DaSH 2022

Data Science with Human-in-the-Loop (Language Advances)

Proceedings of the DaSH Workshop at EMNLP 2022

December 8, 2022

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

With support from

Order copies of this and other ACL proceedings from:

Introduction

We are delighted to welcome to you DaSH 2022, the Fourth Workshop on Data Science with Human-in-the-loop at EMNLP 2022!

The aim of this workshop is to stimulate research on the cooperation between humans and computers within the broad area of natural language processing, including but not limited to information extraction, information retrieval and text mining, machine translation, dialog systems, question answering, language generation, summarization, model interpretability, evaluation, fairness, and ethics. We invite researchers and practitioners interested in understanding how to optimize human-computer cooperation and how to minimize human effort along an NLP pipeline in a wide range of tasks and applications.

This year, the workshop program includes three keynote talks, two invited talks, fourteen accepted papers, and three 'Findings of EMNLP' papers. We received 22 submissions, each of which received at least two reviews from our distinguished program committee. The submissions and accepted papers show a strong mix of participation both from the academia and industry: 50% of the accepted papers have an academic provenance (the primary affiliation of the lead author is a university), while 50% originate in industry labs.

We hope to bring together interdisciplinary researchers from academia, research labs and practice to share, exchange, learn, and develop preliminary results, new concepts, ideas, principles, and methodologies on understanding and improving human-computer interaction in natural language processing. We expect the workshop to help develop and grow a strong community of researchers who are interested in this topic and to yield future collaborations and scientific exchanges across the relevant areas of computational linguistics, natural language processing, data mining, machine learning, data and knowledge management, human-machine interaction, and intelligent user interfaces. We are thankful to IBM research for sponsoring the workshop and best paper awards.

We hope you have a wonderful time at the workshop.

Cheers!

DaSH 2022 Organizers
Eduard Dragut, Temple University
Yunyao Li, Apple Inc.
Lucian Popa, IBM Research
Shashank Srivastava, UNC Chapel Hill
Slobodan Vucetic, Temple University

Organizing Committee

Organizers

Eduard Dragut, Temple University
Yunyao Li, Apple
Lucian Popa, IBM Research - Almaden
Shashank Srivastava, University of North Carolina at Chapel Hill
Slobodan Vucetic, Temple University

Program Committee

Chairs

Eduard Dragut, Temple University
Yunyao Li, Apple
Lucian Popa, IBM Research - Almaden
Shashank Srivastava, University of North Carolina at Chapel Hill
Slobodan Vucetic, Temple University

Program Committee

Maria Antoniak, Allen Institute for Artificial Intelligence
Jeffrey Bigham, Carnegie Mellon University/Apple
Su Lin Blodgett, Microsoft Research
Marina Danilevsky, IBM Research
Aritra Dasgupta, New Jersey Institute of Technology
Nemanja Djuric, Aurora Innovation
Anna Lisa Gentile, IBM Research - Almaden
Sayan Ghosh, University of North Carolina at Chapel Hill
Saptarshi Guha, Apple
Toby Jia-jun Li, University of Notre Dame
Katherine Luna, Apple
Shen Ma, Apple
Dominik Moritz, Carnegie Mellon University
Kun Qian, Apple
Rakesh R. Menon, University of North Carolina at Chapel Hill
Sajjadur Rahman, Megagon Labs
Huan Sun, Ohio State University
Mihai Surdeanu, University of Arizona
Kerem Zaman, University of North Carolina at Chapel Hill
Li Zhang, University of Pennsylvania

Steering Committee

AnHai Doan, University of Wisconsin, Madison
Marti Hearst, University of California, Berkeley
Sunita Sarawagi, Indian Institute of Technology Bombay
Daniel Weld, University of Washington
ChengXiang Zhai, University of Illinois, Urbana-Champaign

Keynote Talk: Decision-making in an uncertain world

Azza Abouzied

NYU Abu Dhabi

Abstract: Decision-makers in a broad range of domains, such as finance, and healthcare, need to derive optimal decisions given a set of constraints and objectives over uncertain data or models. Traditional solutions to such problems are typically complex, do not generalize, and do not scale to our modern-day massive data scales. The emerging research area of prescriptive analytics aims to provide declarative, and scalable approaches. In this talk, I discuss some strategies for addressing key challenges related to usability, scalability, data uncertainty, dynamic environments with changing data and models, and the need to support decision-making agents with emphasis on a real-life application of supporting policy-makers with epidemic intervention planning.

