

CURRICULUM VITAE ET STUDIORUM

Andrea Esuli

Personal Information

- Born December 26th, 1977, in Pisa, Italy.
- Resident in Pisa, Italy.
- Married (2003), two daughters (2006, 2009).



Career Summary

- (September 2011 - current) **Research scientist** at Istituto di Scienza e Tecnologia dell'Informazione, Consiglio Nazionale delle Ricerche (ISTI-CNR) in Pisa.
- (January 2010 - August 2011) **Research fellow** at ISTI-CNR in Pisa.
- (January 2008 - December 2009) **Post-doc** at ISTI-CNR in Pisa.
- (January 2005 - December 2007) **Ph.D.** student at ISTI-CNR in Pisa.
- (January 2002 - December 2004) **Research grant holder** at Dipartimento di Informatica, Università di Pisa.

Education

- **Ph.D. in Information Engineering** (2008)
Università di Pisa, June 2008.
Dissertation: “*Automatic Generation of Lexical Resources for Opinion Mining: Models, Algorithms and Applications*”
Advisors: Prof. Luca Simoncini (Dipartimento di Ingegneria dell'Informazione, Università di Pisa), Dott. Fabrizio Sebastiani (Istituto di Scienza e Tecnologie dell'Informazione, CNR).
- **Master Degree in Computer Science Technologies** (2003)
Laurea Specialistica in Tecnologie Informatiche, 110/110, con lode (full marks, cum laude).
Università di Pisa, Dicembre 2003.
- **Master Degree in Computer Science** (2001)
Laurea in Informatica, 110/110, con lode (full marks, cum laude).
Università di Pisa, October 2001.
Dissertation: “*Progettazione e realizzazione di Interfacce multimediali intelligenti per la presentazione di scenari dinamici*”
“Design and implementation of multimedia intelligent interfaces for the presentation of dynamic scenarios”
Advisors: Prof.ssa Maria Simi (Dipartimento di Informatica, Università di Pisa), Prof. Giuliano Pacini (Accademia Navale di Livorno).

Awards

- **ERCIM Cor Baayen Award 2010**

ERCIM - the European Research Consortium for Informatics and Mathematics, which gathers 21 european national research institutions - awards each year the Cor Baayen Award to a promising young researcher in computer science and applied mathematics. The award was created in 1995 to honour the first ERCIM President. The award consists of a cheque for 5000 Euro together with an award certificate.

Andrea Esuli has been the first Italian to win this award in sixteen editions.

Research projects

- **ASSETS** (ISTI-CNR, started 05/2010)

Advanced Service Search and Enhancing Technological Solutions for the European Digital Library European FP7 ICT PSP (2010-2011).

ASSETS is a 2-year project that aims to improve the usability of the Europeana portal (<http://www.europeana.eu>) by developing, implementing and deploying software services focused on search, browsing and interfaces. All software services created by ASSETS will be available as Open Source code.

Andrea Esuli is responsible for tasks 2.1.1 Metadata Cleaning, 2.1.2 Knowledge Extraction, and 2.1.3 Metadata Classification, which all aim to provide machine learning-based methods for the automatic enrichment of metadata records composing the Europeana's collections.

- **VCS** (ISTI-CNR, started 10/2005)

Andrea Esuli is the technical leader in the development of the Verbatim Coding System (VCS), an automated system for verbatim coding, an important phase of customer relationship management (CRM) which consists in the classification of the customers of a company based on the textual responses that they have returned to questions concerning the satisfaction in the product or service purchased by the company.

VCS has been originall developed for the Egg Group plc, one of the world's largest purely online banks (3.2 million customers), to automate their CRM department.

Currently VCS is the core technology adopted by Language Logic llc, Cincinnati, USA, for the automatic coding of verbatim responses. Language Logic is the world leader in providing verbatim management services, with millions of verbatims processed every month.

VCS has been awarded the Vespucci Prize 2006 in the "Successful innovation in marketing" section. The Vespucci prize, now in its 3rd edition, is being awarded by Consiglio Regionale della Toscana, Confindustria Toscana, Federazione Regionale dei Cavalieri del Lavoro della Toscana, Banca Cassa di Risparmio di Firenze, Ente Cassa di Risparmio di Firenze, and is awarded "to particularly significant and innovative initiatives and projects, capable of combining ethics, invention, research and economic results with quality, development, and innovation."

