

NLP4HR 2024

**The First Workshop on Natural Language Processing for
Human Resources (NLP4HR 2024)**

Proceedings of the Workshop

March 22, 2024

The NLP4HR organizers gratefully acknowledge the support from the following sponsors.



Megagon Labs

©2024 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 979-8-89176-076-9

Introduction

Welcome to the first Workshop on Natural Language Processing for Human Resources!

The field of human resources (HR) encompasses a wide range of tasks where the application of natural language processing (NLP) holds significant promise. Applications such as talent acquisition, career development guidance, performance management, and ongoing education and training involve a substantial amount of unstructured, semi-structured, and structured data. NLP can offer (semi-)automated solutions to address the challenges associated with these tasks. At the same time, the integration of AI in HR applications also presents certain risks and concerns, such as fairness, privacy, reproducibility, controllability, and transparency. There is an enormous opportunity to leverage advanced NLP techniques to address these challenges. These areas not only intersect with established NLP tasks but also overlap with fields like data mining and management, which have inspired workshops across various communities. These workshops include the Workshop on Recommender Systems for Human Resources, the International Workshop on Talent and Management Computing, and the International Workshop AI for Human Resources and Public Employment Services. The NLP4HR workshop specifically aims to bring together NLP researchers and practitioners to advance innovative solutions for HR challenges.

For this inaugural edition of the NLP4HR workshop, we received 11 original research submissions. Thanks to the diligent efforts of our Program Committee members, we curated a collection of 6 original research contributions (an acceptance rate of 55%). Furthermore, we will feature one paper from the EACL Findings during the workshop. We believe that these papers will inspire the research community and lay a solid foundation for ongoing and future endeavors in this area.

In addition to the presentation of research papers, the workshop will feature invited talks and a panel discussion, providing a platform for participants to exchange ideas and experiences through open discussions. We are pleased to have the following invited speakers: Trey Causey (Indeed), David Graus (Randstad), Marko Grobelnik (Jozef Stefan Institute, OECD, European Union Commission), and Barbara Plank (Ludwig-Maximilians-Universität München, IT University of Copenhagen). They will cover a range of critical topics associated with NLP applications in the HR domain. The panel, titled “HR in the Era of Large Pre-trained Models: Sorting out *the Good, the Bad, and the Ugly*,” will shed light on the impacts and potential risks of AI-based approaches in HR tasks.

We extend our congratulations to the authors of the accepted papers and express gratitude to all authors who submitted their work, members of the program committee, mentors and mentees who participated in the mentorship program, and the EACL 2024 workshop chairs.

We would like to give special thanks to **Megagon Labs** for their support of our workshop.

The First NLP4HR Workshop Program Chairs

<https://megagon.ai/nlp4hr-2024/>

Program Committee

Program Chairs

Estevam Hruschka, Megagon Labs
Thom Lake, Indeed and University of Texas at Austin
Tom Mitchell, Carnegie Mellon University
Naoki Otani, Megagon Labs

Reviewers

Aisha Aijaz, Indraprastha Institute of Information Technology, Delhi
Benjamin Clavié, Bright Network
Yoshinari Fujinuma, AWS AI Labs
Nidhi Goyal, Indraprastha Institute of Information Technology, Delhi
Prakhar Gupta, Google
Sairam Gurajada, Megagon Labs
Bo Kang, Ghent University
Mesut Kaya, Aalborg University Copenhagen
Takashi Kodama, Kyoto University
Kaixin Ma, Tencent AI Lab
Seiji Maekawa, Megagon Labs
Takuya Makino, Megagon Labs
Lorenzo Malandri, University of Milan - Bicocca
Hiroshi Matsuda, Megagon Labs
Sajjadur Rahman, Megagon Labs
Weijie Xu, Amazon
Michiharu Yamashita, Pennsylvania State University
Mike Zhang, IT University of Copenhagen

Mentorship Program - Mentors

Nikita Bhutani, Megagon Labs
Lorenzo Malandri, University of Milan - Bicocca
Pouya Pezeshkpour, Megagon Labs
Rob van der Goot, IT University of Copenhagen

Invited Talk: 1

David Graus

Randstad, University of Amsterdam

March 22nd, 2024 – Time: **08:40 - 09:30 CET** – Room: **Gardjola 3 (Corinthia)**

Bio: David Graus is lead data scientist at Randstad, the global leader in the HR services industry. At Randstad he works with his chapter of data scientists on AI-powered products and services, including algorithmic matching of jobs to job seekers and natural language processing for labor market analysis. Prior to his role at Randstad he worked on news personalization for Het Financieele Dagblad, and the award-winning SMART Radio for BNR Nieuwsradio. David obtained his PhD in Information Retrieval in 2017 at the University of Amsterdam under supervision of prof. dr. Maarten de Rijke, where he worked on semantic search and computational methods for automated understanding of large-scale textual digital traces.

