



LREC 2022 Workshop
Language Resources and Evaluation Conference
20-25 June 2022

**The First Computing Social Responsibility Workshop
(CSR-NLP I 2022)**

PROCEEDINGS

Editors:
Mingyu Wan & Chu-Ren Huang

Proceedings of the LREC 2022 workshop on The First Computing Social Responsibility Workshop –NLP Approaches to Corporate Social Responsibilities (CSR-NLP I 2022)

Edited by:
Mingyu Wan & Chu-Ren Huang

ISBN: 979-10-95546-89-4
EAN: 9791095546894

For more information:

European Language Resources Association (ELRA)
9 rue des Cordelières
75013, Paris
France
<http://www.elra.info>
Email: lrec@elda.org



© European Language Resources Association (ELRA)

These workshop proceedings are licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License

Message from the General Chair

This volume documents the Proceedings of the first Workshop on Corporate Social Responsibility (CSR) using NLP methods, held on 25 June 2022 as part of the LREC 2022 conference (International Conference on Language Resources and Evaluation). This workshop is a very first attempt of bridging data resources, language theories, and NLP technologies on CSR in particular. It has called for innovative and practical NLP methods for tackling the CSR/ESG challenges at the cross-disciplinary research, especially with the nowadays' big data language technologies. This very first volume has finally included 8 excellent papers (5 oral presentations + 3 posters) on some core topics pertaining to CSR/ESG.

Message from the Program Chairs

Corporate Social Responsibility (CSR) as a shared grand challenge in business studies and in computational linguistics has not been tackled yet in the recently thriving financial NLP studies. These work so far have been more driven by the NLP downstream technology instead of the theoretical or real-world issues driving studies of economics or business.

Conventional methods usually focus on shared values of companies such as sustainability, carbon footprint, diversity and inclusion, fair-trade, social justice, environmental impact. However, different businesses may breed additional and more specific areas of issues to address, such as pollution/emission, pharmacovigilance, food safety etc.

The goal of the workshop is to identify and develop niche research methodologies that are highly competitive and world-leading for CSR modelling. Tackling the grand challenges of the world by promoting mutual understanding through language and CSR is a necessary step towards tackling other grand challenges that also may be aided by language big data, deep neural networks, linguistic tools and methods, towards some of the trending issues in CSR studies such as environmental degradation and the climate crisis.

In CSR-NLP I 2022, we have provided such a venue for researchers and practitioners worldwide to conduct computational linguistic research and make use of NLP methods to address some of the core issues for CSR or ESG related research. We look forward to more submissions and participation in the next event, hopefully to build up large-scale CSR data resources, tools, or platforms with NLP-facilitated technologies, as well as the language resource evaluation, and to launch a shared task for CSR modeling/prediction for open competition. We believe with this considerable outcome and fruitful discussions for the first computing workshop on CSR, there will be more promising and cheerful results in the next events.

Organizers

Kathleen Ahrens – The Hong Kong Polytechnic University
Emmanuele Chersoni – The Hong Kong Polytechnic University
Chu-Ren Huang – The Hong Kong Polytechnic University
Huyen Nguyen Thi Minh – VNU University of Science
Cindy Sing Bik Ngai – The Hong Kong Polytechnic University
Weiwei Sun – Cambridge University
Rachel Edita O. Roxas – National University, Manila, Philippines
Qi Su – Peking University
Mingyu Wan – The Hong Kong Polytechnic University
Qiang Wu – The Hong Kong Polytechnic University

Program Committee:

Xiaopeng Bai (ECNU)
Maria Bonnafous-Boucher (Novancia Business School)
Jason Chang (National Tsing Hua University)
Chung-Chi Chen (AI Research Center, Japan)
Hsin-Hsi Chen (National Taiwan)
Agnes Cheng (Oklahoma State)
Emmanuele Chersoni (HK PolyU)
Chris Cieri (LDC-UPenn)
Jinghang Gu (HK PolyU)
Shu-kai Hsieh (National Taiwan)
Yu-yin Hsu (HK PolyU)
Ngyun Thi Minh Hyuen (Vietnam National)
Chunyu Kit (CityU HK)
John Lee (CityU HK)
Jing Li (HK PolyU)
Karl Neergaard (University of Macau)
Cindy Sing Bik Ngai (HK PolyU)
Bo Peng (HK PolyU)
Rachel Roxas (National U-Philippines)
Enrico Santus (Bayer)
Qi Su (Peking U)
Mingyu Wan (HK PolyU)
Qiang Wu (HK PolyU)
Rong Xiang (HK PolyU)
Winnie Zeng (HK PolyU)

Sponsors

The workshop is substantially supported by the funding for Projects of Strategic Importance of The Hong Kong Polytechnic University (Project Code: ZE2J, Project title: Artificial Intelligence and Applications of Natural language processing in Business) and sponsored by the Research Centre for Professional Communication in English (RCPCE).

