

Report on the FIRE 2020 Evaluation Initiative

Parth Mehta
Parmonic
parth.mehta126@gmail.com

Thomas Mandl
University of Hildesheim
mandl@uni-hildesheim.de

Prasenjit Majumder
DA-IICT, Gandhinagar
prasenjit.majumder@gmail.com

Surupendu Gangopadhyay
DA-IICT, Gandhinagar
surupendu.g@gmail.com

Abstract

This report gives an overview on the Forum for Information Retrieval Evaluation (FIRE) initiative for South-Asian languages¹. The FIRE conference was conducted online in December 2020. The event combined a conference including keynotes, peer reviewed paper session with an Evaluation Forum. This report will present an overview of the conference and provide insights into the evaluation tracks. Current domains include legal information access, mixed script information retrieval, semantic analysis and social media posts classification. The tasks are discussed and connections to other evaluation initiatives are shown.

1 Introduction

FIRE is a South-Asian evaluation campaign focusing on multilingual information access. Over the last twelve editions, diverse NLP applications and evaluation tasks have been offered for more than 10 Indian languages, besides German, Russian, Arabic, Persian and English. Over the years, many resources have been created for a number of languages, several of which are low resource languages. A substantial percentage of the organizers and participants continue to come from outside the Indian subcontinent.

The 12th annual meeting of the FIRE evaluation initiative took place in December 2020. It was organized by IDRBT Hyderabad and was conducted fully online. FIRE 2020 continued the series of successful evaluation exercises and conferences and was characterized by a large growth including 10 evaluation tracks and the largest number of participants so far. It was the first time in the history of FIRE that over a hundred working notes were submitted across the evaluation tracks. As a comparison, there were 112 working notes this year compared to 65 in FIRE 2019, marking an increase of over 50% from last year. The participants were from academia as well as industry spanning over a dozen countries and six continents, making FIRE 2020 a truly international event. Apart from the keynotes, a peer reviewed conference and evaluation tracks, the program also included a session on Women in Information Retrieval (WIRe).

¹<http://fire.irsir.res.in/fire/2020/home>

2 The FIRE Conference

2.1 Keynotes

Since its inception in 2008, keynotes by leading researchers across the world have been an important aspect of FIRE. Continuing that tradition FIRE 2020 included four scholarly talks.

- **Ellen Voorhees** (NIST, USA) talked on “TREC-COVID: Building a Pandemic IR Test Collection” explaining the Text REtrieval Conference (TREC)² task [Voorhees et al., 2021] designed to build a test collection that captures the information needs of biomedical researchers using the scientific literature during a pandemic. The talk touched upon the challenges faced in a fast-evolving and fluid information needs during the pandemic, the test collection created as a result of this exercise and several other aspects of the task.
- **Paul Clough** (University of Sheffield and Peak Indicators, UK) gave a talk titled “Algorithmic Bias, Transparency and Fairness: What is it, Why Does it Matter and What’s Being Done About it?” He covered an important topic which has been debated extensively in recent times. The talk gave insights into how and why biases enter search algorithms especially in the context of search engines. Furthermore, the talk highlighted the extent to which such bias is present and what can be done to prevent it.
- **Adam Wyner** (Swansea University) spoke about “Developing Resources for Legal Information Retrieval” and gave an overview of the current landscape of legal text mining. The challenges to wider adoption of text mining applications in legal domain, especially at the level of policy making, and what it signifies for researchers as well as companies working in this domain. He also covered in detail recent work on rhetorical role labelling [Bhattacharya et al., 2019b], which led to one of the subtasks in the AILA evaluation track [Bhattacharya et al., 2020].
- **Ajit Balakrishnan** (Rediff.com, India) addressed “Challenges in Machine learning in the Indian Legal Domain”. Compared to Adam’s talk, which outlined issues related to legal tech at a broader level, Ajit’s talk focused more on the scenario in India. He gave special attention to the problems of intricacies in the legal system and other societal aspects such as affordability of legal services and the role of technology for solving these problems.

2.2 Peer Reviewed Session

The peer reviewed session at FIRE 2020 saw 14 submissions in all from both academia and industry, including papers from outside India. Based on the reviews 2 full papers and 4 short papers were accepted for oral presentations. In addition, a short overview from each evaluation track was invited. The conference proceedings are available in the ACM digital library [Majumder et al., 2020].

