

6th Workshop on Indian Language Data: Resources and Evaluation (WILDRE)

Date: Monday, 20th June 2022

Venue: Palais du Pharo, Marseille, France (Organized under [LREC 2022 \(20-25 June 2022\)](#))

Website: <http://sanskrit.jnu.ac.in/conf/wildre6>

WILDRE-6, the 6th Workshop on Indian Language Data: Resources and Evaluation is proposed to be organized in Marseille (France) on 20th June 2022 under the LREC platform. India has a huge linguistic diversity and has seen concerted efforts from the Indian government and industry for developing language resources. European Language Resource Association (ELRA) and its associate organizations have been very active and successful in addressing the challenges and opportunities related to language resource creation and evaluation. It is therefore a big opportunity for resource creators of Indian languages to showcase their work on this platform and also to interact and learn from those involved in similar initiatives all over the world. The broader objectives of the WILDRE will be

- To map the status of Indian Language Resources
- To investigate challenges related to creating and sharing various levels of language resources
- To promote a dialogue between language resource developers and users
- To provide opportunity for researchers from India to collaborate with researchers from other parts of the world

Dates for Short/Long papers and Posters and Demos

March 31, 2022: Paper submissions due

May 03, 2022: Paper notification acceptance

May 23, 2022: Camera-ready papers due

June 20, 2022: Workshop

SUBMISSIONS

Papers must describe original, completed/ in progress and unpublished work. Each submission will be reviewed by three program committee members.

Accepted papers will be given up to 10 pages (for full papers) 5 pages (for short papers and posters) in the workshop proceedings, and will be presented as oral paper or poster.

Papers should be formatted according to the LREC style-sheet, which is provided on the LREC 2022 website (<http://lrec2022.lrec-conf.org/en/submission/authors-kit/>). Please submit papers in PDF format to the LREC website.

We are seeking submissions under the following category

- Full papers (10 pages)
- Short papers (work in progress : 5 pages)
- Posters (innovative ideas/proposals, research proposal of students)
- Demo (of working online/standalone systems)

WILDRE-6 will have a special focus on Demos of Indian Language Technology. In the past few years, as more resources have been developed and made available, there has been an increased activity in developing usable technology using these. WILDRE-6 would like to encourage and widen the Demo track to allow the community to showcase their demos and have mutually beneficial interactions with each other as well as resource developers.

WILDRE-6 is seeking full, short papers, posters and demos on the following topics related to Indian Language Resources:

- Digital Humanities, heritage computing
- Corpora - text, speech, multimodal, methodologies, annotation and tools
- Lexicons and Machine-readable dictionaries
- Ontologies, Grammars
- Language resources for NLP/ IR/Speech tasks, tools and Infrastructure for language resources
- Standards or specifications for language resources application
- Licensing and copyright issues
- Data mining
- Text summarization

Both submission and review processes will be handled electronically. The review process will be double blind. The workshop website will provide the submission guidelines and the link for the electronic submission.

When submitting a paper from the START page, authors will be asked to provide essential information about resources (in a broad sense, i.e. also technologies, standards, evaluation kits, etc.) that have been used for the work described in the paper or are a new result of your research. Moreover, ELRA encourages all LREC authors to share the described LRs (data, tools, services, etc.), to enable their reuse, replicability of experiments, including evaluation ones, etc.

For further information on this initiative, please refer to <http://lrec2022.lrec-conf.org/en/>

Shared Task

Following the success of the five WILDRE workshops, WILDRE-6 will include shared tasks for Indian languages. The organizers of shared tasks will provide datasets and evaluation platforms to evaluate systems developed by the participants. For further information on this initiative, please refer to <http://sanskrit.jnu.ac.in/conf/wildre6>

Workshop Chairs

- Girish Nath Jha, Jawaharlal Nehru University, India
- Kalika Bali, Microsoft Research India Lab, Bangalore, India
- Sobha L, AU-KBC, Anna University, Chennai, India

Workshop Organizing Committee

- Girish Nath Jha, Jawaharlal Nehru University, India
- Kalika Bali, Microsoft Research India Lab, Bangalore, India
- Sobha L, AU-KBC, Anna University, Chennai, India
- Atul Kr. Ojha, National University of Ireland, Galway, Ireland & Panlingua Language Processing LLP, India

Program Committee (to be updated)

