

Pasquale Pavone, Ph.D.

Pagina web: www.textminer.it

Attuale posizione nell'ambito della ricerca:

Sono Docente di **Text Mining** al **Master in Big Data** dell'Università di Roma "La Sapienza"

Dal 2015 sono **membro del CAPP, Centro di Analisi delle Politiche Pubbliche**, Dipartimento di Economia "Marco Biagi" Università di Modena e Reggio Emilia, www.capp.unimore.it nel cui ambito svolgo progetti di ricerca su applicazioni di analisi automatica dei testi in progetti di ricerca economica e sociale di carattere multidisciplinare.

TEMI DI RICERCA

Metodi e tecniche di **Text Mining**; metodologie per il trattamento automatico dei dati testuali; **statistica testuale per l'analisi del contenuto**; trattamento del linguaggio naturale.

Algoritmi per l'estrazione dell'informazione, **classificazione e categorizzazione automatica dei documenti**.

Analisi delle domande aperte in indagini qualitative, focus-group, interviste, analisi del contenuto e del discorso, web-mining e analisi di mercato.

Costruzione di risorse statistico-linguistiche per l'**analisi automatica dei testi**.

FORMAZIONE

DOTTORE DI RICERCA IN ECONOMIA

Universidad de Màlaga, Facultad de Ciencias Económicas.

Titolo della Tesi: Un Método de Text Mining para la categorización Fuzzy de documentos.

Conseguito il 14 marzo 2015.

Direttori: Prof. Sergio Bolasco; Prof. Manuel Jimenez Berrocal.

DIPLOMA DI STUDI AVANZATI

Universidad de Màlaga, Facultad de Ciencias Económicas.

Corso di dottorato in "Gestión y Desarrollo Turístico Sostenible" (Spagna), 2004/2006.

LAUREA IN ECONOMIA E COMMERCIO, ORIENTAMENTO ECONOMICO GENERALE

Università degli Studi di Roma "La Sapienza", Facoltà di Economia.

Tesi: Un' applicazione di Text Mining all'indagine ISTAT sull' "Uso del Tempo"

Voto: 106/110

Relatore: Professor Sergio Bolasco, (A.A 2002-2003)

ASSEGNI DI RICERCA

16 APRILE 2022 – 15 APRILE 2023

Università degli Studi di Modena e Reggio Emilia, Dipartimento di Economia "Marco Biagi".

"Metodi di analisi automatica di dati testuali: applicazioni per la progettazione, l'analisi, il monitoraggio delle politiche pubbliche.

Tutor: Prof.ssa Margherita Russo

15 DICEMBRE 2019 – 14 DICEMBRE 2020

Scuola Universitaria Superiore Sant'Anna di Pisa. Department of Excellence, Economics and Management in the Era of Data Science (EMbeDS).

"Analisi empirica delle dinamiche di innovazione e crescita di impresa: un approccio Data Science".

Ambito scientifico disciplinare di Politica Economia

Tutor: Professor Federico Tamagni

01 MAGGIO 2019 – 14 DICEMBRE 2019

Scuola Universitaria Superiore Sant'Anna di Pisa. Department of Excellence, Economics and Management in the Era of Data Science (EMbeDS).

“Text Mining ed Analisi Multidimensionale di testi in linguaggio naturale”.

Tutor: Prof.ssa Maria Francesca Romano

16 MARZO 2018 – 15 MARZO 2019

Università degli Studi di Modena e Reggio Emilia, Dipartimento di Economia “Marco Biagi”.

“Analisi dei documenti per le Politiche Regionali”. Ambito scientifico disciplinare di Politica Economia.

Tutor: Prof.ssa Margherita Russo

01 FEBBRAIO 2016 – 31 GENNAIO 2017

Università degli Studi di Padova, Dipartimento FISPPA.

“Tracing the History of Words. A Portrait of Discipline Through Analysis of Keyword Count in Large Corpora of Scientific Literature”. Ambito scientifico disciplinare di Statistica Sociale.

Tutor: Prof.ssa Arjuna Tuzzi

15 DICEMBRE 2009 – 30 NOVEMBRE 2012 (per un totale di 28 mesi non consecutivi)

Scuola Superiore di Studi universitari Sant'Anna di Pisa.

“Costruzione di dizionari tematici finalizzato alla individuazione di mappe semantiche su specifiche aree tematiche”. Ambito scientifico disciplinare di Statistica Sociale.

Tutor: Prof.ssa Maria Francesca Romano

COLLABORAZIONI DI RICERCA

SCUOLA UNIVERSITARIA SUPERIORE S.ANNA DI PISA

Department of Excellence, Economics and Management in the Era of Data Science (EMbeDS).

