

Report on the 2nd International Conference on Design of Experimental Search & Information REtrieval Systems (DESIRES 2021)

Omar Alonso
Instacart
oralonso@gmail.com

Stefano Marchesin
University of Padua
stefano.marchesin@unipd.it

Marc Najork
Google
marc@najork.org

Gianmaria Silvello
University of Padua
gianmaria.silvello@unipd.it

Abstract

This is a report on the second edition of the International Conference on Design of Experimental Search & Information REtrieval Systems (*DESIRES* 2021) held at the Department of Information Engineering of the University of Padua (Padua, Italy) from September 15 to September 18, 2021.

Date: 15–18 September, 2021.

Website: <http://desires.dei.unipd.it/>.

1 The DESIRES conference series manifesto

DESIRES is a systems-oriented conference, complementary in its mission to the mainstream Information Access and Retrieval conferences like SIGIR, ECIR, and other symposiums focusing on specific aspects of IR such as ICTIR or CHIIR, emphasizing the innovative technological elements of search and retrieval systems. This conference is inspired by the *Conference on Innovative Data Systems Research* (CIDR)¹ conference series of the database community.

DESIRES is thought as a *biennial* retreat-like event that favors discussion and brainstorming. Nevertheless, due to the COVID-19 pandemics, the 2020 edition of DESIRES, which had to take place in Venice (Italy) has been postponed to 2021 to favor an in-presence participation which is fundamental for DESIRES.

DESIRES gathers researchers and practitioners from both academia and industry to discuss the latest innovative and visionary ideas in the field. The goal of this conference is to provide the IR community with a venue for presenting innovative search systems architectures, as well as a

¹<http://cidrdb.org/>

publication opportunity. DESIRES does not compete with the mainstream conferences presenting rigorous treatises in established areas; instead its goal is to air radically new ideas.

DESIRES mainly encourages papers about innovative and risky information access and retrieval system ideas, systems-building experience and insight, resourceful experimental studies, provocative position statements, multidisciplinary takes on IR, and any new application domains. DESIRES especially welcomes contributions focusing on implementation details, successful or failed reproducibility attempts, technological breakthroughs, and new uses of old ideas.

Traditionally, program committees in our IR field reward scholarship on narrow ideas, operate by consensus, discard “loosey-goosey” papers including half-baked ideas and submissions that resemble war stories from the field. The major IR conferences usually reject such submissions because they are not scientific. However, these are often the very papers that offer long-term value to the field and should be widely disseminated. DESIRES values innovation, experience-based insight, and vision.

DESIRES looks for original work that could not be submitted in parallel to another venue. To encourage authors to submit only their best work, each person could be an author or co-author of just a single full paper. DESIRES also accepts prototype short papers and abstracts presenting ideas that are still in very early stages or challenge the current trends in IR. Any author of a full paper could additionally submit one abstract or one prototype. Abstracts and prototypes are inserted in the conference proceedings as “short papers”.

DESIRES favor open-science, thus its proceedings are gold open access and do not require the payment of any fee from the authors. The presented and employed datasets and software are required to be open and public available to favor reproducibility of results and re-use of the resources.

2 DESIRES 2021

DESIRES 2021 was organized, sponsored and partially sponsored by the Department of Information Engineering² of the University of Padua.³

DESIRES invited four kinds of contributions:

- **Papers (up to 6 pages + unlimited references).** Papers usually lack rigorous frameworks, simulations of performance, or prototype implementations but present a radical departure from conventional approaches that enables new applications.
- **Prototypes (up to 4 pages + unlimited references).** The prototype descriptions generally are a detailed report on successes and mistakes.
- **Abstracts (1 page + unlimited references).** Ideas that are too half-baked for a paper or demo proposal are good candidates for an abstract. Any author of a paper or prototype demo could additionally submit one abstract.
- **Open Problems in IR (up to 2 pages + unlimited references).** This is a track introduced in 2021. We look for abstract-like papers framing an important but unsolved problem in IR. “Open-Problems” papers were expected to have a single author.