²<https://trec.nist.gov>

2.3 Women in IR (WIRe)

A session on Women in IR was organized for the first time. The session was well attended with over 15 participants. WIRe consisted of three parts including a keynote by *Abha Bang Soni*, a Consultant Psychiatrist, who talked about the challenges faced by female researchers during the time of pandemic and possible solutions for them. The other two parts included a talk on *Different scholarships opportunities for Women scientists*, and finally six short presentations by participating researchers allowing them to showcase their works.

3 The FIRE Evaluation Tracks

The evaluation tracks have been selected after a review process. Ten tracks were chosen and offered in 2020. This year the focus of tracks was on legal information access, mixed script information retrieval, semantic analysis and classification of social media posts, event identification and Authorship identification for Source code to name a few. For all tracks, the overview and participant papers are available in online working notes [Mehta et al., 2021]. Below, we present a short overview of the tracks. We also indicate how their tasks are related to other initiatives.

3.1 Hate Speech and Offensive Content Identification

Hate Speech and Offensive Content Identification (HASOC) in Indo-European Languages and was offered for the second round. HASOC addresses the issue of hateful content in social media which is a serious problem for rational discourse and deliberation in open societies. Offensive and problematic content including insulting, hurtful, derogatory or obscene user contributions are pervasive in social media. Societies need to develop adequate response mechanisms in order to find a balance between freedom of expression on one hand and the ability to life without oppressive remarks on the other hand. A requirement for any response is robust technology for identifying problematic content automatically. HASOC provides a forum for developing and testing text classification systems for various languages. Many initiatives have been developing resources for similar tasks [Poletto et al., 2020]. HASOC is partially building upon a task organized at SemEval and its predecessors [Zampieri et al., 2020]. They initiate with a binary classification between problematic and regular content. Based on that, further detailed tasks are derived.

Despite the fact that Hindi is spoken by several hundreds of millions of speakers, it is still a low resource language. That is also true for the identification of offensive language. In the first edition (HASOC 2019), resources for German, English and Hindi have been assembled [Modha et al., 2019]. With a slightly different collection policy but the same annotation guidelines, the second edition (HASOC 2020) offered the same three languages [Mandl et al., 2020]. The core task is a binary classification. As a secondary task, the type of the offensive language needs to be classified which results in a lower performance by the systems. In 2020, HASOC was extended to include Dravidian languages using Youtube comments [Chakravarthi et al., 2020a]. Another challenge has been offered for Dravidian languages at the Workshop on Speech and Language Technologies for Dravidian Languages within the EACL conference [Chakravarthi et al., 2021].

HASOC is growing for the 2021 edition of FIRE and is offering four subtracks. The two subtracks on *Hate Speech in English and Hindi* as well as *Dravidian Languages* will continue.

In addition, two new subtracks will be offered: *Arabic Misogyny Identification* and *Threatening and Abusive Language in Urdu*. HASOC 2021 offers various opportunities to test systems across languages and across genres of texts.

3.2 Artificial Intelligence for Legal Assistance

Artificial Intelligence for Legal Assistance (AILA) is a series of shared tasks aiming to solve some challenging problems in legal domain using broad range of IR and NLP techniques. Compared to the Western countries where legal document processing has been an area of research for a long time, the legal IR/NLP community is very nascent in Indian subcontinent. While there were a few attempts, including some at previous editions of FIRE, they failed to receive sufficient attention. The AILA series is a fresh attempt at developing legal text processing technology which is specific for India. While AILA has focused so far exclusively on English documents, the novelty comes from the fact that most publicly available records are unstructured, unlike several other countries. This makes the role of AILA more important, as it has generated benchmark corpora consisting of judgements delivered by the Supreme court of India.

Problems which broadly fall under legal text retrieval and processing have been offered at several venues in the past. A broad range of topics have been explored ranging from e-discovery, that was focus of TREC for several years [Cormack et al., 2010], to patent mining [Piroi et al., 2013] at the Conference and Labs of the Evaluation Forum (CLEF)³. Competition on Legal Information Extraction/Entailment (COLIEE)⁴ is another venue that has been offering several shared tasks including legal text entailment identification and legal question answering among others [Rabelo et al., 2019]. One of the tasks offered at recent editions of COLIEE, precedent and statute retrieval, is closely related to one of the AILA subtasks. The main focus of COLIEE is on Japanese and English documents.