Invited Talk: 2

Marko Grobelnik

Jozef Stefan Institute, OECD, European Union Commission

March 22nd, 2024 – Time: **11:00 - 11:50 CET** – Room: **Gardjola 3 (Corinthia)**

Bio: Marko Grobelnik is an expert researcher in the field of Artificial Intelligence (AI). Focused areas of expertise are Machine Learning, Data/Text/Web Mining, Network Analysis, Semantic Technologies, Deep Text Understanding, and Data Visualization. Marko co-leads the Department for Artificial Intelligence at Jozef Stefan Institute, co-founded UNESCO International Research Center on AI (IRCAI), and is the CEO of Quintelligence.com specialized in solving complex AI tasks for the commercial world. He collaborates with major European academic institutions and major industries such as Bloomberg, British Telecom, European Commission, Microsoft Research, New York Times. Marko is co-author of several books, co-founder of several start-ups and is/was involved into over 50 EU funded research projects in various fields of Artificial Intelligence. Marko represents Slovenia in the OECD AI Committee (AIONE) and in the Council of Europe Committee on AI (CAHAI). In 2016 Marko became Digital Champion of Slovenia at the European Commission.

Invited Talk: 3

Barbara Plank

Ludwig-Maximilians-Universität München, IT University of Copenhagen

March 22nd, 2024 – Time: **13:00 - 13:50 CET** – Room: **Gardjola 3 (Corinthia)**

Bio: Barbara Plank is Full Professor and Chair for AI and Computational Linguistics at Ludwig-Maximilians-Universität München (LMU Munich) and Full professor at the IT University of Copenhagen. Her research focuses on NLP and includes learning under sample selection bias (domain adaptation, transfer learning), annotation bias (human disagreements and human uncertainty), learning from beyond the text, and in general learning under limited supervision. Barbara is the recipient of an ERC Consolidator grant, a Sapere Aude Research Leader grant and an Amazon Research Award. Barbara regularly serves in *ACL committees and is currently president of the Northern European Association for Language Technology.

Invited Talk: 4

Trey Causey

Indeed

March 22nd, 2024 – Time: 16:00 - 16:50 CET – Room: Gardjola 3 (Corinthia)

Bio: Trey Causey is Head of Responsible AI and Senior Director of Data Science at Indeed, where he and his team work to ensure Indeed's use of AI is beneficial to job seekers, employers, and society. As the author of Indeed's responsible AI strategy, he leads an interdisciplinary group of data scientists, engineers, and researchers in tackling the sociotechnical issues of algorithmic, human, and systemic bias. He also is a member of Indeed's Environmental, Social, and Governance leadership team, where he and his colleagues are changing the way the world hires, for good. Prior to Indeed, Trey has led data science teams, machine learning engineering teams, and product teams across the technology industry. Trained as a sociologist at the University of Washington, he confronts challenges in Responsible AI with both a sociological and a technological approach.

Table of Contents

<i>Deep Learning-based Computational Job Market Analysis: A Survey on Skill Extraction and Classification from Job Postings</i>	
Elena Senger, Mike Zhang, Rob van der Goot and Barbara Plank	1
<i>Aspect-Based Sentiment Analysis for Open-Ended HR Survey Responses</i>	
Lois Rink, Job Meijdam and David Graus	16
<i>Rethinking Skill Extraction in the Job Market Domain using Large Language Models</i>	
Khanh Cao Nguyen, Mike Zhang, Syrielle Montariol and Antoine Bosselut	27
<i>JobSkape: A Framework for Generating Synthetic Job Postings to Enhance Skill Matching</i>	
Antoine Magron, Anna Dai, Mike Zhang, Syrielle Montariol and Antoine Bosselut	43
<i>HR-MultiWOZ: A Task Oriented Dialogue (TOD) Dataset for HR LLM Agent</i>	
Weijie Xu, Zicheng Huang, Wenxiang Hu, Xi Fang, Rajesh Kumar Cherukuri, Naumaan Nayyar, Lorenzo Malandri and Srinivasan H. Sengamedu	59
<i>Big City Bias: Evaluating the Impact of Metropolitan Size on Computational Job Market Abilities of Language Models</i>	
Charlie Campanella and Rob van der Goot	73

Program

Friday, March 22, 2024

- 08:30 - 08:40 *Opening Remarks*
- 08:40 - 09:30 *Invited Talk 1*
- 09:30 - 10:30 *Oral Presentation Session 1*
- 10:30 - 11:00 *Break*
- 11:00 - 11:50 *Invited Talk 2*
- 11:50 - 12:20 *Oral Presentation Session 2*
- 12:20 - 13:00 *Lunch*
- 13:00 - 13:50 *Invited Talk 3*
- 13:50 - 14:40 *Panel*
- 14:40 - 15:30 *Poster Session*
- 15:30 - 16:00 *Break*
- 16:00 - 16:50 *Invited Talk 4*
- 16:50 - 17:00 *Closing Remarks*