²<https://www.dei.unipd.it/en/>

³<https://www.unipd.it/en/>

DESIRES received a total of 38 submissions in two main categories: 22 full papers and 16 short papers (prototypes and abstracts). From the first edition DESIRES increased the number of submissions and unique authors passing from 33 to 38 submissions, 26 to 28 accepted contributions, 49 to 58 unique authors and from 33 to 58 participants.

All full papers were reviewed by at least three members of an international Program Committee formed by experts from industry and academia. Of the full papers submitted to the conference, 18 were accepted for oral presentation. All the abstracts were reviewed by at least two reviewers, and ten were accepted for oral presentation. The accepted contributions represent state of the art in information retrieval, cover a diverse range of topics, propose new uses for IR techniques, querying, personal search, expert search, visual systems, IR research tools, teaching in IR, evaluation, NLP and collaborative search.

There were 58 unique authors with papers accepted at *DESIRES 2021* with the following geographical distribution (affiliation): 15 authors from USA (25.9%), 10 authors from Germany (17.2%), 6 authors from The Netherlands (10.31%), 5 authors from Austria, Canada and Italy (8.6% each), 2 authors from Belgium, Japan, Norway and Switzerland (3.4% each) and 1 author from Australia, Denmark, France, Portugal and UK (1.7%).

There were three keynotes:

- Brian Johnson, Head of Search Engineering at Apple (USA) on *Apple Media Products Search*.
- Hinrich Schütze, Professor at the University of Munich (Germany) on *Humans Learn From Task Descriptions and So Should Our Models*,
- Fabian M. Suchanek, Professor at Télécom Paris University (France) on *A hitchhiker's guide to Ontology*.

DESIRES 2021 would not have been possible without the financial support from the ExaMode H2020 EU project (Gold sponsor), Bloomberg and Google (Silver sponsors) and the Padua convention and visitor bureau (Bronze sponsor).

The proceedings of DESIRES 2021 are published open-access with CEUR-WS [Alonso et al., 2021a]. Detailed information about the overall organization of DESIRES, its scope and the advisory and program committees are available in the proceedings preface [Alonso et al., 2021b].

Thirty-two talks have been recorded and are publicly available online.⁴

DESIRES is part of the “Friends of SIGIR” program who financed **two best student paper awards** assigned to Sophia Althammer for her single author paper “RUDI: Real-Time Learning to Update Dense Retrieval Indices” [Althammer, 2021] and to Nadja Geisler for her single author paper “Quest: A framework for query-driven explainers on tabular data” [Geisler, 2021]. The **best paper award** was assigned to Krisztian Balog for “Conversational AI from an Information Retrieval Perspective: Remaining Challenges and a Case for User Simulation” [Balog, 2021] and to Roger B. Bradford for his work “Lessons Learned from 20 Years of Implementing LSI Applications” [Bradford, 2021].

The next edition, **DESIRES 2022**, will take place most likely in North America in August or September 2022 to go back to the planned scheduling (every even year). The submission deadline for full papers, abstracts and demos will be in late Spring 2022. Bookmark the DESIRES website⁵ and follow @DESIRES_IR on Twitter to keep posted.

⁴https://youtube.com/playlist?list=PL4yRbSc0b_NprvmT9KHyUwZarMRs20SJI

⁵<http://desires.dei.unipd.it/>

3 Keynotes

Apple Media Products Search. Brian Johnson presented the work by the Apple Media Products Search team which powers search for App Store, iTunes, Apple Music, Apple TV, Books, and Podcasts across multiple platforms. The talks covered three main topics: Query understanding, ranking beyond text, and evaluation. About “query understanding” Brian discussed why is necessary to have a deep understanding of developers, their inventory, users, and their intents and interests. Brian mentioned the robust and scalable query understanding systems they employ to detect the intent of a user’s query. About “ranking beyond text”, Brian described how the previously described rich understanding of their users, queries, and inventory are used to build next generation search ranking for Apple Media Products. Some of mentioned challenges included language processing, shifting inventory, and fraud prevention. Finally, about “evaluation”, Brian described the robust evaluation platform which helps Apple to understand how well they are doing by using widely-used IR metrics for precision and recall. Overall, they perform rigorous offline and online evaluations to predict how changes perform in the wild.

