A Software System to Support Tagging Large Collection of Images

Rayhan Hossain, Sara Mousavi, Dr. Audris Mockus
rhossai2@vols.utk.edu, mousavi@vols.utk.edu, audris@utk.edu

Motivation and Aims

Large image collections are common in many domains. Relevant information is, however, concealed within the images. The content needs to be extracted in order to use such collections in education, research, and business applications.

Aim: to design and implement an extensible, collaborative platform to extract and use domain specific information from images.

Approach

- Implement a system that is capable of:
  - Browsing or searching the collection using metadata
  - Manual tagging of images (extracting content)
  - Validating tags
  - Collaborating (sharing/editing tags and image subsets)
- Support automation of the tagging process [1]

Research Question: How to design a system that supports tagging and supervised learning on large image collections?

Related Work

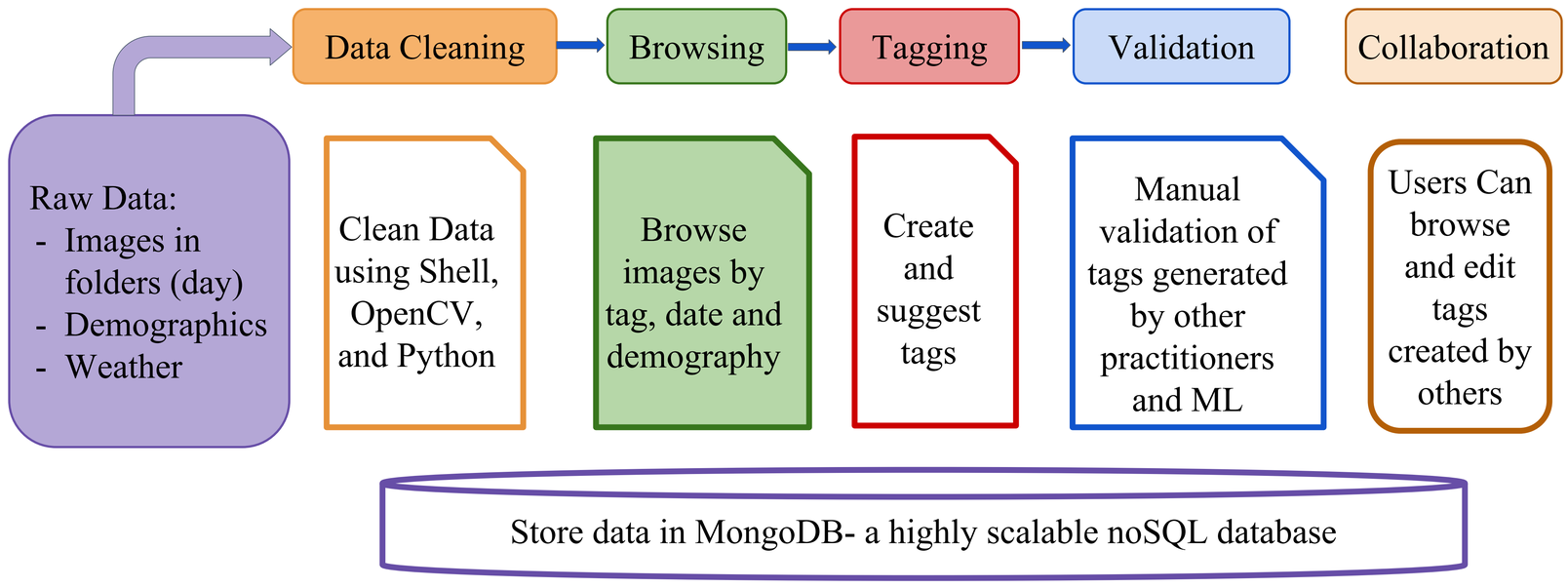
- LabelMe: Manual image annotation tool
- LabelImg: Graphical image annotation tool
- Fluid Annotation: Human-machine collaboration interface for image annotation.
- Prodi.gy: Image annotation and validation interface

Methodology and Prototype Development

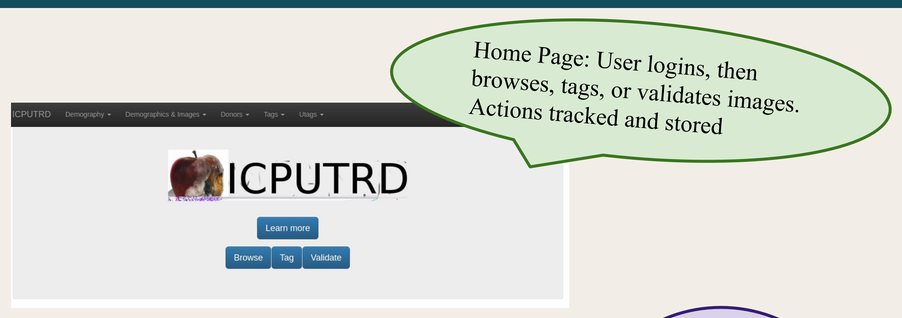
Data

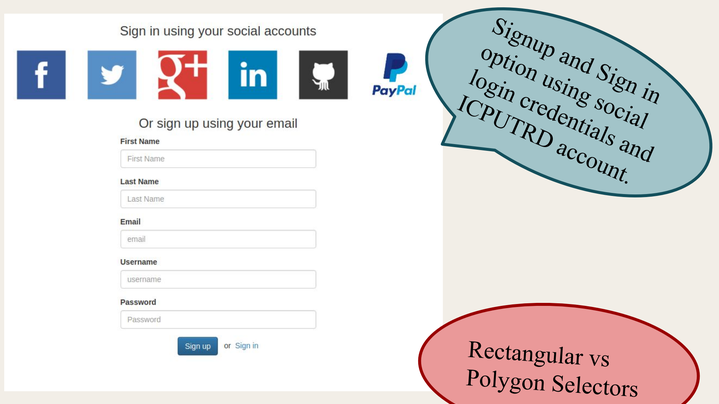
- Data Source: More than one million photos of decomposing bodies at Anthropological Research Facility (ARF) of the University of Tennessee, Knoxville
- Data Size: ~4TB
- Data Description:
  - Daily photos
  - Up to 40/day
  - Taken using hand-held camera

