

Practical D2T 2024

**The 2nd Workshop on Practical
LLM-assisted Data-to-Text Generation**

Proceedings of the Workshop

September 23, 2024

©2023 The 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 979-8-89176-126-1

Preface

We present the Proceedings of The 2nd Workshop on Practical LLM-assisted Data-to-Text (Practical D2T). This year's Practical D2T takes place at INLG 2024 on Sept 23 in Tokyo, Japan. We would like to thank the INLG organisers for their support.

Natural Language Generation (NLG) has been an active area of research for decades, both academically and industrially. Data-to-text (D2T) generation is the NLG task where a system describes structured data in natural language. Traditionally, commercial D2T systems have been based on symbolic approaches, i.e. handcrafted rules or templates. More experimental approaches to D2T, such as E2E and Transformer-based systems have been limited to research because of well-known issues like knowledge gaps, lack of factuality, and hallucination.

The recently introduced instruction-tuned, multi-task Large Language Models (LLMs) promise to become a viable alternative to rule-based D2T systems. They exhibit the ability to capture knowledge, follow instructions, and produce coherent text from various domains. However, even the best LLMs still suffer from well-known issues of neural models, such as lack of controllability and risk of producing harmful text. Recent research thus proposed various approaches to improve the semantic accuracy of LLMs D2T, including prompt tuning, targeted fine-tuning, Retrieval Augmented Generation (RAG), external tool integration, and neuro-symbolic approaches.

Practical D2T 2024 aims to build a space for researchers to discuss and present innovative work on D2T systems using LLMs.

This year, we are excited to present two keynotes covering the use of LLMs in D2T and related tasks. The keynote speakers are:

- Craig Thomson, Dublin City University / ADAPT, UK
- Marco Valentino, Idiap Research Institute, Switzerland

Practical D2T hosts a hackathon (for the second consecutive time), which this year is focused on the evaluation and semantic accuracy of D2T using LLMs. The hackathon will allow participants to explore the challenges of using LLMs for both generating textual summaries of structured data and text span error annotation of them.

Finally, the workshop features a panel of experts on D2T who will discuss the use of LLMs for generating text from data. They will cover the main challenges involved and share insights on the latest developments in this area.

The Practical D2T 2024 program chairs,
Simone Balloccu (lead), Charles University
Zdeněk Kasner, Charles University
Ondřej Plátek, Charles University
Patricia Schmidtová, Charles University
Kristýna Onderková, Charles University
Mateusz Lango, Charles University
Ondřej Dušek, Charles University
Lucie Flek, University of Bonn
Ehud Reiter, University of Aberdeen
Dimitra Gkatzia, Edinburgh Napier University
Simon Mille, ADAPT Centre

The organisation of Practical D2T 2024 was partially funded by the European Union (ERC, NG-NLG, 101039303).

Organizing Committee

Workshop Chairs

Simone Balloccu (lead), Charles University
Zdeněk Kasner, Charles University
Ondřej Plátek, Charles University
Patrícia Schmidtová, Charles University
Kristýna Onderková, Charles University
Mateusz Lango, Charles University
Ondřej Dušek, Charles University
Lucie Flek, University of Bonn
Ehud Reiter, University of Aberdeen
Dimitra Gkatzia, Edinburgh Napier University
Simon Mille, ADAPT Centre

Table of Contents

Beyond the Hype: Identifying and Analyzing Math Word Problem-Solving Challenges for Large Language Models

Romina Soledad Albornoz-De Luise, David Arnau, Pablo Arnau-González and Miguel Arevalillo-Herráez 1

Enhancing Situation Awareness through Model-Based Explanation Generation

Konstantinos Gavriilidis, Ioannis Konstas, Helen Hastie and Wei Pang 7

