

## Gianluca DEMARTINI

School of ITEE  
University of Queensland  
GP South building, Staff House Road, 4072 St Lucia, Australia.

*E-mail:* demartini@acm.org  
<http://www.gianlucademartini.net>

*Academic Employment*    **School of IT and Electrical Engineering**, The University of Queensland, Australia.  
Associate Professor in Data Science. Aug 2019 – now.  
Senior Lecturer in Data Science. Aug 2017 – Aug 2019.

**Information School**, The University of Sheffield, UK.  
Senior Lecturer in Data Science. Nov 2015 – Jun 2017.  
Lecturer (equivalent to Assistant Professor) in Data Science. Aug 2014 – Oct 2015.

**eXascale Infolab**, University of Fribourg, Switzerland.  
Post-doctoral researcher.  
May 2011 – July 2014.

**L3S Research Center**, University of Hannover, Germany.  
Junior Researcher.  
Mar 2006 – Apr 2011.

*Internships and Visits*    **AMPLab, UC Berkeley**, CA, USA. Summer 2012.  
With Michael J. Franklin, Tim Kraska, Beth Trushkowsky on crowdsourced query understanding.

**Yahoo! Research**, Barcelona, Spain. Summer 2009.  
With Hugo Zaragoza and Roi Blanco on Time-aware Entity Ranking in news articles.

*Academic Education*    Leibniz University of Hannover, Germany.  
**Ph.D. in Computer Science** (Dr. rer. nat.)  
Supervisors: Wolfgang Nejdl, Arjen P. de Vries.  
Title: "From People to Entities: Typed Search in the Enterprise and the Web".  
Apr 2006 - Apr 2011.

University of Sheffield, UK.  
Post-graduate Certificate in Learning and Teaching, School of Education, Mar 2016.  
**Fellow of the Higher Education Academy** (FHEA), UK.

University of Udine, Italy.  
**M.Sc. (*cum laude*) in Computer Science**  
Supervisor: Stefano Mizzaro.  
Dec 2003 - Dec 2005.  
Summer School participation: "Information Retrieval" (2005).

University of Udine, Italy.  
**B.Sc. in Computer Science**  
Supervisor: Alberto Policriti.  
Oct 2000 - Dec 2003.

*Research Interests*    Crowdsourcing for Human Computation, Information Retrieval, AI for Public Good.  
*Awards*

- Selected as a Google Cloud Research Innovator, 2022.
- "Methods Recognition" at the CSCW 2021 conference, intended to recognize significant methodological advances or prime examples of good methods implementation.
- **ACM Senior Member**, 2020.
- Selected for the **Google** AI for Social Good "Impact Scholars" program, 2021.
- **Best Short Paper Award** at the ECIR 2020 conference. **Top 0.6%** of all submissions.
- **Honorable Mention** Award at the CSCW 2020 conference. **Top 2%** of all submissions.

- **ACM Senior Member**, 2020.
- **Outstanding Student Paper** in the field of Consumer and Pervasive Health Informatics, American Medical Informatics Association Annual Symposium (AMIA) in 2019 for the paper “Health Cards to Assist Decision Making in Consumer Health Search”.
- The Information Retrieval Distinguished Board Member Award for 2018 and 2019.
- Italian National Scientific Habilitation (Abilitazione Scientifica Nazionale) to Full Professor with validity 2018-2024 in sectors 09/H1 (Computer Engineering) and 01/B1 (Computer Science).
- **Best Paper Award** at the HCOMP 2018 conference. Top 1% of all submissions.
- Ranked in top **7% of EPSRC College members** for peer review activities during the year 2016/17.
- Italian National Scientific Habilitation (Abilitazione Scientifica Nazionale) to Associate Professor with validity 2017-2023 in sectors 09/H1 (Computer Engineering) and 01/B1 (Computer Science).
- **Best Paper Award** at the ECIR 2016 conference. **Top 0.5%** of all submissions.
- Outstanding Early Career **Research Achievement Award**, U. Sheffield Faculty of Social Sciences, 2015.
- Selected for the “**Future Academic Leaders Programme**” at the University of Sheffield, 2015.
- Nominated for a University of Sheffield Faculty Learning and Teaching Award, 2015.
- Nominated for **Best Research Paper** award at the ISWC 2013 conference. **Top 2%** of all submissions.
- Third place at the “**Semantic Web Challenge 2013**” with “B-hist: Entity-Centric Search over Personal Web Browsing History”. Top 18% of all submissions.
- Selected for the 1st **Heidelberg Laureate Forum**, 2013. 10% acceptance rate.
- “Crowdsourcing for Multimedia Ideas Competition” award at **CrowdMM** 2013.
- **Best Demo** Award at the ISWC 2011 conference. **Top 1%** of all submissions.
- 2nd **most Outrageous Idea** Award at the ISWC 2011 conference. **Top 8%** of all submissions.
- **PhD dissertation award** – “Diploma Thesis Award 2011”, Semiramis Research and Service Unit.
- 3rd Best Paper Award at the LA-WEB 2009 conference.
- Selected for the **IBM EMEA** Best Student Recognition Event 2006.
- World Finalist (representing Italy) at **Microsoft** Imagine Cup 2005 Software Eng competition.
- Finalist at the Italian **Microsoft** Imagine Cup 2004, 2005, 2006, 2007.

