

Kevin Roitero

Current Position: Assistant Professor (RTD/a) at University of Udine.

Email: kevin.roitero@outlook.com | kevin.roitero@uniud.it

Web: <https://kevinroitero.com/>

ACADEMIC EMPLOYMENT

- **Assistant Professor (RTD/a)** Udine, Italy
University of Udine 10 Jan 2022 – 09 Jan 2025
Founded PON 2014–2020, Azione IV.6 - Tematiche Green.
Role: Principal Investigator.
Scientific Sector 09/H1 “Sistemi di elaborazione delle informazioni”.
Project: “Data-Driven Multiutility Grid: Decision Support to Guarantee Sustainability from Real Time to Long Term”.
- **Post Doctoral Research Fellow** Udine, Italy
University of Udine 16 Oct 2021 – 09 Jan 2022
Project: “ Study and Development of Artificial Intelligence Techniques for Automatic Analysis of Texts and Time Series on Social Media”.
- **Post Doctoral Research Fellow** Udine, Italy
University of Udine 01 Jan 2021 – 15 Oct 2021
Project: “Crowdsourcing and Machine Learning: Computational tools to identify medical misinformation (“fake news”) on social media”.
- **(Post Doctoral) Research Fellow** Udine, Italy
University of Udine 01 Jan 2020 – 31 Dec 2020
Project: “Analysis and development of computer science methods for detecting fake news in the medical area on social media”.
- **Research Fellow** Udine, Italy
University of Udine Nov 2019 – Dec 2019
Project: “Audit e Feedback in Emergenza: data analytics su tempi di soccorso territoriale”, carried on in “SOC Istituto di Igiene ed Epidemiologia Clinica”. The project deals with the analysis of the times and the data collected on the ambulances operating in the FVG region.

PLANNED FUTURE VISITS

- **Visiting Scholar** Cambridge, U.K.
University of Cambridge 25 June 2023 – 16 June 2023
Visit to Andreas Vlachos at the the Department of Computer Science and Technology at the University of Cambridge, in the Natural Language Processing and Machine Learning research group to work on fact checking.

INTERNSHIPS AND VISITS

- **Visiting Scholar** Copenhagen, Denmark
Pioneer Center for Artificial Intelligence & University of Copenhagen 4 June 2023 – 18 June 2023
Visit to Isabelle Augenstein at the Department of Computer Science, in the Copenhagen Natural Language Understanding research group to work on fact checking.
- **Research Intern** London, U.K.
Spotify Sept 2018 – Dec 2018
Project: Statistical modeling, machine and transfer learning techniques to model internal processes. Manager: Mounia Lalmas-Roelleke (Director, Tech Research).
- **Visiting Researcher** Brisbane, Australia
The University of Queensland, School of Information Technology and Engineering. 23 July 2018 – 30 July 2018
Short visiting period (as PhD Student).
- **Visiting Researcher** Melbourne, Australia
RMIT University, Information Storage, Analysis and Retrieval (ISAR) Group. Mar 2018 – July 2018
Visiting period (as PhD Student). Grant by the ISAR Group.

- **Visiting Researcher**

University of Sheffield, Information School.

Visiting period (as PhD Student). Grant by ERASMUS+ Traineeship.

Sheffield, U.K.

Apr 2017 – June 2017

- **Visiting Student**

Alpen-Adria-Universität

Visiting period (as Master Student). Grant by University of Udine and ERASMUS.

Klagenfurt, Austria

Sept 2015 – Feb 2016

EDUCATION

- **Ph.D., University of Udine**

Ph.D., Department of Mathematics, computer Science, and Physics.

Focus: Information Retrieval, crowdsourcing, machine and transfer learning, data mining and analysis, and statistical modeling. Scholarship provided by University of Udine. Title of the Thesis: “Cheap IR Evaluation: Fewer Topics, No Relevance Judgements, and Crowdsourced Assessments.”. Supervisor: Stefano Mizzaro. Grade: Cum Laude.

Graduation Date: 19 March 2020.

Udine, Italy.

Nov 2016 – Nov 2019

- **MS, University of Udine & Alpen-Adria-Universität**

Master Joint Degree in International Computer Science.

Grade: cum laude (110 cum laude/110). Graduation date: 22 July 2016.

Udine, Italy & Klagenfurt, Austria

Sept 2014 – July 2016

- **BS, University of Udine**

Bachelor Degree in Computer Science.

Grade: 102/110. Graduation Date: 17 October 2014.

Udine, Italy

Sept 2011 – Oct 2014

PUBLICATIONS

- **Google Scholar:** <https://scholar.google.com/citations?hl=en&user=1xd52jMAAAAJ>

- **Scopus:** <https://www.scopus.com/authid/detail.uri?authorId=57191042663>

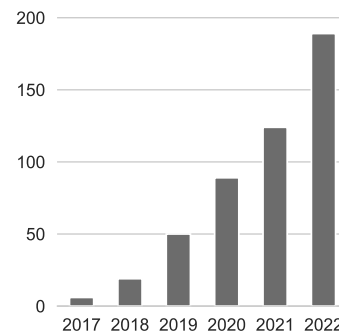
- **ACM DL:** <https://dl.acm.org/profile/99659204806>

- **DBLP:** <https://dblp.uni-trier.de/pers/hd/r/Roitero:Kevin>

Metrics (at May 1, 2023):

	Scholar	Scopus	ACM DL
Articles	56	46	30
Citations	556	330	176
Citations/article	9.93	7.17	5.87
H-index	13	11	–
i10-index	19	–	–