In 2007 VCS has been awarded with the "Best New Thinking Award" of the Market Research Society, the greatest association of professionals in market research.

Relevant publications: (ECIR2009b), (ECIR2009a), (ICTIR2009), (IR2008), (SPIRE2006a), (SPIRE2006b).

- **MiningForLife** (ISTI-CNR, started 05/2009, ended 04/2010)

Technological transfer project, funded by NoemaLife Spa.

The focus of the project is the development of methods for the automatic processing of medical reports of patients, in order to allow the automatic population of structured DBs and simplify the data exchange between different medical specialists.

Andrea Esuli is the workpackage leader of the “automatic text classification” and the “information extraction” workpackages.

- **SAPIR** (ISTI-CNR, started 01/2007, ended 06/2009)
European project, IST FP6 “SAPIR: Search In Audio Visual Content Using Peer-to-peer IR”. Andrea Esuli contributed to the development and analysis of the CoPhiR test collection, the largest image metadata collection currently available for research purposes, with 106 million images. Andrea Esuli individually invented a novel similarity search algorithm that could effectively and efficiently work on such large collections.
Relevant publications: (CORR2009), (VLDL2009), (SISAP2009), (LSDSIR2009).
- **OntoText** (ISTI-CNR, started 10/2004, ended 03/2008)
Research project funded by the Provincia Autonoma di Trento. Theme: “From Text to Knowledge for the Semantic Web”. Andrea Esuli has been the technical leader for ISTI-CNR, working on the workpackage on “Opinion Mining”.
Relevant publications: (CIKM2005), (EACL2006), (LREC2006), (LTC2007a), (LTC2007b), (ACL2007), (LREC2008).
- **OpTex** (ISTI-CNR, started 12/2004, ended 12/2006)
Curiosity-driven project funded by ISTI-CNR on sentiment classification.
Relevant publications: (CIKM2005), (EACL2006), (LREC2006).
- **ClickWorld** (Università di Pisa, started 01/2002, ended 12/2003)
Industrial research project funded by the Ministero dell’Istruzione, dell’Università e della Ricerca.
Theme: Customized Proactive Web. Andrea Esuli worked on the theme “techniques for indexing, automatic classification, linguistic analysis, and information extraction, for the selection and proactive presentation of customized Web content, based on user profiles”.
Relevant publications: (WWW2004).
- **ECD** (Università di Pisa, started 01/2002, ended 12/2005)
“Technologies and Services for Enhanced Content Delivery”, funded by “Fondo Speciale Innovazione 2000”, Ministero dell’Istruzione, dell’Università e della Ricerca.
Andrea Esuli has taken part to the “Action 2: Mine the Web”, working on the theme: “Participatory search”.
Relevant publications: (WWWI2004).

Professional Service

- PC member of **SAC 2012**, “Information Access and Retrieval” track, 27th Symposium On Applied Computing, Riva del Garda (Trento), Italy, March 26-30, 2012.
- Tutorial Chair of **SPIRE 2011**, 18th Symposium on String Processing and Information Retrieval, Pisa, Italy, October 17-21, 2011.
- PC member of **ACL-HTL 2011**, 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Portland, Oregon, USA, June 19-24, 2011.
- PC member of **SPIRE 2010**, 17th Symposium on String Processing and Information Retrieval, Los Cabos, Mexico, October 11-13, 2010.
- PC member of **SIGIR 2010**, 33rd Annual ACM SIGIR Conference, Geneva, Switzerland, July 19-23, 2010.

