Annotating images

Workflow for diagram detection – April 2023

1. Digitizations of the manuscripts or printed books are sent to the database by the researchers, either as IIIF manifests or PDF files ;
2. A flash message informs the user of the processing time for the extraction of the images and creation of the EiDA manifest ;
3. A new IIIF manifest is generated with the images extracted from the original manifests or PDF files ;
4. The URL of the manifests of the documents we want to extract diagrams from are manually listed in a `.txt` file ;
5. The `.txt` file is transferred to the GPU (dishas-ia) using SCP ;
6. Diagram objects are extracted by the `detect_vhs` script :
   a. The manifest list is taken as an input,
   b. The images are extracted from the manifests by the [iiif-downloader](#) script and saved on the GPU in a directory containing one subdirectory per manifest.
   The name of the subdirectory corresponds to the id of the witness (`ms00`), and the name of the images file corresponds to the number of the view (`0000.jpg`)
   c. The `.jpg` files are taken as input to run inference on by the `detect_vhs` script ;
   d. The output is a `00.txt` (the number being the id of the witness) file for each manifest containing a list of the images and the detected objects for each image in the `x y width height` format.
7. The `00.txt` file is transferred to EiDA using SCP and saved in `vhs/vhs-platform/mediafiles/manuscripts/annotation` ;
8. The users can view their annotated images and edit the annotations..