

Q&A Session Answers

- *Course*: Big Data and Language Technologies
- *Date*: 11.07.2022

Group work

How should it be specified who did what? See <https://www.elsevier.com/authors/policies-and-guidelines/credit-authorstatement> for guidance on writing a contribution statement.

Will there be different grades for different team members depending on who did what?
Yes, we will have to keep that option open.

Experiments

Should statistical measures be reported (confidence, significance)? Yes, but please make sure you know what they mean (esp. with significance).

How many CPUs may we utilize by multiprocessing? As many as is reasonable.

Coding

Is the code graded? Yes.

Should we send the repository while working on it or just in the end? We'll have a look at the git history, so feel free to send it in the end.

Final report

Should we document the technical implementation/coding in the paper? How detailed?

Rough outline: Yes. Also introduce the pipeline, and maybe the used libraries.

Everything else: Probably best in a Readme file in the code repository.

Scientific writing

What should be the structure of the paper? Suggestion: Introduction, Related Work, Data, Experiments, Results, Conclusion (incl. discussion). Deviate if it makes sense.

How many pages? There are conflicting infos on the website and in the slides. See this PDF for the valid formal requirements: <https://temir.org/teaching/big-data-and-language-technologies-ss22/materials/601-project-fair.pdf>