Collaborazione di ricerca per lo svolgimento di attività di analisi del contenuto testuale del corpus di brevetti mondiali del database EPO-PATSTAT, nell'ambito del progetto GROWINPRO-Horizon2020. 04/01/2021 – 03/05/2021

Tutor: Professor Federico Tamagni

UNIVERSITÀ DEGLI STUDI DI ROMA “LA SAPIENZA”

Dipartimento di Scienze Statistiche

Contratto di collaborazione nell'ambito della ricerca sul tema *Web Attivo contro la Violenza* per la “Costruzione di un vocabolario della violenza utilizzando tecniche di analisi automatica di dati testuali non strutturati”. Responsabile Scientifico del Progetto di ricerca: Prof.ssa Fiorenza Deriu. 23/01/2019 – 23/04/2019

BRUNEL UNIVERSITY LONDON

Research Assistant for “Text Analysis based on techniques of textual statistics and tools of Text Mining”

Principal investigator: Dr. Ainurul Rosli. 15 gennaio - 15 febbraio 2019

UNIVERSITÀ DEGLI STUDI DI ROMA TRE

Contratto di collaborazione coordinata e continuativa.

Analisi con tecniche statistico-testuali della letteratura scientifica su terrorismo e radicalizzazione, nell'ambito del **Progetto H2020 “Trivalent”**. 01/12/2017 – 31/01/2018

SCUOLA UNIVERSITARIA SUPERIORE S.ANNA DI PISA

Analisi semantica di un Corpus di Sentenze. Progetto di Ricerca: “Text Mining e Big Data”.

1-30 novembre 2017.

EUROPEAN COMMISSION, JOINT RESEARCH CENTER (SEVILLA)

Text Mining for clustering and network analysis of techno Economic Segments Studies. 1-31 marzo 2017.
Coordinatore scientifico del gruppo di ricerca: Prof. Margherita Russo. Dipartimento "Marco Biagi",
Università di Modena e Reggio Emilia. 1-28/02/2017

UNIVERSITY OF WESTMINSTER

Research Assistant on the British-Academy Leverhulme funded project "Knowledge co-creation between universities and small and medium-sized enterprises: drivers and impact"
Principal investigator: Dr. Ainurul Rosli. 1-28 Febbraio 2017.

BORSA DI RICERCA, 16 FEBBRAIO – 15 AGOSTO 2015

Università degli Studi di Modena e Reggio Emilia.

"Analisi degli effetti sociali, economici e sanitari del sisma in Emilia" sul tema "Energie per la ricostruzione".
Dipartimento di Economia "Marco Biagi".

Tutor: Prof.ssa Margherita Russo.

UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA, progetti di ricerca del CAPP.

- STIP Covid watch
maggio-luglio 2021
- Electrive.com
aprile –novembre 2021
- Indagine studenti Unimore
maggio luglio 2020
- Indagine studenti EIDI
aprile-maggio 2020
- Analisi delle componenti testuali in occasione delle interviste all'indagine "Energie_400", da approfondire tramite analisi statistica multidimensionale. Progetto "Energie Sisma Emilia", 1-28 febbraio 2017.
- Analisi linguistica e statistica dei risultati delle attività volte alla realizzazione del Progetto "Poli.in-Poli di innovazione in Toscana", 15 gennaio 15 febbraio 2015.
- Analisi linguistica e statistica dei risultati delle attività volte alla realizzazione del Progetto "LEVE – Reti di competenze, istruzione e innovazione nella Provincia di Livorno", 1 gennaio – 31 marzo 2014;
- Analisi linguistica e statistica all'interno del Progetto "Metalnet" – Officina Emilia, 1 marzo – 30 aprile 2013.

CONSORZIO PER LO SVILUPPO DELLE METODOLOGIE E DELL'INNOVAZIONE DELLE PUBBLICHE AMMINISTRAZIONI (MIPA)

- Classificazione di bandi di gara e appalti della regione Puglia mediante la costruzione di variabili testuali con tecniche di Text Mining. Consorzio Mipa di Roma, 15 ottobre – 15 dicembre 2010.
- Costruzione di variabili testuali con tecniche di Text Mining sui bandi di gara e gli appalti relativi a lavori stradali della regione Puglia. Consorzio Mipa di Roma, 15 giugno – 15 luglio 2006.

UNIVERSITÀ DEGLI STUDI DI ROMA "LA SAPIENZA"

- Costruzione di risorse statistico-linguistiche per l'analisi automatica di un testo. Dipartimento di Studi Geoeconomici, Linguistici, Statistici, Storici per l'analisi regionale, 1 ottobre – 30 novembre 2006
- Elaborazione di dati con tecniche di Text Mining. Dipartimento di Studi Geoeconomici, Linguistici, Statistici, Storici per l'analisi regionale, 15 novembre – 15 dicembre 2005.

CONSULENZE

OECD – Organisation for Economic Co-Operation and Development. Directorate for Science, Technology and Innovation.

CONTRACT FOR INTELLECTUAL SERVICES 2021/STI/STP RELATIVE TO WORK ON SEMANTIC ANALYSIS AS PART OF THE WORKING PARTY ON INNOVATION AND TECHNOLOGY POLICY (TIP) COLLABORATIVE TRANSITIONS PROJECT.

31/05/2022-ONGOING

FONDAZIONE GIACOMO BRODOLINI

Creazione del corpus electrive.com, analisi statistica e creazione di strumenti per l'interpretazione dei risultati, nell'ambito della ricerca "Geopolitical and technological challenges for the automotive global value chains.