- PC member of **SIGIR 2010 Poster and Demo session**, 33rd Annual ACM SIGIR Conference, Geneva, Switzerland, July 19-23, 2010.
- PC member of **ACL 2010**, 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, July 11-16, 2010.
- PC member of **NAACL-HLT 2010**, “Sentiment” area, 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics, Los Angeles, USA, June 1-6, 2010.
- PC member of **NAACL 2010 CAAGET Workshop**, workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, USA, June 5, 2010.
- PC member of **TextGraph-5 2010**, 5th ACL workshop on Graph-based Methods for Natural Language Processing.
- PC member of **SAC 2010**, “Information Access and Retrieval” track, del 25th ACM Symposium On Applied Computing, Crans-Montana, Switzerland, March 22-26, 2010.
- PC member of **EMNLP 2009**, “Sentiment Analysis, Opinion Mining and Text Classification” area, Conference on Empirical Methods in Natural Language Processing, Singapore, August 6-7, 2009.
- PC member of **EMNLP 2009**, “NLP applications” area, Conference on Empirical Methods in Natural Language Processing, Singapore, August 6-7, 2009.
- PC member of **ACL-IJCNLP 2009**, “Sentiment Analysis, Opinion Mining and Text Classification” area, 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Singapore, August 2-7, 2009.
- PC member of **CAW2.0 2009**, Content Analysis 2.0 Workshop, co-located with the ACM WWW Conference, Madrid, Spain, April 21, 2009.
- PC member of **SIGIR 2009**, 32nd Annual ACM SIGIR Conference, Boston, USA, July 19-23, 2009.
- PC member of **SAC 2009**, “Information Access and Retrieval” track, 24th ACM Symposium On Applied Computing, Honolulu, Hawaii, USA, March 8 - 12, 2009.
- PC member of **ECIR 2008**, 30th European Conference on Information Retrieval, Glasgow, Scotland, 30th March - 3rd April, 2008.
- PC member of **ECIR 2008 Poster session**, 30th European Conference on Information Retrieval, Glasgow, Scotland, 30th March - 3rd April, 2008.
- PC member of **SAC 2008**, “Information Access and Retrieval” track, 23th ACM Symposium On Applied Computing, Cear, Brazil, March 16 - 20, 2008.
- PC member of **SAC 2007**, “Information Access and Retrieval” track, 22th ACM Symposium On Applied Computing, Seoul, Korea, March 11 - 15, 2007.
- PC member of **SIGIR 2007 Poster and Demo session**, 30th Annual International ACM SIGIR Conference 23-27 July 2007, Amsterdam.
- PC member of **COLING ACL 2006**, Joint Conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics, Sydney, Australia, July 17-21, 2006.

Reviewing activity

- International project proposal reviewer:
 - **IWT**, *agentschap voor Innovatie door Wetenschap en Technologie* (research and innovation funding agency of the Flemish government).
- International journals:
 - **LRE**, *Language Resources and Evaluation*, Springer, ISSN: 1574-020X.
 - **IPM**, *Information Processing & Management*, Elsevier, ISSN: 0306-4573.
 - **IR**, *Information Retrieval*, Springer, ISSN: 1386-4564.
 - **IEEE TPAMI**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
 - **IEEE TAFFC**, *IEEE Transactions on Affective Computing*, ISSN: 0162-8828.
 - **IEEE TKDE**, *IEEE Transactions on Knowledge & Data Engineering*, ISSN: 1041-4347.
 - **IEEE IS**, *IEEE Intelligent Systems*, ISSN: 1541-1672.
 - **JIS**, *Journal of Information Science*, SAGE, ISSN: 0165-5515.

Advisor activity

- Ph.D. students:
 - (2010-current) Giacomo Berardi. Topics in Automatic Text Classification.
 - (2009-current) Stefano Baccianella. Topics in Sentiment Analysis.
- M.S. and B.S. students:
 - (2009) Diego Marcheggiani (M.S.). Active Learning method for Information Extraction
 - (2008) Stefano Baccianella (M.S.). Ordinal Regression methods for Sentiment Analysis.
 - (2007) Claudio Petroni (B.S.). Semi-supervised Boosting algorithms.
 - (2006) Sebastiano Brafa (B.S.). Random-walk method for Sentiment Analysis.