Table of Contents

<i>An NLP Approach for the Analysis of Global Reporting Initiative Indexes from Corporate Sustainability Reports</i>	
Marco Polignano, Nicola Bellantuono, Francesco Paolo Lagrasta, Sergio Caputo, Pierpaolo Pontandolfo and Giovanni Semeraro	1
<i>Tracking Changes in ESG Representation: Initial Investigations in UK Annual Reports</i>	
Matthew Purver, Matej Martinc, Riste Ichev, Igor Lončarski, Katarina Sitar Šuštar, Aljoša Valentinčič and Senja Pollak	9
<i>A Corpus-based Study of Corporate Image Represented in Corporate Social Responsibility Report: A Case Study of China Mobile and Vodafone</i>	
Xing Chen and Liang Xu	15
<i>Framing Legitimacy in CSR: A Corpus of Chinese and American Petroleum Company CSR Reports and Preliminary Analysis</i>	
Jieyu Chen, Kathleen Ahrens and Chu-Ren Huang	24
<i>MobASA: Corpus for Aspect-based Sentiment Analysis and Social Inclusion in the Mobility Domain</i>	
Aleksandra Gabryszak and Philippe Thomas	35
<i>Detecting Violation of Human Rights via Social Media</i>	
Yash Pilankar, Rejwanul Haque, Mohammed Hasanuzzaman, Paul Stynes and Pramod Pathak ..	40
<i>Inclusion in CSR Reports: The Lens from a Data-Driven Machine Learning Model</i>	
Lu Lu, Jinghang Gu and Chu-Ren Huang	46
<i>Towards Classification of Legal Pharmaceutical Text using GAN-BERT</i>	
Tapan Auti, Rajdeep Sarkar, Bernardo Stearns, Atul Kr. Ojha, Arindam Paul, Michaela Comerford, Jay Megaro, John Mariano, Vall Herard and John P. McCrae	52

Conference Program

- 14:00–15:00 Roundtable discussion by PC members and CSR/ESG experts**
- 15:05–18:00 Oral Presentations**
- 15:05–15:30 *An NLP Approach for the Analysis of Global Reporting Initiative Indexes from Corporate Sustainability Reports*
Marco Polignano, Nicola Bellantuono, Francesco Paolo Lagrasta, Sergio Caputo, Pierpaolo Pontrandolfo and Giovanni Semeraro
- 15:35–16:00 *Tracking Changes in ESG Representation: Initial Investigations in UK Annual Reports*
Matthew Purver, Matej Martinc, Riste Ichev, Igor Lončarski, Katarina Sitar Šuštar, Aljoša Valentinčič and Senja Pollak
- 16:00–16:30 Coffee break**
- 16:35–17:00 *A Corpus-based Study of Corporate Image Represented in Corporate Social Responsibility Report: A Case Study of China Mobile and Vodafone*
Xing Chen and Liang Xu
- 17:05–17:30 *Framing Legitimacy in CSR: A Corpus of Chinese and American Petroleum Company CSR Reports and Preliminary Analysis*
Jieyu Chen, Kathleen Ahrens and Chu-Ren Huang
- 17:35–18:00 *MobASA: Corpus for Aspect-based Sentiment Analysis and Social Inclusion in the Mobility Domain*
Aleksandra Gabryszak and Philippe Thomas
- 18:05–18:45 Poster presentations**
- 18:05–18:15 *Detecting Violation of Human Rights via Social Media*
Yash Pilankar, Rejwanul Haque, Mohammed Hasanuzzaman, Paul Stynes and Pramod Pathak
- 18:20–18:30 *Inclusion in CSR Reports: The Lens from a Data-Driven Machine Learning Model*
Lu Lu, Jinghang Gu and Chu-Ren Huang
- 18:35–18:45 *Towards Classification of Legal Pharmaceutical Text using GAN-BERT*
Tapan Auti, Rajdeep Sarkar, Bernardo Stearns, Atul Kr. Ojha, Arindam Paul, Michaela Comerford, Jay Megaro, John Mariano, Vall Herard and John P. McCrae

No Day Set (continued)