01/11/2021 – 26/11/2021

IRS – ISTITUTO PER LA RICERCA SOCIALE

Contratto di collaborazione all'interno del progetto CO.N.T.R.O *COunter Narratives AgainsT Racism Online* **coordinato da UNAR – Ufficio per la promozione della parità di trattamento e la rimozione delle discriminazioni fondate sulla razza o sull'origine etnica, della Presidenza del Consiglio dei ministri.**

Attività di analisi della metodologia attualmente in uso presso UNAR per l'identificazione e l'analisi dei messaggi d'odio a sfondo razziale online, al fine di predisporre un report per l'aggiornamento di detta metodologia.

30/05/2019 – 31/01/2020

AGT ENGINEERING, SOCIETÀ FORNITRICE DI SERVIZI DI INTERNAZIONALIZZAZIONE

Contratto di collaborazione per l'Analisi del mercato aerospaziale attraverso la definizione di un modello di estrazione di informazioni con tecniche di Text Mining da un set di schede aziendali provenienti dal web, dicembre 2014 – gennaio 2015.

DIPARTIMENTO DELLA FUNZIONE PUBBLICA DELLA PRESIDENZA DEL CONSIGLIO DEI MINISTRI DEL GOVERNO ITALIANO

Contratto di collaborazione per l'analisi automatica delle email ricevute in risposta ai quesiti posti dal Presidente del Governo, Matteo Renzi, e dal Ministro per la Semplificazione e la Pubblica Amministrazione, Marianna Madia.

Coordinatore scientifico del gruppo di ricerca Prof. Sergio Bolasco, Dipartimento Memotef dell'Università di Roma "La Sapienza", maggio-giugno 2014.

UNIVERSITÀ DEGLI STUDI DI MILANO "BICOCCA"

Contratto di collaborazione per la Content/textual analysis di bilanci di aziende europee per la valutazione della disclosure del capitale intellettuale. Dipartimento Di.Sea.De, ottobre - dicembre 2013.

GRUPPI E PROGETTI DI RICERCA

Partecipazione al Progetto di Ricerca scientifica (2021-2022) dal titolo "Semantic analysis of innovation policies", Resp. Scien. Prof. Margherita Russo (DEMB Unimore) con Caroline Paunov (OECD) e Dirk Meissner (HSE University).

Partecipazione al Progetto di Ricerca scientifica (2021), dal titolo "Electric Vehicles A-A Space", work package del progetto "Geopolitical and technological challenges for the automotive global value chains (GTC-auto)", Resp. Scien. Prof. Annamaria Simonazzi, Fondazione Brodolini. Finanziato da Institute for New Economic Thinking.

Partecipazione al Progetto di Ricerca scientifica (2020), dal titolo "Covid-19 e studenti UNIMORE: come l'emergenza cambia lo studio e l'esperienza universitaria". Resp. Scien. Prof. Margherita Russo, DEMB Unimore. Progetto finanziato dal CAPP (Centro Analisi delle Politiche Pubbliche)

Partecipazione al Progetto di Ricerca scientifica (2018-2021), dal titolo "Innovation intermediaries and the Fourth Industrial Revolution", Resp. Scien. Federica Rossi, Birckbeck University, London (UK). Progetto finanziato da BA/Leverhulme Small Research Grants, SRG18R1\180931

Partecipazione al Progetto di Ricerca scientifica (2019-2020), Universidad Nacional Autónoma de México, Faculty of Economy, dal titolo "Transformación digital en la cadena de suministro automotriz en México" (Digitalisation in the Automotive Supply Chain in Mexico). (Resp. Scien. Prof. Jorge Carreto Sanginés).

Partecipazione al Progetto di Ricerca Scientifica (2019-2020), Scuola Universitaria Superiore Sant'Anna di Pisa. Department of Excellence, Economics and Management in the Era of Data Science (EMbeDS), dal titolo "HARMONization and integrative analysis of regional, national and international Cohorts on primary Sjögren's Syndrome (pSS) towards improved stratification, treatment and health policy making disease". (Resp. scien. Prof.ssa Seghieri Chiara).

Partecipazione al Progetto di Ricerca Scientifica (2019), Università degli Studi di Roma "La Sapienza", Dipartimento di Scienze Statistiche, dal titolo Web Attivo contro la Violenza, (Resp. Prof.ssa Fiorenza Deriu)

Partecipazione al Progetto di Ricerca Scientifica (2018-2019), Università degli studi di Modena e Reggio Emilia (Resp. Prof.ssa Margherita Russo) dal titolo "R&I_Mapping", in the Alp-Gov EU Project.

Partecipazione al Progetto di Ricerca Scientifica (2016), Università degli studi di Modena e Reggio Emilia (Resp. Prof.ssa Margherita Russo) dal titolo "CLUES - CLUstering and network analysis of techno-Economic Segments Studies". Research project commissioned by JRC-Digital Economy Unit in the EU PREDICT Project

Partecipazione al Progetto di Ricerca Scientifica (2016), Università degli studi di Padova (Resp. Prof.ssa Arjuna Tuzzi) dal titolo "Tracing the History of Words. A portrait of a Discipline Through Analyses of Keyword Counts in Large Corpora of Scientific Literature. Progetto di Ricerca di Ateneo.

