ReadMe Instructions

This was created to provide an easy and quick way for researchers to test image preprocessing techniques and visualize the OCR outputs of Textract and Cloud Vision. To see an example of how this works, look at the Example\_Usage.pdf file.

Guides are thoroughly written to make this accessible for beginning users with no coding experience.

For those who do not have Python installed on their machines, an easy way to get started is to download and install Anaconda from <https://www.anaconda.com/download>

For those who do have Python installed, or installed in different ways, you will need to make sure you can use Jupyter Notebooks.

Once you have done this, you need to set up API (Application Programming Interface) credentials for both AWS and Google Cloud. Follow the instructions found in:

* Setup\_AWS\_Root
* Setup\_Google\_Cloud

To ensure that your output works exactly as shown, you can use the virtual environment provided. Follow the instructions found in:

* Setup\_Virtual\_Environment

Finally, you are now ready to use the program on whatever files you may have. Using Jupyter Notebooks open up the OCR\_Python/PreProcess\_OCR.ipynb file and follow along!