Solutions Engineering Interview

Coding Segment

Early in an engagement, you are asked to prototype what some customer data would look like in a graph. You are given some structured tabular data ( [data-records](https://drive.google.com/open?id=1IHnFPZtnkfQvFQR6zQnRmEuL_uqyprBLCbUsNIOIX1U&authuser=0) ) and some schema-less, but structured document data ( [patient-details.yml](https://drive.google.com/open?id=0B_UFjTNXLHRBRnItUFQxRk9sOGc&authuser=0), [physician-details.yml](https://drive.google.com/open?id=0B_UFjTNXLHRBdDRHekVLVHlCb2s&authuser=0) ).

Your task is to set up some system to get the data into a Neo4J database.

Automate as much or as little as you deem prudent given a timebox of 4-6 hours.

Provide your submission as a github repository or a zipfile. It should contain at least a readme file describing how we would populate a running remote Neo4J database. Any code provided should come with necessary files for building ( e.g. a gradle script for Java ) or running ( e.g. requirements.txt for a python virtual environment. )

The graph schema you adopt is up to you.

Keep in mind that our organization’s pre-sales team often sketches out data in this way, and that they are very likely to add new records to these files. Although less likely, they may also add new tabs to the data-records spreadsheet and corresponding document files.

It is difficult to arrive complete solution to this problem with 6 hours; you will need to decide what the most important aspects of this project are, and outline the feature requests you’d file for future work.