#### *Leadership Roles*

- AAAI HCOMP **Steering Committee** member, 2022-2024.
- **PC Chair** for the Short Paper track at the ACM International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR**), 2022.
- **General Chair** for ACM Int. Conference on Information and Knowledge Management (**CIKM**) 2021.
- ACM SIGIR Awards Chair, 2020-now.
- Co-Chair for the Australasian Document Computing Symposium (ADCS), 2019-2020.
- External Faculty at the Human Computation Institute, 2020-now.
- **EU Project coordinator** of H2020-FashionBrain (2017-2019), leading a research team spread across 6 organizations from UK, DE, CH, NL.
- Co-Chair of the “Crowdsourcing and Human Computation for the Web” track at the Web Conference (**WWW**), 2018.
- Semantic Web Challenges co-Chair at the International Semantic Web Conference (ISWC) 2019.
- External Adviser for Data Science MSc and BSc programs at the University of Chichester, UK, 2017.
- Program Director for “B. Computer Science” 2018-2022, U. Queensland.
- Program Director for “MSc Data Science” 2016-2017, U. Sheffield.
- Deputy program coordinator for “MSc Data Science” 2014-2016, U. Sheffield.
- Strategy group member 2015-2017, Information School, U. Sheffield.
- Ethics application reviewer, Information School, 2014-2017, U. Sheffield.

#### *Selected Publications*

##### **Crowdsourcing Platforms and Hybrid Human-Machine Systems for Data Analysis**

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 Bias. In: The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR 2020**). Xi'an, China, July 2020.

Alessandro Checco, Jo Bates, and Gianluca Demartini. Adversarial Attacks on Crowdsourcing Quality Control. In: Journal of Artificial Intelligence Research (**JAIR**), 2020.

Lei Han, Kevin Roitero, Ujwal Gadiraju, Cristina Sarasua, Alessandro Checco, Eddy Maddalena, and Gianluca Demartini. The impact of task abandonment in crowdsourcing. In: IEEE Transactions on Knowledge and Data Engineering (**TKDE**), October 2019.

Kevin Roitero, Eddy Maddalena, Gianluca Demartini, and Stefano Mizzaro. On Fine-Grained Relevance Scales. In: 41st Annual ACM SIGIR Conference (**SIGIR 2018**), Ann Harbor, Michigan, USA, July 2018. (Acceptance rate: 21%).

Djellel Eddine Difallah, Michele Catasta, Gianluca Demartini, Panagiotis G. Ipeirotis, and Philippe Cudré-Mauroux. The Dynamics of Micro-Task Crowdsourcing -- The Case of Amazon MTurk. In: 24th International Conference on World Wide Web (**WWW 2015**), Research Track. Firenze, Italy, May 2015. (Acceptance rate: 14.1%)

Ujwal Gadiraju, Ricardo Kawase, Stefan Dietze, and Gianluca Demartini. Understanding Malicious Behaviour in Crowdsourcing Platforms: The Case of Online Surveys. In: Proceedings of the ACM Special Interest Group on Computer Human Interaction (**CHI 2015**). Seoul, South Korea, April 2015. (Acceptance rate: 23%).

Djellel Eddine Difallah, Gianluca Demartini, and Philippe Cudré-Mauroux. Pick-A-Crowd: Tell Me What You Like, and I'll Tell You What to Do. In: 22nd International Conference on World Wide Web (**WWW 2013**), Rio de Janeiro, Brazil, May 2013. (Acceptance rate: 15%).

Gianluca Demartini, Djellel Eddine Difallah, and Philippe Cudré-Mauroux. ZenCrowd: Leveraging Probabilistic Reasoning and Crowdsourcing Techniques for Large-Scale Entity Linking. In: the 21st International Conference on World Wide Web (**WWW 2012**), Lyon, France, April 2012. (Acceptance rate: 12%).

### **Entity-Centric Information Access**

Kevin Roitero, Alessandro Checco, Stefano Mizzaro, and Gianluca Demartini. Preferences on a Budget: Prioritizing Document Pairs when Crowdsourcing Relevance Judgments. In: 31st ACM International World Wide Web Conference (**WWW 2022**). April 2022. (Acceptance rate: 17.7%).

Alberto Tonon, Michele Catasta, Roman Prokofyev, Gianluca Demartini, Karl Aberer, and Philippe Cudré-Mauroux. Contextualized Ranking of Entity Types based on Knowledge Graphs. In: **Journal of Web Semantics**, Elsevier. 2016. (Journal impact factor at publication time: 2.550).

Roman Prokofyev, Gianluca Demartini, and Philippe Cudré-Mauroux. Effective Named Entity Recognition for Idiosyncratic Web Collections. In: 23rd International Conference on World Wide Web (**WWW 2014**). Seoul, South Korea, April 2014. (Acceptance rate: 12.9%).

Alberto Tonon, Michele Catasta, Gianluca Demartini, Philippe Cudré-Mauroux, and Karl Aberer. TRank: Ranking Entity Types Using the Web of Data. In: The 12th International Semantic Web Conference (**ISWC 2013**), Sydney, Australia, October 2013. (Acceptance rate: 21.5%). Nominated for the Best Research Paper Award. (**Top 2%**).

Alberto Tonon, Gianluca Demartini, and Philippe Cudré-Mauroux. Combining Inverted Indices and Structured Search for Ad-hoc Object Retrieval. In: 35th Annual ACM SIGIR Conference (**SIGIR 2012**), Portland, Oregon, USA, August 2012. (Acceptance rate: 20%).

