrea Rossi: *In Codice Ratio: Scalable Transcription of Historical Handwritten Documents*. SEBD 2017

Studenti di dottorato

Conclusi

2020 – Federico Piai: “Instance-level attribute alignment for heterogeneous product sources”.
Posizione attuale: post doc presso Università Roma Tre

2018 – Matteo Cannaviccio: “Knowledge Graphs Augmentation with Natural Language Evidence”.
Posizione attuale: Machine Learning Engineer at Diffbot (San Francisco, California)

2015 – Disheng Qiu: “Crowdsourcing Large scale Data Extraction from the Web: Bridging Unsupervised and Supervised Approaches”. Posizione attuale: co-founder e CTO di Wanderio spa

2012 – Mirko Bronzi: “Exploiting Information Redundancy to Collect Data from the Web”.
Posizione attuale: Applied Research Scientist at Mila - Quebec Artificial Intelligence Institute (Montreal, Canada)

2011 – Lorenzo Blanco: “Extraction, integration and probabilistic characterization of web data”.
Posizione attuale: Staff Software Engineer at DeepMind (London, UK)

In corso

2016-2021 Andrea Rossi. Thesis: Neural Knowledge graph augmentation

2017-2020 Elena Nieddu. Thesis: Automatic transcription of medieval manuscripts

ATTIVITA' DI GESTIONE

Finanziamenti

2018-2020 **Finanziamento:** Erasmus+ KA2 - Cooperation for innovation and the exchange of good practices in Entrepreneurship Development **Nome progetto:** “STARTED - STArtup skills for Researchers and innovaTors”. **Ruolo:** PI Roma Tre

2018-2019 **Finanziamento:** Programma Strategico Regionale per la ricerca, l'innovazione ed il trasferimento tecnologico 2017-2019 Legge Regionale n. 13 del 04/08/2008. **Nome progetto:** “In Codice Ratio”. **Ruolo:** PI

2018-2019 **Finanziamento:** FAMI - Fondo Asilo, Migrazione e Integrazione 2014-2020. **Nome progetto:** “Integr-Azioni: conoscere, gestire, comunicare”. **Ruolo:** PI Roma Tre

2014-2015 **Finanziamento:** POR FESR Lazio 2007/2013. **Nome progetto:** “SocialEyers: Sentiment Analysis Techniques for Real Time Social Media Monitoring”. **Ruolo:** PI Roma Tre

2011-2013 **Finanziamento:** POR FESR Lazio 2007/2013. **Nome progetto:** “MORNING:

Monitoraggio della Reputazione su social Network e Informazione Generata da utenti. **Ruolo:** PI Roma Tre

2010-2012 **Finanziamento:** MIUR - PRIN2008. **Nome progetto:** "EASE: Entity Aware Search Engines". **Ruolo:** PI Roma Tre unit

2004-2006 **Finanziamento:** MIUR - PRIN 2004. **Nome progetto:** "WISDOM: Intelligent Web Search based on Domain Ontologies". **Ruolo:** PI Roma Tre

Autorizzo il trattamento dei miei dati personali ai sensi dell'art. 13 d. lgs. 30 giugno 2003 n°196 – "Codice in materia di protezione dei dati personali" e dell'art. 13 GDPR 679/16 – "Regolamento europeo sulla protezione dei dati personali".

Roma, 28 luglio 2020

A handwritten signature in blue ink, appearing to read "Rob. Pistole".