Bio: Azza Abouzied's research work focuses on designing intuitive data querying tools. Today's technologies are helping people collect and produce data at phenomenal rates. Despite the abundance of data, it remains largely inaccessible due to the skill required to explore, query and analyze it in a non-trivial fashion. While many users know exactly what they are looking for, they have trouble expressing sophisticated queries in interfaces that require knowledge of a programming language or a query language. Azza designs novel interfaces, such as example-driven query tools, that simplify data querying and analysis. Her research work combines techniques from various research fields such as UI-design, machine learning, and databases. Azza Abouzied received her doctoral degree from Yale in 2013. She spent a year as a visiting scholar at UC Berkeley. She is also one of the co-founders of Hadapt – a Big Data analytics platform.

Keynote Talk: Measuring what matters for human-AI teams

Saleema Amershi

Microsoft Research

Abstract: There is a significant discrepancy between the success metrics driving the AI industry and what people value in the real world. In the AI-assisted programming scenario, for example, a key value proposition is the potential for code generation models to dramatically improve developer productivity. Yet, offline metrics used to inform model development decisions and gate which models are deployed to people in the real world currently focus on generation correctness rather than correctness or effort with a developer-in-the-loop. Similarly, online metrics currently focused on acceptance rates overlook interaction costs to developers in prompting, reviewing, and editing generated code. In this talk, I will describe ongoing work from the HAX team at Microsoft Research to develop metrics and measurement tools that more faithfully reflect the needs and effectiveness of human-AI teams.

Bio: Saleema Amershi is a Senior Principal Research Manager at Microsoft Research where she leads the Human-AI eXperiences (HAX) team, building tools for creating responsible AI experiences. She also currently chairs Microsoft's Aether Working Group on Human-AI Interaction and Collaboration. Aether is Microsoft's advisory committee on responsible and ethical AI. Saleema holds a PhD in Computer Science & Engineering from the Paul G. Allen School of Computer Science and Engineering. Prior to UW, she completed a MSc in Computer Science and a BSc in Computer Science and Mathematics at the University of British Columbia.

Keynote Talk: Data Preparation with human in the loop - The Case for a commodity crowdsourcing system

Mourad Ouzzani

Qatar Computing Research Institute, Hamad Bin Khalifa University, Qatar

Abstract: Data science is becoming central to many industries as it helps extract insights from large amounts of data that in turn lead to critical actions in many applications. An important aspect of data science is data preparation that we broadly define to include data cleaning, extraction, matching, merging, labeling, and so on. It is often the case that these tasks are performed collaboratively by multiple users. Such collaborative work is often referred to as crowdsourcing. While much work has been published in this area, most focus on developing algorithmic solutions for point problems. There is very little work on how to build crowdsourcing systems that can be used widely. In this talk, I will describe our current efforts to build Cymphony, a commodity collaborative management system that effectively supports various collaborative data preparation tasks. We aim to build a platform that enables a uniform specification as well as efficient execution of data preparation tasks that require collaboration among a group of users. While existing systems may cover some aspects of data collaboration, they do not provide easy ways to express and execute the many collaboration scenarios that are often found in practice. In a nutshell, we define a simple workflow consisting of three basic operators, namely Assign, Annotate, and Aggregate. Multiple simple workflows along with basic data manipulation operations (e.g., SQL queries, sampling) can be composed to build more complex workflows. In addition to in-house users who can connect to Cymphony, we also provide integration with public crowdsourcing platforms such as Amazon Mechanical Turk. Early experiments show that Cymphony allows the easy specification and automatic execution of many practical workflows without having to write code to run them or manually connect parts of these workflows.

Bio: Mourad Ouzzani is a Research Director with the Qatar Computing Research Institute at Hamad Bin Khalifa University, Qatar Foundation. Before joining QCRI, he was a research associate professor at Purdue University. Mourad has played a key role in establishing the data analytics group within QCRI. He is now leading the Research Engineering Group whose mission is to productize QCRI's research. Mourad's research interests lie in the fields of data management and analytics with a focus on data integration, data cleaning, and collaborative data science. He was the project lead of Rayyan, the leading systematic reviews web and mobile app, which is now being used by more than 250k users worldwide. Rayyan has since graduated from QCRI to a start-up, Rayyan Systems Inc. His work has led to numerous publications in top tier venues including PVLDB, TKDE, SIGMOD, and ICDE. He holds a PhD from Virginia Tech, and a BSc and MSc from USTHB, Algiers, all in computer science.

