



## Credo Head of Engineering Assessment

- Time box the work to 2 hours
- It's OK to use AI tools
- Feel free to ask any questions about this in our shared slack channel

Implement a simulated [temporal](#) workflow (we prefer temporal's python sdk and find it good for this use case but it's ok to use any python based orchestration framework you're familiar with such as celery) with dummy activities that should just sleep and log start and finish timestamps and activity name and arguments (in csv file or any other shared log). The workflow should orchestrate communication with database, document storage and 2 REST services: DocumentExchange and TaskTracker.

The workflow should accept a patient ID as an argument and do these things:

- **Get patient first and last name from database**
- Send these requests to DocumentExchange: send **“search documents”** request providing patient first and last name and get job id in response, poll for **“job id status”**, when search job has finished, status request returns url of documents archive, **“download”** it and store extracted files into storage service.
- Send a request to TaskTracker to **“create new entry”** with patient id, first and last name, receive task item ID in response.
- **“Convert”** each downloaded file to pdf and put into storage service.
- If no documents were found, **“update patient status”** in database and TaskTracker item to DOCUMENTS\_NOT\_FOUND, if found - to DOCUMENTS\_FOUND
- **“Generate a report”** with list of converted document names (if found) and **“append”** it to the TaskTracker item.

Design workflow to work within these resource limits:

- DocumentExchange can perform up to 10 “search documents” or “download” requests per MINUTE.
- DocumentExchange “job id status” can be polled up to 10 requests per SECOND.
- Up to 10 requests per second can be made to TaskTracker
- There can be 0-3000 files in documents archive.
- File conversion is single-threaded and each file might take up to 1gb of RAM. The worker has 8 cores and 12gb of RAM.
- Storage service and database can be considered unlimited resource (unlimited request rate and connections).
- DocumentExchange and TaskTracker will occasionally have outages, describe (don't code) how you would handle that

It should be possible to start thousands of instances of such workflow at once and it should complete without exhausting resource limits.