Invited Talks

- **Panel on WordNet relations, Global WordNet Conference 2008**
Event: “Panel on WordNet relations”, Global WordNet Conference, Szeged, Hungary, 22-25 January 2008.
Presentation: “Annotating WordNet Synsets by Sentiment-Related Information: Issues and Potential Solutions”.
- **Reading Group su Language and Intelligence, Ph.D. course in Computer Science, Università di Pisa**
Event: Reading Group on Language and Intelligence, Università di Pisa, invited seminar, 14/6/2006.
Presentation: “Opinion Mining”.

H-Index

12, measured on 10/09/2011 on Google Scholar.

Publications

- **International Journals** (peer-reviewed)

- (IPM2011) Andrea Esuli
Use of permutation prefixes for efficient and scalable approximate similarity search
Information Processing & Management, Available online 8 January 2011, ISSN 0306-4573, 10.1016/j.ipm.2010.11.011.
- (IJMR10) Andrea Esuli and Fabrizio Sebastiani
Machines that Learn how to Code Open-Ended Survey Data
International Journal of Market Research, 52(6):775-800, 2010
- (IEEEIS2010) Andrea Esuli, Fabrizio Sebastiani
Sentiment Quantification
Invited contribution to IEEE Intelligent Systems, 25(4), pages 72-75, IEEE Computer Society, Los Alamitos, CA, USA, ISSN 1541-1672, 2010.
- (IR2008) Andrea Esuli, Tiziano Fagni and Fabrizio Sebastiani
Boosting multi-label hierarchical text categorization
Information Retrieval, 11(4), Springer Netherlands, pages 287-313, ISSN 1386-4564, 2008.

- **Book chapters** (peer-reviewed)

- (LNCS2011) Andrea Esuli and Fabrizio Sebastiani
Enhancing Opinion Extraction by Automatically Annotated Lexical Resources (Extended Version)
Lecture Notes in Computer Science n. 6562, “Human Language Technology. Challenges for Computer Science and Linguistics”, Springer Verlag, Heidelberg, DE, 2011, pp. 500-511.
- (LNAI2009) Shlomo Argamon, Kenneth Bloom, Andrea Esuli and Fabrizio Sebastiani
Automatically Determining Attitude Type and Force for Sentiment Analysis
Lecture Notes in Computer Science, n. 5603, “Human Language Technology. Challenges of the Information Society”, Springer Berlin/Heidelberg, DE, pages 218-231, ISBN 978-3-642-04234-8, 2009.

- **Proceeding of international conferences** (peer-reviewed)

- (CLEF2010) Andrea Esuli and Fabrizio Sebastiani
Evaluating information extraction
Proceedings of the Conference on Multilingual and Multimodal Information Access Evaluation, CLEF 2010, Padova, IT, 2010.
Lecture Notes in Computer Science 6360, pages 100-111, Springer Verlag, Heidelberg, DE.
- (SAC2010) Stefano Baccianella, Andrea Esuli and Fabrizio Sebastiani
Feature Selection for Ordinal Text Classification
Proceedings of the 25th Symposium On Applied Computing, Crans-Montana, Switzerland, 2010.
ACM, New York, NY, USA, pages 1750-1756, ISBN 978-364204416-8, 2010.
- (LREC2010) Stefano Baccianella, Andrea Esuli and Fabrizio Sebastiani
SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining
Proceedings of the 7th Conference on Language Resources and Evaluation, LREC 2010, Valletta, MT, 2010.
European Language Resources Association (ELRA), Paris, FR, 2010.

