

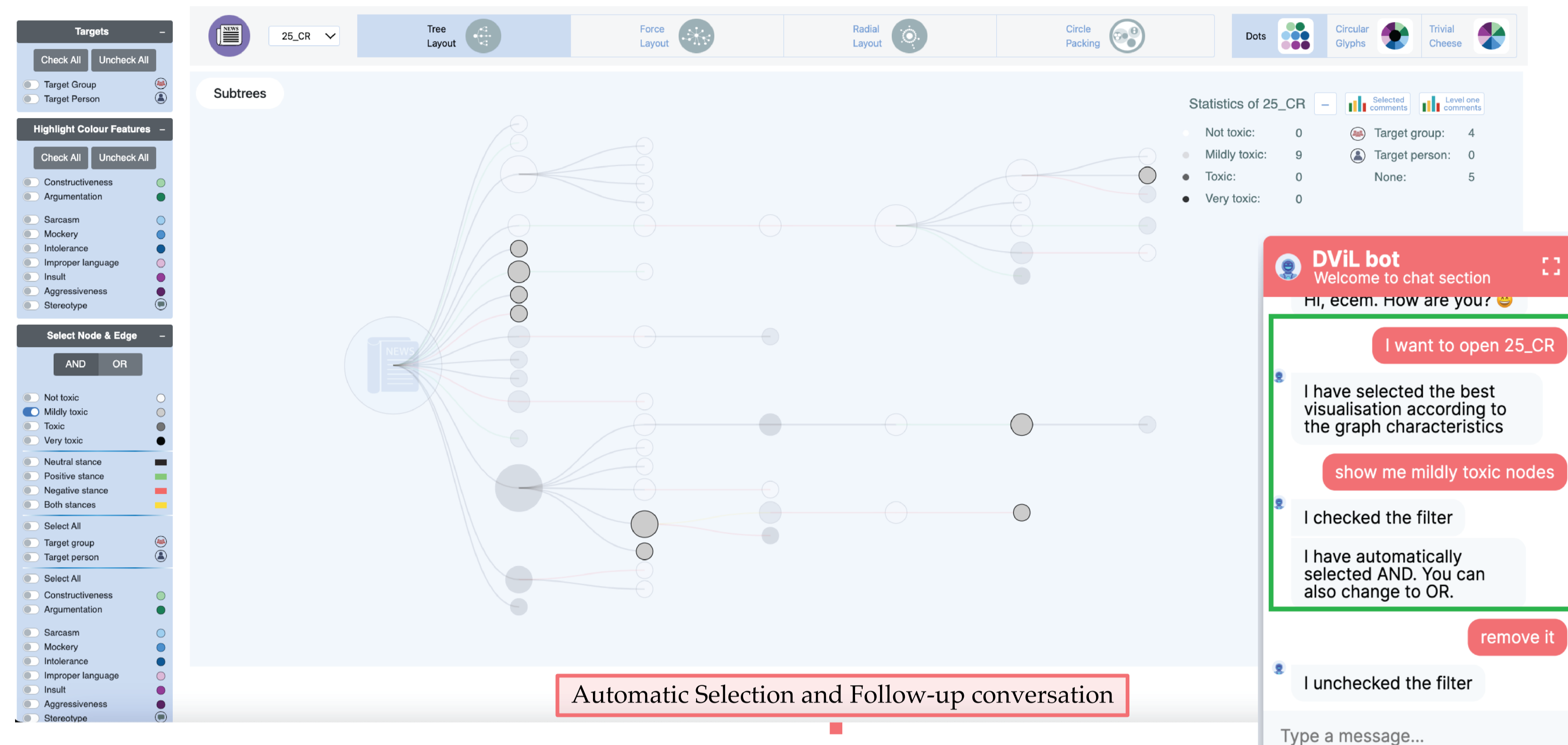
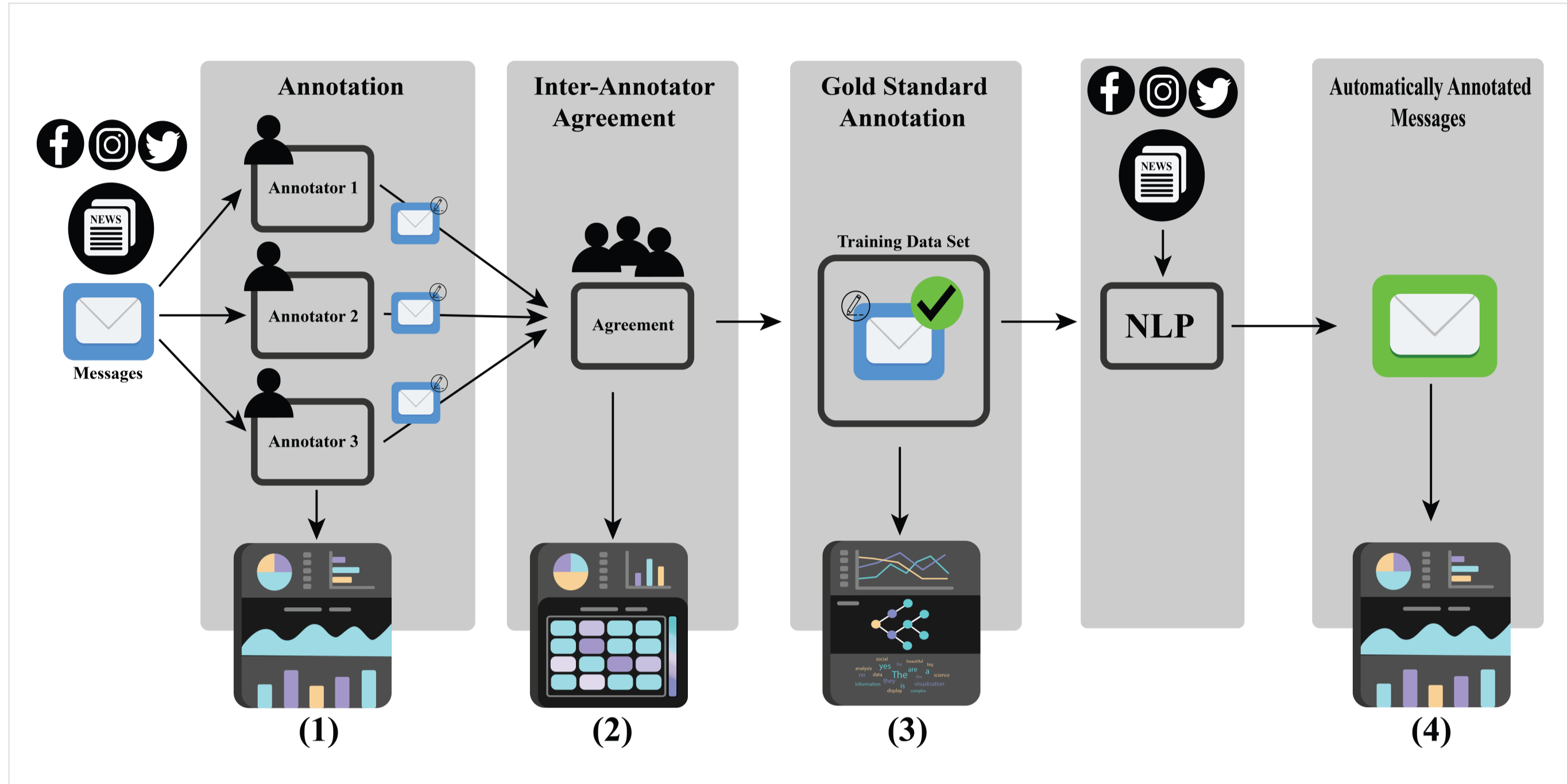
A CONVERSATIONAL DATA VISUALISATION PLATFORM FOR HIERARCHICAL MULTIVARIATE DATA