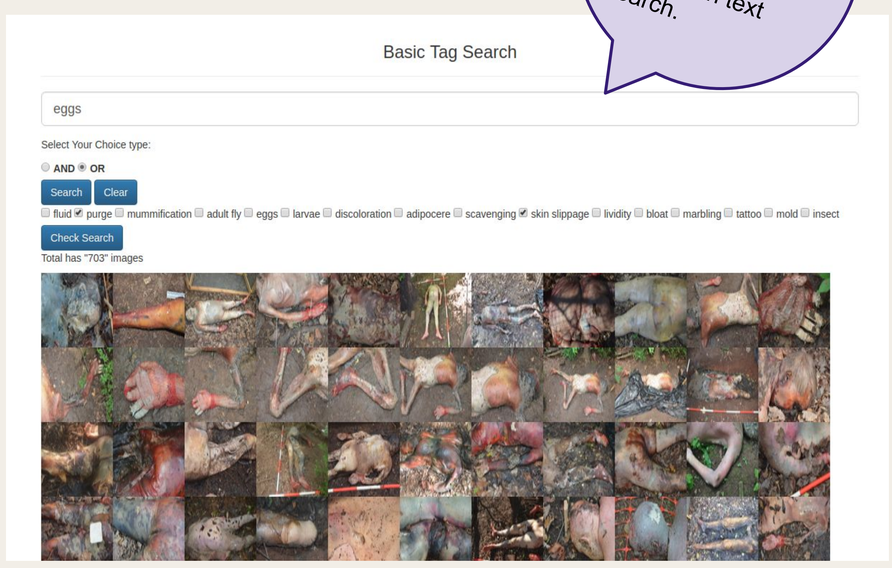
Architecture of the Solution

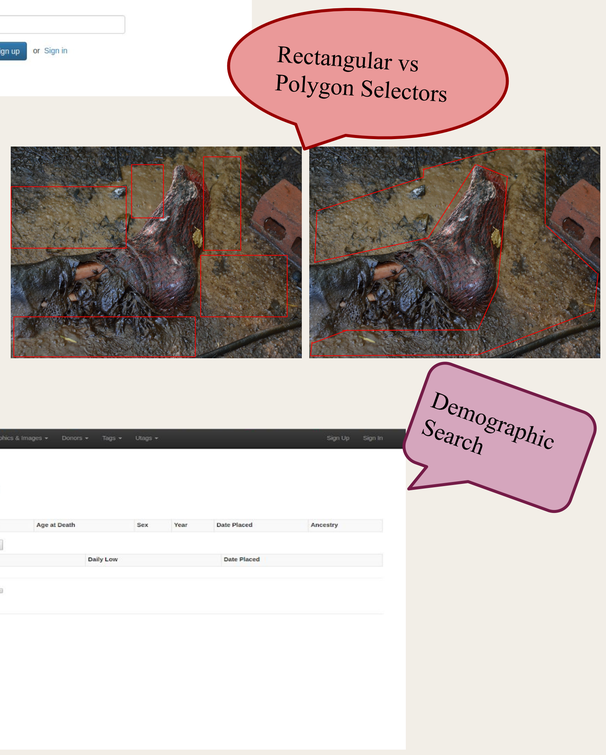


Implementation and Findings

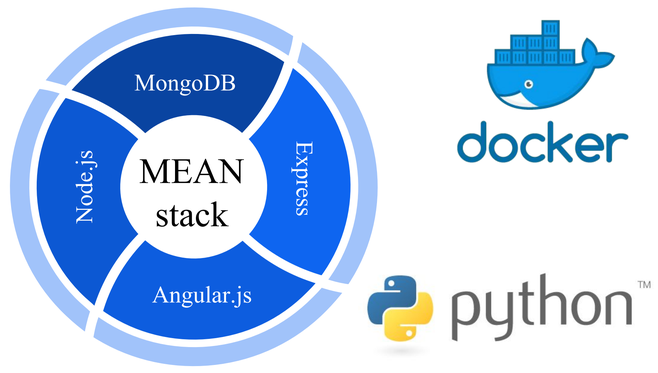








Tools and Technologies



Challenges

- Requirements need to be collected from domain, computer, and ML experts
- What content to extract: nomenclature
- Problematic data: needs cleaning

Future Work

Speed up tagging:
- Use ML to find images likely to contain specific features
- Provide tagging suggestions for an untagged image

THE UNIVERSITY OF TENNESSEE KNOXVILLE

References:
1. Machine-assisted annotation of forensic imagery, Mousavi, Nabati, Kleeschulte, Mockus [2019]
2. https://www.docker.com/
3. http://meanjs.org/