AILA 2020 offered two subtracks [Bhattacharya et al., 2020]. One of the problems offered this year was retrieving precedents and statutes related to a given legal scenario. This was in fact a continuation from previous edition of AILA [Bhattacharya et al., 2019a]. The other task was introduced this year and focused on providing structure to court judgements by identifying rhetorical categories within the judgement texts. AILA will continue in 2021 with the rhetorical role labelling task. It will also offer a new subtask on legal document summarization.

3.3 Sentiment Analysis of Dravidian Languages in Code-Mixed Text

Dravidian languages are low-resource languages and require more research and development in order to generate better systems [Chakravarthi et al., 2020c]. Dravidian languages are spoken in the South of India. Examples are Tamil, Malayalam, and Kannada. They do not belong to the Indo-European language family and pose several specific challenges. Also Code-Mixed scripts are a specific challenge of social media in India. Sometimes users do not change the keyboard setting while typing on a smart phone and mix Western (Latin) and local (e.g. Devangari) scripts.

Sentiment analysis on social media have been a productive research area. Sentiment analysis technology is useful for companies to monitor opinions on products and for political stakeholder

³<https://clef2020.clef-initiative.eu/>

⁴<https://sites.ualberta.ca/~rabelo/COLIEE2020/>

to analyze attitudes toward social issues. This track encourages the analysis of code-mixed text in Dravidian languages (Tamil-English and Malayalam-English). It is a message level polarity classification task. The classes are positive, mixed, negative, neutral and not in the language [Chakravarthi et al., 2020b].

3.4 Fake News Detection in the Urdu Language

Fake News Detection in the Urdu Language (UrduFake) task focused on identifying fake news articles. Misinformation is another serious problem of online media. Many forms of misinformation have been described [Froehlich, 2017] and some have been picked up by shared tasks. During the COVID crisis, the problem has called for special attention [Shahi et al., 2021]. So far, very heterogeneous data sources have been used for research collections for identifying misinformation. Several have focused on social media. Within SemEval-2019, the task RumourEval for English used Twitter and Reddit data [Gorrell et al., 2019]. Within TREC, a domain specific track has been initiated which focuses on health information. The TREC Health Misinformation Track uses a English Web collection and is developing COVID related topics⁵.

Within CLEF, the CheckThat! lab has offered tasks on misinformation [Nakov et al., 2021]. Task 3 of CheckThat! is focused on news articles which disseminate wrong information, which has been spotted by fact checking agencies. The task Fake News Detection in the Urdu Language at FIRE offers a collection for a low resource language, which is spoken by more than 100 million people. In addition, it introduces methodological novelty within the collection creation. Professional journalists are asked to create the fake news articles. This ensures that they cannot be easily found by Web search. The articles are organized in balanced topic categories. The main task is a classification task into the classes fake and not fake.

3.5 Authorship Identification of SOURCE CODE

Authorship Identification of SOURCE CODE (AI-SoCo) task was organized in cooperation with PAN⁶ [Bevendorff et al., 2021] and aimed at uncovering the author who wrote a particular piece of code. AI-SoCo is a part of the SoCo series with several other related tasks offered in previous editions of FIRE like detection of Source Code Reuse in single language [Flores et al., 2014] and cross-language [Flores et al., 2015] setups as well as Personality Recognitions using Source Codes [Pardo et al., 2016]. The task uses a dataset of 100,000 publicly available source codes from 100 different authors [Fadel et al., 2020].

3.6 Retrieval from Conversational Dialogues

Retrieval from Conversational Dialogues (RCD) task address the problem of contextualization of concepts in a conversation which one of the participants might not be well versed in [Ganguly et al., 2020]. There are two subtasks, one identifies the concepts which might need additional context and the second focuses on retrieving documents that can provide this additional context. Thus, this task contributes to the growing research area of chatbots. A task was also organized

⁵<https://trec-health-misinfo.github.io/>

⁶<https://pan.webis.de/>

TREC. Another related task on analyzing the quality of chatbot communication is offered at NII Testbeds and Community for Information Access Research (NTCIR)⁷ [Zeng et al., 2020].

