Questionnaire

1. Are you in academia or in industry? [academia, industry, other] (If you have a dual affiliation, please respond with your dominant affiliation in mind.)
2. How many years have you been working in NLG? [options]

Definition:

- We define “error analysis” as a formalised procedure (similar to annotation) in which errors in the output of an NLG system are identified and categorised, after which the frequencies for the different kinds of errors are reported.
- Error analyses are different from “error mentions”, which give an impression of the kinds of errors that are made by an NLG system, but are less formal and don’t quantify the distribution of errors.

Illustration:
[table from a published paper]

3. Have you ever read NLG papers that include an error analysis? [yes, no]
   If yes: did you find the error analysis useful?
      If yes: what made the error analysis useful?
      If not: why not?
   If not: is this surprising to you? Why (not)?

4. Have you ever carried out an error analysis? [yes, no]
   If yes: what did you find difficult about the process? Did you feel like there was a lack of resources/reference material?
   If not:
      (a) have you ever considered carrying out an error analysis?
      (b) what was the reason you didn’t carry out an error analysis?
      (c) are you willing to carry out an error analysis?

5. For what kinds of papers do you think error analyses may be useful? [open question]

6. I would be more likely to carry out an analysis in a conference paper if... [agreement on a scale from 1 to 5]
   - There was a higher page limit.
   - There would be an existing error taxonomy that I could use.

    - There would be dedicated annotation tools for error analysis that I could use.
    - There would be a crowdsourcing template for carrying out error analyses.
    - Reviewers paid more attention to error analyses.
    - There were an available pool of annotators or crowd workers
    - I had more time.
    - I had more money.
    - I had more collaborators.

7. Are there any other factors that would make it more likely for you to carry out an error analysis? [open question]

8. Please indicate whether you agree or disagree with the following statements [agreement on a scale from 1 to 5]:
    - There should be more error analyses in the NLG literature
    - Error analyses are a valuable part of a paper.
    - Carrying out an error analysis is enjoyable.
    - Carrying out an error analysis is boring/tedious.
    - Error analyses are necessary to fully evaluate the performance of an NLG system.
    - Knowing what errors a system makes is helpful for future research.
    - Knowing what errors a system makes is helpful for practitioners/ NLG in industry.
    - If you publish at a **conference**, and you present an NLG system as one of your main contributions, you should include an error analysis.
    - If you publish in a **journal**, and you present an NLG system as one of your main contributions, you should include an error analysis.

9. I am … likely to include an error analysis in a journal article than/as I would be for a conference publication. [more, less, equally]

10. Are there currently enough resources to support error analysis? [yes,no]
    If not: what is still missing?

11. Are there other barriers that prevent you from carrying out an error analysis? [open question]

We believe that it is essential for authors of error analyses to include a table with the distribution of errors in the output of their system. This data should be based on a formalised annotation procedure, with at least two annotators, so that the paper can also report inter-annotator agreement to

gauge the reliability of the analysis.

12. What would you recommend that authors should include in an error analysis? [open question]

13. Is there anything else you would like to comment on? [open question]