**Final report**

The final report describes the research you have done, reporting on the method and results, relating the research to other work in the field, and offering conclusions and directions for future work. The report should be about 2000 words long, not counting the references; reports that substantially exceed this length will be penalized. The structure is similar to the proposal, but with more detail, and two additional sections following the method section.

* Title for the project.
* Names, USC IDs, and USC emails of the team members.
* **Introduction.** Motivate the problem that you have worked on, which involves the use of real-world natural language data. There is no need to motivate Natural Language Processing in general, but rather your specific application. Describe the problem, why it is interesting or challenging, what existing work has been done, and how your contribution relates to that.
* Method.
  + **Materials.** Identify the source data that you use, such as a specific corpus that you accessed or collected yourself. Describe the data in some detail, including the source, the amount of data, what kinds of annotation it has, and what annotations needed to be added (if any).
  + **Procedure.** Describe the *experimental procedure*, that is what methods you use to process the data. This may include algorithms, features, and tools, and any annotations you made. Well-known methods (such as Naive Bayes or Conditional Random Fields) do not need to be explained, but you do need to explain how you use them, for example the features you choose. If you created a tool (such as an annotation or visualization tool), describe it in some detail.
  + **Evaluation.** Describe your method for evaluating the system’s performance, and for evaluating your annotation procedure (if any). Include a description of the specific measures will you took, and the baseline to which you compare your system.
* **Results.** Report how your system performs, and how it compares to the baseline or to other comparable work. Discuss what it gets right, what it gets wrong, and why.
* **Discussion.** Discuss conclusions that can be drawn from the research, implications of your findings, the overall contribution to the general NLP community, and directions for future research.
* References cited. You may choose your preferred style for in-text citations (for example, numerical or author-year) and for the reference listing, but please keep it consistent across the document. The reference listing should contain all the information required for accessing the reference – author(s), year, title, and publication information (such as the conference, journal, volume etc.).
* Division of labor between the teammates.
* Word count for the document, excluding references.

The six main content sections (introduction, materials, procedure, evaluation, results, and discussion) carry equal weight. Therefore, they should be of similar lengths – this means reserving about 300–350 words for each section. This is only a general guideline, as you may find that some sections require more text than others. However, if you find you have more to say than fits within the length requirement, then you’ll need to concentrate on the more important aspects of your project.

When giving examples of text in languages other than English, please use the following multi-line format, to make the examples readable to English speakers. Below is an example for how to present a sentence in Hindi.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **1. The original text in its native script:** | किस | ने | दवाई | को | खरीदा |
| **2. A transcription into Latin script:** | kis | ne | davaaii | ko | khariidaa |
| **3. A word-by-word gloss:** | who | ERG | medicine | ACC | bought |
| **4. A translation into English:** | ‘Who bought the medicine?’ | | | | |

Line 2 is not needed if the language natively uses a version of the latin script. Also, the line numbers and explanations on the left are not needed in the report.