Partecipazione al Progetto di Ricerca Scientifica (2015), Università degli studi di Modena e Reggio Emilia (Resp. Prof.ssa Margherita Russo) dal titolo "Energie Sisma Emilia" www.energie.unimore.it, Call "Ricerca applicata", cofunded by FCRMO

Partecipazione al Progetto di Ricerca Scientifica (2014) Dipartimento Memotef università degli Studi "La Sapienza" di Roma (Resp.Prof. Sergio Bolasco) per l'analisi automatica delle email ricevute dal Presidente del Governo, Matteo Renzi, e dal ministro per la Semplificazione e la Pubblica Amministrazione, Marianna Madia. Dipartimento della Funzione Pubblica della Presidenza del Consiglio dei ministri del Governo italiano.

Partecipazione al Progetto di Ricerca Scientifica (2013) "Poli.in_Analisi e modellizzazione", inter-institutional research project Regione Toscana and Dipartimento di Economia Marco Biagi (Resp. Prof. Margherita Russo)

Partecipazione al Progetto di Ricerca Scientifica (2013) "LEVE_Reti di competenze, istruzione e innovazione nella Provincia di Livorno", Università degli studi di Modena e Reggio Emilia (Resp. Prof.ssa Margherita Russo).Call on "Ricerche, studi ed analisi a supporto dell'osservatorio provinciale sull'economia ed il mercato del lavoro", Provincia di Livorno, Lotto 5 "Ricerca sulla generazione delle reti di competenza e lo sviluppo locale" www.leve-livorno.unimore.it

Partecipazione al Progetto di Ricerca Scientifica (2013) "Metalnet" – Officina Emilia, Università degli studi di Modena e Reggio Emilia, (Resp. Prof.ssa Margherita Russo).
<http://www.metalnet.unimore.it/>

Partecipazione al Progetto di Ricerca Scientifica (2013) Di.Sea.De Università degli studi di Milano Bicocca (Resp. Prof. Angelo Di Gregorio) dal titolo "Content/Textual analysis di bilanci di aziende europee per la valutazione della disclosure del capitale intellettuale"

Partecipazione al Progetto di Ricerca Scientifica (2006; 2010), presso il MIPA Consorzio per lo Sviluppo e delle innovazioni nelle Pubbliche Amministrazioni, (Resp. Dott.ssa Siriana Salvi), per la Classificazione di bandi di gara e appalti della regione Puglia.

DOCENZE

Titolare di moduli di didattica

Master in Big Data

Università degli Studi Roma "La Sapienza"

Dipartimento di Scienze Statistiche

Titolare della docenza per il corso di "Metodi di Analisi automatica dei testi e soluzioni di Text Mining".

AA 2022-2023 (16 ore);

AA 2021-2022 (16 ore);

AA 2020-2021 (16 ore);

AA 2019-2020 (16 ore);

AA 2018-2019 (20 ore);

AA 2017-2018 (20 ore).

Coordinatrice del Master: Prof.ssa Fiorenza Deriu

Master Interateneo di II livello in "in Metodologia e Tecniche Avanzate di Ricerca Sociale"

Università degli Studi Roma "La Sapienza"

Dipartimento di Comunicazione e Ricerca Sociale

Titolare della docenza per il corso di "Laboratorio avanzato di archiviazione e analisi di Big Data testuali I e II".

AA 2021-2022 (8 Ore);

Coordinatore del Master: Prof. Sergio Mauceri

Master di I livello in "Data Science"

Università degli Studi Roma "Tor Vergata"

Titolare della docenza per il corso di "Applicazioni di Text Mining".

AA 2016-2017 (12 Ore);

AA 2015-2016 (10 ore).

Coordinatrice del Master: Prof.ssa Domenica Fioredistella Iezzi

Master in Fonti Strumenti e Metodi per la Ricerca Sociale

Università degli Studi Roma "La Sapienza"

Titolare della docenza per il corso di "Text Mining e analisi dei dati testuali".

AA 2013-2014 (12 ore);

AA 2012-2013 (12 ore);

AA 2011-2012 (12 ore).

Coordinatore del Master: Prof. Giovanni B. Sgritta.

Scuola Internazionale di Analisi dei Dati Testuali – Text Mining

Titolare della docenza per il corso di "Introduzione all'analisi automatica dei testi e ai Metodi e Modelli di Text Mining"

Dipartimento Memotef, Facoltà di Economia, Università degli studi di Roma "La Sapienza".

Terza edizione: **novembre 2013;**

Seconda edizione **febbraio 2012.**

Coordinatore del Corso: Prof. Sergio Bolasco.

Responsabile di moduli/attività formative a supporto della didattica

Tutor in Cina del "China Study Tour 2018" (Beijing and Shanghai, China)

Collaborazione nella conduzione del China Study Tour 2018, organizzato dal Dipartimento di Economia Marco Biagi, Università degli studi di Modena e Reggio Emilia.

Coordinatrici Prof.sse Cinzia Parolini e Margherita Russo.

