

2022 Creative Natural Language Generation minutes

The Multi3Generation 2022 Training School was held in Venice, at Ca' Foscari University (19-23 September 2022).

The school was successful, thanks to the support of the local organizer *in loco* (Federico Boschetti), seven trainers, and three invited speakers. Fifteen students from eight countries attended the training school enthusiastically, taking part in discussion slots included in the programme (especially on the last day) but also proactively posing questions and making comments during lessons and invited talks. On the 19, they also discussed collectively small projects that are suitable for short-term scientific missions, an opportunity offered by COST action projects that everyone is interested in seizing.

On the first day (September 19), after registration, the chair, Anabela Barreiro proposed a short introduction about the Multi3Generation project.

The invited talk in the morning by Carlo Strapparava about *Computational Exercises of Creative Language* was the overlapping event between the training school and the MC meeting in the afternoon. Trainees and MC delegates indeed attended it, and it was a good starting point to frame the topic of the school, leaving the audience with enough open questions to be explored in the following days.

On the second day (September 20), Mika Hämäläinen and Khalid Alnajjar, from University of Helsinki & Rootroo Ltd, proposed a gentle introduction to creative NLG. They briefly listed all the available methodologies and discussed the role of hybrid models (rule-based + deep learning) for generating various instances of creative languages (rap lyrics, poems, metaphors). Their lessons included short working group sessions that required discussing hot topics between students. The small groups of students were then invited to report their discussion to the whole audience. Thanks to this exercise, several recurring themes that would also characterize the final debate emerged.

In the afternoon, a hands-on session about metaphor generation made students aware of the multiple ways linguistic resources can be exploited in generation experiments.

The last part of the afternoon was devoted to the second invited talk by Mario Monteleone from University of Salerno about rule-based text generation after formalizing morphosyntactic rules. He introduced NooJ as a framework for rule-based natural language generation that also enables the generation of novels' plots. The discussion after the talk was lively since, in the audience, not everybody was persuaded that a purely rule-based approach is, at the moment, the best approach for natural language generation.

On the third day, a short course about NooJ was given by its creator, Max Silberztein, from Université de Franche-Comté and Cristina Mota, from INESC-ID Lisboa. The lesson specifically explained how to develop NooJ linguistic resources for Natural Language Generation. In the afternoon, a hand-on session on paraphrasing in NooJ made the students aware of the technical aspect of NooJ.

On the fourth day, Federico Boschetti from CNR-ILC & VeDPH and Ioannis Pavlopoulos from Athens University of Economics and Business presented a set of experiments about annotating empathy in texts and proposed an annotation session about the Iliad in the hands-on session. This lesson made students aware of emotions as a crucial aspect of creative language, something that is, at the moment, difficult to handle in automatic generation but central in natural language processing.

In the afternoon, Francesca Frontini from ILC-CNR and CLARIN ERIN introduces CLARIN ERIC, a distributed infrastructure for language resources and technologies and its national node CLARIN-IT. CLARIN indeed supports the training school. CLARIN has, among its mission, to promote the use of research infrastructures in digital humanities research. During the hands-on sessions, the students were invited to explore the resources included at the moment in the CLARIN Virtual Language Observatory and Switchboard.

There was no invited talk at the end of the day because the speaker Alexandra Balahur had some last-minute problems.

On the fifth day, one hour and a half discussion about small project ideas, previously submitted through an interface by students, were exposed in front of the audience. It was a fruitful discussion that

improved the focus of the ideas. Hopefully, some of them will become short Short-Term Scientific Missions for Multi3Generation in the near future. In the morning, a round table about open questions in creative NLG, moderated by Marco Guerini from FBK, was characterized by many additional questions and lively discussions among students and invited speakers. It was clear that, exposed in the previous days to creative NLG open issues, the students could mature their partial but personal views on several of these topics.

The last invited talk by Tony Veale from University College Dublin presented several experiments about multimodal creative NLG and hybrid techniques for the generation of language that seems, at the moment, the most promising avenue for future research in the field.