Your team's initial plan and first progress report should be at most two pages long. Your team's second progress report and final report will each be a video you will post on YouTube (share privately with the instructors and TAs).[1] The former can be up to **two** minutes long; the latter can be up to **six** minutes long.

Your initial plan should be one-two pages long, and it needs to contain a discussion of how you are going to approach the project at a high level, and how you are going to divide your effort up. Some examples of things you should discuss are:

- How are you going to use the development data to improve your system?
- Is there going to be a relationship between the asking and answering components of your system, or are you going to implement them independently?
- What tools are you going to use?
- How are you going to share code, data inside your team?
- How are you going to coordinate development inside the team?
- What technical approaches are you going to take?

The final report will include the high-level discussion and analysis of your approach to solving the problem, the motivation for major design choices, and sufficient detail for someone else who took this class to approximately replicate your system. We want to see that you've used what you've learned in class to build your system. As always, your presentation must clearly cite prior work and any tools you used, and acknowledge anyone who helped you.

Code needs to compile (if it's in a compiled language) and run on the `ghc.andrew.cmu.edu` Linux machines.

You will create two programs.

- The "asking" program: `./ask article.txt nquestions`
  where `article.txt` is a text file containing a Wikipedia article and `nquestions` is an integer > 0 telling how many questions to generate. Output is a sequence of `nquestions` English language questions about the content of the article that a human could answer, given the article. Each question should be followed by a newline character. Your goal is to generate questions that are **fluent** and **reasonable**.
- The "answering" program: `./answer article.txt questions.txt`
  where `article.txt` is a text file containing a Wikipedia article and `questions.txt` is a text file containing questions (one per line) about the content of the article. Output is a sequence of answers to the questions, one per line. Your goal is to make the answers **fluent**, **correct**, and **concise**.

You are permitted to use existing, freely available NLP components to help build your system. You must acknowledge these in your reports. However, your system must include a **substantial** novel portion that was implemented by your team. If you're not sure, come ask us.

[1] For last semester's prize-winning video, see `https://www.youtube.com/watch?v=wPQIqiWVDl0`.

You are not permitted to discuss any aspect of your project with **anyone** not on your team, with the exception of the instructors and TA. See the cheating policy.

Grades. The project is 30% of your grade in 11-411. The breakdown is as follows.

Deadline	**What's Graded**	**How Much**
01/09/2017	Initial plan	2 points
TBA	Progress report 1	2 points
TBA	Instructor meeting	2 points
TBA	Progress report 2	2 points
TBA	Dry run system	2 points
TBA	Final working system	6 points
TBA	Final project report	14 points
TBA	Awards	
	Total	**30 points**

Everyone on your team shares a grade out of 30. At the end of the project, you will have the opportunity to anonymously reward team-members that you feel deserve extra credit for extra effort on the project. Note that the generation of questions and answers, as well as evaluations, are folded into homework assignments and graded individually.