Ecem Kavaz, Inmaculada Rodríguez, Anna Puig, and Eduard Vives

Departament de Matemàtiques i Informàtica, University of Barcelona, Spain

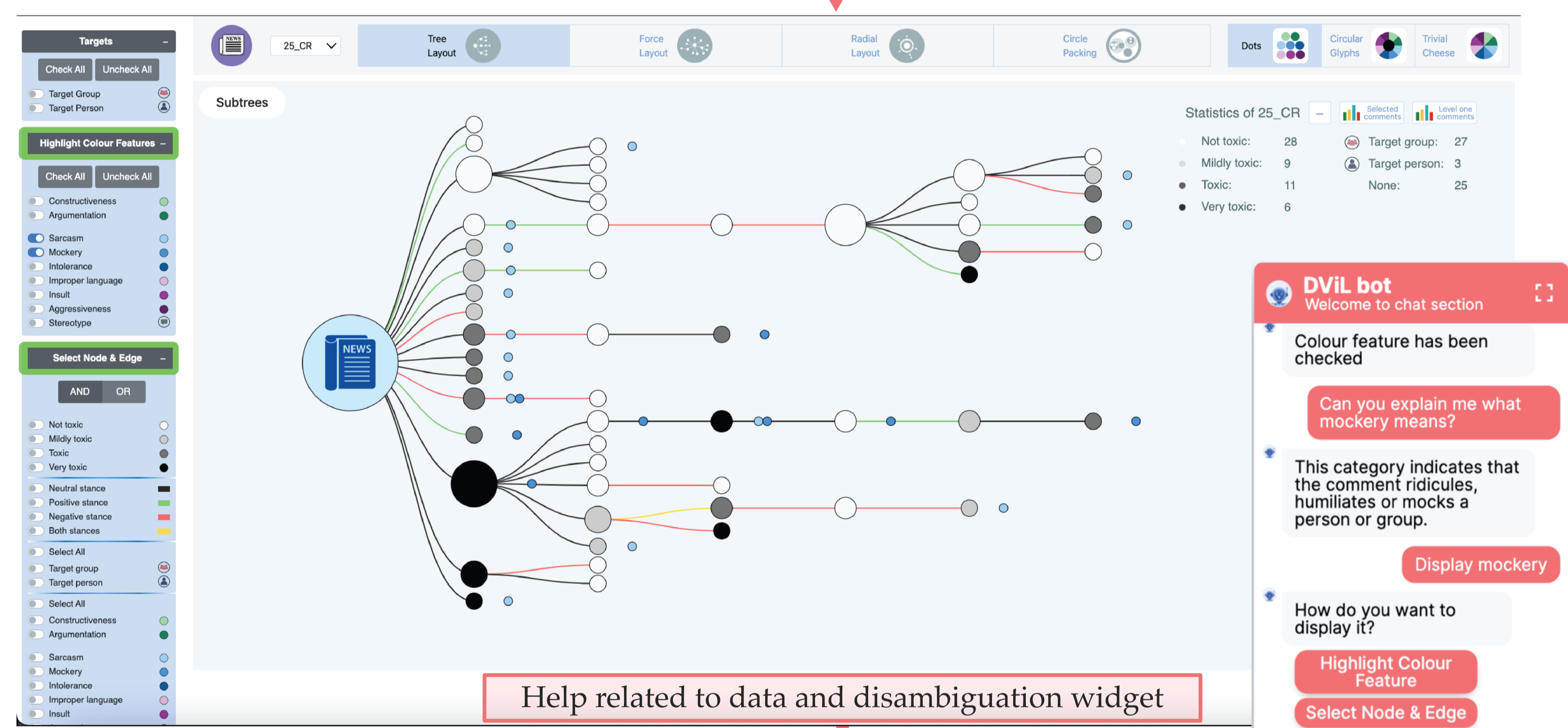
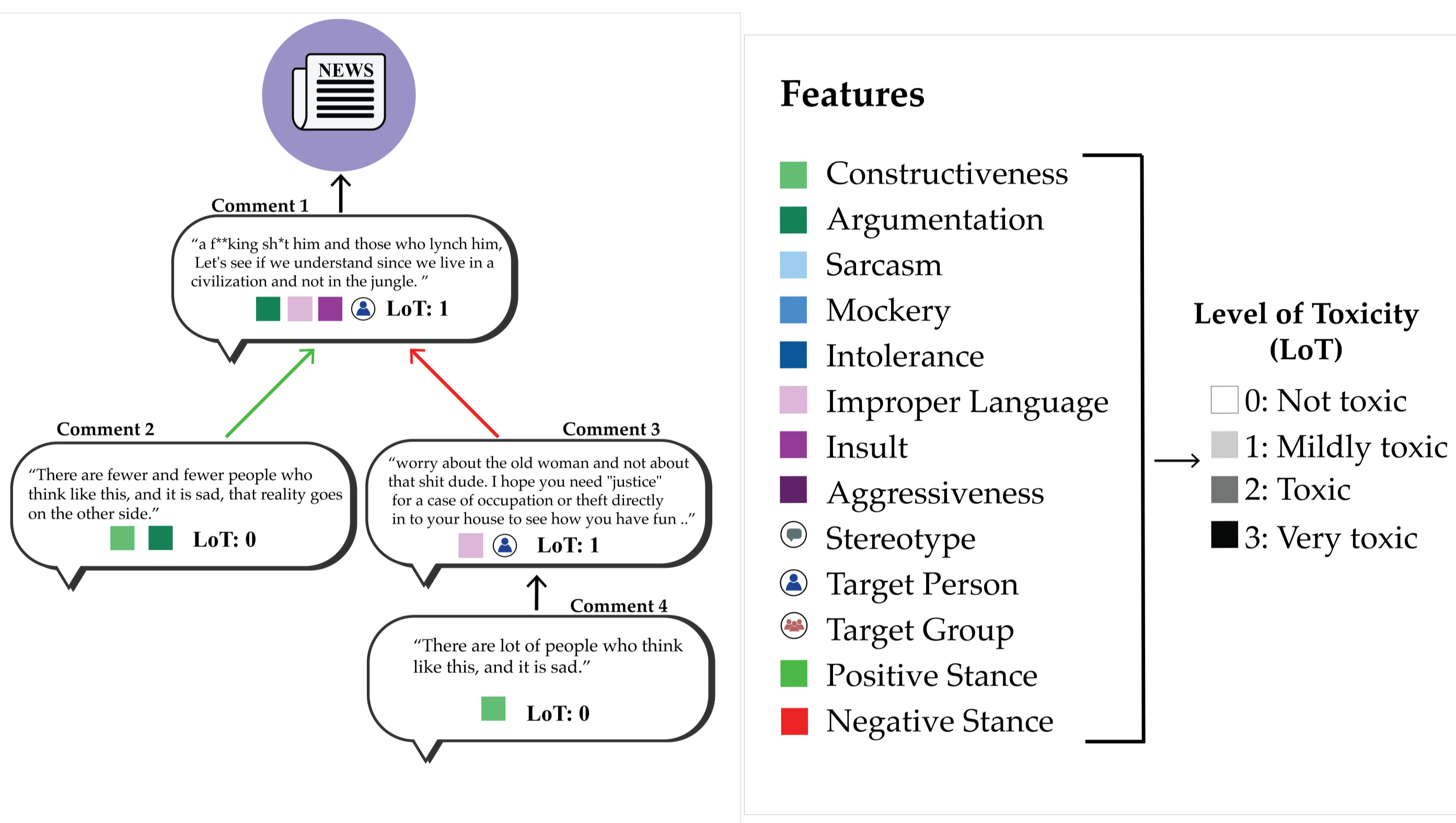
Introduction