Keynote Talk: Enabling domain experts to create their own NLP models: Notes from our journey

Yannis Katsis

IBM Research - Almaden

Abstract: Creating AI models for Natural Language Processing (NLP) tasks remains a daunting task for domain experts that have the need but lack the technical expertise and resources to create such models. To address this issue, at IBM Research we have been designing and developing human-in-the-loop systems that enable domain experts to interactively build their own NLP models. In this talk I will summarize our experience with building two such systems focusing on text classification and information extraction, respectively. After a brief description of the systems, I will focus on the lessons learned during this process and interesting research challenges that lie ahead in our journey to lowering the barrier of entry to NLP.

Bio: Yannis Katsis is a Senior Research Scientist at IBM Research - Almaden with expertise in the management, integration, and extraction of knowledge from structured, semi-structured, and unstructured data. In his recent work, Yannis focuses on lowering the barrier of entry to knowledge extraction by designing, analyzing, and building human-in-the-loop systems that enable domain experts to interactively generate knowledge extraction AI models that serve their needs. Yannis received his PhD in Computer Science from UC San Diego. His work has appeared in top conferences and journals in the areas of data management, natural language processing, and human-computer interaction, and has been leveraged for multiple IBM products.

Keynote Talk: Understanding and Addressing Uncertainty in Crowd Annotation on Subjective Tasks

Amy Zhang
University of Washington

Abstract: Uncertainty is an important factor that is crucial when it comes to understanding human judgments in subjective and nuanced domains. Whether it's annotators producing data used to train and evaluate machine learning systems for subjective tasks or online communities adjudicating moderation actions on nuanced content, many groups and individuals need to account for and address the uncertainty that comes along with conducting judgments. This issue is exacerbated as the application of computing technology expands to more areas that depend on social and cultural factors. I will present an overview of work consisting of a set of novel annotation tools and workflow designs that support capturing, distinguishing, and addressing uncertainty throughout each step involved in making group judgments. Specifically, I will talk about: (1) how we can use ranges to better capture and distinguish sources of uncertainty in scalar rating tasks; (2) how we can use precedents to interact with uncertainty in categorical decision tasks; (3) how we can address disagreements to reduce uncertainty through pairwise multi-turn deliberation; and (4) how we can dynamically select targeted interventions for reducing uncertainty.

Bio: Amy X. Zhang is an assistant professor at University of Washington's Allen School of Computer Science and Engineering. Previously, she was a 2019-20 postdoctoral researcher at Stanford University's Computer Science Department after completing her Ph.D. at MIT CSAIL in 2019, where she received the George Sprowls Best Ph.D. Thesis Award at MIT in computer science. During her Ph.D., she was an affiliate and 2018-19 Fellow at the Berkman Klein Center at Harvard University, a Google Ph.D. Fellow, and an NSF Graduate Research Fellow. Her work has received a best paper award at ACM CSCW, a best paper honorable mention award at ACM CHI, and has been profiled on BBC's Click television program, CBC radio, and featured in articles by ABC News, The Verge, New Scientist, and Poynter. She is a founding member of the Credibility Coalition, a group dedicated to research and standards for information credibility online. She has interned at Microsoft Research and Google Research. She received an M.Phil. in Computer Science at the University of Cambridge on a Gates Fellowship and a B.S. in Computer Science at Rutgers University, where she was captain of the Division I Women's tennis team.

Table of Contents

Program

Thursday, December 8, 2022

08:55 - 09:00   *Opening Remarks*

09:00 - 09:40   *Keynote Talk 1*

09:40 - 10:40   *Paper Session 1*

10:40 - 11:00   *Coffee Break*

11:00 - 11:40   *Keynote Talk 2*

11:40 - 12:30   *Paper Session 2*

12:30 - 14:00   *Lunch Break*

14:00 - 14:20   *Invited Talk 1*

14:20 - 14:40   *Invited Talk 2*

14:40 - 15:40   *Paper Session 3*

15:30 - 16:00   *Coffee Break*

16:00 - 16:40   *Keynote Talk 3 (Hybrid)*

16:40 - 17:30   *Panel Discussion*