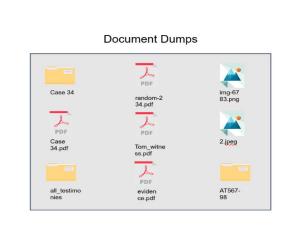
PRESENTER

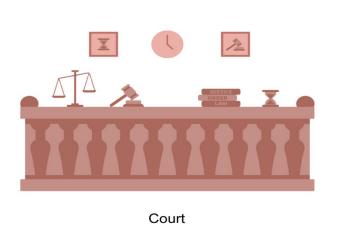
Hellina Hailu Nigatu

Background

Domain experts like journalists and public defenders have to do a lot of manual data cleaning and processing when working with large document dumps. We built a document organization tool and tested three programming paradigms to identify what works for our set of users.





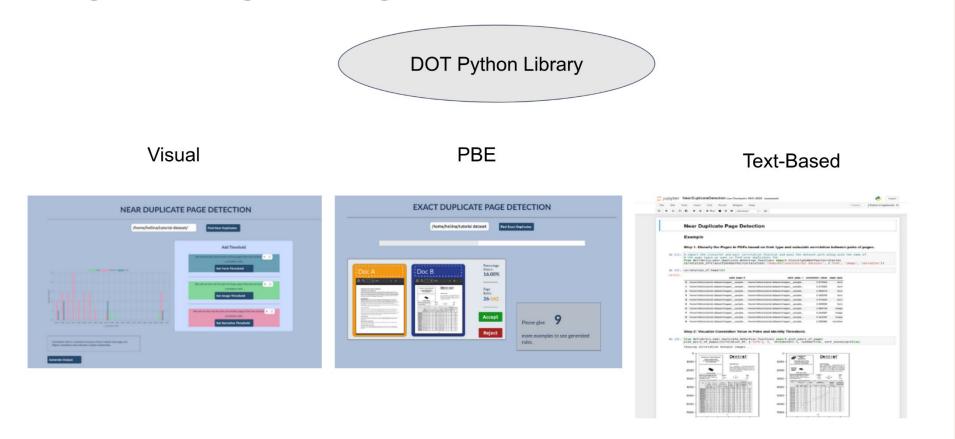


Methods

DOT Library features



Programming Paradigms

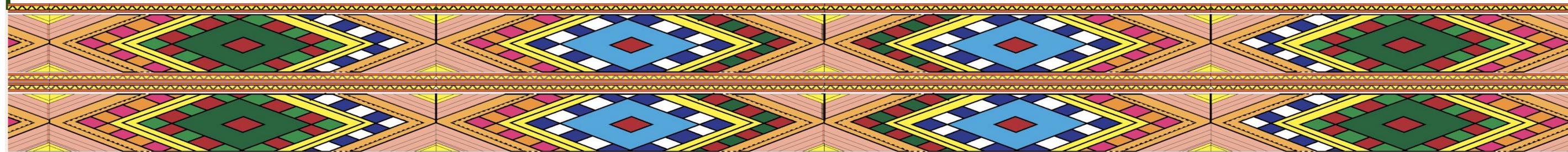


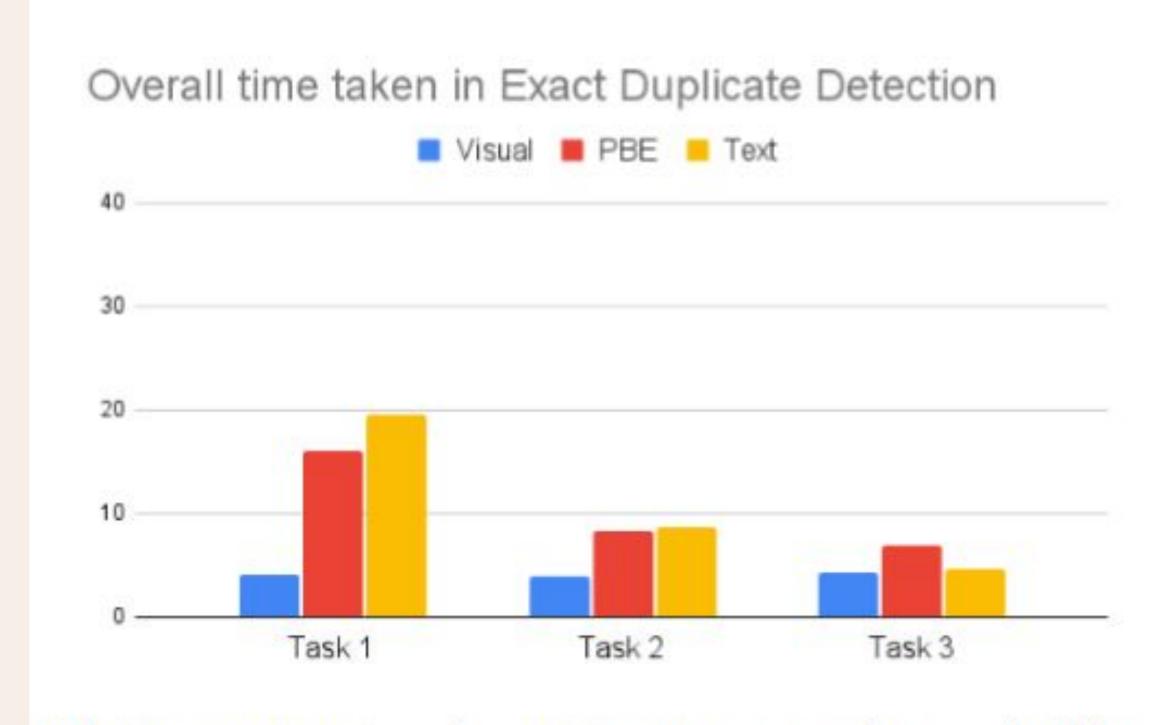