AA 2017-2018 03/05/2018-02/06/2018 (attività equivalente a un insegnamento di 9 CFU)

Dipartimento di Studi Linguistici e letterari

Università degli Studi di Padova

Docenza formativa di supporto per l'insegnamento di Metodi Statistici di analisi dei testi

AA 2017-2018 (20 ore).

Titolare del Corso: Prof.ssa Arjuna Tuzzi

Corsi di formazione sul software Taltac2

Dipartimento Memotef, Facoltà di Economia, Università degli studi di Roma "La Sapienza".

Corsi base e avanzato, dal **2007 al 2013**.

Coordinatore dei corsi di formazione: Prof. Sergio Bolasco.

Cultore della materia

Università degli Studi di Roma "La Sapienza", Facoltà di Economia.

Membro della commissione di esami e verifiche del profitto per la materia:

Metodi Esplorativi per l'analisi dei dati e laboratorio di Data Mining. **AA 2013-2014**.

Titolare del corso: Prof. Sergio Bolasco.

PUBBLICAZIONI

Articoli in rivista con procedure di revisione (in neretto sono indicate le riviste di Classe A per il settore concorsuale dell'Area 13/D3, Statistica Sociale e Demografia)

- Morlini I., Russo M., Alboni F., Pavone P., Sartori L. (2022). *Learning from university students on changes prompted by emergency remote teaching: exploratory analysis of an empirical survey*, in *Statistica Applicata-Italian Journal of Applied Statistics* Vol. 34 (1). doi.org/10.26398/IJAS.0034-002.
- Russo M., Alboni F., Bonifati G., Carreto-Sanginés J., Pavone P., Simonazzi A. *Agents and artefacts in the emerging electric vehicle space*, in *International Journal of Automotive Technology and Management*, Vol.22 N.2. https://doi.org/10.1504/IJATM.2022.124369.
- Russo M., Caloffi A., Colovic A., Pavone P., Romeo S., Rossi F. (2021). *Mapping regional strengths in a key enabling technology: The distribution of Internet of Things competences across European regions*, in **Papers in Regional Science**. DOI: 10.1111/pirs.12679.
- Russo M., Pavone P. (2021). *Evidence-based portfolios of innovation policy mixes: A cross-country analysis*, in **Technological Forecasting & Social Change**. doi.org/10.1016/j.techfore.2021.120708
- Romano M.F. , Morgante G., Di Vetta G., Pavone P., Baldassarini A. *La misurazione della corruzione attraverso le sentenze della magistratura: una proposta metodologica con strumenti di text mining*, in *Federalismi.it – ISSN 1826-3534* |n. 33/2020.
- Pavone P., Pagliacci F., Russo M., Righi S., Giorgi A. (2020). *Multidimensional clustering of EU regions. A contribution to orient public policies in reducing regional disparities*. Special Issue "Data Science & Social Research", **Social Indicator Research**. DOI: 10.1007/s11205-020-02324-9
- Pagliacci F., Pavone P., Russo M., Giorgi A. (2019). *Regional structural heterogeneity: evidence and policy implications for RIS3 in macro-regional strategies*, in **Regional Studies** pp.1-11, DOI: 10.1080/00343404.2019.1635689.

- Russo M., Pavone P., Cetrulo A., (2019). *Conflict and Participation in Bargaining at Company Level: The Lamborghini Case*, in "Economia & lavoro, Rivista di politica sindacale, sociologia e relazioni industriali" 1/2019, pp. 53-74, DOI: 10.7384/93830.
- Doni F., Gasperini A., Pavone P. (2016). *Early adopters of integrated reporting: The case of the mining industry in South Africa*. African Journal of Business Management, Vol.10(9), pp. 187-208, May 2016. DOI: 10.5897/AJBM2016.8029. ISSN: 1993-8233. Article number 326647358341.
- Pavone P., Romano M.F., (2013). *Models of Text Mining to measure improvements to doctoral courses suggested by "STELLA" PHD survey respondents*. In Statistica, pp 463-475, anno LXXIII n°4, 2013, Bologna. DOI: 10.6092/issn.1973-2201/4500.

Capitoli in volumi

- Russo M., Pagliacci F., Pavone P., Giorgi A. (2019) *RIS3 in macro-regional strategies: tools to design and monitor integrated territorial development paths*. In: Building the next generation of research on territorial development. ESPON Scientific Report. ESPON 2019. ISBN: 978-99959-55-90-8.
- Pavone P. (2018). *Automatic Multiword Identification in a Specialist Corpus*. In: Tracing the Life Cycle of Ideas in the Humanities and Social Sciences, a cura di Tuzzi A., Springer, pp. 151-166. DOI 10.1007/978-3-319-97064-6. e-book ISBN 978-3-319-97064-6. ISSN 2199-0956.
- Russo M., Pavone P., Bigarelli D., Baracchi M. (2017). *La filiera automotive in Emilia-Romagna*. In: Osservatorio sulla componentistica automotive italiana 2017. Edizioni Ca'Foscari, 2017. pp. 145-172. DOI 10.14277/6969-193-5/RIIA-2-4 | Submission 2017-10-02. ISBN [ebook] 978-88-6969-193-5.
- Bolasco S, Pavone P. (2010). *Automatic dictionary and rule-based systems for extracting information from text*. In: Data Analysis and Classification, A cura di Palumbo F., Lauro C.N., Greenacre M.J., p. 189-198, BERLIN-HEIDELBERG: Springer, ISBN/ISSN: 978-3-642-03738-2, DOI: 10.1007/978-3-642-03739-9_22.
- Bolasco S., D'Avino E., Pavone P. (2007). *Analisi dei diari giornalieri con strumenti di statistica testuale e text mining*, in: "I tempi della vita quotidiana. Un approccio multidisciplinare all'analisi dell'uso del tempo." Maria Clelia Romano (a cura di). Roma, ISTAT, pp. 309-340. ISBN 978-88-458-1554-6.