## Research Funding

### External Research Grants

- **Meta AI** Request for Proposals: Mephisto Dataset Collection Tool “Monitoring and Modeling Human Annotator Behaviors in Mephisto.” (PI, 36,228USD), 2022-2023.
- **Swiss National Science Foundation Sinergia** “Large-Scale Political Participation: Issue Identification, Deliberation, and Co-creation” (co-I, **2.4M CHF total cost**, 550K AUD funding to UQ), 2022-2026.
- **Google AI Research** “Human-in-the-loop Chimpanzee Identification” (PI, 10K USD), 2021-2022.
- ARC ITTC 2020 “ARC Training Centre on Information Resilience” (co-I and Theme Leader leading the 'Data Curation at Scale' theme - one of the five themes in the centre - **13.8M AUD total cost**, 4.9M AUD ARC contribution), 2020-2025.
- **Facebook Research**, “Crowdsourcing an Online Safety Benchmark for Fake News Identification” (PI, 41K USD funding to UQ), 2019-2020.
- ARC **Discovery Project 2019** “Building crowd sourced data curation processes” (co-I, **900K AUD total cost**, 440K AUD funding to UQ), 2019-2021.
- Industry funding by Neusoft Research of Intelligent Healthcare Technology, Co Ltd “Collaborative Lab of Health Informatics with Neusoft” (Co-I, **1M AUD total cost**, 469K AUD funding to UQ), 2019-2022.
- EU **H2020-ICT-14-2016 (PI, 2.9M EUR total cost**, 1.7M EUR EC Contribution, 420K EUR U Sheffield share) call “Big Data PPP: cross-sectorial and cross-lingual data integration and experimentation”, project “FashionBrain: Understanding Europe’s Fashion Data Universe”, 6 partners, 2017-2019.
- UK **EPSRC First Grant (PI, 120K GBP total cost**, 100K GBP EPSRC Contribution), project “BetterCrowd: Human Computation for Big Data”, individual grant, 2016-2017.
- UK **AHRC Collaborative Doctoral Partnership (Co-I, 70K GBP)** “Investigating entity-centric methods for information exploration and discovery in big digital archives”, 2015-2018.
- **Amazon AWS in Education Grants (PI, 2'500 USD)**, 2015.

### Internal Research Grants

- U Queensland, UQ Cyber Security Transdisciplinary Research Seed Funding (PI, 73K AUD), “A Data-Driven Analysis of Disinformation in Australian Social Media Campaigning”, 2021.
- U Queensland, UQ Early Career Researcher grant (PI, 22K AUD), “Artificial Intelligence with Humans in the Loop”, 2019.
- U Queensland, School of ITEE (PI, 30K AUD), “Virtual reality devices for the Data Science Human Computation Laboratory”, 2018.
- U Queensland, Faculty new staff start-up grant (PI, 50K AUD), “Understanding the Dynamics of a Crowdsourced Knowledge Graph”, 2017.
- U Sheffield iSchool Internal Seed Grant (PI, 700 GBP) “OpenTurk: A Task Recommender System for Crowd Workers”, 2016.
- U Sheffield iSchool Internal Seed Grant (Co-I, 1971 GBP) “Cultures of Data Science Practice”, 2016.
- Sheffield Methods Institute Fund (PI, 720 GBP) “Crowdsourcing Social Science Data Analysis”, 2015.
- U Sheffield iSchool Internal Seed Grant (PI, 2'300 GBP) “Crowdsourcing for Big Data”, 2015.

### External Travel Grants

- Universities Australia – DAAD Joint Research Co-operation Scheme “Human-in-the-loop Natural Language Processing in GLAM” (PI, 24'000 AUD), 2022-2023.
- Uni. Vienna non-EU teaching mobility, 2022.
- ESF ELIAS Scientific meeting (PI, 7'500 EUR) Dagstuhl meeting titled “Crowdsourcing Research - Transcending Disciplinary Boundaries”, 2016.
- ESF ELIAS STSM Travel grant for a research visit at U Udine, Italy, 2015.
- MUMIA STSM Travel grant for a research visit at CWI, Amsterdam, The Netherlands, 2013.
- International Short Visits grant from the Swiss National Science Foundation (SNSF), 2012.
- ACM Student Travel grant to attend SIGIR 2010.

## Community activities

### Journal Editorial Board

- Editorial Board Member for the **Information Retrieval Journal**, Springer, 2017-now.
- **Area Editor** for the Journal of **Web Semantics**, Elsevier, 2016-2023.
- **Co-Editor in Chief** for the Human Computation Journal, 2018-2021.

- Special Issue on eDemocracy & eGovernment (SIEDEG) of the IEEE Transactions on Emerging Topics in Computing, 2016-2017.
- Editorial Board Member for the Journal of **Web Semantics**, Elsevier, 2015-2016.

#### **Senior Roles in Program Committees**

- Senior PC member, International Joint Conference on Artificial Intelligence (**IJCAI**), 2023.
- Short Paper track **PC Chair**, the ACM International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR**), 2022.
- Senior PC member, AAAI Conference on Artificial Intelligence (**AAAI**), 2021.
- Senior PC member, the ACM International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR**), 2020-2022.
- Senior PC member, the ACM International Conference on Web Search and Data Mining (**WSDM**), 2020, 2022-23.
- Senior PC member, the “Semantics and Knowledge” research track at the Web Conference (**WWW**), 2023.
- Senior PC member, the “Crowdsourcing and Human Computation” research track at the Web Conference (**WWW**), 2020, 2023.
- Chair for the “Web and the Society” track at the ACM **Hypertext** conference, 2020.
- Senior PC member, the International Semantic Web Conference (**ISWC**), 2019-2020.
- Co-Chair of the “Crowdsourcing and Human Computation for the Web” track at the Web Conference (**WWW**), 2018.
- Senior PC member, AAAI Conference on Human Computation and Crowdsourcing (**HCOMP**) 2016, 2020
- Senior PC member, the International Conference on Web Engineering (**ICWE**), 2016-2017.
- Senior PC member, the ACM International Conference on Information and Knowledge Management (**CIKM**), 2015-2018, 2020, 2022.

