**Question/Need:**

* What is the question behind your analysis or model?
  + From the *VizWiz Image Captioning Challenge 2021:* can a computer predict an accurate caption, given an image?
* What practical impact will your work have?
  + People who are blind have technological needs that diverge from the status quo, especially when it comes to photography. By working with a data set composed of images taken by people who are blind, a model can be created to predict captions for such images, which may be of great utility to a user who is blind. Additionally, a model of this sort could theoretically be trained to caption any image (such as one from a satellite) and flag captions with concerning topics, serving as a kind of environmental monitoring system.
* Who is your client and how will that client benefits from exploring this question or building this model/system?
  + I hesitate to say that my client is any computer user who is blind, however it is the subset of computer users who would benefit the most from this model because it could eliminate the need for human-based captioning services, granting a great deal of digital independence.

**Data Description:**

* What dataset(s) do you plan to use, and how will you obtain the data? Please include a link! (The link can be to the dataset you’re downloading, the site you’re scraping, etc.)
  + I am working with the [VizWiz-Captions Dataset](https://vizwiz.org/tasks-and-datasets/image-captioning/)
  + ~23,000 training images + ~117,000 training captions
  + 7,750 validation images + ~38,750 validation captions
  + 8,000 test images + 40,000 test captions
* What is an individual sample/unit of analysis in this project?
  + A single image
* What characteristics/features do you expect to work with? In other words, what are your columns of interest?
  + The image id
  + Captions 1 – 5
  + ‘text\_detected’: set to true for the image if it is set to true for at least three of the five crowd-sourced results and false otherwise
* If modeling, what will you predict as your target?
  + The caption for a given image.

**Tools:**

* How do you intend to meet the tools requirement of the project?
  + Construction of one or more neural networks with **keras**
  + Possible use of visualization libraries if circumstances allow
* Are you planning in advance to need or use additional tools beyond those required?
  + Not at this time! That is always open to change, though.

**MVP Goal:**

* What would a [minimum viable product (MVP)](https://app.thisismetis.com/courses/156/assignments/1526) look like for this project?
  + Construction of a baseline neural network model
  + Baseline performance metrics
  + Sample output predictions of initial model along with the image captioned by the model.