Humans Learn From Task Descriptions and So Should Our Models. Hinrich Schütze thoroughly describe his late-breaking results in the field of computational linguistics. He mentioned that in many types of human learning, task descriptions are a central ingredient and they are usually accompanied by a few examples. Nevertheless, there is very little human learning that is based on examples only. In contrast, the typical learning setup for NLP tasks lacks task descriptions and is supervised with 100s or 1000s of examples. This is even true for so-called few-shot learning, a term often applied to scenarios with tens of thousands of “shots”. Inspired by the GPT models, Hinrich presented Pattern-Exploiting Training (PET). PET reformulates task descriptions as close questions that can be effectively processed by pre-trained language models. In contrast to GPT, PET combines task descriptions with supervised learning. We show that PET learns well from as little as ten training examples and outperforms GPT-3 on GLUE even though it has 99.9% fewer parameters.

A Hitchhiker’s Guide to Ontology. Fabian M. Suchanek presented his vast body of work on knowledge base creation, access and management [Suchanek, 2021]. Knowledge bases are an important asset in many of today’s AI-based applications, including personal assistants and search engines. In his talk, Fabian gave an overview of the most recent work in this area. He first talked about one of his main projects, the YAGO knowledge base. Fabian presented different methods of information extraction, in particular also the extraction of beliefs and causal relationships. Afterwards, Fabian discussed the incompleteness of knowledge bases and introduced some late results from his group proposing several techniques to estimate how much data is missing in a knowledge base, as well as rule mining methods to derive that data. Moreover, he presented their work on efficient querying of knowledge bases. Finally, he described some applications of knowledge bases in the domain of computational creativity and the digital humanities for concluding with some remarks about the methods they developed for explainable AI.

4 Scientific program

4.1 Systems

Serverless BM25 Search and BERT Reranking [Anand et al., 2021]. In his presentation (online), Jimmy Lin focused on the retrieve and re-rank pipeline – a well-established architecture for search applications – that typically pairs first-stage retrieval using keyword search with reranking using a neural model. By deploying such an architecture in the cloud, developers have to put a considerable effort to resource provisioning and management. In this paper, the authors introduce a serverless prototype for the retrieve and re-rank pipeline by employing the Amazon Web Services (AWS). This prototype comprises BM25 for first-stage retrieval using Lucene followed by a re-ranking component employing the monoBERT model with Hugging Face Transformers. Since its first presentation, three years ago at DESIRES 2018, Jimmy further developed his ideas and proposals about this topic. In particular, the work is thought for a research-oriented community, where batches of queries are usually used to perform experiments – rather than answering real users needs.

GeeseDB: A Python Graph Engine for Exploration and Search [Kamphuis and de Vries, 2021]. Chris Kamphuis (in-presence) presented GeeseDB, a Python toolkit for IR research problems that leverages on graph data structures. The goal of GeeseDB is simplifying IR research by allowing researchers to easily formulate graph queries through a structured query language. GeeseDB is available as an easy-to-install Python package. With just few lines of code users can create a first stage retrieval ranking using BM25. Queries read and write Numpy arrays and Pandas dataframes, at zero or negligible data transformation cost (dependent on base datatype). Overall, the audience was pretty interested for GeeseDB; for instance some asked if tutorials were available to see the application running. A tutorial was available and presented off-line during the coffee break.

Contemporary Web-Based Interaction Logging Infrastructure: Discussing the Achievements and Remaining Developmental Challenges of LogUI [Maxwell and Hauff, 2021]. Studies involving user interfaces often require the capturing and recording of key user interactions between the user and the system under examination. Nevertheless, anecdotal evidence suggests that researchers often implement their own logging infrastructure, which can lead to numerous implementation mistakes (due to misunderstanding or ignoring differences between web browsers, for example). Efforts have been made to develop interaction logging solutions for experimentation and commercial use. However, many solutions either use obsolete technology, which are prohibitively expensive, complex to use, or have no source code available. To address these issues, David Maxwell and Claudia Hauff have developed LogUI, an easy-to-use yet powerful interaction logging framework that can capture virtually any user interaction within a web-based environment. David Maxwell (online) presented LogUi and he focused on the process of looking backwards to analyze and eventually fix all the different mistakes made in research. David talked about his PhD and all the mistakes students and early researchers do that he wanted to avoid/fix.