Citation Trend (Scholar)



Articles

- David La Barbera, Kevin Roitero, and Stefano Mizzaro. Combining Human Intelligence and Machine Learning for Fact-Checking: Towards a Hybrid Human-In-The-Loop Framework. *Intelligenza Artificiale*, 2023. Journal Rank: Scimago Q3 (2022).
- Kevin Roitero, Beatrice Portelli, Giuseppe Serra, Vincenzo Della Mea, Stefano Mizzaro, Gianni Cerro, Michele Vitelli, and Mario Molinara. Detection of Wastewater Pollution through Natural Language Generation with a Low-Cost Sensing Platform. *IEEE Access*, 2023. Journal Rank: Scimago Q1 (2022).
- Gianluca Demartini, Kevin Roitero, and Stefano Mizzaro. Data Bias Management. *Communications of the ACM*, 2023. Journal Rank: Scimago Q1 (2022).
- Kevin Roitero, David La Barbera, Michael Soprano, Gianluca Demartini, Stefano Mizzaro, and Tetsuya Sakai. How Many Crowd Workers Do I Need? On Statistical Power When Crowdsourcing Relevance Judgments. *ACM Transactions on Information Systems (TOIS)* 2023. Journal Rank: Scimago Q1 (2022).
- Sara Ceschia, Kevin Roitero, Gianluca Demartini, Stefano Mizzaro, Luca Di Gaspero, and Andrea Schaerf. Task Design in Complex Crowdsourcing Experiments: Item Assignment Optimization. *Computers and Operations Research (CAOR)* 2022. Journal Rank: Scimago Q1 (2021).

- Yunke Qu, David La Barbera, Kevin Roitero, Damiano Spina, Stefano Mizzaro, and Gianluca Demartini. Combining Human and Machine Confidence in Truthfulness Assessment. *ACM Journal of Data and Information Quality (JDIQ)* 2022. Journal Rank: Scimago Q2 (2021).
- Erik Brand, Kevin Roitero, Michael Soprano, Afshin Rahimi, and Gianluca Demartini. A Neural Model to Jointly Predict and Explain Truthfulness of Statements. *ACM Journal of Data and Information Quality (JDIQ)* 2022. Journal Rank: Scimago Q2 (2021).
- Kevin Roitero, Beatrice Portelli, Mihai Horia Popescu, and Vincenzo Della Mea. DiBERT: Cheap Embeddings for Disease Related Medical NLP. *IEEE Access. Journal Rank: Scimago Q1 (2020)*.
- Michael Soprano, Kevin Roitero, David La Barbera, Davide Ceolin, Damiano Spina, Stefano Mizzaro, and Gianluca Demartini. The Many Dimensions of Truthfulness: Crowdsourcing Misinformation Assessments on a Multidimensional Scale. *Information processing and Management (IPM)* 2021. Journal Rank: Scimago Q1 (2020).
- Kevin Roitero, Michael Soprano, Beatrice Portelli, Massimiliano De Luise, Damiano Spina, Vincenzo Della Mea, Giuseppe Serra, Stefano Mizzaro, and Gianluca Demartini. Can the Crowd Judge Truthfulness? A Longitudinal Study on Recent Misinformation about COVID-19. *Personal and Ubiquitous Computing (PAUC)* 2021. Journal Rank: Scimago Q2 (2020).
- Kevin Roitero, Eddy Maddalena, Stefano Mizzaro, and Falk Scholer. On the Effect of Relevance Scales in Crowdsourcing Relevance Assessments for Information Retrieval Evaluation. *Information processing and Management (IPM)* 2021. Journal Rank: Scimago Q1 (2020).
- Kevin Roitero. Cheap IR Evaluation: Fewer Topics, No Relevance Judgements, and Crowdsourced Assessments. *SIGIR Forum. December 2020, Volume 54, Number 2*.
- Vincenzo Della Mea, Mihai Horia Popescu, and Kevin Roitero. Underlying Cause of Death Identification from Death Certificates using Reverse Coding to Text and a NLP Based Deep Learning Approach. *Informatics in Medicine Unlocked*. Journal Rank: Scimago Q3 (2020).
- David La Barbera, António Polónia, Kevin Roitero, Eduardo Conde-Sousa, and Vincenzo Della Mea. Detection of HER2 from Haematoxylin-Eosin Slides Through a Cascade of Deep Learning Classifiers via Multi-Instance Learning. *Journal of Imaging, Special Issue "Deep Learning on Medical Image Analysis"*. Journal Rank: Scimago Q2 (2020).
- Kevin Roitero, Andrea Brunello, Giuseppe Serra, and Stefano Mizzaro. Effectiveness Evaluation Without Human Relevance Judgments: A Systematic Analysis of Existing Methods and of their Combinations. *Information Processing and Management (IPM)*. Journal Rank: Scimago Q1 (2020).
- Lei Han, Kevin Roitero, Ujwal Gadiraju, Cristina Sarasua, Alessandro Checco, Eddy Maddalena, and Gianluca Demartini. The Impact of Task Abandonment in Crowdsourcing. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*. Journal Rank: Scimago Q1 (2019).
- Kevin Roitero, J. Shane Culpepper, Mark Sanderson, Falk Scholer, and Stefano Mizzaro. Fewer Topics? A Million Topics? Both?! On Topics Subsets in Test Collections. *Information Retrieval Journal (IRJ)*. Journal Rank: Scimago Q2 (2020).
- Kevin Roitero, Marco Passon, Giuseppe Serra, and Stefano Mizzaro. Reproduce. Generalize. Extend. On Information Retrieval Evaluation without Relevance Judgments. *ACM Journal of Data and Information Quality (JDIQ)*. Journal Rank: Scimago Q2 (2018).
- Kevin Roitero, Michael Soprano, Andrea Brunello, and Stefano Mizzaro. Reproduce and Improve: An Evolutionary Approach to Select a Few Good Topics for Information Retrieval Evaluation. *ACM Journal of Data and Information Quality (JDIQ)*. Journal Rank: Scimago Q2 (2018).

