

Report on the 43rd European Conference on Information Retrieval (ECIR 2021)

Raffaele Perego and Fabrizio Sebastiani
Istituto di Scienza e Tecnologie dell'Informazione
Consiglio Nazionale delle Ricerche
56124 Pisa, Italy
firstname.lastname@isti.cnr.it

Abstract

The 43rd European Conference on Information Retrieval (ECIR 2021), organized under the auspices of the Information Retrieval Specialist Group of the British Computer Society (BCS IRSG), took place between March 28 and April 1, 2021. As sadly customary in these dark times, the conference was held entirely online, due to the COVID-19 pandemic. According to the original plans, it should have instead taken place in Lucca, a small Italian town in Tuscany, Italy, which enjoys a beautiful, extremely well-preserved historic centre.

1 Structure of the Conference

ECIR 2021¹ had the traditional structure of previous ECIR conferences, i.e., its program consisted of five days, with the first day (March 28) devoted to tutorials and the Doctoral Consortium, the three days that followed (March 29 to 31) occupied by the main conference, and the last day (April 1) devoted to satellite workshops and the industry day.

On the first day, two full-day (FD) tutorials and six half-day (HD) tutorials were offered, on themes ranging from “Reinforcement Learning for Information Retrieval” (FD – by Alexander Kuhnle, Miguel Aroca-Ouellette, Murat Sensoy, John Reid, and Dell Zhang), “IR From Bag-of-words to BERT and Beyond through Practical Experiments” (FD – by Sean MacAvaney, Craig Macdonald, and Nicola Tonellotto), “Fake News, Disinformation, Propaganda, Media Bias, and Flattening the Curve of the COVID-19 Infodemic” (HD – by Preslav Nakov and Giovanni Da San Martino), “Operationalizing Treatments against Bias” (HD – by Ludovico Boratto and Mirko Marras), “Large-Scale Information Extraction under Privacy-Aware Constraints” (HD – by Rajeep Gupta and Ranganath Kondapally), “Search Among Sensitive Content” (HD – by Graham Mcdonald and Douglas Oard), “Adversarial Learning for Recommendation” (HD – by Vito Walter Anelli, Yashar Deldjoo, Tommaso Di Noia, and Felice Antonio Merra), to “Biomedical Data, Text

¹<https://www.ecir2021.eu/>

Retrieval and Semantic Processing” (HD – by Francisco Couto). The 2nd tutorial in this list was the best-attended, with 149 attendees.

In parallel to the tutorials, the Doctoral Consortium was held, which offered five PhD students from around the world the opportunity to discuss their research plan with individually chosen mentors.

The three days of the main conference featured 3 keynote talks, 50 full paper presentations (out of 211 submissions, for a 23.7% Acceptance Rate), 11 presentations from the Reproducibility Track (47.8% AR), 7 presentations of papers recently published on the Information Retrieval Journal, 12 short-paper presentations of the CLEF 2021 Labs, 39 poster presentations of short papers (28.5% AR), 15 demos (48.4% AR), and a panel. Submissions were received from 40 different countries, with China, India, Germany, the US, and the UK, ranking (in this order) as the top 5 countries in terms of number of submissions. The 3 keynotes were by Ricardo Baeza-Yates (on “Ethics in AI” – the best-attended session in the entire conference, with 214 attendees), Ahmed H. Awadallah (winner of the Karen Sparck Jones Award – on “Learning with Limited Labeled Data: The Role of User Interactions”), and Ophir Frieder (on “Untraditional (Computer) Medicine”). We are particularly grateful to Ricardo Baeza-Yates for accepting to give his keynote with a one-day advance notice, as a replacement of the originally planned keynote by Francesca Rossi, who had a last-minute impediment.