Our data visualisation platform integrates both direct manipulation and conversational interaction styles for analysing hierarchical multivariate data. We show a real-life case study for analysing hate speech in online news.



Conversational Data

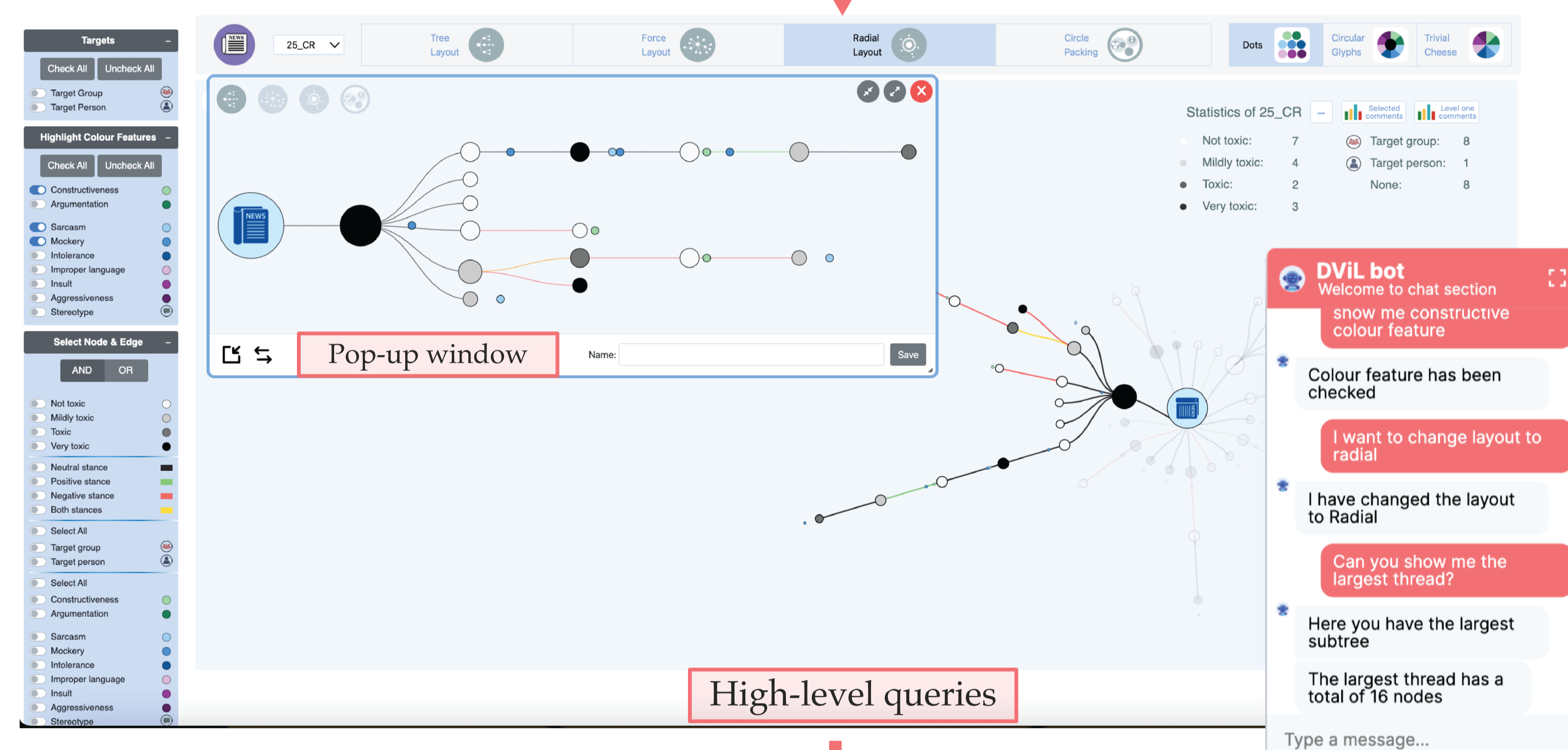
Data is collected from comment sections of online news articles. Comments are written either directly under news article, or under other comments creating a hierarchical structure. All comments are annotated with various features (e.g., insult) to detect level of toxicity (LoT), making data multivariate.



User Interface

Our platform includes different visualisation layouts (tree, radial, force, circle packing), visual elements representing annotated data (dots, glyphs, icons), and complementary charts (histograms, pie). Users can interact with both the chatbot and the WIMP to filter data, change visual mapping, enable/disable glyphs, and visualise statistics.

The chatbot automatically selects the layout according to data's hierarchical structure [KPR*23] and it can perform high-level queries that are not possible with the WIMP. Moreover, the chatbot is able to answer follow-up questions, provide help related to data or tool with additional visual feedback, display disambiguation widgets, and some direct questions.



Interacting with Targets and Glyphs

The chatbot window shows interactions with targets and glyphs, including a request to show target group icons and a response that the target group has been checked.

Help related to tool and direct questions

What's next?

Improve the training data, integrate more high level queries, and conduct a user evaluation with linguists to test the effectiveness of our chatbot.

Acknowledgements

This work was supported by SGR project (2021-SGR00313 CLIC-SGR, funded by the Generalitat de Catalunya, FairTransNLP-Language: Analysing Toxicity and Stereotypes in Language for Unbiased, Fair and Transparent Systems (PID2021-124361OB-C33) funded by Ministerio de Ciencia e Innovación (Spain). And CI-SUSTAIN: Grant PID2019-104156GB-I00 funded by MCIN/AEI/10.13039/501100011033.

Our publications

- [KPR*23] KAVAZ E., PUIG A., RODRÍGUEZ I., CHACÓN R., DE-LA-PAZ D., TORRALBA A., NOFRE M., TAULE M.: Visualisation of hierarchical multivariate data: Categorisation and case study on hate speech. Information Visualization 22, 1 (2023), 31–51.
- [KPR*21] KAVAZ E., PUIG A., RODRIGUEZ I., TAULE M., NOFRE M.: Data visualization for supporting linguists in the analysis of toxic messages