Conference Papers

- Davide Ceolin, Kevin Roitero, and Furong Guo. Predicting Crowd Workers Performance: An Information Quality Case. *Proceedings of the 23rd International Conference on Web Engineering, ICWE 2023, Alicante, Spain. June 6–9 2023*. Conference Rank: CORE B, GGS B-.
- Kevin Roitero, Andrea Martinuzzi, Maria Teresa Armellin, Gabriella Paparella, Alberto Maniero, and Vincenzo Della Mea. Automated ICF Coding of Rehabilitation Notes for Low-Resource Languages via Continual Training of Language Models. *Proceedings of Medical Informatics Europe 2023, EFMI MIE 2023, Göteborg, Sweden. May 22–25 2023*.

- Francesca Da Ros, Luca Di Gaspero, David La Barbera, Vincenzo Della Mea, Kevin Roitero, Laura Deroma, Sabrina Licata, and Francesca Valent. A Multi-objective Biased Random-key Genetic Algorithm for the Optimal Siting of Emergency Vehicles. *Proceedings of the 14th Metaheuristics International Conference, MIC 2022, Ortigia-Syracuse, Italy. July 11–14 2022.*
- Tim Draws, David La Barbera, Michael Soprano, Kevin Roitero, Davide Ceolin, Alessandro Checco, Stefano Mizzaro. The Effects of Crowd Workers Biases in Fact-Checking Tasks. *Proceedings of the fifth ACM Conference on Fairness, Accountability, and Transparency, FAccT 2022, Seoul, South Korea. June 21-24 2022.*
- Kevin Roitero, Alessandro Checco, Stefano Mizzaro, and Gianluca Demartini. Preferences on a Budget: Prioritizing Document Pairs when Crowdsourcing Relevance Judgments. *Proceedings of the ACM Web Conference, WWW 2022, Lyon, France. April 25-29, 2022.* Conference Rank: CORE A*, GGS A++.
- Michael Soprano, Kevin Roitero, Francesco Bombassei De Bona, and Stefano Mizzaro. Crowd_Frame: A Simple and Complete Framework to Deploy Complex Crowdsourcing Tasks Off-the-shelf. *Proceedings of the 15th ACM International Conference on Web Search and Data Mining, WSDM 2022, Phoenix, Arizona. February 21-25, 2022.* Conference Rank: CORE A*, GGS A+.
- Erik Brand, Kevin Roitero, Michael Soprano, and Gianluca Demartini. E-BART: Jointly Predicting and Explaining Truthfulness. *Conference for Truth and Trust Online (TTO2021).*
- Yunke Qu, Kevin Roitero, Stefano Mizzaro, Damiano Spina, and Gianluca Demartini. Human-in-the-Loop Systems for Truthfulness: A Study of Human and Machine Confidence. *Conference for Truth and Trust Online (TTO2021).*
- David La Barbera, Kevin Roitero, Stefano Mizzaro, Vincenzo Della Mea, and Francesca Valent. A Software Simulator for Optimizing Ambulance Location and Response Time: A Preliminary Report. *IEEE International Conference on Digital Health (ICDH 2021).*
- Mihai Horia Popescu, Kevin Roitero, Stefano Travasci, and Vincenzo Della Mea. Automatic Assignment of ICD-10 Codes to Diagnostic Texts using Transformers Based Techniques. *Proceedings of the 9th IEEE International Conference on Healthcare Informatics (ICHI2021). Victoria, British Columbia, Canada. August 9th-12th, 2021.*
- Kevin Roitero, Michael Soprano, Beatrice Portelli, Damiano Spina, Vincenzo Della Mea, Giuseppe Serra, Stefano Mizzaro, and Gianluca Demartini. The COVID-19 Infodemic: Can the Crowd Judge Recent Misinformation Objectively? *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM2020). Online. October 19-23, 2020.* Conference Rank: CORE A, GGS A+.
- Vincenzo Della Mea, Mihai Horia Popescu, Kevin Roitero. Underlying Cause of Death Identification from Death Certificates via Categorical Embeddings and Convolutional Neural Networks. *Proceedings of the 8th IEEE International Conference on Healthcare Informatics (ICHI2020). Oldenburg, Germany. Nov 30 – Dec 3 2020.*
- Kevin Roitero, Michael Soprano, Shaoyang Fan, Damiano Spina, Stefano Mizzaro, and Gianluca Demartini. Can The Crowd Identify Misinformation Objectively? The Effects of Judgments Scale and Assessor’s Background. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. Xi’an, China. July 25-30, 2020.* Conference Rank: CORE A*, GGS A++.
- David La Barbera, Kevin Roitero, Damiano Spina, Gianluca Demartini, and Stefano Mizzaro. Crowdsourcing Truthfulness: The Impact of Judgment Scale and Assessor Bias. *Proceedings of the 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14-17, 2020.* Conference Rank: CORE A, GGS A-. **Best Short Paper Award (Top 0.6%).**
- Lei Han, Eddy Maddalena, Alessandro Checco, Cristina Sarasua, Ujwal Gadiraju, Kevin Roitero, and Gianluca Demartini. Crowd Worker Strategies in Relevance Judgment Tasks. *Proceedings of the 13th ACM International Conference on Web Search and Data Mining, WSDM 2020, Houston, Texas. February 3-7, 2020.* Conference Rank: CORE A*, GGS A+.
- Kevin Roitero, Andrea Brunello, Julián Urbano and Stefano Mizzaro. Towards Stochastic Simulations of Relevance Profiles. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM2019). Beijing, China. November 3-7, 2019.* Conference Rank: CORE A, GGS A+.
- Lei Han, Kevin Roitero, Eddy Maddalena, Stefano Mizzaro and Gianluca Demartini. On Transforming Relevance Scales. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM2019). Beijing, China. November 3-7, 2019.* Conference Rank: CORE A*, GGS A+.