3.7 Cause-Effect Relation Extraction from Text

Cause-Effect Relation Extraction from Text (CEREX) task proposed two sub-tasks related to causal relation extraction [Sinha et al., 2020]. First task focuses on automatically identifying whether a given sentence has causal information. The second task further identifies the cause, effect and causal connectives within a sentence containing causal information.

3.8 Causality-driven Ad hoc Information Retrieval

Causality-driven Ad hoc Information Retrieval (CAIR) focused on building causal search systems capable of identifying information on the likely causes leading to a query event [Datta et al., 2020]. Specifically, the queries for this task were events that were likely to be caused by other past events. The participants were expected to identify documents containing information about these past events that caused the query event.

3.9 Anaphora Resolution from Social Media Text in Indian Languages

The task offered Anaphora Resolution for three Indian languages Hindi, Tamil and Malayalam [Devi, 2020]. Specifically the task was targeted towards anaphora resolution in social media posts which pose a unique set of challenges compared to regular text.

3.10 Event Detection from News in Indian Languages

Event Detection from News in Indian Languages (EDNIL) focused on developing algorithms for identifying events in Indian language news articles [Dave et al., 2020]. The track was offered as two sub-tasks and encompassed four Indian Languages Hindi, Bengali, Marathi, Tamil as well as English. While the first task was identifying event, the second task constituted of identifying finer details related to the event like type, time and place among other.

4 Outlook to FIRE 2021

The 13th edition of the Forum for Information Retrieval Evaluation (FIRE 2021) is scheduled for December 2021 and will be organized fully online. It will include the peer reviewed track and evaluation tracks. The following six tracks will be offered⁸:

- Hate Speech and offensive content detection (HASOC) *continued*
- Artificial Intelligence for Legal Assistance (AILA) *continued*

⁷<http://research.nii.ac.jp/ntcir>

⁸See also <http://fire.irs.res.in/fire/2021/home>

-
- Sentiment Analysis of Dravidian Languages in Code-Mixed Text; *continued*
 - Fake News Detection in the Urdu Language (UrduFake); *continued*
 - Causality-driven Ad hoc Information Retrieval (CAIR); *continued*
 - Information Retrieval from Microblogs during Disasters (IRMiDis); *new*

The IRMiDis track has been organized at FIRE previously and is being reintroduced, with a new subtask, after a gap of two years. IRMiDis supports research on crisis information [Basu et al., 2020]. It originated from a collection of tweets related to the earthquake in Nepal in 2015. The information needs and information dissemination change dramatically during crisis situations. Microblog content has also been often used for disaster response. The new edition of IRMiDis is extending the focus on the quality and the reliability of the tweets. For that purpose, a collection of news articles about the earthquake has been added. The participating systems need to find fact-checkable tweets and in a second task, they need to find supporting news articles and sentences within them. IRMiDis also introduces a new subtask on identifying the public sentiment toward vaccination during the current pandemic. It aims to identify pro and anti vaccination sentiments from social media posts.

Participation at the tracks and at the conference is welcome. Proposals for hosting future editions of FIRE (limited to locations in India), or offering evaluation tracks and tutorials can be sent to fire@isical.ac.in.

Acknowledgements

FIRE has received continued support from ACM-SIGIR and we are grateful for that. A special thanks to the organizing team and the program committee who worked hard to ensure a successful FIRE 2020 even during these challenging times. We look forward to your continued association to future editions of FIRE.

References

- Moumita Basu, Kripabandhu Ghosh, and Saptarshi Ghosh. Information retrieval from Microblogs during disasters: In the light of IRMiDis task. *SN Computer Science*, 1(1):61, 2020. URL <https://doi.org/10.1007/s42979-020-0065-1>.
- Janek Bevendorff, Berta Chulvi, Gretel Liz De la Peña Sarracén, Mike Kestemont, Enrique Manjavacas, Iliia Markov, Maximilian Mayerl, Martin Potthast, Francisco Manuel Rangel Pardo, Paolo Rosso, Efstathios Stamatatos, Benno Stein, Matti Wiegmann, Magdalena Wolska, and Eva Zangerle. Overview of PAN 2021: Authorship verification, profiling hate speech spreaders on twitter, and style change detection - extended abstract. In *43rd European Conference on IR Research, ECIR, Virtual Event, March 28 - April 1*, volume 12657 of *LNCS*, pages 567–573. Springer, 2021. URL https://doi.org/10.1007/978-3-030-72240-1_66.

