

# Image Captioning

Team Members: Emily Anderson, Omar Elsegeiny, Bohdan Ivanyshyn, Ian King, Mason Myles  
Sponsor: Ovative Group

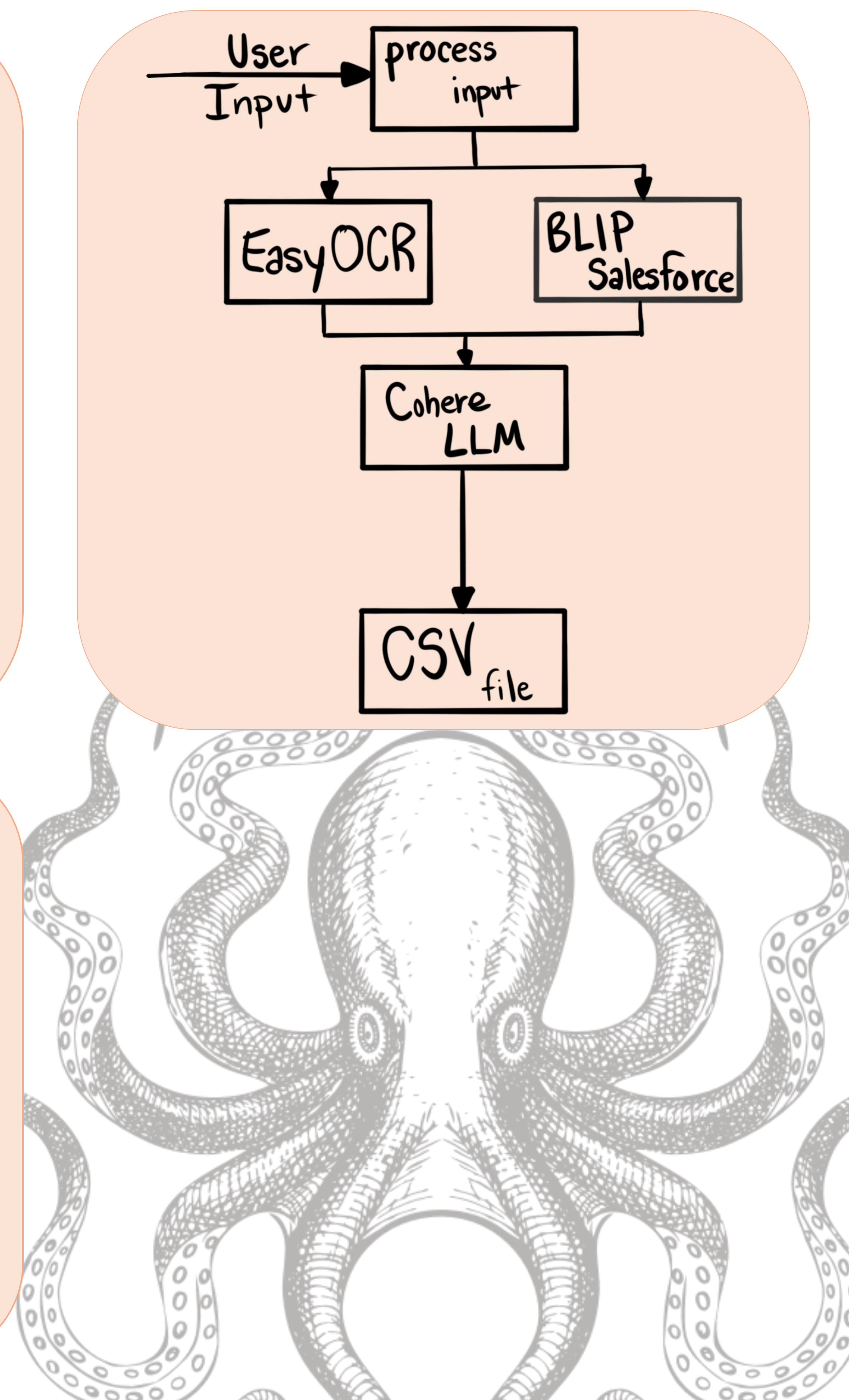
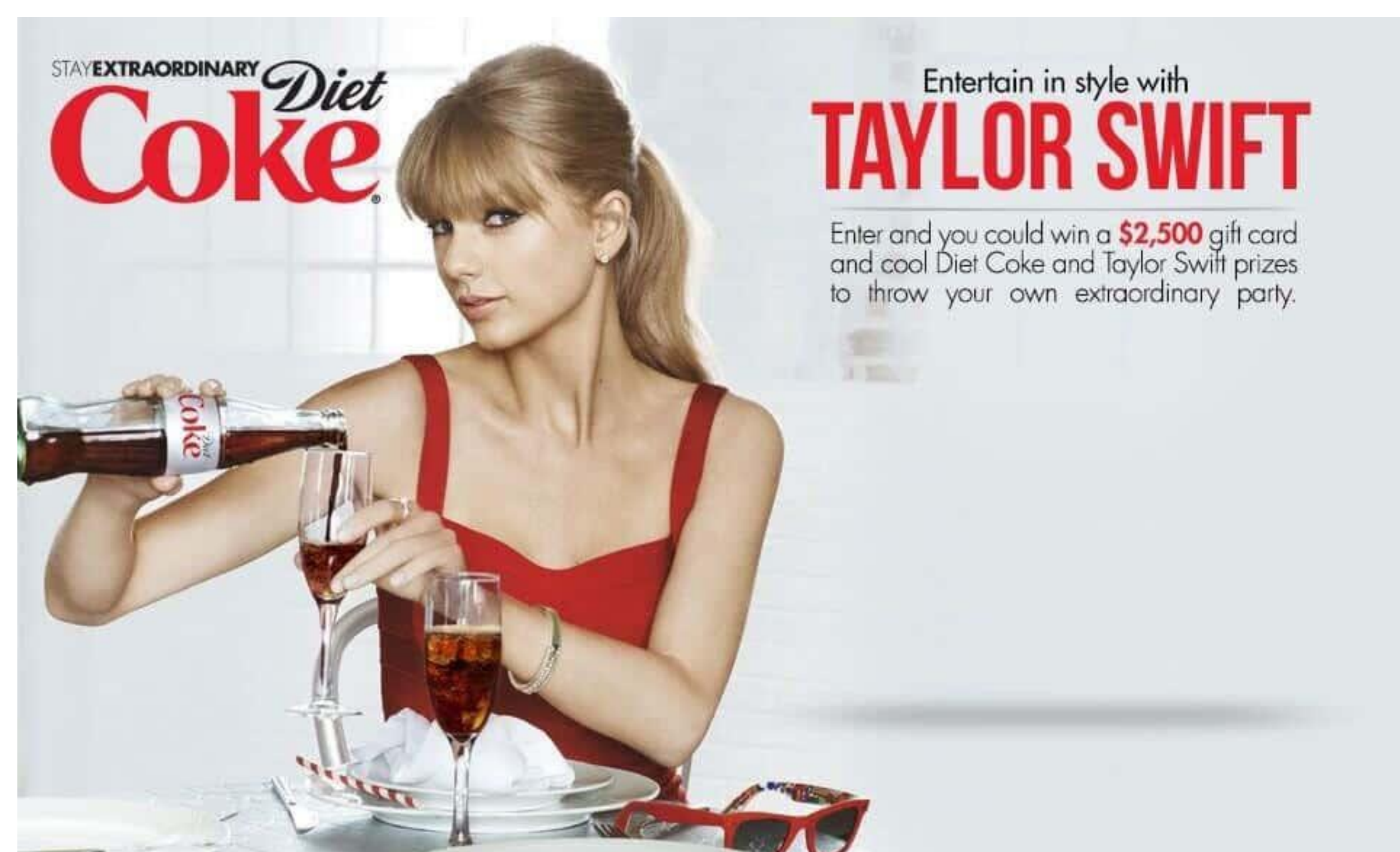


## The Problem

is a digital media and marketing firm that provides analysis of their clients' ads. However, in order to analyze the ads, they need to describe them. This is often time intensive. They came to us with the request to design an Image Captioning program that would allow the advertisements to be analyzed easier.

## Our Solution

A CLI that intakes one or multiple images from a URL(s) or a directory. It then runs the images through several models and returns those captions in a CSV file. This CSV file, once created, then compiles all future captions in one place to allow for easier analysis.



## Output

The advertisement seems to be for a Coke campaign featuring Taylor Swift. The ad seems to mention throwing your own extraordinary party and the chance to win prizes and a \$2,500 gift card.

## Dev Phase 1

- Researched image captioning models
- Created basic terminal application with a single model
- Researched additional methods of captioning to improve accuracy

## Dev Phase 2

- Implemented OCR to add additional captioning context
- Added LLM to compile together captions into a single coherent caption

## Models Used