The panel, held in a plenary session on the 3rd day of the conference, was on the theme of “Open Access and IR Literature”. It was coordinated by Djoerd Hiemstra, with invited panelists Hassina Aliane (CERIST, Algeria), Ralf Gerstner (Springer – the publisher of the ECIR proceedings for the last 20 years), Min-Yen Kan (National University of Singapore – former director of the ACL Anthology), Haiming Liu (University of Bedfordshire – representative of BCS IRSG), João Magalhães (Universidade Nova de Lisboa – General co-Chair of ECIR 2020), Hussein Suleman (University of Cape Town), and Min Zhang (Tsinghua University – ACM TOIS Editor-in-Chief). The discussion dwelled on the growing request from the scientific community to publish scientific literature (and, specifically of relevance to the present context, the ECIR proceedings) in an open access way (see also [Hiemstra et al., 2020]), and touched upon several sensitive issues, such as the fact that many research institutions simply cannot afford subscribing to the digital libraries that host IR-related publications, that this prevents young researchers in those institutions (and in entire countries) to take up IR as a subject (instead of, say, NLP, whose literature is almost 100% free – with no cost for authors and readers alike – thanks to the ACL Anthology), and that experiences indicating that this state can be changed abound (e.g., the proceedings of top NLP conferences such as ACL, NAACL, EACL, EMNLP, and journals such as JAIR and JMLR). A straw poll conducted among the attendees at the end of the panel indicated that these attendees believe that the ECIR proceedings should go open access (97% of respondents voted in this way), that the model to imitate should be that of the ACL Anthology, i.e., zero cost for authors, zero cost for readers (59% of respondents), and that the costs of running such an endeavour should be covered via a combination of volunteer work, sponsor money, and conference registration fees (65% of respondents); concerning the last question, only 24% of respondents believe that article processing charges should be the solution.

Finally, the last day of the conference hosted the Industry Day (organized by Roi Blanco and Fabrizio Silvestri) and five satellite workshops, on themes ranging from “Algorithmic Bias in Search and Recommendation” (organized by Ludovico Boratto, Stefano Faralli, Mirko Marras, and

Giovanni Stilo), “Bibliometric-Enhanced Information Retrieval” (organized by Ingo Frommholz, Philipp Mayr, Guillaume Cabanac, and Suzan Verberne), “Mixed-Initiative Conversational Systems” (organized by Ida Mele, Cristina Muntean, Mohammad Aliannejadi, and Nikos Voskarides), “Reducing Online Misinformation through Credible Information Retrieval” (organized by Fabio Saracco and Marco Viviani), to “Narrative Extraction from Texts” (organized by Ricardo Campos, Alípio Jorge, Adam Jatowt, Sumit Bhatia, and Mark Finlayson).

The Industry Day presented interesting success stories by IR researchers working in prominent companies, and featured two keynote speeches. The first keynote was on “Search and Discovery for Finance, in Practice” by Edgar Meij (Bloomberg), while the second, “Geometric Deep Learning: From Euclid to Drug Design”, was given by Michael Bronstein (Twitter and Imperial College).

The proceedings of ECIR 2021 are published by Springer Nature in its Lecture Notes in Computer Science series [[Hiemstra et al., 2021](#)].

2 The Online Experience

Organizing ECIR 2021 was challenging, since, like for many other online conferences, it required making preparations and bookings for a regular, in-presence conference (which we believed, or wanted to believe, would be the case until 6 months before the event), and eventually setting up things for the conference to be run completely online. Concerning the latter, the bar had been set very high by the organizers of ECIR 2020, who had managed to set up an online-only conference of astounding quality with very little advance notice, given that Europe had started to realize the impact that pandemic would have, barely weeks before ECIR 2020 took place.

Three online platforms, i.e., Whova, Zoom Webinars, and Gather Town, were used for the conference. As for many other conferences, Whova was used to generate an online, mobile-friendly conference schedule to be used by all attendees as the only entry point to the conference. The Whova schedule pointed to Zoom webinars, that were used as the platform for running all ECIR 2021 events other than the short paper presentations and the demos, and to Gather Town, which was used for the short paper presentations and the demos. Gather Town provides a simulated environment where attendees can meet and interact, and was thus also used during coffee breaks and lunch breaks for attendees to meet spontaneously, discuss, and network.