-
- Paheli Bhattacharya, Kripabandhu Ghosh, Saptarshi Ghosh, Arindam Pal, Parth Mehta, Arnab Bhattacharya, and Prasenjit Majumder. Overview of the FIRE 2019 AILA track: Artificial intelligence for legal assistance. In *Working Notes of FIRE - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15*, volume 2517, pages 1–12. CEUR-WS.org, 2019a. URL <http://ceur-ws.org/Vol-2517/T1-1.pdf>.
- Paheli Bhattacharya, Shounak Paul, Kripabandhu Ghosh, Saptarshi Ghosh, and Adam Wyner. Identification of rhetorical roles of sentences in Indian legal judgments. In *Proc. International Conference on Legal Knowledge and Information Systems (JURIX)*, 2019b.
- Paheli Bhattacharya, Parth Mehta, Kripabandhu Ghosh, Saptarshi Ghosh, Arindam Pal, Arnab Bhattacharya, and Prasenjit Majumder. Overview of the FIRE 2020 AILA track: Artificial intelligence for legal assistance. In *Working Notes of FIRE - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20*, volume 2826, pages 1–11. CEUR-WS.org, 2020. URL <http://ceur-ws.org/Vol-2826/T1-1.pdf>.
- Bharathi Raja Chakravarthi, Anand Kumar M, John P. McCrae, B. Premjith, K. P. Soman, and Thomas Mandl. Overview of the track on HASOC-offensive language identification-DravidianCodeMix. In *Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20, 2020*, volume 2826, pages 112–120. CEUR-WS.org, 2020a. URL <http://ceur-ws.org/Vol-2826/T2-2.pdf>.
- Bharathi Raja Chakravarthi, Ruba Priyadharshini, Vigneshwaran Muralidaran, Shardul Suryawanshi, Navya Jose, Elizabeth Sherly, and John P. McCrae. Overview of the track on sentiment analysis for Dravidian languages in Code-Mixed text. In *Forum for Information Retrieval Evaluation, FIRE 2020*, page 21–24, New York, NY, USA, 2020b. ACM. ISBN 9781450389785. URL <https://doi.org/10.1145/3441501.3441515>.
- Bharathi Raja Chakravarthi, Ruba Priyadharshini, Vigneshwaran Muralidaran, Shardul Suryawanshi, Navya Jose, Elizabeth Sherly, and John P. McCrae. Overview of the track on sentiment analysis for Dravidian languages in Code-Mixed text. In *Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20*, volume 2826, pages 480–489. CEUR-WS.org, 2020c. URL <http://ceur-ws.org/Vol-2826/T4-1.pdf>.
- Bharathi Raja Chakravarthi, Ruba Priyadharshini, Navya Jose, Anand Kumar M, Thomas Mandl, Prasanna Kumar Kumaresan, Rahul Ponnusamy, Hariharan R L, John P. McCrae, and Elizabeth Sherly. Findings of the shared task on offensive language identification in Tamil, Malayalam, and Kannada. In *Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages*, pages 133–145, Kyiv, April 2021. ACL. URL <https://www.aclweb.org/anthology/2021.dravidianlangtech-1.17>.
- Gordon V. Cormack, Maura R. Grossman, Bruce Hedin, and Douglas W. Oard. Overview of the TREC 2010 legal track. In *Proceedings of The Nineteenth Text REtrieval Conference, TREC 2010, Gaithersburg, Maryland, USA, November 16-19*, volume 500-294 of *NIST Special Publication*. National Institute of Standards and Technology (NIST), 2010. URL <https://trec.nist.gov/pubs/trec19/papers/LEGAL10.OVERVIEW.pdf>.