Extra results

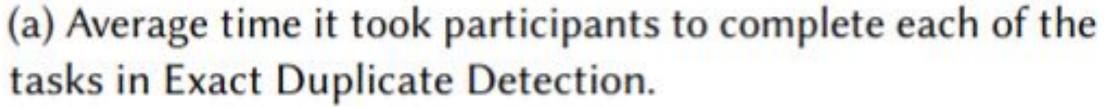
- Text-based gives users low level control and exploration outside of designer provided abstraction.
- Visual gives designers the opportunity to provide information that users could not uncover on their own.
- PBE puts the focus on the data rather than the program structure.

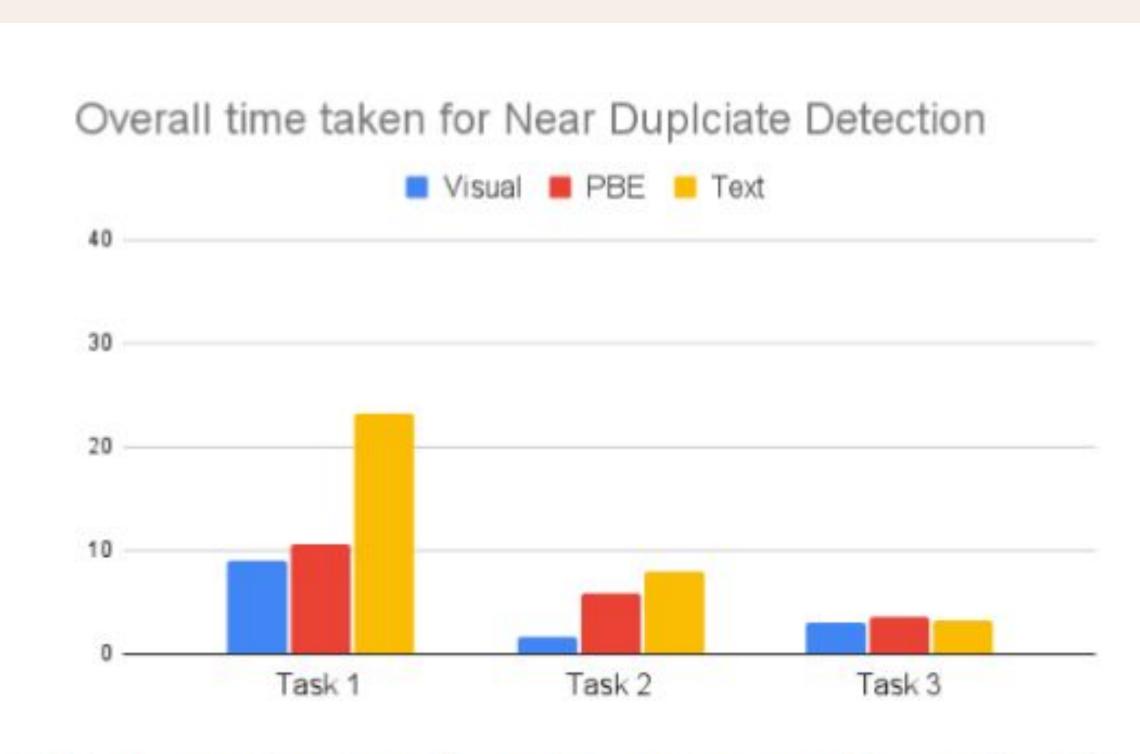


It is possible to make Visual, PBE and Text-Based programming paradigms work relatively equally for domain experts.









(b) Average time it took participants to complete each of the tasks in Near Duplicate Detection.