- (ISDA2009) Stefano Baccianella, Andrea Esuli and Fabrizio Sebastiani
Evaluation Measures for Ordinal Text Classification
 Proceedings of the 9th IEEE International Conference on Intelligent Systems Design and Applications, ISDA 2009, Pisa, IT, 2009.
IEEE Computer Society, Los Alamitos, CA, USA, pages 283-287, ISBN 978-076953872-3, 2009.
- (LTC2009) Andrea Esuli and Fabrizio Sebastiani
Enhancing Opinion Extraction by Automatically Annotated Lexical Resources
 Proceedings of the 4th Language & Technology Conference, LTC 2009, Poznan, PL, November 6-8, 2009.
Wydawnictwo Poznanskie Sp. z o. o., Poznan, PL, pages 224-228, ISBN 978-83-7177-746-2, 2009.
- (ECIR2009c) Stefano Baccianella, Andrea Esuli and Fabrizio Sebastiani
Multi-facet Rating of Product Reviews
 Advances in Information Retrieval, Proceedings of the 31th European Conference on IR Research, ECIR 2009, Toulouse, France, April 6-9, 2009.
Lecture Notes in Computer Science, n. 5478, Springer Verlag, pages 461-472, ISBN 978-3-642-00957-0, 2009.
- (ECIR2009b) Andrea Esuli and Fabrizio Sebastiani
Encoding Ordinal Features into Binary Features for Text Classification
 Advances in Information Retrieval, Proceedings of the 31th European Conference on IR Research, ECIR 2009, Toulouse, France, April 6-9, 2009.
Lecture Notes in Computer Science, n. 5478, Springer Verlag, pages 771-775, ISBN 978-3-642-00957-0, 2009.
- (ECIR2009a) Andrea Esuli and Fabrizio Sebastiani
Active Learning Strategies for Multi-Label Text Classification
 Advances in Information Retrieval, Proceedings of the 31th European Conference on IR Research, ECIR 2009, Toulouse, France, April 6-9, 2009.
Lecture Notes in Computer Science, n. 5478, Springer Verlag, pages 102-113, ISBN 978-3-642-00957-0, 2009.
- (ICTIR2009) Andrea Esuli and Fabrizio Sebastiani
Training Data Cleaning for Text Classification
 Advances in Information Retrieval Theory, Proceedings of the Second International Conference on the Theory of Information Retrieval, ICTIR 2009, Cambridge, UK, September 10-12, 2009.
Lecture Notes in Computer Science, n. 5766, Springer Verlag, pages 29-41, ISBN 978-3-642-04416-8, 2009.
- (LREC2008) Andrea Esuli, Fabrizio Sebastiani and Ilaria Clara Urciuoli
Annotating Expressions of Opinion and Emotion in the Italian Content Annotation Bank
 Proceedings of the Sixth International Language Resources and Evaluation, LREC 2008, Marrakech, Morocco, May 28-30, 2008.
European Language Resources Association (ELRA), Paris, FR, ISBN 2-9517408-4-0, 2008.
- (ACL2007) Andrea Esuli and Fabrizio Sebastiani
PageRanking WordNet Synsets: An Application to Opinion Mining
 Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, ACL 2007, Prague, Czech Republic, June 23-30, 2007.
The Association for Computer Linguistics, East Stroudsburg, PA, USA, pages 424-431, ISBN 978-1-932432-86-2, 2007.
- (LTC2007b) Shlomo Argamon, Kenneth Bloom, Andrea Esuli and Fabrizio Sebastiani
Automatically Determining Attitude Type and Force for Sentiment Analysis

Proceedings of the 3rd Language & Technology Conference, LTC 2007, Poznan, PL, October 5-7, 2007.

Wydawnictwo Poznanskie Sp. z o. o., Poznan, PL, pages 369-373, ISBN 978-83-7177-407-2, 2007.