Atti di convegno

- De Gasperis G., Pavone P. Bolasco S. (2022). *La Risorsa di Italiano Standard ad alta variabilità linguistica per misurare la peculiarità di un corpus*. JADT 2022 - PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON STATISTICAL ANALYSIS OF TEXTUAL DATA. Isbn 979-12-80153-30-2. Napoli, Italy.
- Romano M.F., De Gasperis G., Pavone P., Bolasco S. (2022). *Potenzialità di TaLTaC nella anonimizzazione di sentenze della magistratura*. JADT 2022 - PROCEEDINGS OF THE 16TH INTERNATIONAL CONFERENCE ON STATISTICAL ANALYSIS OF TEXTUAL DATA. Isbn 979-12-80153-30-2. Napoli, Italy.
- Alboni F., Pavone P., Russo M. (2022). *What do we learn by applying multiple methods in topic detection? A comparative analysis on a large online dataset about mobility electrification*. Solicited Session, SIS 2022 – 51st Scientific Meeting of the Italian Statistical Society. Caserta, Italy.
- Pavone P., Martinelli A., Tamagni F. (2021). *Text Mining on large corpora using Taltac4: An explorative analysis of the USPTO patents database*. SIS 2021 – Scientific Meeting of the Italian Statistical Society. Pisa, Italy.
- Deriu F., Navigli R., Pavone P. (2020). *Sentieri semantici della violenza: un algoritmo per l'individuazione di potenziali vittime*. JADT2020, Ed. Lexicometrica.
- Roma F., Comandè G., Amram D., Pavone P. (2020). *The challenges of legal analysis between text mining and machine learning*. JADT2020, Ed. Lexicometrica.
- Cetrulo A., Pavone P., Russo M. (2020). *From Wage Bargaining to the Negotiation on "Cognitive Surplus Value": reflections on a laboratory for Industry 4.0 in Lamborghini*. SASE, Society for the Advancement of Socio-economics, 32nd Annual Meeting.
- Antonelli F., Musolino S., Sodano P., Pavone P. (2017). *Literature Review: a meta analysis on Scientific Journals*. Conferenza sul tema "The challenges of radicalization: the role of the institutions, religious communities and civil society". Conference organized within the horizon project terrorism

prevention via radicalisation counternarrative (trivalent) coordinated by Centro di Eccellenza "Altiero Spinelli" Università di Roma tre, 24/11/2017.

- Rossi F., De Silva M., Rosli A., Yip N., Pavone P. (2017). *A text mining approach to appraising and analysing the impact of academic engagement*. Conference "Resource Competition, Research Assessment and Institutional Affiliations in Academia", Technical University Munich, Munich, Germany, 25-27 September 2017.
- Pavone P., Righi R., Righi S., Russo M., (2016). *Text mining and network analysis to support improvements in legislative action. The case of the earthquake in Emilia-Romagna*. JADT2016, 7-10 giugno 2016, Nizza, Francia.
- Pavone P., Dolcetti F., Canzonetti A., D'amen B., (2016). *Riforma della Pubblica Amministrazione italiana. Text mining e sentiment della consultazione pubblica*. JADT2016, 7-10 giugno 2016, Nizza, Francia. pp. 203-214. ISBN 978-2-7466-9067-7.
- Natilli M, Pavone P., Romano M.F., (2013). *The language of tourists in Wine and Food Blog*. In *Wine and Food Tourism – First European Conference*, Romano M. F., Natilli M, (eds). Collana Strumenti e Metodi Quantitativi, Edizioni ETS - Pisa, pp 267-286; ISBN 978-884673529-4.
- Doni F., Gasperini A., Pavone P. (2013). *The integrated report in the South African mining companies listed on the JSE Stock Exchange: analysis of non-financial information and impacts on external disclosure*. 36th Annual Congress European Accounting (EAA) (Paris, France 6-8 Maggio 2013).
- Doni F., Gasperini A., Pavone P. (2013). *Integrated report in South African mining companies listed on the Johannesburg stock exchange (JSE): Analysis of non-financial information*. In *EIASM 8th Interdisciplinary workshop on Intangibles, Intellectual Capital & extra financial information* (pp.1-38). Grenoble, Francia il 27-28 settembre 2012
- Bolasco S., Pavone P. (2012). *How the text mining measures complex phenomena in official statistics*, Proceedings of the XLVI Scientific Meeting of the Italian Statistical Society, Roma, 20-22 giugno 2012; ISBN 978-88-6129-882-8
- Bolasco S., Pavone P. (2011). *A Two-Level Fuzzy Classification between Different Profiles of the Gastronomic Lexicon*, in *CLADAG 2011. Book of abstract*, University of Pavia – Pavia University Press, Editoria Scientifica. ISBN: 978-88-96764-22-0
- Pavone P., Natilli M., Romano M.F. (2011). *The language of tourists in wine and food blog*, in *Wine and Food Tourism – First European Conference, Proceedings*. Edizioni ETS, pp. 67-69. ISBN 978-884672967-5
- Pavone P. (2010). *Sintagmazione del testo: una scelta per disambiguare la terminologia e ridurre le variabili di un'analisi del contenuto di un corpus*, in *Statistical Analysis of Textual Data. Proceedings of 10th International Conference*, JADT 2010. Edizioni Universitarie di Lettere Economia Diritto, vol 1, pp. 131-140. ISBN 978-88-7916-450-9
- Bolasco S., Pavone P. (2008). *Multi-class categorization based on cluster analysis and TFIDF*, in S. Heiden & B. Pincemin (eds.). *JADT2008*, Presses Universitaires de Lyon, vol. 1, pp. 209-218. ISBN 978-2-7297-0810-8.
- Bolasco S., Pavone P. (2007). *Automatic dictionary and rule-based systems for extracting information from text*, in *Classification and Data Analysis 2007. Book of short papers CLADAG2007*. EUM - Edizioni Università di Macerata, pp. 255-258. ISBN 978-88-6056-020-9.

