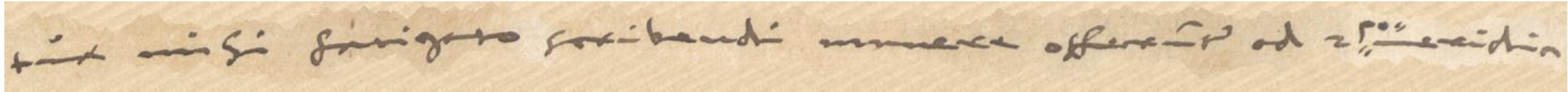# Historical Image Explorers: Jawad Huzaifa, Nimol Thuon, Florent Meyer



- Topic: Hand written recognition for small datasets
- Our model: TrOCR + Binary + RGB + Data augmentations

- Preprocessing: binarize on manual threshold to remove the background
- Generating syntactic datasets
- Recommend ideas for training on small dataset or small machines.
- Additional ideas to try:
  - Find a model pretrained on the dataset languages
    to benefit from a better language model
  - Combine preprocessing with postprocessing