- (LTC2007a) Andrea Esuli and Fabrizio Sebastiani
Random-Walk Models of Term Semantics: An Application to Opinion-Related Properties
Proceedings of the 3rd Language & Technology Conference, LTC 2007, Poznan, PL, October 5-7, 2007.
Wydawnictwo Poznanskie Sp. z o. o., Poznan, PL, pages 221-225, ISBN 978-83-7177-407-2, 2007.
- (SPIRE2006b) Andrea Esuli, Tiziano Fagni and Fabrizio Sebastiani
MP-Boost: A Multiple-Pivot Boosting Algorithm and Its Application to Text Categorization
Proceeding of the 13th International Conference on String Processing and Information Retrieval, SPIRE 2006, Glasgow, UK, October 11-13, 2006.
Lecture Notes in Computer Science, n. 4209, Springer Berlin/Heidelberg, DE, pages 1-12, ISBN 978-3-540-45774-9, 2006.
- (SPIRE2006a) Andrea Esuli, Tiziano Fagni and Fabrizio Sebastiani
TreeBoost.MH: A Boosting Algorithm for Multi-label Hierarchical Text Categorization
Proceeding of the 13th International Conference on String Processing and Information Retrieval, SPIRE 2006, Glasgow, UK, October 11-13, 2006.
Lecture Notes in Computer Science, n. 4209, Springer Berlin/Heidelberg, DE, pages 13-24, ISBN 978-3-540-45774-9, 2006.
- (EACL2006) Andrea Esuli and Fabrizio Sebastiani
Determining Term Subjectivity and Term Orientation for Opinion Mining
Proceedings of 11st Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of the Conference, EACL 2006, Trento, Italy, April 3-7, 2006.
The Association for Computer Linguistics, East Stroudsburg, PA, USA, pages 193-200, ISBN 1-932432-59-0, 2006.
- (LREC2006) Andrea Esuli and Fabrizio Sebastiani
SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining
Proceedings of the 5th Conference on Language Resources and Evaluation, LREC 2006, Genova, IT, 24-26 May, 2006.
European Language Resources Association (ELRA), Paris, FR, pages 417-422, 2006.
- (CIKM2005) Andrea Esuli and Fabrizio Sebastiani
Determining the semantic orientation of terms through gloss analysis
Proceedings of the 14th ACM international conference on Information and knowledge management, CIKM 2005, Bremen, Germany, October 31 - November 5, 2005.
ACM, New York, NY, USA, pages 617-624, ISBN 1-59593-140-6, 2005.
- (TREC2004) Giuseppe Attardi, Andrea Esuli and Patel Chirag
Using Clustering and Blade Clusters in the TeraByte task
Proceedings of the Thirteenth Text REtrieval Conference, TREC 2004, Gaithersburg, Maryland, November 16-19, 2004.
National Institute of Standards and Technology (NIST), Special Publication 500-261, 2004.
- (WWWI2004) Giuseppe Attardi, Andrea Esuli, Lucian Hancu and Maria Simi
Participatory Search
Proceedings of the IADIS International Conference WWW/Internet 2004, Madrid, Spain, October 6-8, 2004. **IADIS**, Lisboa, PT, pages 643-650, ISBN 972-99353-0-0, 2004.
- (WWW2004) Giuseppe Attardi, Andrea Esuli and Maria Simi
Best bets: thousands of queries in search of a client

Proceedings of the 13th international conference on World Wide Web - Alternate Track Papers & Posters, WWW 2004, New York, NY, USA, May 17-20, 2004.
ACM, New York, NY, USA, pages 422-423, ISBN 1-58113-912-8, 2004.

- (WEB3D2003) Andrea Esuli, Antonio Cisternino, Giuliano Pacini, and Maria Simi
Multimodal presentation of dynamic object scenarios on the web
Proceeding of the Eighth International Conference on 3D Web Technology, Web3D 2003, Saint Malo, France, March 9-12, 2003
ACM, New York, NY, USA, pages 147-153, ISBN 1-58113-644-7, 2003.
- (AH2002) Andrea Esuli, Antonio Cisternino, Giuliano Pacini and Maria Simi
Adaptive Real Time Comment Generation for Sail Racing Scenarios
Adaptive Hypermedia and Adaptive Web-Based Systems, Second International Conference, AH 2002, Malaga, Spain, May 29-31, 2002.
Lecture Notes in Computer Science, n. 2347, Springer Berlin/Heidelberg, DE, pages 392-396, ISBN 978-3-540-43737-6, 2002.

• **Proceedings of international workshops** (peer-reviewed)