Working Paper

- Russo, M., F., Alboni, S., Colombini, I., Morlini, P., Pavone e L., Sartori. "Covid-19 e studenti UNIMORE: come l'emergenza cambia lo studio e l'esperienza universitaria" Working paper, DEMB WORKING PAPER SERIES n°173, Dipartimento di Economia Marco Biagi - Università degli Studi di Modena e Reggio Emilia, 2020. https://doi.org/10.25431/11380_1203690
- Russo M., Pavone P., (2019) *What innovation policy mix does matter for which countries? Patterns emerging from multidimensional analysis on STIP Compass platform*. DEMB WORKING PAPER SERIES n.159, ISSN: 2281-440X online.
- Carreto Sanginés J., Pavone P., Russo M., Simonazzi A., (2019). *Digital upgrade in the automotive supply chain in Mexico: issues and challenges*. DEMB WORKING PAPER SERIES n.153, ISSN: 2281-440X online.

- Pavone P., Pagliacci F., Russo M., Giorgi A., (2019). "R&I smart specialisation strategies: classification of EU regions' priorities. Results from automatic text analysis," DEMB WORKING PAPER SERIES n.148, ISSN: 2281-440X online.
- Russo M., Pagliacci F., Pavone P., Giorgi A. (2019). RIS3 in macro-regional strategies: tools to design and monitor integrated territorial development paths. DEMB WORKING PAPER SERIES n.145. ISSN: 2281-440X online
- Russo M., Pavone P., Pagliacci F., Righi S., Giorgi A. (2019). Detecting multidimensional clustering across EU regions. Focus on R&I smart specialisation strategies and on socio-economic and demographic conditions. DEMB WORKING PAPER SERIES n.142. ISSN: 2281-440x online.
- Russo M., Pavone P., Cetrulo A. (2018). Cinquanta anni di contrattazione di secondo livello: che cosa impariamo dal caso Lamborghini? DEMB WORKING PAPER SERIES n.130. ISSN: 2281-440X online.
- Russo M., Pavone P. (2017). WPTIP 1993-2017: a longitudinal analysis of issues, connections and impact of the discussion on technology and innovation policy at OECD. DEMB WORKING PAPER SERIES n. 119. ISSN: 2281-440X online.
- Pavone P., Russo M. (2017). Strategie, pubbliche e private, in azione per ri-costruire meglio. Analisi dei testi di quattro interviste. DEMB-Università di Modena e Reggio Emilia. DEMB Working Paper Series n. 115. ISSN: 2281-440X online
- Pavone P., Russo M. (2017). Clusters of specializations in the automotive supply chain in Italy. An empirical analysis using text mining. DEMB-Università di Modena e Reggio Emilia. DEMB WORKING PAPER SERIES n. 116. ISSN: 2281-440X online
- Pavone P., Russo M. (2015). Analisi lessico testuale delle ordinanze del Commissario Delegato alla ricostruzione in Emilia-Romagna: un contributo alla legge nazionale su emergenza e ricostruzione, DEMB Working Paper Series n. 72, pp. 28, ISSN on line: 2281-440X

PRESENTAZIONI SU INVITO A WORKSHOP E SEMINARI

Relatore al workshop dal titolo "Semantic analysis of innovation policies: new tools and new research questions", Organizzato da OECD-WPTIP, 10 settembre 2021, evento online.

Relatore al seminario dal titolo "Big Data e Violenza di Genere. Metodologie a supporto dell'analisi del sentiment" su invito della Dott.ssa Claudia Villante, presso la sede dell'ISTAT. Roma, 7/02/2020.

