**HME-1**

By Sarah Klioris

# Introduction

As someone who used to tutor, I know how repetitive and tiresome it is to correct the same maths equation 30 times. This dataset is primarily aimed at data scientists who want to improve educational resources for teachers. However, other use cases are most welcome.

# HME-1

Hierarchical Mathematical Equations 1—known as HME-1—is an open-source dataset created for use in Computer Vision (CV) and Optical Character Recognition (OCR) algorithms. It is being used to create MathOCR and is being sold on OCEAN protocol’s data marketplace (https://market.oceanprotocol.com/).

# Data formats

* JSON
* DarkNet
* EAST (Efficient and Accurate Scene Text Detector)
* CRNN
* STN-net/SEE

# Predecessor Datasets

The JSON metadata structure of HME-1 is heavily based on the metadata structure of COCO datasets, specifically the COCO-text dataset and the general COCO dataset.

* COCO-text dataset: <https://vision.cornell.edu/se3/coco-text-2/>
* COCO general computer vision dataset: <https://towardsdatascience.com/getting-started-with-coco-dataset-82def99fa0b8>