Going entirely online also means entirely reconsidering the registration fee structure. We opted for asking a flat, nominal registration fee (150 EUR) to authors only, in a one-author-per-paper fashion. Given that the about 125 paying authors, together with the contributions from the sponsors (Bloomberg Engineering, SIGIR, Amazon Science, eBay, Google, Signal, TextKernel, Springer), allowed us to cover our costs, we decided to grant a free registration to all other attendees, with the goal of maximizing participation, especially from developing countries. This generated a high number of registrants (more than 1,100) from no less than 63 countries, with Germany, the US, the UK, the Netherlands, and Italy, ranking (in this order) as the top 5 countries in terms of registrants. Of the more than 1,100 registrants, 1,028 logged into the conference at least once; 40% of them used the mobile app at least once while 85% of them used the web app at least once.

We all wish that future ECIRs will be back to the in-presence format, where speakers, attendees, and sponsor representatives will be able to meet face-to-face and network without the somehow clumsy barriers generated by the online platforms.

The online-only format does not have negative aspects only, though; it's just a different experience from in-presence conferences. While many “regular” attendees of ECIRs obviously complained that interaction and the experience is not the same as when we meet in presence with our favourite drink in hand, it is quite clear that many other attendees would simply not have been there if the conference had been of the traditional kind. This includes participants from developing countries (it was particularly heart-warming to receive “Thank you” notes from delegates from such countries), participants from far away countries whose travel expenses would probably be too high to justify participation in a “physical” ECIR, and participants who, although they can in principle afford the money to travel to Europe, would not do this in practice, e.g., because they do not have a paper in the program, or because work or family commitments would not allow them to be away for a week.

It was very easy to notice, instead, that the type of participation that the online format encourages is very different from that of in-presence conferences. When we travel to a “physical” conference, it is often the case that more than, say, 80% of our time in the day is devoted to the conference (attending sessions, networking with other attendees on or off the venue, etc.). For an online conference, this is different, and our participation is part-time: we attend a session we are interested in, we skip the next because we need to attend a departmental meeting, we then attend a keynote, and then we skip the other talks of the day since we need to pick up the kids at school and go grocery shopping. We realized this would be the case weeks before the event, when an unusually high number of registered attendees declined our invitations to be session chairs; because of meetings, family commitments, classes to be taught, and other commitments that would simply not be there when attending an in-presence conference. Indeed, this phenomenon was also clear from the data: while more than 1,100 people registered to attend, the Zoom sessions were never altogether attended by more than 200 people, a typical number for in-presence ECIRs. This means that online ECIRs generate a very high number of part-time attendees, instead of a much smaller number of (almost) full-time attendees.

We also think we learnt some lessons that might be useful for organizers who might be faced again with the challenge of going completely online. The most important such lesson, organization-wise, is that in-presence conferences and online conferences are very different in terms of the cash flow they generate. In the case of in-presence conferences, the bulk of expenses is often incurred just before or just after the event takes place, when bills (e.g., venue hire, catering, buses, printing services) can usually be paid; at that time, most of the money from registration fees is in already, most of the money from sponsors is too, and there is thus money in the cash for paying the bills. In the case of online conferences, instead, the most expensive service providers are the online platforms (in our case: Whova, Zoom, Gather Town), and these want to be paid in full at the time of booking the service, which is often way before registration fees have started to trickle in, and before sponsor money has also started to be paid. This generates a problem for those organizers who, due to administrative rules in place at their organizations, are unable to borrow money from their own grants for coping with these early expenses. The suggestion for prospective organizers of online events is not to overlook this problem.

3 Let's meet all at ECIR 2022!

ECIR 2022 will be in beautiful Stavanger, Norway, organized by Krisztian Balog and Kjetil Nørvåg, and will be the northernmost ECIR ever! We are all looking forward to meeting there and enjoying all the pleasures associated with being all together in a real place, with real people.

References

- Djoerd Hiemstra, Marie-Francine Moens, Raffaele Perego, and Fabrizio Sebastiani. Transitioning the information retrieval literature to a fully open access model. *SIGIR Forum*, 54(1), 2020.
- Djoerd Hiemstra, Marie-Francine Moens, Josiane Mothe, Raffaele Perego, Martin Potthast, and Fabrizio Sebastiani, editors. *Proceedings of the 43rd European Conference on Information Retrieval Research (ECIR 2021)*, 2021. Springer Nature, Cham, CH. Lecture Notes in Computer Science, Volumes 12656 and 12657.