Relatore al seminario dal titolo "Metodologie di individuazione e analisi automatica dei contenuti di odio online" su invito dell'IRS (Istituto per la Ricerca Sociale), presso la sede dell'UNAR – Ufficio Nazionale Antidiscriminazioni Razziali della Presidenza del Consiglio dei Ministri. Palazzo Chigi, Roma, 29/01/2020.

Relazione su invito al Workshop su "RIS3 in macro-regional strategies: building a comparative framework to learn from other regions ", EUSALP Action Group 1, Palazzo Lombardia, Milano, 25/06/2019.

Relazione su invito al Workshop su "Semantic analysis for Innovation Policies", OECD CST and WPTIP Paris, France, 12-13/03/2018.

Relazione su invito al Workshop su "Techno-economic segments (TES) analysis", EU project PREDICT, JRC-European Commission, Seville, Spain, 14/02/2018.

Seminari su invito riguardo l'Analisi automatica dei Testi:

- University of International Business and Economics, Beijing, China Study Tour, 20/05/2018.
- Istituto Zooprofilattico Sperimentale delle Venezie. Padova. 26-27-28/09/2016.
- Dipartimento di Scienze Economiche e Statistiche, Università "Federico II" di Napoli. 03/12/2015. Su invito della Prof.ssa Simona Balbi.
- Dipartimento di Economia "Marco Biagi", Università di Modena e Reggio Emilia. 21/04/2015. Su invito della Prof.ssa Margherita Russo.
- Scuola Superiore di Studi Universitari e di Perfezionamento Sant'Anna di Pisa, 17/01/2011

Su invito della Prof.ssa Maria Francesca Romano.

PRESENTAZIONI COME RELATORE A CONVEGNI

Contributi presentati dal 2005 al 2020

- Con De Gasperis G., Bolasco S. *La Risorsa di Italiano Standard ad alta variabilità linguistica per misurare la peculiarità di un corpus*. JADT 2022 - Napoli, Italy, July 6-8 2022.
- Con Alboni F., Russo M., *What do we learn by applying multiple methods in topic detection? A comparative analysis on a large online dataset about mobility electrification*. Solicited Session, SIS 2022 – 51st Scientific Meeting of the Italian Statistical Society. Caserta, Italy, June 22-24 2022.
- Con Martinelli A., Tamagni F. *Text mining su corpora di grandi dimensioni utilizzando Taltac4: Un'analisi esplorativa del database dei brevetti USPTO*. SIS2021, 21-25 giugno 2021, Università di Pisa.
- Con Russo M., Righi S., Caloffi A., Rossi F., (2019). *Mapping specialisations in the automotive international trade: A multilayer network analysis*. ARS Social Network Analysis. University of Salerno, 28-31/10/2019.
- Con Russo M., Pagliacci F., Righi S., Giorgi A., (2019). *Detecting multidimensional clustering across EU regions*. DSSR Data Science & Social Research. Milano, 4-5/02/2019.
- Con Righi R., Righi S., Russo M., (2016). *Text mining and network analysis to support improvements in legislative action. The case of the earthquake in Emilia-Romagna*. JADT2016, 7-10 giugno 2016, Nizza, Francia.
- Con Dolcetti F., Canzonetti A., D'amen B. (2016). *Riforma della Pubblica Amministrazione italiana. Text mining e sentiment della consultazione pubblica*. JADT2016, 7-10 giugno 2016, Nizza, Francia.
- Con Bolasco S., (2012). *How the text mining measures complex phenomena in official statistic*. XLVI Scientific Meeting of the Italian Statistical Society, Roma, 20-22 giugno 2012;
- Con Romano M.F., Natilli M., (2011). *The language of tourists in Wine and Food Blog*. In Wine and Food Tourism – First European Conference, Pisa. Volterra, 13-15/04/2011
- Con Bolasco S., (2011). *A Two-Level Fuzzy Classification between Different Profiles of the Gastronomic Lexicon*, CLADAG 2011, University of Pavia. 7-9/09/2011 Pavia.
- (2010). *Sintagmazione del testo: una scelta per disambiguare la terminologia e ridurre le variabili di un'analisi del contenuto di un corpus*, JADT 2010, 09-11/06/2010 Roma.
- Con Bolasco S., (2008). *Multi-class categorization based on cluster analysis and TFIDF*, JADT2008, 12-14/03/2008 Lyon, Francia
- Con Bolasco S., (2007). *Automatic dictionary and rule-based systems for extracting information from text*, in Classification and Data Analysis 2007. CLADAG2007. EUM, 12-14/09/2007 Macerata
- Con Bolasco S. D'Avino E., (2005). *Analisi lessicale dei diari giornalieri con strumenti di statistica testuale e text mining*. Relazione su invito al convegno sul tema "I tempi della vita quotidiana", Istat, Roma, 20 dicembre 2005.

Competenze informatiche specifiche

Nel corso degli anni di ricerca, sono state acquisite competenze avanzate riguardo i strumenti informatici:

- TaLTaC
- Spad
- IramuteQ
- Gephi
- TXM
- Tableau
- Lexico
- R
- Python