Offline Corpus Curation Workflow