- (SEMEVAL2010) Andrea Esuli, Diego Marcheggiani and Fabrizio Sebastiani
ISTI@SemEval-2 Task #8: Boosting-Based Multiway Relation Classification
Proceedings of the 5th International Workshop on Semantic Evaluation (SEMEVAL10), Uppsala, SE, 2010.
- (VLDDL2009) Paolo Bolettieri, Andrea Esuli, Fabrizio Falchi, Claudio Lucchese, Raffaele Perego, Fausto Rabitti
Enabling Content-Based Image Retrieval in Very Large Digital Libraries
Second Workshop on Very Large Digital Libraries, VLDDL 2009, Corfu, Greece, October 2, 2009.
DELOS: an Association for Digital Libraries, Pisa, Italy, pages 43-50, ISBN 978-888850685-2, 2009.
- (LSDSIR2009) Andrea Esuli
PP-Index: Using Permutation Prefixes for Efficient and Scalable Approximate Similarity Search
Proceedings of the 7th Workshop on Large-Scale Distributed Systems for Information Retrieval co-located with ACM SIGIR 2009, LSDS-IR 2009, Boston, USA, July 23, 2009.
CEUR-WS, Vol-480, pages 17-24, 2009.
- (SISAP2009) Andrea Esuli
MiPai: Using the PP-Index to Build an Efficient and Scalable Similarity Search System
Proceedings of the 2nd International Workshop on Similarity Search and Applications, SISAP 2009, Prague, Czech Republic, August 29-30, 2009.
IEEE Computer Society, Los Alamitos, CA, USA, pages 146-148, ISBN 978-0-7695-3765-8, 2009.

• **Proceedings of national workshops** (peer-reviewed)

- (IIR2011) Giacomo Berardi, Andrea Esuli, Fabrizio Sebastiani and Fabrizio Silvestri
Blog Distillation via Sentiment-Sensitive Link Analysis
Proceedings of the 2nd Italian Information Retrieval Workshop (IIR11), Milano, IT, 2011.
- (IIR2011b) Stefano Baccianella, Andrea Esuli and Fabrizio Sebastiani
Using Micro-Documents for Feature Selection: The Case of Ordinal Text Classification
Proceedings of the 2nd Italian Information Retrieval Workshop (IIR11), Milano, IT, 2011.
- (SEBD2010) Andrea Esuli
PP-Index: Using Permutation Prefixes for Efficient and Scalable Similarity Search

Proceedings of the Eighteenth Italian Symposium on Advanced Database Systems (SEBD 2010), Rimini, IT, 2010, ISBN: 978-88-7488-369-1, pages 318-325.

(IIR2010) Andrea Esuli, Diego Marcheggiani, Fabrizio Sebastiani
Sentence-Based Active Learning Strategies for Information Extraction
Proceedings of the First Italian Information Retrieval Workshop, Padova, IT, January 27-28, 2010.
CEUR-WS, Vol-560, pages 41-45, 2010.

(IIR2010b) Stefano Baccianella, Andrea Esuli, Fabrizio Sebastiani
Selecting Features for Ordinal Text Classification
Proceedings of the First Italian Information Retrieval Workshop, Padova, IT, January 27-28, 2010.
CEUR-WS, Vol-560, pages 13-14, 2010.

- **Books**

(VDM2010) Andrea Esuli
Automatic Generation of Lexical resource for Opinion Mining
VDM Verlag, Saarbrücken, Germany, ISBN 978-3-8364-7333-0, 2010.

- **Magazines and other publications**

(ERCIM2010) Andrea Esuli, Diego Marcheggiani and Fabrizio Sebastiani
Extracting Information from Free-text Mammography Reports
ERCIM News, 82 R&D and Technology Transfer, pages 80-81.

(ERCIM2009) Stefano Baccianella, Andrea Esuli and Fabrizio Sebastiani
Multi-Faceted Rating of Product Reviews
ERCIM News, 77, “Future Internet Technology”, 2009.

(CORR2009) Paolo Bolettieri, Andrea Esuli, Fabrizio Falchi, Claudio Lucchese, Raffaele Perego, Tommaso Piccioli, Fausto Rabitti
CoPhIR: a Test Collection for Content-Based Image Retrieval
CoRR (Computing Research Repository), abs/0905.4627, arXiv, Cornell University Library, 15 pages, 2009.

(SIGIRFOR2008) Andrea Esuli
Automatic generation of lexical resources for opinion mining: models, algorithms and applications
SIGIR Forum, 42(2), ACM, New York, NY, USA, pages 105-106, ISSN 0163-5840, 2008.

- **Ph.D. Thesis**

(PhD2008) Andrea Esuli
Automatic Generation of Lexical Resources for Opinion Mining: Models, Algorithms and Applications
PhD in Information Engineering, PhD School “Leonardo da Vinci”, University of Pisa, 2008.