- Adil Amin Kak, Kashmir University
- Anil Kumar Singh, IIT BHU, Benaras
- Anupam Basu, Director, NIIT, Durgapur
- Anoop Kunchukuttan, Microsoft AI and Research, India
- Arul Mozhi, University of Hyderabad
- Asif Iqbal, IIT Patna, Patna
- Atul Kr. Ojha, National University of Ireland Galway, Ireland & Panlingua Language Processing LLP, India
- Bharathi Raja Asoka Chakravarthi, National University of Ireland Galway, Ireland
- Bogdan Babych, Heidelberg University, Germany
- Chao-Hong Liu, Potamu Research Ltd., Ireland
- Claudia Soria, CNR-ILC, Italy
- Dafydd Gibbon, Universität Bielefeld, Germany
- Daan van Esch, Google, USA
- Dan Zeman, Charles University, Prague, Czech Republic
- Delyth Prys, Bangor University, UK
- Dipti Mishra Sharma, IIIT, Hyderabad
- Diwakr Mishra, Amazon-Bangalore, India
- Dorothee Beermann, Norwegian University of Science and Technology (NTNU)
- Elizabeth Sherley, IIITM-Kerala, Trivandrum
- Esha Banerjee, Google, USA
- Eveline Wandl-Vogt, Austrian Academy of Sciences, Austria
- Georg Rehm, DFKI, Germany
- Girish Nath Jha, Jawaharlal Nehru University, New Delhi
- Jan Odijk, Utrecht University, The Netherlands

- John P. McCre, National University of Ireland Galway, Ireland
- Jolanta Bachan, Adam Mickiewicz University, Poland
- Joseph Mariani, LIMSI-CNRS, France
- Jyoti D. Pawar, Goa University
- Kalika Bali, MSRI, Bangalore
- Khalid Choukri, ELRA, France
- Lars Hellan, NTNU, Norway
- M J Warsi, Aligarh Muslim University, India
- Malhar Kulkarni, IIT Mumbai
- Manji Bhadra, Bankura University, West Bengal
- Marko Tadic, Croatian Academy of Sciences and Arts, Croatia
- Massimo Monaglia, University of Florence, Italy
- Monojit Choudhary, MSRI Bangalore
- Narayan Choudhary, CIIL, Mysore
- Nicoletta Calzolari, ILC-CNR, Pisa, Italy
- Niladri Shekhar Dash, ISI Kolkata
- Panchanan Mohanty, GLA, Mathura
- Pinky Nainwani, Cognizant Technology Solutions, Bangalore
- Pushpak Bhattacharya, IIT Mumbai
- Rajeev R R, ICFOSS, Trivandrum
- Ritesh Kumar, Agra University
- Shantipriya Parida, Idiap Research Institute, Switzerland
- Vijay Kumar, TDIL, MEITY, Govt of India
- S.S. Agrawal, KIIT, Gurgaon, India
- Sachin Kumar, EZDI, Ahmedabad
- Santanu Chaudhury, Director, IIT Jodhpur
- Shivaji Bandhopadhyay, Director, NIT, Silchar
- Sobha L, AU-KBC Research Centre, Anna University
- Stelios Piperidis, ILSP, Greece
- Subhash Chandra, Delhi University
- Swaran Lata, TDIL, MCIT, Govt of India
- Virach Sornlertlamvanich, Thammasat University, Bangkok, Thailand
- Vishal Goyal, Punjabi University, Patiala
- Zygmunt Vetulani, Adam Mickiewicz University, Poland

Workshop contact:

Atul Kr. Ojha, National University of Ireland, Galway, Ireland & Panlingua Language Processing
LLP, India, shashwatup9k@gmail.com

Identify, Describe and Share your LRs

Describing your LRs in the LRE Map is now a normal practice in the submission procedure of LREC (introduced in 2010 and adopted by other conferences). To continue the efforts initiated at LREC 2014 about “Sharing LRs” (data, tools, web-services, etc.), authors will have the possibility, when submitting a paper, to upload LRs in a special LREC repository. This effort of sharing LRs, linked to the LRE Map for their description, may become a new “regular” feature for conferences in our field, thus contributing to creating a common repository where everyone can deposit and share data.

As scientific work requires accurate citations of referenced work so as to allow the community to understand the whole context and also replicate the experiments conducted by other researchers, LREC 2022 endorses the need to uniquely Identify LRs through the use of the International Standard Language Resource Number (ISLRN, www.islrn.org), a Persistent Unique Identifier to be assigned to each Language Resource. The assignment of ISLRNs to LRs cited in LREC papers will be offered at submission time.