#### **Program Committee Member**

- The 14th International Conference on Ambient Systems, Networks and Technologies (ANT), 2023.
- Extended Semantic Web Conference (ESWC) PhD Symposium, 2021.
- **AAAI** Conference on Artificial Intelligence, 2020.
- The IEEE International Conference on Distributed Computing Systems (ICDCS), 2020.
- AAAI workshop on Knowledge Discovery from Unstructured Data in Financial Services (KDF), 2020-22.
- Semantics & Knowledge research track at the Web Conference (**WWW**), 2020.
- International Workshop on Knowledge Graphs for Online Discourse (KnOD) at ESWC, 2020.
- IEEE Workshop on Human-in-the-Loop Methods and Future of Work in BigData (HMDData) at IEEE Bigdata 2020-2022.
- ACM Conference on Fairness, Accountability, and Transparency (FAT\*), 2019.
- Workshop on Fairness, Accountability, Confidentiality, Transparency, and Safety in Information Retrieval at SIGIR, 2019.
- Australasian Database Conference (ADC), 2019-2020.
- International Workshop on Intelligent User Interfaces for Algorithmic Transparency in Emerging Technologies at ACM IUI, 2019.
- Workshop on Quality of Open Data (QOD) at BIS, 2019.
- The IEEE International Conference on Data Mining (**ICDM**), 2018.
- International Conference on Database Systems for Advanced Applications (DASFAA), 2018.
- Social Informatics conference (SocInfo), 2018, 2020.
- ACM International Web Search and Data Mining (WSDM) Conference, Demo track, 2018.
- ACM International Web Search and Data Mining (WSDM) Conference, Doctoral Consortium, 2018.
- AAAI Conference on Human Computation and Crowdsourcing (**HCOMP**), 2018-2021.
- Research track of the International Semantic Web Conference (**ISWC**), 2012, 2017.
- Crowdsourcing research track at the World Wide Web Conference (**WWW**), 2017-2020.
- ACM International Conf. on Research and Development in Information Retrieval (**SIGIR**), 2012-2020.
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**), 2017-2018.
- Australasian Document Computing Symposium (ADCS), 2017-2018.
- Poster track at the World Wide Web Conference (**WWW**), 2017-2018.
- International ACM Web Science Conference (**WebSci**), Poster Track, 2017, 2019.
- iConference, 2017-2018.

- Workshop on Crowdsourcing and Multiagent Systems at the International Conference on Autonomous Agent and Multiagent Systems (ICAAMAS), 2017.
- Workshop on Data Science for Intelligent Decision Support (DS4IDS) at the 5th International Conference on Future Internet of Things and Cloud (FiCloud), 2017.
- Workshop on Knowledge Graphs and Semantics for Text Retrieval and Analysis (KG4IR) at SIGIR 2017.
- International Joint Conference on Artificial Intelligence (**IJCAI**), 2013, 2016, 2019, 2020, 2022.
- The International AAI Conference on Weblogs and Social Media (**ICWSM**) 2014, 2017-2018.
- Content Analysis research track at the World Wide Web Conference (**WWW**), 2016.
- Doctoral Consortium at the International Semantic Web Conference (**ISWC**), 2016-2019.
- Resources Track at the International Semantic Web Conference (**ISWC**), 2016.
- Special Session on Collective Intelligence, Crowd Sourcing, and Collaboration (CICS) at the International Conference on Collaboration Technologies and Systems (CTS), 2016.
- Russian Summer School in Information Retrieval (RussIR), 2016.
- Conference on Artificial Intelligence and Natural Language (AINL), 2016-2017.
- IEEE Conference on Business Informatics (CBI), 2016.
- Italian Conference on Computational Linguistics (CLIC-IT), 2016.
- International Workshop on Benchmarks for Ubiquitous Crowdsourcing: Metrics, Methodologies, and Datasets (Crowdbench) at IEEE PERCCOM 2016.
- Workshop on Data Science for Intelligent Decision Support (DS4IDS) at the International Conference on Computational Science and Applications (ICCSA), 2016.
- Workshop on LINKed Education at ISWC, 2015.
- Empirical Studies and Experiments track of the International Semantic Web Conference (**ISWC**), 2015.
- Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR) at CIKM, 2015-2016.
- ACM SIGIR International Conference on the Theory of Information Retrieval (**ICTIR**), 2015-2017.
- International Symposium on String Processing and Information Retrieval (**SPIRE**), 2015.
- Symposium on Data-Driven Process Discovery and Analysis (SIMPDA), 2015.
- International Conference on Knowledge Discovery and Information Retrieval (KDIR), 2015.
- International Workshop on the Theory and Practice of Social Machines (SOCM) at WWW, 2015-2016.
- Social Web Applications track at the International Conference on Web Engineering (**ICWE**) 2015.
- Social networks and graph analysis track at the World Wide Web Conference (**WWW**), 2013-2014.
- Replication, Benchmark and Data Track of the International Semantic Web Conference (**ISWC**), 2014.
- Extended Semantic Web Conference (**ESWC**), 2013-2016, 2018-2021.
- ACM International Conference on Information and Knowledge Management (**CIKM**), 2011, 2014.
- International Workshop on Social Media in Crowdsourcing and Human Computation at SocInfo, 2014.
- International Conference on Advances in Semantic Processing (SEMAPRO), 2014-2016.
- “Crowdsourcing” Session at the Int. Conf. on Communications and Electronics (ICCE), 2014-2015.
- “Computational Science and Computational Intelligence” Session at the Int. Conf. on Advanced Technologies for Communications (ATC), 2014.
- CIKM 2012 City Data Management 2012 Workshop.
- SIGIR 2012 1st Joint International Workshop on Entity-oriented and Semantic Search, 2012.
- SIGIR Workshop on Time-aware Information Access (TAIA), 2012-2015.
- Artificial Intelligence meets the Web of Data Workshop at ECAI 2012 and ESWC 2013.
- International Workshop on Learning and Education with the Web of Data at WWW (LILE), 2012-2016.
- European Conference on Information Retrieval (**ECIR**), 2012-2014, 2022.
- First International Workshop on Entity-Oriented Search (EOS) at SIGIR, 2011.
- Posters and Demonstrations Track of the 9th Intern. Semantic Web Conference (ISWC), 2010.
- Personal Information Management (PIM) Workshop 2009, 2012.