4.2 Search Domains and Applications

Towards Supporting Complex Retrieval Tasks Through Graph-Based Information Retrieval and Visual Analytics [Bobic et al., 2021]. The retrieval result analysis approaches of existing IR solutions tend to be either too simple, with too few features for exploring results, or very specific and restricted. Aleksandar Bobic (in-presence) presented an enhanced approach that attempts to address these issues and help the wider community to get more insights from their retrieved data. To this end, the authors proposed an enhanced graph-based retrieval prototype built on the “Collaboration Spotting” platform, combining IR and visual analytics to provide an advanced solution for data retrieval and exploration.

Netted?! How to Improve the Usefulness of Spider & Co. [Hättasch et al., 2021]. Natural language interfaces for databases (NLIDBs) are intuitive ways to access and explore structured data, which is one of the reasons making systems like “Spider” valuable, as they handle a series of approaches for NL-to-SQL-translation. However, the available solutions are still missing some key features. Benjamin Hättasch and Nadia Geisler (both in-presence) presented a prototypical implementation called UniverSQL that makes NL-to-SQL-translation approaches easier to use in information access systems. Finally, they discussed what can be done to improve future benchmarks and shared tasks for NLIDBs. The audience suggested to develop the modules such that they can be easily used by the community – for instance, one of the reasons behind the success of Machine Learning is that you can “pip” install most of the models you require to perform experiments. Thus, one suggestion was to focus on the easiness to deploy&re-use to facilitate usability and reproducibility.

Getting to Know You: Search Logs and Expert Grading to Define Children’s Search Roles in the Classroom [Landoni et al., 2021]. Monica Landoni (in-presence) presented the roles children play when using web search engines in the classroom context by revisiting a seminal work set in the home context. Monica described how they juxtaposed performance indicators inferred from a combination of search logs and expert grading of completed inquiry assignments to discern emerging search roles among children aged nine to eleven. The authors also explored differences when a traditional classroom is replaced by online instruction at home, due – for instance – to COVID-19. Finally, they discussed future research directions that are pivotal to advance research in IR to, and for, children. The audience was engaged and asked questions about how to operationalize the presented work in a TREC-like fashion. How to do ground truth creation? Traditional assessors do not look very suitable to grade relevance in this scenario. Monica said that school people would be the best assessors, users, etc. – and they actually have already them on board. They did a small study to understand the concept of relevance that children have, because it is one of the directions to pursue to properly evaluate these kind of works. Another question asked if they looked into the type of content that made children distracted or explorative. Monica answered that children are more distracted when multiple sources are competing in the same SERP – cartoons, videos – or they are even distracted by simply typing. Some children are more prone to distraction than others. overall, emotionally-type tasks were more engaging for children, who were way less distracted and way more involved.

4.3 Explorative search and querying

Exploring Dataset via Cell-Centric Indexing [Heflin et al., 2021]. Jeff Heflin (online) presented a novel approach to dataset search and exploration. Cell-centric indexing is a unique indexing strategy enabling a powerful, new interface. The strategy treats individual cells of a table as the indexed unit and, combining this with a number of structure-specific fields, it enables queries that cannot be answered by a traditional indexing approach.

What Makes a Query Semantically Hard? [Faggioli and Marchesin, 2021]. Guglielmo Faggioli (in presence) presented a work focusing on the semantic gap, a long-standing issue in IR. In particular, the authors focused on the differences between lexical and semantic models and worked on a methodology to understand if a query is better answered by a lexical or a semantic model. Finally, Guglielmo described a labeling strategy to classify queries into semantically hard or easy, and described a prototype classifier to discriminate between them.