- Shane Culpepper, Oren Kurland, Stefano Mizzaro, Kevin Roitero and Fabio Zampieri. On Topic Difficulty in IR Evaluation: The Effect of Corpora, Systems, and System Components. *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Paris, France. July 21-25, 2019. Conference Rank: CORE A*, GGS A++.
- Lei Han, Kevin Roitero, Ujwal Gadiraju, Cristina Sarasua, Alessandro Checco, Eddy Maddalena and Gianluca Demartini. All Those Wasted Hours: On Task Abandonment in Crowdsourcing. *ACM International Conference on Web Search and Data Mining, WSDM 2019, Melbourne, Australia*. Conference Rank: CORE A*, GGS A+.
- Kevin Roitero, Eddy Maddalena, Gianluca Demartini and Stefano Mizzaro. On Fine-Grained Relevance Scales. *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*. Ann Arbor Michigan, U.S.A. July 8-12, 2018. Conference Rank: CORE A*, GGS A++.
- Stefano Mizzaro, Josiane Mothe, Kevin Roitero, and Md Zia Ullah. Query Performance Prediction and Effectiveness Evaluation Without Relevance Judgments: Two Sides of the Same Coin. *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*. Ann Arbor Michigan, U.S.A. July 8-12, 2018. Conference Rank: CORE A*, GGS A++.
- Kevin Roitero, Eddy Maddalena, Ponte Yannick and Stefano Mizzaro. IRevalOO: An Object Oriented Framework for Retrieval Evaluation. *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*. Ann Arbor Michigan, U.S.A. July 8-12, 2018. Conference Rank: CORE A*, GGS A++.
- Kevin Roitero, Michael Soprano and Stefano Mizzaro. Effectiveness Evaluation with a Subset of Topics: A Practical Approach. *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*. Ann Arbor Michigan, U.S.A. July 8-12, 2018. Conference Rank: CORE A*, GGS A++.
- Kevin Roitero. Cheap and Engineered Evaluation of Retrieval Systems. *Doctoral Consortium of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*. Ann Arbor Michigan, U.S.A. July 8-12, 2018. Conference Rank: CORE A*, GGS A++.
- Alessandro Checco, Kevin Roitero, Eddy Maddalena, Gianluca Demartini and Stefano Mizzaro. Let's Agree to Disagree: Fixing Agreement Measures for Crowdsourcing. *Proceedings of the The fifth AAAI Conference on Human Computation and Crowdsourcing, AAAI HCOMP2017*. Quebec City, Canada. October 24-26 2017.
- Eddy Maddalena, Kevin Roitero, Gianluca Demartini and Stefano Mizzaro. Considering Assessor Agreement in IR Evaluation. *Proceedings of the 3rd ACM International Conference on the Theory of Information Retrieval, ICTIR2017*. Amsterdam, Netherlands. October 1-4 2017.
- Kevin Roitero, Eddy Maddalena, and Stefano Mizzaro. Do Easy Topics Predict Effectiveness Better Than Difficult Topics? *Proceedings of the 39th European Conference on IR Research, ECIR 2017, Aberdeen, UK, April 8-13, 2017*. Conference Rank: CORE A, GGS A-.

Workshop Papers

- David La Barbera, Michael Soprano, Kevin Roitero, Eddy Maddalena, and Stefano Mizzaro. Fact-Checking at Scale with Crowdsourcing: Experiments and Lessons Learned. *13th Italian Information Retrieval Workshop (IIR 2023)*. June 08 – 09, 2023, Pisa, Italy.
- Michael Soprano, Kevin Roitero, Vincenzo Della Mea, and Stefano Mizzaro. Towards a Conversational-Based Agent for Health Services. *Workshop AI for Public Administration. Co-located with the 3rd CINI Conference on Artificial Intelligence (Ital-IA 2023)*. May 29 – 31, 2023, Pisa, Italy.
- Mihai Horia Popescu, Kevin Roitero, Vincenzo Della Mea. Explainable Classification of Medical Documents Through a Text-to-Text Transformer. *1st Workshop on Artificial Intelligence For Healthcare (HC). Co-located with the AIXIA conference*. November 28 – December 2, 2022, Udine, Italy.
- Kevin Roitero, Cristina Gattazzo, Andrea Zancola, Vincenzo Della Mea, Stefano Mizzaro. Causal Text-to-Text Transformers for Water Pollution Forecasting. *2nd Italian Workshop on Artificial Intelligence and Applications for Business and Industries (AIABI). Co-located with the AIXIA conference*. November 28 – December 2, 2022, Udine, Italy.
- David La Barbera, Kevin Roitero, and Stefano Mizzaro. A Hybrid Human-In-The-Loop Framework for Fact Checking. *6th Workshop on Natural Language for Artificial Intelligence (NL4AI). Co-located with the AIXIA conference*. November 28 – December 2, 2022, Udine, Italy. **Best Paper Award (Top 7%)**.