-
- Suchana Datta, Debasis Ganguly, Dwaipayan Roy, Derek Greene, Charles Jochim, and Francesca Bonin. Overview of the causality-driven adhoc information retrieval (CAIR) task at FIRE-2020. In *Forum for Information Retrieval Evaluation*, pages 14–17, 2020. URL <https://doi.org/10.1145/3441501.3441513>.
- Bhargav Dave, Surupendu Gangopadhyay, Prasenjit Majumder, Pushpak Bhattacharya, Sudeshna Sarkar, and Sobha Lalitha Devi. Overview of the FIRE 2020 EDNIL track: Event detection from news in indian languages. In *Working Notes of FIRE - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20*, volume 2826, pages 806–816. CEUR-WS.org, 2020. URL <http://ceur-ws.org/Vol-2826/T10-1.pdf>.
- Sobha Lalitha Devi. Anaphora resolution from social media text in indian languages (SocAnaRes-IL)- overview. In *FIRE 2020: Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20, 2020*, pages 9–13. ACM, 2020. URL <https://doi.org/10.1145/3441501.3441512>.
- Ali Fadel, Husam Musleh, Ibraheem Tuffaha, Mahmoud Al-Ayyoub, Yaser Jararweh, Elhadj Benkhelifa, and Paolo Rosso. Overview of the PAN@FIRE 2020 task on the authorship identification of SOURCE CODE. In *Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, Dec. 16-20*, volume 2826, pages 649–676. CEUR-WS.org, 2020. URL <http://ceur-ws.org/Vol-2826/T5-1.pdf>.
- Enrique Flores, Paolo Rosso, Lidia Moreno, and Esaú Villatoro-Tello. On the detection of source code re-use. In *Proceedings of the Forum for Information Retrieval Evaluation, FIRE 2014, Bangalore, India, December 5-7, 2014*, pages 21–30. ACM, 2014. URL <https://doi.org/10.1145/2824864.2824878>.
- Enrique Flores, Paolo Rosso, Esaú Villatoro-Tello, Lidia Moreno, Rosa Alcover, and Vicente Chirivella. PAN@FIRE: Overview of CL-SOCO track on the detection of cross-language source code re-use. In *Post Proceedings of the Workshops at the 7th Forum for Information Retrieval Evaluation, Gandhinagar, India, December 4-6*, volume 1587, pages 1–5. CEUR-WS.org, 2015. URL <http://ceur-ws.org/Vol-1587/T1-1.pdf>.
- Thomas J Froehlich. A not-so-brief account of current information ethics: The ethics of ignorance, missing information, misinformation, disinformation and other forms of deception or incompetence. *BiD*, (39), 2017.
- Debasis Ganguly, Dipasree Pal, Manisha Verma, and Procheta Sen. Overview of RCD-2020, the FIRE-2020 track on retrieval from conversational dialogues. In *Forum for Information Retrieval Evaluation*, pages 33–36, 2020. URL <https://doi.org/10.1145/3441501.3441518>.
- Genevieve Gorrell, Elena Kochkina, Maria Liakata, Ahmet Aker, Arkaitz Zubiaga, Kalina Bontcheva, and Leon Derczynski. SemEval-2019 task 7: RumourEval, determining rumour veracity and support for rumours. In *Proceedings of the 13th International Workshop on Semantic Evaluation*, pages 845–854, Minneapolis, Minnesota, USA, June 2019. ACL. URL <https://www.aclweb.org/anthology/S19-2147>.

Prasenjit Majumder, Mandar Mitra, Surupendu Gangopadhyay, and Parth Mehta, editors. *FIRE 2020: Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20, 2020*, 2020. ACM. ISBN 978-1-4503-8978-5. URL <https://doi.org/10.1145/3441501>.

Thomas Mandl, Sandip Modha, Gautam Kishore Shahi, Amit Kumar Jaiswal, Durgesh Nandini, Daksh Patel, Prasenjit Majumder, and Johannes Schäfer. Overview of the HASOC track at FIRE 2020: Hate speech and offensive content identification in Indo-European Languages. In *Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20*, volume 2826, pages 87–111. CEUR-WS.org, 2020. URL <http://ceur-ws.org/Vol-2826/T2-1.pdf>.

Parth Mehta, Thomas Mandl, Prasenjit Majumder, and Mandar Mitra, editors. *Working Notes of FIRE 2020 - Forum for Information Retrieval Evaluation, Hyderabad, India, December 16-20, 2020*, volume 2826 of *CEUR Workshop Proceedings*, 2021. CEUR-WS.org. URL <http://ceur-ws.org/Vol-2826>.

Sandip Modha, Thomas Mandl, Prasenjit Majumder, and Daksh Patel. Overview of the HASOC track at FIRE 2019: Hate speech and offensive content identification in Indo-European Languages. In *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15*, volume 2517, pages 167–190. CEUR-WS.org, 2019. URL <http://ceur-ws.org/Vol-2517/T3-1.pdf>.