Lessons Learned from 20 Years of Implementing LSI Applications [Bradford, 2021]. Roger Bradford (online) summarized the lessons learned, over a period of 20 years, from implementing information systems employing LSI. The data presented is drawn from 63 projects undertaken over the period 1999 through 2019. The presentation generated many industry-related questions. In particular, one question was about whether people are re-inventing the wheel with transformers-based models given all the work Roger did in improving LSI applications – which are, to some extent, kind of an improved version of LSI representations. Roger replied that it's not that way, and that the two things are not competing, but rather targeting different aspects. An interesting work and a compelling presentation that granted Roger the best paper award.

4.4 Conversational Search

Conversational AI from an Information Retrieval Perspective: Remaining Challenges and a Case for User Simulation [Balog, 2021]. Krisztian Balog (online) focused on conversational information access, a problem that is uniquely suited to be addressed by IR community. Despite the significant research activity in this area, progress is mostly limited to component-level improvements. A disconnection between current efforts and truly conversational information access systems remains. Apart from the inherently challenging nature of the problem, the lack of progress, can be attributed in large part to the shortage of appropriate evaluation methodology and resources. Krisztian highlighted the challenges that render both offline and online evaluation methodologies unsuitable for this problem, and discussed the use of user simulation as a viable solution. Krisztian delivered a compelling presentation that generated many questions and large interest in the crowd that granted him the best paper award.

Toward Conversational Query Reformulation [Kiesel et al., 2021]. In traditional web search interfaces, information seekers reformulate their queries by editing the terms in the search box to guide the retrieval process. Such a kind of editing is at odds with the natural language interaction paradigm in conversational interfaces, and it is impossible for purely voice-based interfaces. Conversational search studies reveal that participants describe their changes to a query;

however, the principles of such “editing conversations” have not been analyzed in depth, yet. Johannes Kiesel (in presence) presented a formalization of the conversational query reformulation problem. The authors cast reformulations as meta-queries that imply operations on the original query and categorize the operations following the standard CRUD terminology. Based on this formalization they crowdsourced a dataset with 2694 human reformulations across four search domains. There were many questions and we report two in particular. One question was if the types of presented reformulations are realistic; Johannes answered that all of the reformulations will happen, it’s just a matter of how often it will happen. It’s more a question of trust to the system. The second question was about the connection between keyword and conversational reformulations; Johannes answered that in a conversation, people do not think about the queries. In this way, people formulate different queries. But in the end, from the system point of view, the problem can be broken down to traditional query reformulation.

Towards System-Initiative Conversational Information Seeking [Wadhwa and Zamani, 2021]. Currently, most conversational information seeking (CIS) systems work in a passive manner – i.e., user-initiative engagement. Samin Wadhwa (online) discussed the importance of developing CIS systems capable of system-initiative interactions. Furthermore, the authors presented various aspects of such interactions in CIS systems and introduced a taxonomy of system-initiative interactions based on three orthogonal dimensions: initiation moment, initiation purpose, and initiation means.

“Do Users Need Human-like Conversational Agents?” – Exploring Conversational System Design Using Framework of Human Needs [Ghosh and Ghosh, 2021]. In this work, Ghosh and Ghosh (online presentation) ask themselves if we need a human-like conversational system? Before engaging the complex endeavor of implementing human-like characteristics, it should be debated if the pursuit of such a system is logical and ethical. The authors analyze some of the system-level characteristics and discuss their merits and potential harms. Some questions from the audience aimed at investigating the human-like aspects the authors described and proposed: “Do you think we need conversational human-like systems?” to which the authors answered that they just talked about pros and cons, but up to the user to decide. Moreover, somebody asked “how we are going to implement human-like conversational systems? ”, to which the authors relied that sentiment analysis has developed a lot, so we can inject that into IR conversational systems.

4.5 Search and user-generated content

Enhancing SNS Profile Writing with a Search-Based Assistant System [Nagase and Joho, 2021]. In Social Networking Services (SNS), user profiles, which often consist of an image, text, and other items, have an important role to connect with other users. However, in a preliminary study with 3,193 sample profiles on Twitter, Nagase (who presented the paper online) and Joho found that the average length of profile texts was 40 characters, where the maximum length is 160. This suggests that many SNS users are missing potential opportunities to expand their social network due to the short profile texts. Therefore, in this work, they proposed a search-based interactive system to support the writing of profile texts in SNS.