- Francesca Da Ros, Luca Di Gaspero, David La Barbera, Vincenzo Della Mea, Kevin Roitero, Laura Deroma, Sabrina Licata, and Francesca Valent. Simheuristic approach to the ambulance location problem. *48th Annual Meeting of the Euro Working Group on Operational Research Applied to the Health Services, ORAHS 2022, Bergamo, Italy, July 17-22, 2022.*
- Gianluca Demartini, Kevin Roitero, and Stefano Mizzaro. Bias Management in Crowdsourced Data: Moving Beyond Bias Removal. *Investigating and Mitigating Biases in Crowdsourced Data. Workshop at the 24th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 21).*
- Kevin Roitero, Cristian Bozzato, Vincenzo Della Mea, Stefano Mizzaro, and Giuseppe Serra. Twitter goes to the Doctor: Detecting Medical Tweets using Machine Learning and BERT. *Proceedings of Semantic Indexing and Information Retrieval for Health (SIIRH2020) Workshop, in conjunction with the 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14-17, 2020.*
- Kevin Roitero, Ben Carterette, Rishabh Mehrotra and Mounia Lalmas. Leveraging Behavioral Heterogeneity Across Markets for Cross-Market Training of Recommender Systems. *Proceedings of the The 10th International Workshop on Location and the Web in conjunction with The Web Conference.*
- Michael Soprano, Kevin Roitero, and Stefano Mizzaro. Bias and Fairness in Effectiveness Evaluation by Means of Network Analysis and Mixture Models. *Proceedings of the 10th Italian Information Retrieval Workshop, Padua, Italy, September 16-18, 2019.*
- Michael Soprano, Kevin Roitero, and Stefano Mizzaro. HITS Hits Readersourcing: Validating Peer Review Alternatives Using Network Analysis. *Proceedings of the 4th Joint Workshop on Bibliometric-enhanced Information Retrieval and Natural Language Processing for Digital Libraries (BIRNDL 2019).*
- Kevin Roitero, Damiano Spina, Gianluca Demartini, and Stefano Mizzaro. How Many Truth Levels? Six? One Hundred? Even more? Validating Truthfulness of Statements via Crowdsourcing. *RDSM 2018 : Workshop on Rumours and Deception in Social Media.*
- Kevin Roitero, Stefano Mizzaro, and Giuseppe Serra. Economic Evaluation of Recommender Systems: A Proposal. *Proceedings of the 8th Italian Information Retrieval Workshop, Switzerland. June 05-07, 2017.*
- Kevin Roitero and Stefano Mizzaro. Improving the Efficiency of Retrieval Effectiveness Evaluation: Finding a Few Good Topics with Clustering? *Proceedings of the 7th Italian Information Retrieval Workshop, Venezia, Italy, May 30-31, 2016.*

Technical Reports

- David La Barbera, Kevin Roitero, Joel Mackenzie, Damiano Spina, Gianluca Demartini, and Stefano Mizzaro. BUM at CheckThat! 2022: a Composite Deep Learning Approach to Fake News Detection Using Evidence Retrieval *Proceedings of CLEF 2022 – Conference and Labs of the Evaluation Forum, September 5–8, 2022, Bologna, Italy.*

TALKS

- **Invited Talks:**
 - **28 April 2023.** Practical Machine Learning. *Guest lecture to the graduate school for the “Radiodiagnostics” and “Clinical Pathology and Clinical Biochemistry” courses. University of Udine, Udine. Italy.*
 - **18 September 2022.** ProMa²L²: PROstate MAlignancy Localization by MACHine Learning . *Invited Talk to Automated Gleason Grading Challenge (AGGC2022) Event at the 25th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI2022). Singapore.*
 - **23 June 2022.** Fact-checking Online Information Using The Crowd: Lessons Learned and Future Directions. *Invited Talk to the Meeting of the AI, Media and Democracy Lab, Cluster 2 (Making the news with AI). Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands.*
 - **9 June 2022.** Practical Machine Learning Using Weka. *Guest lecture to the graduate school for the “Radiodiagnostics” and “Clinical Pathology and Clinical Biochemistry” courses. University of Udine, Udine. Italy.*
 - **13 April 2022.** A Deep Learning Approach to Automatic Encoding of Death Certificates (and other applications). *Invited talk to the BSc/MSc Course “Data and Techniques for E-Health”. University of Udine, Udine. Italy.*
 - **3 August 2021.** Incorporating Signals from Human Annotators to Build Effective Systems. *Invited research talk at Amazon, London, United Kingdom.*
 - **24 June 2021.** Leveraging Human Signals to Build Effective Systems. *Invited research talk at the Faculty Of Engineering, Architecture & Info Tech, The University of Queensland, Brisbane, Australia.*
 - **16 April 2020.** Fewer topics? A million topics? Both?! On topics subsets in test collections. *Invited talk for the presentation of IRJ Paper at the European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14-17, 2020.*