### **Organization Roles**

- Tutorial co-Chair for the 17th International AAI Conference on Web and Social Media (ICWSM), Cyprus, 2023.
- Co-organizer of the ‘Workshop On Bias in Automatic Knowledge Graph Construction’ at the Conference on Automated Knowledge Base Construction (AKBC), 2020.
- Co-organizer of the ACM International Conference on Information and Knowledge Management (CIKM) 2020 Doctoral Consortium.
- Metadata Chair for the International Semantic Web Conference (**ISWC**), 2020.
- Co-organizer of Semantic Web Challenges at the International Semantic Web Conference (**ISWC**), 2019.

- Co-organizer of the ‘International Workshop On Job Knowledge Discovery on the Web & Social Media’ (JOBNOW) at **ACM HyperText**, 2019.
- Co-organizer of the ‘Augmenting Intelligence with Bias-aware Humans-in-the-Loop’ (HumBL2018) Workshop at **The Web Conference**, 2019.
- Co-organizer of the ‘CrowdBias: Disentangling the Relation Between Crowdsourcing and Bias Management’ (CrowdBias2018) Workshop at **HCOMP**, 2018.
- Publicity co-chair for the AAAI Conference on Human Computation and Crowdsourcing (HCOMP), 2018.
- Co-organizer of the ‘Augmenting Intelligence with Humans-in-the-Loop’ (HumL2018) Workshop at **WWW** and **ISWC**, 2018.
- Student Coordinator for the International Semantic Web Conference (**ISWC**), 2017.
- External Liaison Chair at the 2017 Workshop on Hybrid Human-Machine Computing (HHMC 2017).
- Co-organizer of the workshop ‘Weaving Relations of Trust in Crowd Work: Transparency and Reputation across Platforms’ at WebScience 2016.
- Human Computation and Crowdsourcing Track at the Extended Semantic Web Conference (**ESWC** 2015)
- Member of the Local Organization team for CLEF 2014, Sheffield.
- Entity Ranking Track at the 9th Initiative for the Evaluation of XML Retrieval (INEX 2009).
- Entity Ranking Track at the 8th Initiative for the Evaluation of XML Retrieval (INEX 2008).

#### **Best Paper Award Selection Committee Member**

- International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR) Best short paper award, 2020.
- ACM International Web Search and Data Mining (WSDM) Conference, 2019.
- ACM Conference on Fairness, Accountability, and Transparency (ACM FAT\*), 2019.

#### **Reviewer for Journals**

- Scientific Reports, Springer Nature, 2021.
- IEEE Access, 2020.
- EPJ Data Science, 2018.
- Knowledge and Information Systems (KAIS), 2018.
- IEEE Transactions on Knowledge and Data Engineering (**TKDE**), 2012, 2017-2018.
- ACM Transactions on Internet Technology (TOIT), 2017.
- Foundations and Trends in Information Retrieval (**FnTIR**), 2017.
- Computer Supported Cooperative Work Journal (CSCWJ), Springer, 2017.
- Journal of Information Science (JIS), 2017.
- Engineering Applications of Artificial Intelligence, Elsevier, 2017.
- Communications of the ACM, 2016.
- The Computer Journal, Oxford University Press, 2016.
- Online Information Review, Emerald, 2016.
- IEEE Transactions on Services Computing (TSC), 2015-2016.
- Information Processing & Management (**IPM**), 2015, 2019.
- Information Research, 2015.
- Information Systems, Elsevier, 2015, 2018.
- Knowledge-Based Systems, Elsevier, 2015.
- Engineering Applications of Artificial Intelligence, Elsevier, 2015.
- ACM TIST Special Issue on Crowd in Intelligent Systems, 2015.
- Semantic Web journal (SWJ) Special Issue on Human Computation and Crowdsourcing (HC&C), 2015
- Semantic Web journal (SWJ) Special Issue on Ontology and Linked Data Matching, 2014.
- Journal of Machine Learning Research (**JMLR**), 2014.
- International Journal of Human-Computer Studies (**IJHCS**), 2014, 2016.
- ACM Transactions on Information Systems (**TOIS**), 2014, 2017.
- Special Issue of the Journal of Distributed and Parallel Databases (DAPD) on “Databases and Crowdsourcing”, 2014.
- Journal of Computer Science and Technology, Springer, 2014.
- IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**), 2014-2016.
- ACM Transactions on the Web (**TWEB**), 2013-2016.
- Special Issue of the Journal of **Web Semantics** on “Semantic Search”, Elsevier, 2013.
- Special Issue of the Journal of **Web Semantics** on “Semantic Web Challenge 2013”, Elsevier, 2014.