EXAM: How to Evaluate Retrieve-and-Generate Systems for Users Who Do Not (Yet) Know What They Want [Sander and Dietz, 2021]. In this work, Sander and Dietz propose EXAM, an evaluation paradigm that uses held-out exam questions and an automated question-answering system to evaluate how well generated responses can answer follow-up questions, without knowing the exam questions in advance. The audience was engaged by the presentation and there were many questions ranging from the robustness of EXAM with different Q&A systems to the potential biases due to the used language. Moreover, the authors mentioned that they might consider the employment of active learning to improve the system in the future. David Sander and Laura Dietz presented their work online and they were very active and present in the discussions contributing to make DESIRES a successful hybrid event.

TAR on Social Media: A Framework for Online Content Moderation [Yang et al., 2021]. Content moderation is one tool social networks use to fight problems such as harassment and disinformation. Manually screening all content is usually impractical given the scale of social media data, and the need for nuanced human interpretations makes fully automated approaches infeasible. Eugene Yang (online) presented content moderation from the perspective of technology-assisted review as a human-in-the-loop active learning approach developed for high recall retrieval problems in civil litigation and other fields.

4.6 Neural Models

A Bayesian Neural Model for Documents' Relevance Estimation [Purpura and Susto, 2021]. In this work, Alberto Purpura (in presence) proposed QLFusion, an approach based on Quantification Learning (QL) to improve rank fusion performance in Information Retrieval. The key aspect of the work is the use of a Quantification Learning model based on a Bayesian Neural Network to estimate the proportion of relevant documents in a ranked list.

Combining Lexical and Neural Retrieval with Longformer-based Summarization for Effective Case Law Retrieval [Askari and Verberne, 2021]. In this paper, Arian Askari (in-presence) and Suzan Verberne combine lexical and neural ranking models for case law retrieval. In case law retrieval, the query is a full case document, and the candidate documents are prior cases that are potentially relevant to the current case. The authors created summaries to be used as queries and then combine them with five different ranking models ranging from lexical to neural retrieval models. A thorough evaluation is presented and provided in the paper.

4.7 Short Papers

DESIRES short papers are presented in a gong-show fashion: 10 minutes for the presentation plus up to 5 minutes for questions. The presentation were divided in three slots, one per day. Short papers touched upon several hot and new research topics in IR at large comprising the creation of a test collection to evaluate public library document access [Canu, 2021], dense retrieval indexes and real-time learning [Althammer, 2021], the separation between logical and physical ranking models [Lin et al., 2021], exploration of test collection via structured queries [Hättasch, 2021], temporal relation extraction [Sousa, 2021], timeline as a ranking unit in news search [Jatowt, 2021],



Figure 1: Padua, 17 September 2021. *DESIRES 2021* in-presence participants. Picture taken in front of the historical Café Pedrocchi and courtesy of Ricardo Baeza-Yates.

tabular data [Geisler, 2021], users knowledge gain [Zein, 2021], search and smart cities [Ahlens, 2021] and neuroscience and IR [Moshfeghi, 2021].

5 Final remarks

DESIRES 2021 has been the first hybrid IR event in the post-pandemic era with a strong in-presence social program. The integration between in presence and online participants worked well and there were vibrant discussions. One drawback is that you cannot have a real interaction between the online and the offline during the informal and social events and this might affect the experience of the online participants.

Acknowledgments

We desire to thank all the people that encouraged us and helped us to make DESIRES happen. In particular, we would like to thank the top-tier program committee members for the timely and accurate reviews and the advisory board in the person of Maristella Agosti, Ricardo Baeza-Yates, Bruce Croft, Shane Culpepper, Susan Dumais, Norbert Fuhr, Donna Harman, Kalervo Järvelin, Jan Pedersen and Gerhard Weikum for all the work, support and suggestions they provided.