- **19 December 2019.** Statistical Modelling for Recommendation Systems Evaluation. *Invited talk to the MSc Course “Data Analytics”.* University of Udine, Udine. Italy.
- **14 November 2019.** Magnitude Estimation for Relevance Assessment (& Other Applications). *Invited talk at the European Network for Combining Language Learning with Crowdsourcing Techniques, enetCollect.* Sliema, Malta.
- **7 June 2019.** Economical Evaluation of Systems: Users as a Signal. *Invited research talk at Spotify NYC.*
- **15 March 2019.** Let’s Agree to Disagree: Fixing Agreement Measures for Crowdsourcing. *Invited talk at the European Network for Combining Language Learning with Crowdsourcing Techniques, enetCollect.* Lisbon.
- **14 November 2018.** “Statistical modelling for the evaluation of RecSys”. *Internal Research Talk at Spotify.* London, Great Britain.
- **17 May 2018.** iPython notebooks for data analysis and plotting. *Invited research talk done while visiting RMIT University, Melbourne, Australia.*
- **10 May 2018.** Crowdsourcing relevance judgments and considering agreement in IR evaluation (with a new agreement measure) *Invited research talk done while visiting RMIT University, Melbourne, Australia.*
- **12 May 2016.** Improving the Efficiency of Retrieval Effectiveness Evaluation: Finding a Few Good Topics with Clustering and Dimensionality Reduction. *Invited talk to the MSc Course “Web Information Retrieval”.* University of Udine, Udine. Italy.
- **Conference Presentations:**
 - **11 November 2021.** Presentation of the paper “A Quantum Approach to Crowdsourcing”. *3rd Symposium on Biases in Human Computation and Crowdsourcing, BHCC 2021.*
 - **27 July 2020.** Presentation of the paper “Can The Crowd Identify Misinformation Objectively? The Effects of Judgments Scale and Assessor’s Bias”. *ACM SIGIR Conference on Research and Development in Information Retrieval.*
 - **14 April 2020.** Presentation of the paper “Twitter goes to the Doctor: Detecting Medical Tweets using Machine Learning and BERT”. *SSIRH Workshop at the European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14-17, 2020.*
 - **25 July 2019.** Presentation of the paper “HITS Hits Readersourcing: Validating Peer Review Alternatives Using Network Analysis”. *ACM SIGIR Conference on Research and Development in Information Retrieval.* Paris, France.
 - **11 July 2018.** Presentation of the paper “On fine Grained Relevance Scales”. *ACM SIGIR Conference on Research and Development in Information Retrieval.* Ann Arbor Michigan, U.S.A.
 - **6 June 2017.** Presentation of the paper “Economic Evaluation of Recommender Systems: A Proposal”. *IIR2017, Lugano, Switzerland.*
 - **10 April 2017.** Presentation of the paper “Do Easy Topics Predict Effectiveness Better Than Difficult Topics?” *ECIR2017, Aberdeen. UK.*
 - **30 May 2016.** Presentation of the paper “Improving the Efficiency of Retrieval Effectiveness Evaluation: Finding a Few Good Topics with Clustering?” *IIR 2016, Venezia. Italy.*

TEACHING

Academic Year: 2022–2023

- **Information Analysis and Processing for training (Sistemi di Analisi e Elaborazione delle Informazioni per l’Allenamento).** Role: Course Co-Leader. University of Udine (2022/2023). “*Scienze e Tecniche delle Attività Motorie Preventive ed Adattate*” course, MSc. I was one of the two main instructors of the course. I designed and instructed my part of the course. Classes consisted of lectures, with evaluation based on exam. The course had a total duration of 24 hours, mine was 12 hours.
- **Recommendation Systems.** Role: Course Leader. University of Udine (2022/2023). “*Computer Science*” course, MSc. I was the main instructor of the course. I designed and instructed the course. Classes consisted of lectures, with evaluation based on an oral exam and a final project. The course had a total duration of 48 hours.
- **Object Oriented Analysis and Design.** Role: Course Leader. University of Udine (2022-2023). “*Computer Science*” course, MSc. I was one of the two main instructors of the course. I co-designed and co-instructed the course. Classes consisted of lectures, with evaluation based on an oral examination. The course had a total duration of 48 hours, mine was 12 hours.
- **Programming Methodologies for Data Analysis.** Role: Course Leader. University of Padova (2021-2022). *PhD Program (XXXVIII cycle) in “Statistical Sciences”.* I was one of the two main instructors of the course. I co-designed and co-instructed the course. Classes consisted of lectures, with evaluation based on practicums and a final project. The course had a total duration of 30 hours, mine was 12 hours.

Academic Year: 2021–2022

- **Web Technologies for the Cloud and Laboratory.** Role: Teacher. University of Udine (2021-2022). “*Computer Science*” course, BSc. I provided laboratory assistance; furthermore, I prepared slides and provided lectures on the topic “Docker”; total duration of 16 hours. The course had about 60 students and a total duration of 48 hours, mine was 16 hours. Instructor: Vincenzo Della Mea.
- **Information Systems and Data Warehouse.** Role: Teacher. University of Udine (2021-2022). “*Computer Science*” course, MSc. I prepared slides and provided lectures for a total of 12 hours on the topic “Data Warehouse”. The course had about 20 students and a total duration of 96 hours, mine was 12 hours. Instructors: Paolo Coppola and Carlo Tasso.

