# DDD of Document Analysis System



**Document read and clean**
- Remove special chars.
- Read file
- Doc. Read and clean
- Remove newlines
- Convert to lowercase

**Collecting data**
- Get user permission
- Save file for further training
- Train other ML model

**Document Analysis**
- Unique word with frequency
- Text to word
- Analysis
- Calculate stop word %
- Calculate doc. similarity

**Machine learning**
- Train other models
- Evaluate result
- Train other ML model