

1. Install NEO4J

At <https://neo4j.com/download/>

Download the neo4j-community-3.5.17 server, and unzip to location on your pc

Start the Neo4j service at \$server/bin

- Neo4j install-service
- Neo4j start

2. Load the document

Move the text file used on midterm to the neo4j/import folder

C:\neo4j-community-3.5.17\import

Load the text file as words into neo4j, with the following transformation

- Load each word as a node
- Build the relationship between words as the sequence of the text
- Merge the same words as distinct node
- Convert to lower case
- Remove the defined stop words
- Remove the punctuation symbols

```
load csv from "file:///2019-Nov-13-LK.OQ-139494797335-Transcript.txt" as row fieldterminator ","
with row
unwind row as text
with reduce(t=tolower(text), delim in [",", ".", "!", "?", "", ":", ";", "'", "-"] | replace(t,delim,"")) as normalized
with [w in split(normalized," ") | trim(w)] as words
with [w in words WHERE NOT w IN ["the","an","on","to"]] as words
unwind range(0,size(words)-2) as idx
MERGE (w1:Word {name:words[idx]})
MERGE (w2:Word {name:words[idx+1]})
MERGE (w1)-[r:NEXT]->(w2)
```

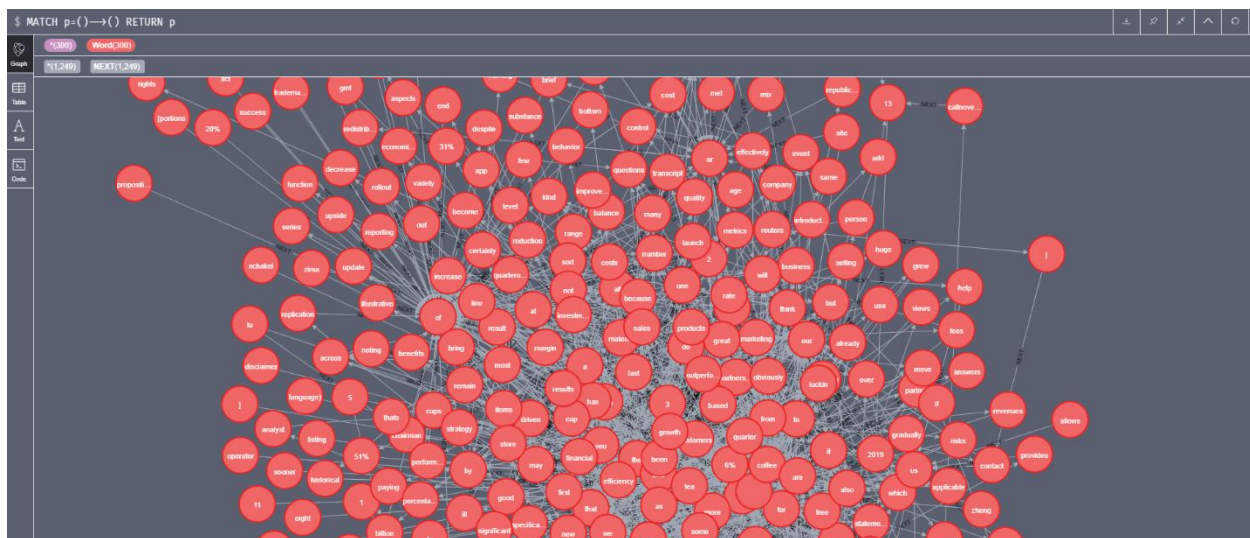
After the run, it created 1191 nodes and 3961 relationships out of 7550 words in the text file. The screenshot of the output is displayed below

```
$ load csv from "file:///2019-Nov-13-LK.0Q-139494797335-Transcript.txt" as row fieldterminator "," with row unwind row as text with r...
```

Added 1191 labels, created 1191 nodes, set 1191 properties, created 3865 relationships, completed after 13030 ms.

Added 1191 labels, created 1191 nodes, set 1191 properties, created 3865 relationships, completed after 13030 ms.

3. Use the graph database, to search the relationship of the words visually



4. Prediction with the key word "risks"

```
MATCH path = (w:Word {name:"risks"})-[:NEXT*..20]->()
RETURN [n in nodes(path) | n.name] as phrase
LIMIT 20;
```

The output of the prediction seems ironic giving that this is based on the 2019 Q3 earning statement, and the recent news about the company scandal was in the recent month.

```
["risks", "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "a", "little", "bit", "further", "decreased", "cny", "21", "billion"]
```

<pre>\$ MATCH path = (w:Word {name:"risks"})-[:NEXT*..20]->() RETURN [n in nodes(path) n.name] as phrase LIMIT 20;</pre>	
Table	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", ""]
Text	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", ""]
Code	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", "a"]
	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", "a", "little"]
	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", "a", "little", "bit"]
	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", "a", "little", "bit", "further"]
	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", "a", "little", "bit", "further", "decreased"]
	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", "a", "little", "bit", "further", "decreased", "cny"]
	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", "a", "little", "bit", "further", "decreased", "cny", "21"]
	[risks, "is", "a", "higher", "frequency", "and", "obviously", "not", "limited", "capital", "markets", "", "a", "little", "bit", "further", "decreased", "cny", "21", "billion"]
Started streaming 20 records in less than 1 ms and completed in less than 1 ms.	

5. Clean up

If you want to keep the neo4j installation, you can clean up the objects with the following statement.

```
match (n:word) optional match (n:word)-[r]-() delete n
```

Or remove the Neo4j from your desktop

You can stop and service and uninstall the Neo4j database

Remove the Neo4j service at \$server/bin

- Neo4j stop
- Neo4j uninstall-service

Remove the Neo4j software folder