Preslav Nakov, Giovanni Da San Martino, Tamer Elsayed, Alberto Barrón-Cedeño, Rubén Míguez, Shaden Shaar, Firoj Alam, Fatima Haouari, Maram Hasanain, Nikolay Babulkov, Alex Nikolov, Gautam Kishore Shahi, Julia Maria Struß, and Thomas Mandl. The CLEF-2021 CheckThat! lab on detecting check-worthy claims, previously fact-checked claims, and fake news. In *43rd European Conference on IR Research, ECIR Virtual Event, March 28 - April 1*, pages 639–649. Springer, 2021. URL https://doi.org/10.1007/978-3-030-72240-1_75.

Francisco M. Rangel Pardo, Fabio A. González, Felipe Restrepo-Calle, Manuel Montes-y-Gómez, and Paolo Rosso. PAN@FIRE: Overview of the PR-SOCO track on personality recognition in SOURCE CODE. In *Text Processing - FIRE 2016 International Workshop, Kolkata, India, December 7-10, 2016*, volume 10478 of *LNCS*, pages 1–19. Springer, 2016. URL https://doi.org/10.1007/978-3-319-73606-8_1.

Florina Piroi, Mihai Lupu, and Allan Hanbury. Overview of CLEF-IP 2013 lab - information retrieval in the patent domain. In *Information Access Evaluation. Multilinguality, Multimodality, and Visualization - 4th International Conference of the CLEF Initiative, CLEF, Valencia, Spain, September 23-26. Proceedings*, volume 8138 of *LNCS*, pages 232–249. Springer, 2013. URL https://doi.org/10.1007/978-3-642-40802-1_25.

Fabio Poletto, Valerio Basile, Manuela Sanguinetti, Cristina Bosco, and Viviana Patti. Resources and benchmark corpora for hate speech detection: a systematic review. *Language Resources and Evaluation*, pages 1–47, 2020. URL <https://doi.org/10.1007/s10579-020-09502-8>.

Juliano Rabelo, Mi-Young Kim, Randy Goebel, Masaharu Yoshioka, Yoshinobu Kano, and Ken Satoh. A summary of the COLIEE 2019 competition. In *New Frontiers in Artificial Intelligence -*

JSAI-isAI International Workshops, JURISIN, AI-Biz, LENLS, Kansei-AI, Yokohama, Japan, November 10-12, volume 12331 of *LNCS*, pages 34–49. Springer, 2019. URL https://doi.org/10.1007/978-3-030-58790-1_3.

Gautam Kishore Shahi, Anne Dirkson, and Tim A. Majchrzak. An exploratory study of COVID-19 misinformation on twitter. *Online Soc. Networks Media*, 22:100104, 2021. URL <https://doi.org/10.1016/j.osnem.2020.100104>.

Manjira Sinha, Tirthankar Dasgupta, and Lipika Dey. CEREX@ FIRE-2020: Overview of the shared task on cause-effect relation extraction. In *Forum for Information Retrieval Evaluation*, pages 18–20, 2020. URL <https://doi.org/10.1145/3441501.3441514>.

Ellen Voorhees, Tasmee Alam, Steven Bedrick, Dina Demner-Fushman, William R Hersh, Kyle Lo, Kirk Roberts, Ian Soboroff, and Lucy Lu Wang. TREC-COVID: constructing a pandemic information retrieval test collection. *ACM SIGIR Forum*, 54(1):1–12, 2021. URL <https://doi.org/10.1145/3451964.3451965>.

Marcos Zampieri, Preslav Nakov, Sara Rosenthal, Pepa Atanasova, Georgi Karadzhov, Hamdy Mubarak, Leon Derczynski, Zeses Pitenis, and Çagri Çöltekin. Semeval-2020 task 12: Multilingual offensive language identification in social media (OffensEval 2020). *CoRR*, abs/2006.07235, 2020. URL <https://arxiv.org/abs/2006.07235>.

Zhaohao Zeng, Sosuke Kato, Tetsuya Sakai, and Inho Kang. Overview of the NTCIR-15 dialogue evaluation (DialEval-1) task. In *Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies*, 2020. URL <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings15/pdf/ntcir/01-NTCIR15-0V-DIALEVAL-ZengZ.pdf>.