- Journal of Data and Information Quality (JDIQ), 2013, 2015.
- Majlesi Journal of Electrical Engineering (MJEE), 2013.
- Springer book series: "Theory and Applications of Natural Language Processing", 2012.
- **Information Retrieval**, Springer, 2012, 2015-2017.
- Interactive Learning Environments, Routledge, 2012.
- Journal of Web Engineering (JWE), Rinton Press, 2012.
- IEEE Computational Intelligence Magazine, 2011.
- Knowledge and Information Systems Journal (KAIS), Springer, 2011.
- Artificial Intelligence Journal (AIJ), Elsevier, 2010-2011.
- Journal of the American Society for Information Science and Technology (**JASIST**), 2010, 2016.
- Journal of Network and Computer Applications, Elsevier, 2010.
- International Journal of Web Information Systems, Emerald, 2009.

#### **External Reviewer for Conferences**

- ACM CHI, 2020.
- ACM CSCW, 2019.
- ACM UIST, 2019.
- ACM CHI 2019 Late Breaking Work track, 2019.
- ACM CHI 2015 alt.chi track, 2015.
- International Conference on Information Systems (ICIS) 2014, IC Design Science Track.
- Research track of the 12th International Semantic Web Conference ISWC 2013.
- AAAI Conference on Artificial Intelligence (AAAI-13).
- 1st International Workshop on "Crowdsourcing the Semantic Web" at ISWC 2013.
- Demo Track of the 38th International Conference on Very Large Databases VLDB 2012.
- Semantic Web Track of the 21st International World Wide Web Conference WWW 2012.
- Research Track of International Conference on Extending Database Technology EDBT 2012, 2013.
- In-Use Track of the 10th International Semantic Web Conference ISWC 2011.
- Poster Track of the International World Wide Web Conference WWW 2008, 2013.
- Poster Track of the 31st International ACM SIGIR 2008 Conference.

#### **Session Chairing at Conferences**

- "Evaluation and User Studies" Session at SIGIR 2022.
- "Graph Search" Session at WWW 2022.
- "Bias and Counterfactual Learning 2" Session at SIGIR 2021.
- "Entities and Relations" Session at CIKM 2020.
- "Advanced methodology and measurement" Session at WSDM 2020.
- "Learning to be Efficient" Session at HCOMP 2016.
- "Crowdsourcing & Web Science" Session at ESWC 2015.
- "Crowdsourcing 1" Session at WWW 2014.
- "Social and Collaborative Semantics" Session at ISWC 2012.

#### **Service for Funding Bodies (Reviewer for funding applications)**

- Australian Research Council (ARC), 2018-2021.
- Czech Science Foundation, 2020-2021.
- H2020 SoBigData project funding review committee, 2019.
- UK Engineering and Physical Sciences Research Council **EPSRC Full College member**, 2017-now.
- Netherlands Organisation for Scientific Research (NWO), 2017-2018.
- UK Engineering and Physical Sciences Research Council EPSRC Associate College member, 2016-2017.
- Commonwealth Scholarship Commission (CSC), 2017.
- UK Science and Technology Facilities Council (STFC), 2017.
- UK Engineering and Physical Sciences Research Council (EPSRC), 2014-2015.
- UK Arts & Humanities Research Council (AHRC), 2014.
- The German Academic Exchange Service (DAAD), 2015.

#### *Invited Talks & Tutorials (Selected)*

#### **Invited Talks**



- “The impact of human assessors on judgements, labels, supervised models, and evaluation results” - **Dagstuhl Seminar** 23031 on Frontiers of Information Access Experimentation for Research and Education, 2023.
- “Bias in Human-In-The-Loop Artificial Intelligence”, Invited talk at the Crowd Science workshop at **NeurIPS** 2020.
- “Knowledge Graph Quality Management” – Invited Keynote at the Semantic Web solutions for large-scale biomedical data analytics (SeWeBMeDA-2019) Workshop at ISWC 2019.
- “Content moderators: the gatekeepers of social media” - **TEDxUQ**, Brisbane, Australia.
- “Human Factors and Bias in Crowdsourced Information Retrieval Evaluation”, University of Zurich, Switzerland, 2018.
- “Human Factors and Bias in Crowdsourced Information Retrieval Evaluation”, University of Fribourg, Switzerland, 2018.
- “Human Computation for Entity-Centric Information Access”, Università della Svizzera Italiana (USI), Lugano, Switzerland, 2017.
- “The Power of Big Data” - ACM Distinguished Speaker talk as **keynote** at the Second International “Digital Transformation & Global Society” Conference (DTGS’17), St Petersburg, 2017.
- “Crowdsourcing for Data Management”, invited talk at the PhD School of the Australasian Database Conference (ADC) 2017, Brisbane, 2017.
- “Human Factors in Crowdsourcing” – Research Seminar at the University of Fribourg, 2017.
- “The Power of Big Data” - ACM Distinguished Speaker talk at Accenture Latvia, 2017.
- “Human Computation for Entity-centric Information Access” – Research Seminar at U Glasgow, 2016.
- “Human Computation for Entity-centric Information Access” – Research Seminar at U Leeds, 2016.
- “Introduction to Crowdsourcing” and “Quality at Scale” – Invited talks at **Facebook**, London, UK, 2016.
- “Human Computation for Entity-centric Information Access” – Research Seminar at U Twente, 2016.
- “Human Computation for Entity-Centric Information Access” – Research Seminar at UCL, 2016.
- “Crowdsourcing for Entity-centric Information Access” – **Keynote** at AINL-ISMW, St Petersburg, 2015.
- “Leveraging n-gram statistics for grammatical correction and entity recognition” – Invited talk at **Facebook**, London, UK, 2015.
- “Crowdsourcing for Entity-centric Information Access” - **Keynote** at KDWeb, Cagliari, Italy, 2015.
- “The randomForest package” - Sheffield R Users Group, 2015.
- “MSc Data Science at the University of Sheffield” - Open Data Science Meetup, U Sheffield, 2014.
- “Entity-Centric Information Access” – NLP Group research seminar, U Sheffield, 2014.
- “Human Computation: Machines are not enough” – Ignite Academy, U Sheffield, 2014.
- “Crowdsourcing is for the tail” – **Dagstuhl Seminar** 14282 - Crowdsourcing and the Semantic Web, 2014.
- “Human Computation for Big Data” – CUSO Seminar on Big Data, Fribourg, 2014.
- “Entities, Graphs, and Crowdsourcing for better Web Search” – UvA, Amsterdam, The Netherlands, 2013.
- “Workshop on Crowd-Sourcing” - Heidelberg Laureate Forum, Heidelberg, Germany, 2013.
- “Next-Generation Micro-task Crowdsourcing Platforms” - **Dagstuhl Seminar** 13361 on Crowdsourcing: From Theory to Practice and Long-Term Perspectives, 2013.
- “Entities, Graphs, and Crowdsourcing for better Web Search” – USI, Lugano, Switzerland, 2013.
- “Getting Semantics from the Crowd” - **Dagstuhl Seminar** 12171 on Semantic Data Management, 2012.
- “Current Approaches to Search Result Diversification” - Search Computing Workshop, Milano, Italy, 2010.