References

- Dirk Ahlers. Searching in the smart city? In [Alonso et al. \[2021a\]](#), pages 191–192. URL <http://ceur-ws.org/Vol-2950/paper-19.pdf>.
- Omar Alonso, Stefano Marchesin, Marc Najork, and Gianmaria Silvello, editors. *Proceedings of the Second International Conference on Design of Experimental Search & Information REtrieval Systems, Padova, Italy, September 15-18, 2021*, volume 2950 of *CEUR Workshop Proceedings*, 2021a. CEUR-WS.org. URL <http://ceur-ws.org/Vol-2950>.
- Omar Alonso, Marc Najork, and Gianmaria Silvello. Preface. In [Alonso et al. \[2021a\]](#), pages I–III. URL <http://ceur-ws.org/Vol-2950/preface-01.pdf>.
- Sophia Althammer. RUDI: real-time learning to update dense retrieval indices. In [Alonso et al. \[2021a\]](#), pages 173–175. URL <http://ceur-ws.org/Vol-2950/paper-20.pdf>.
- Mayank Anand, Jiarui Zhang, Shane Ding, Ji Xin, and Jimmy Lin. Serverless BM25 search and BERT reranking. In [Alonso et al. \[2021a\]](#), pages 3–9. URL <http://ceur-ws.org/Vol-2950/paper-01.pdf>.
- Arian Askari and Suzan Verberne. Combining lexical and neural retrieval with longformer-based summarization for effective case law retrieval. In [Alonso et al. \[2021a\]](#), pages 162–170. URL <http://ceur-ws.org/Vol-2950/paper-02.pdf>.
- Krisztian Balog. Conversational AI from an information retrieval perspective: Remaining challenges and a case for user simulation. In [Alonso et al. \[2021a\]](#), pages 80–90. URL <http://ceur-ws.org/Vol-2950/paper-03.pdf>.
- Aleksandar Bobic, Jean-Marie Le Goff, and Christian Gütl. Towards supporting complex retrieval tasks through graph-based information retrieval and visual analytics. In [Alonso et al. \[2021a\]](#), pages 30–37. URL <http://ceur-ws.org/Vol-2950/paper-04.pdf>.
- Roger B. Bradford. Lessons learned from 20 years of implementing LSI applications. In [Alonso et al. \[2021a\]](#), pages 70–79. URL <http://ceur-ws.org/Vol-2950/paper-05.pdf>.
- Serena Canu. It’s messy out there: Dbc’s journey towards its first test collection. In [Alonso et al. \[2021a\]](#), pages 171–172. URL <http://ceur-ws.org/Vol-2950/paper-21.pdf>.

