**Project proposal 1: Enable people in remote areas to collect data in exchange for benefit or compensation.** This is in collaboration with a nonprofit organization called Action for Cheetahs in Kenya, enabling local residents in cheetah habitats to collect data about wildlife and environment such as sighting of wildlife, surveys about vegetation, counting of domestic livestock etc. There is an existing opensource app called iNatralist https://www.inaturalist.org/ which is opensource app that’s used for this purpose. And the data collected is open to public download already. This project is to use the existing datasets, and potentially adding additional datasets, then make predictions about cheetah population trend in those areas, such as predicted to decline 2-3% per year, also identify the main contributor for the decline using ML model.

**Problem to solve**: enable people living in remote areas to make profit of the data they have access to, in this case data about wildlife, hence reduce human-wildlife conflict.

**Challenge**: people in remote areas are often not educated, how to design a system that is easy to use, and have control over data quality.

**Data sources**: iNatralist app (open to public download) and may be other data sources.

**Project proposal 2: Trainable quality control system used in manufacturing production line.** For example using pictures of potato chips taken from manufacturing line, to train neural network model to distinguish difference between good and bad chips. Bad chips include chips that are too short, or too dark colored. This can be any manufacturing production line, for items that can be visually detected, potato chips is just an example. Data used can include more than just image, it can also include data from sensors in the productions such as temperature, humidity etc.

**Problem to solve:** using machine learning model to assist human workers in detecting defects.

**Challenge:** finding proper datasets for product quality for one kind of specific product.

**Data sources:** To be determined.

**Project proposal 3: Restaurant popularity prediction based on data from Yelp and Google traffic.** Existing restaurant reviews often consists of old reviews and new reviews, and it is hard to guarantee that a restaurant is good based on all of them, also hard to search for newly opened good restaurant. Using Yelp data and google traffic data, we can make a prediction for which restaurant is trading up, that will cover newly opened or currently popular restaurant, and eliminate no longer popular ones.

**Problem to solve:** surfacing most current restaurant popularity information to people.

**Challenge:** accessing restaurant guest seating rate data.

**Data sources:** Yelp and Google api.