### **Tutorials**

- Tracking User Behaviours: Laboratory-Based and In-The-Wild User Studies, Tutorial at **DASFAA** 2019, Apr 2019, Chiang Mai, Thailand.
- “Crowdsourcing for Data Management” - course at the PhD School of the Australasian Database Conference (ADC) 2017, Brisbane, Australia.
- Course (external consultancy) on Data Mining for Lhasa Limited, May 2017, Leeds, UK.
- “Using Crowdsourcing Effectively for Social Media Research” - Tutorial at **ICWSM** 2016, Cologne, Germany with Ujwal Gadiraju, Djellel E. Difallah, and Michele Catasta.
- “It’s Getting Crowded! How to Use Crowdsourcing Effectively for Web Science Research” - Tutorial at **Web Science** 2016, Hannover, Germany with Ujwal Gadiraju, Djellel E. Difallah, and Michele Catasta.
- “Crowdsourcing for Data Processing and Search” - Half-day tutorial at Search Solutions 2015, London, UK.
- “Improving the quality of crowdsourced annotations” tutorial at IEEE AINL-ISMW 2015, St.Petersburg, Russia, 2015.

- “Leveraging **Knowledge Graphs** for Web Search” - RuSSIR 2015 course (8 hours), St Petersburg, 2015.
- “Micro-task Crowdsourcing” - PhD course (14 hours) at U Udine, 2015.
- Full-day tutorial “Microtask crowdsourcing to solve Semantic Web problems” at **ISWC** 2013 with Elena Simperl and Maribel Acosta. (<https://sites.google.com/site/microtasktutorial/>).
- Full-day tutorial “Crowdsourcing for the Semantic Web” at **ESWC** 2013 with Elena Simperl and Maribel Acosta. (<https://sites.google.com/site/crowdsourcingtutorial/>).
- Full-day tutorial “From Expert Finding to **Entity Search** on the Web” at **ECIR** 2012 with Peter Mika, Thanh Tran, and Arjen P. de Vries. (<http://diuf.unifr.ch/main/xi/EntitySearchTutorial/>).

#### **Public Engagement Activities**

- “The Gig Economy: challenges and opportunities”, Sheffield Festival of Science and Engineering, British Science Week, March 2017.
- “The power of Big Data”, Café Scientifique, Sheffield, 2015.
- “What is Big Data and why should we care?” talk for Mobile University, Sheffield, 2015.  
Invited talk at the University of Sheffield Ignite Academy titled “Human Computation: Machines are not enough” for the Festival of the Mind 2014 (<http://festivalofthemind.group.shef.ac.uk>).

#### *Teaching*

- Course leader of “**DATA2001 Foundations of Data Science**” 2022-now for B Comp Sc, U. Queensland.
- Course leader of “**DATA7201 Data Analytics at Scale**” 2018-now for M Data Science, U. Queensland.
- Course leader of “**INFS3200 Advanced Database Systems**” 2018-2019, U. Queensland.
- Course leader of “**Big Data Analytics**” 2016/2017 for MSc Data Science, U. Sheffield.
- Course leader of “**Data Mining and Visualisation**” 2016/2017 for MSc Data Science, U. Sheffield.
- Teacher for Course “**Introduction to Data Science**” 2016/2017 for MSc Data Science, U. Sheffield.
- Course leader of “**Data Mining and Visualisation**” 2015/2016 for MSc Data Science, U. Sheffield.
- Teacher for Course “**Introduction to Data Science**” 2015/2016 for MSc Data Science, U. Sheffield.
- Course leader of “**Data Mining and Visualisation**” 2014/2015 for MSc Data Science, U. Sheffield.
- Teacher for Course “**Introduction to Data Science**” 2014/2015 for MSc Data Science, U. Sheffield.
- Leader for Course “**Social Computing**” 2013/2014 for the MSc in Computer Science, U. of Fribourg.
- Leader for Course “**Social Computing**” 2012/2013 for the MSc in Computer Science, U. of Fribourg.
- Teacher for Course “**Social Computing**” 2011/2012 for the MSc in Computer Science, U. of Fribourg.
- Guest teacher for Course “**Web Science**” 2010/2011 at the Leibniz University of Hannover.
- TA for “Artificial Intelligence II” 2008/2009 at the Leibniz University of Hannover.