-
- Guglielmo Faggioli and Stefano Marchesin. What makes a query semantically hard? In [Alonso et al. \[2021a\]](#), pages 61–69. URL <http://ceur-ws.org/Vol-2950/paper-06.pdf>.
- Nadja Geisler. Quest: A query-driven explanation framework for black-box classifiers on tabular data. In [Alonso et al. \[2021a\]](#), pages 187–188. URL <http://ceur-ws.org/Vol-2950/paper-24.pdf>.
- Souvick Ghosh and Satanu Ghosh. ”do users need human-like conversational agents?” - exploring conversational system design using framework of human needs. In [Alonso et al. \[2021a\]](#), pages 117–127. URL <http://ceur-ws.org/Vol-2950/paper-07.pdf>.
- Benjamin Hättasch. Wannadb: Ad-hoc structured exploration of text collections using queries. In [Alonso et al. \[2021a\]](#), pages 179–180. URL <http://ceur-ws.org/Vol-2950/paper-23.pdf>.
- Benjamin Hättasch, Nadja Geisler, and Carsten Binnig. Netted?! how to improve the usefulness of spider & co. In [Alonso et al. \[2021a\]](#), pages 38–43. URL <http://ceur-ws.org/Vol-2950/paper-08.pdf>.
- Jeff Hefflin, Brian D. Davison, and Haiyan Jia. Exploring datasets via cell-centric indexing. In [Alonso et al. \[2021a\]](#), pages 53–60. URL <http://ceur-ws.org/Vol-2950/paper-09.pdf>.
- Adam Jatowt. Timeline as information retrieval and ranking unit in news search. In [Alonso et al. \[2021a\]](#), pages 184–186. URL <http://ceur-ws.org/Vol-2950/paper-25.pdf>.
- Chris Kamphuis and Arjen P. de Vries. Geesedb: A python graph engine for exploration and search. In [Alonso et al. \[2021a\]](#), pages 10–18. URL <http://ceur-ws.org/Vol-2950/paper-11.pdf>.
- Johannes Kiesel, Xiaoni Cai, Roxanne El Baff, Benno Stein, and Matthias Hagen. Toward conversational query reformulation. In [Alonso et al. \[2021a\]](#), pages 91–101. URL <http://ceur-ws.org/Vol-2950/paper-12.pdf>.
- Monica Landoni, Theo Huibers, Mohammad Aliannejadi, Emiliana Murgia, and Maria Soledad Pera. Getting to know you: Search logs and expert grading to define children’s search roles in the classroom. In [Alonso et al. \[2021a\]](#), pages 44–52. URL <http://ceur-ws.org/Vol-2950/paper-10.pdf>.
- Jimmy Lin, Xueguang Ma, Joel Mackenzie, and Antonio Mallia. On the separation of logical and physical ranking models for text retrieval applications. In [Alonso et al. \[2021a\]](#), pages 176–178. URL <http://ceur-ws.org/Vol-2950/paper-26.pdf>.
- David Maxwell and Claudia Hauff. Developing contemporaryweb-based interaction logging infrastructure: The design and challenges of logui. In [Alonso et al. \[2021a\]](#), pages 19–29. URL <http://ceur-ws.org/Vol-2950/paper-13.pdf>.
- Yashar Moshfeghi. Neurasearch: Neuroscience and information retrieval. In [Alonso et al. \[2021a\]](#), pages 193–194. URL <http://ceur-ws.org/Vol-2950/paper-27.pdf>.

-
- Kousuke Nagase and Hideo Joho. Enhancing SNS profilewriting with a search-based assistant system. In [Alonso et al. \[2021a\]](#), pages 128–135. URL <http://ceur-ws.org/Vol-2950/paper-14.pdf>.
- Alberto Purpura and Gian Antonio Susto. A bayesian neural model for documents’ relevance estimation. In [Alonso et al. \[2021a\]](#), pages 156–161. URL <http://ceur-ws.org/Vol-2950/paper-15.pdf>.
- David P. Sander and Laura Dietz. EXAM: how to evaluate retrieve-and-generate systems for users who do not (yet) know what they want. In [Alonso et al. \[2021a\]](#), pages 136–146. URL <http://ceur-ws.org/Vol-2950/paper-16.pdf>.
- Hugo O. Sousa. Temporal relation extraction: The event ordering task. In [Alonso et al. \[2021a\]](#), pages 181–183. URL <http://ceur-ws.org/Vol-2950/paper-28.pdf>.
- Fabian M. Suchanek. A hitchhiker’s guide to ontology. In [Alonso et al. \[2021a\]](#), pages 1–2. URL <http://ceur-ws.org/Vol-2950/keynote-02.pdf>.
- Somin Wadhwa and Hamed Zamani. Towards system-initiative conversational information seeking. In [Alonso et al. \[2021a\]](#), pages 102–116. URL <http://ceur-ws.org/Vol-2950/paper-17.pdf>.
- Eugene Yang, David D. Lewis, and Ophir Frieder. TAR on social media: A framework for online content moderation. In [Alonso et al. \[2021a\]](#), pages 147–155. URL <http://ceur-ws.org/Vol-2950/paper-18.pdf>.
- Dima El Zein. User knowledge and search goals in information retrieval: A benchmark and study on the evolution of users’ knowledge gain. In [Alonso et al. \[2021a\]](#), pages 189–190. URL <http://ceur-ws.org/Vol-2950/paper-22.pdf>.