- **Object Oriented Analysis and Design.** Role: Teacher. University of Udine (2021-2022). “*Computer Science*” course, MSc. I prepared slides and provided lectures for a total of 8 hours on the topic “Refactoring”. The course had about 30 students and a total duration of 48 hours, mine was 8 hours. Instructor: Stefano Mizzaro.
- **Web Information Retrieval.** Role: Teacher. University of Udine (2021-2022). “*Computer Science*” course, MSc. I prepared slides and provided lectures for a total of 4 hours on the topic “Machine and Deep Learning Retrieval Methods”. The course had about 50 students and a total duration of 48 hours, mine was 4 hours. Instructor: Stefano Mizzaro.
- **Programming Methodologies for Data Analysis.** Role: Course Leader. University of Padova (2021-2022). *PhD Program (XXXVII cycle) in “Statistical Sciences”*. I was one of the two main instructors of the course. I co-designed and co-instructed the course. Classes consisted of lectures, with evaluation based on practicums and a final project. The course had about 20 students and a total duration of 30 hours, mine was 12 hours.

Academic Year: 2020–2021

- **Programming Methodologies for Data Analysis.** Role: Course Leader. University of Padova (2020-2021). *PhD Program (XXXVI cycle) in “Statistical Sciences”*. I was one of the two main instructors of the course. I co-designed and co-instructed the course. Classes consisted of lectures, with evaluation based on practicums and a final project. The course had about 15 students and a total duration of 30 hours, mine was 12 hours.

Academic Year: 2019–2020

- **Web Applications.** Role: Course Leader. University of Udine (2019-2020). “*Electronic Engineering*” and “*Industrial Systems Engineering and Management*” courses, MSc. I was the main instructor of the course. I designed and instructed the course. Classes consisted of lectures, with evaluation based on an oral exam and a final project. The course had about 80 students and a total duration of 48 hours.
- **Social Computing.** Role: Teaching Assistant. University of Udine (2019/2020). “*Internet of Things, Big data & Web*” course, MSc. I prepared slides and provided lectures. Instructor: Stefano Mizzaro.

Academic Year: 2016-2017

- **Operating Systems.** Role: Teaching Assistant. University of Udine (2016-2017). “*Computer Science*” course, BSc. I prepared slides and instructed the laboratory part of the course. Classes consisted of lectures, with evaluation based on a written exam and a final project. The course had more than 100 students. Instructor: Ivan Scagnetto.

ACADEMIC ACTIVITY

- **Ph.D. Supervision:**
 - David La Barbera, University of Udine. PhD Course in Computer Science and Artificial Intelligence. Dates: November 2021 – now. Role: Co-supervisor.
 - Michael Soprano, University of Udine. PhD Course in Mathematics, Computer Science, and Physics Dates: November 2019 – May 2023. Role: Co-supervisor.
- **Mentoring:**
 - I co-supervised the thesis of 11 Master (MSC), 10 Bachelor (BSC) Students, and various students advanced projects and seminars. Many of the thesis lead to published results. One Master Thesis received the prize FIDAPA: “Non ci sono mestieri da donne”.
- **Journal Editorial Board:**
 - Guest editor in The Information Processing & Management (IPM) journal, “Machine and Human Factors in Misinformation Management” special issue. 2022.
- **Organizational Roles in Conferences:**
 - Publicity and Networking Chair of “The eleventh AAAI Conference on Human Computation and Crowdsourcing held together with The ACM Collective Intelligence Conference 2023, AAAI HCOMP 2023 / ACM CI 2023”. November 6–10 2023, Delft, Netherlands.
 - Organizer of “MisFAccT: The 1st Workshop on Fairness, Accountability, and Transparency in Misinformation Management” at the 23rd International Conference on Knowledge Engineering and Knowledge Management, EKAW 2022. September 26-29, 2022 - Bozen-Bolzano, Italy.
 - General Chair at the 4th Symposium on Biases in Human Computation and Crowdsourcing, BHCC 2022. October 12–13, 2022 - Rome, Italy.
 - Track organizer at the Information Processing & Management Conference, IPMC 2022; “Machine and Human Factors in Misinformation Management” track. October 20-21, 2022 – Online Event.
 - Demo Chair at the 45th ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2022. July 11–15, 2022. Madrid, Spain.
 - Proceedings Chair at the 30th ACM International Conference on Information and Knowledge Management, CIKM 2021. November 1–5, 2021. Gold Coast, Brisbane, Australia.