#### *Internal Service*

- Program Group Convenor for Computer Science and Data Science degrees at the School of ITEE, University of Queensland, 2019-2022.
- Credit assessment for BCompSci at the School of ITEE, University of Queensland 2018-now.
- Student Advisor for BInfTech at the School of ITEE, University of Queensland 2017-2018.
- Program Convenor for “**B. Computer Science**” 2018-now, U. Queensland.
- Program Coordinator for “**MSc Data Science**” 2016-2017, U. Sheffield.
- Member of the Strategy Committee of the Information School, University of Sheffield 2015-2017.
- Member of the Teaching Committee of the Information School, University of Sheffield 2014-2017.
- Member of the ethics review board of the Information School, University of Sheffield 2014-2017.

#### **Ph.D. Supervision**

##### *Ongoing.*

- Linh Le, University of Queensland, 2020-now. (Main supervisor)
- Shaoyang Fan, University of Queensland, 2019-now. (Main supervisor)
- Shaochen Yu, University of Queensland, 2020-now. (Secondary supervisor)
- Jiechen Xu, University of Queensland, 2020-now. (Secondary supervisor)
- Hatim Lahza, University of Queensland, 2020-now. (Secondary supervisor)
- Tianwa Chen, University of Queensland, 2019-now. (Secondary supervisor)

##### *Completed.*

- Lei Han, University of Queensland, 2018-2021. (Main supervisor)
- Kexuan Xin, University of Queensland, 2019-2022. (Secondary supervisor)
- Jimmy, University of Queensland, 2019. (Secondary supervisor)

- Humaira Ehsan, University of Queensland, 2019. (Secondary supervisor)
- Rehab Qarout, University of Sheffield, 2018. (Secondary supervisor)
- Mengdie Zhuang, University of Sheffield, 2018. (Secondary supervisor); now faculty member at U. Sheffield, UK.
- Wasim Ahmed, "Using Twitter data to provide qualitative insights into pandemics and epidemics", University of Sheffield, 2018. (Secondary supervisor); now faculty member at U. Newcastle, UK.

#### **External Examiner for Ph.D.**

- Markus Zlabinger, "Efficient and Effective Manual Corpus Annotation", TU Wien, Austria, July 2021.
- Danula Hettiachchi, "Task Assignment using Worker Cognitive Ability and Context to Improve Data Quality in Crowdsourcing", University of Melbourne, Australia, May 2021.
- Alfian Farizki Wicaksono, "Modelling Search and Session Effectiveness", University of Melbourne, Australia, December 2020.
- Ziyang Yang, "Improving the Reliability and Robustness of Information Retrieval Evaluation", University of Melbourne, Australia, November 2019.
- Yuan Jin, "Modelling Techniques and Inference Procedures for Improving Quality Control in Crowdsourcing Applications", Monash University, Australia, August 2018.
- Dilesha Seneviratne, "Patent Link Discovery", Queensland University of Technology, Australia, July 2018.
- Evi Yulianti, "Answer-biased Summarization", RMIT, Australia, May 2018.
- Panagiotis Mavridis, "Using Hierarchical Skills for Optimized Task Selection in Crowdsourcing", Rennes, France, November 2017.
- Ujwal Gadiraju, "It's Getting Crowded! Improving the Effectiveness of Microtask Crowdsourcing", Leibniz Universität Hannover, Germany, October 2017.
- Kinda El-Maarry, "Socially-aware Quality Control for Crowdsourcing", TU Braunschweig, Germany October 2017.
- Alberto Tonon, "Leveraging Entity Types and Properties for Knowledge Graph Exploitation", U Fribourg, Switzerland, 2017.
- S.M. Khelghati, "Deep Web Content Monitoring", U Twente, Netherlands, 2016.
- Michele Catasta, "Memory-based Information Systems", EPFL, Switzerland, 2015.
- Djellel E. Difallah, "Quality of Service in Crowd-Powered Systems", U Fribourg, Switzerland, 2015.

#### **M.Sc. Thesis Supervision**

- Supervisor of 50+ undergraduate and post-graduate dissertations at the University of Queensland, Australia during 2018-2022.
- Supervisor of 11 MSc dissertations (8 in Data Science, 2 in Information Management, 1 in Information Systems Management) at the University of Sheffield, UK during 2015/2016.
- Supervisor of 10 MSc dissertations (6 in Data Science, 2 in Information Management, 1 in Information Systems, 1 in Information Systems Management) at the University of Sheffield, UK during 2014/2015.
- Supervisor of 7 MSc Computer Science dissertations at the University of Fribourg, Switzerland during 2011-2014.

#### *EU Project Participation*

- EU H2020 FashionBrain project (<http://www.fashionbrain-project.eu>), project coordinator.
- EU FP7 Arcomem project (<http://www.arcomem.eu>), WP leader.
- EU FP7 LivingKnowledge project (<http://livingknowledge.europarchive.org/>), WP leader.
- EU FP7 Okkam project (<http://fp7.okkam.org>): "Enabling the Web of Entities", WP leader.
- EU FP6 Nepomuk project (<http://nepomuk.semanticdesktop.org>): "The Social Semantic Desktop". WP leader.