- Proceedings Chair at the 3rd Symposium on Biases in Human Computation and Crowdsourcing, BHCC 2021. November 10–12, 2021, Delft, Netherlands.
- **Reviewer for Journals:**
 - ACM Transactions on Information Systems (TOIS).
 - Information Processing and Management (IPM).
 - ACM Journal of Data and Information Quality (JDIQ).
 - International Journal of Digital Libraries (IJDL).
- **Program Committee Roles:**
 - ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR): 2023 (full and short paper tracks), 2022 (full, short, and demo paper tracks), 2021 (full paper track).
 - ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR): 2022 (demonstration & resource track).
 - International Conference on Information and Knowledge Management (CIKM): 2022 (full paper track), 2021 (full paper track), 2020 (full paper track).
 - European Conference on Information Retrieval (ECIR): 2023 (full and short paper tracks), 2022 (full and short paper tracks), 2021 (full paper track), 2018 (additional reviewer).
 - The Web Conference (formerly WWW): 2023 (full paper track), 2022 (additional reviewer), 2021 (full paper track), 2020 (full paper track).
 - AAAI Conference on Human Computation and Crowdsourcing (HCOMP): 2022 (main track).
 - Annual Conference on Innovative Applications of Artificial Intelligence (IAAI): 2023 (main track).
 - CLEF: 2022 (Lab on Fighting the COVID-19 Infodemic and Fake News Detection – CheckThat! track)
 - ACM Conference on Hypertext and Social Media (HT): 2021 (main paper track), 2020 (web and society track). item
 - Symposium on Biases in Human Computation and Crowdsourcing (BHCC): 2021 (full, short, and abstract paper tracks)
 - CLIC-it: 2020 (Information Extraction, Information Retrieval and Question Answering track).
 - CIRCLE: 2020 (main track).
 - Italian Information Retrieval Workshop (IIR): 2018 (main track).
- **Session Chairing at Conferences:**
 - Chair of “Special Track Session 3: MISINFO” at the Information Processing & Management Conference, IPMC 2022 for the “Machine and Human Factors in Misinformation Management” track.
 - Chair of multiple sessions at the 4th Symposium on Biases in Human Computation and Crowdsourcing, BHCC 2022.
 - Chair of “5E: Urban Computing” and “1D: Dealing with Difficult Data (Data Quality, Data Curation and De-noising)” sessions at the 30th ACM International Conference on Information and Knowledge Management, CIKM 2021.
 - Chair of “Paper Session III” at the 3rd Symposium on Biases in Human Computation and Crowdsourcing, BHCC 2021.
- **Affiliations:**
 - (2021 – present) affiliated with the Cybersecurity National Lab, Udine node. see <https://cybersecnatlab.it/?lang=en>.
 - (2019 – present) part of the SMDC Lab (Social, Mobile, Data, Crowd) at the University of Udine, Italy, see <http://smdc.uniud.it/smdc/membri/?lang=en>.
 - (2019 – present) part of the MITEL (Medical Informatics, Telemedicine & eHealth Lab) at the University of Udine, Italy, see http://mitel.dimi.uniud.it/mitel2/?page_id=62.
 - (2019 – present) external-collaborator of the AILab (Artificial Intelligence Laboratory) at the University of Udine, Italy, see <http://ailab.uniud.it/members/>.

SKILLS

- **Networking:** I worked with various interdisciplinary research groups, building many research partnerships.
- **Technical:** Proficient in: Python, NumPy, SciPy, Pandas, PyTorch, SQL, Javascript, Node.js, \LaTeX . Familiar with: Java, HTML, CSS, C, WEKA, PyMC3, Angular, R, Keras, TensorFlow, Fast.AI, BigQuery; Basic knowledge of: MATLAB, Scala.
- **Platforms/tools:** GitHub, Google Cloud Suite, Jupyter Notebooks, Slack, Amazon AWS, Amazon Mechanical Turk, Docker, Office Suite, Linux Shell, Dropbox, Google Drive.
- **Languages:** Italian (Mother tongue), English (Listening: C1, Reading: C1, Spoken Interaction: C1, Spoken Production: C1, Writing: C1).
- **Personal Interests:** eager reader, decent runner, film and series watcher, music enthusiast; trying to be the nerd the world needs; keen to minimalism, the art of giving importance to things that really matter. Pronouns: He/Him/His.
- **Other:** During the 3rd year of the Ph.D. I was accepted for an Internship at Apple Cambridge (UK) in the “NLP SIRI understanding team”.

HONORS, AWARDS, AND GRANTS

- **Best Paper Award (Top 5%)** at “NL4AI 2022”, for the paper “A Hybrid Human-In-The-Loop Framework for Fact Checking”
- Award of \$500 USD in Huawei Cloud Credits for the 6th place in the “Automated Gleason Grading Challenge (AGGC2022)”
- Grant supporting the application to the “Horizon Europe” call for the funding round 2022/2023. Action “ERC Starting Grant”. Funded amount: 8.000 Euro.
- Seal of honor of the Friuli Venezia Giulia Region Council (Sigillo del Consiglio Regionale del Friuli Venezia Giulia).
- Winner of the prize “con.Scienze2020” for the **best Ph.D. thesis** in Computer Science discussed in 2020 in Italy.
- **Best Short Paper Award (Top 0.6%)** at “ECIR2020”, for the paper “Crowdsourcing Truthfulness: The Impact of Judgment Scale and Assessor Bias”.
- Participation in the 7th **Heidelberg Laureate Forum** (top 200 young researchers in mathematics and computer science). Provided by Heidelberg Laureate Forum Foundation (HLFF).
- Allowance to visit RMIT ISAR Group at RMIT University, Melbourne, AUS. Provided by RMIT ISAR (Temporary Activity subclass 408 VISA).
- Allowance to visit Information School at University of Sheffield, UK. Provided by ERASMUS+ Traineeship.
- Grant to attend TTO 2021 (provided by TTO), SIGIR 2020 (provided by ACM SIGIR), ECIR 2020 (Provided by ECIR), ASIRF 2019 (Provided by ASIRF), SIGIR 2019 (Provided by ACM SIGIR), SIGIR 2018 (Provided by ACM SIGIR), ICTIR 2017 (Provided by ACM SIGIR), ECIR 2017 (Provided by BCS IRGS).
- Scholarship for the Ph.D. program of the University of Udine. Provided by University of Udine.
- Allowance to visit Alpen-Adria-Universität Kalgenfurt, AU. Provided by University of Udine, and ERAMUS Project.

REFERENCES

Available on request.

PRIVACY

In compliance with the GDPR and the Italian Legislative Decree no. 196 dated 30/06/2003, I hereby authorize you to use and process my personal details contained in this document.

Autorizzo il trattamento dei miei dati personali ai sensi del GDPR 679/16 e del Decreto Legislativo 30 giugno 2003 n. 196 (“Codice in materia di protezione dei dati personali”) ai fini di attività di ricerca e selezione del personale e contatti lavorativi.