Experimentations for KG creation with different LLMs

1. Current affairs article: https://download.oliveboard.in/pdf/December_2023_Bolt.pdf

Banking & Finance

1. RBI Fines Bank Of America, N.A., HDFC Bank

- RBI has imposed a penalty of Rs 10,000 each on Bank of America, N.A. and HDFC Bank Ltd for violation of certain norms.
- The penalty on Bank of America, N.A. has been imposed for violation of Reserve Bank of India's instructions on reporting requirements under Liberalised Remittance Scheme of FEMA 1999.
- Brian Thomas Moynihan, current chairman and CEO of Bank of America.
- The penalty on HDFC Bank is for violation of directions on acceptance of deposits from non-residents.
- Besides, penalties have been imposed on five cooperative banks for non-compliance with various regulatory norms.
- The cooperative banks which have been penalised are The Patliputra Central Cooperative Bank, Bihar; The Balasore Bhadrak Central Co-operative Bank, Odisha; The Dhrangadhra People's Co-operative Bank, Gujarat; Patan Nagarik Sahakari Bank Ltd., Patan, Gujarat; and The Mandal Nagarik Sahakari Bank, Gujarat.
- In all cases, the RBI said the penalties are based on the deficiencies in regulatory compliance and is not intended to pronounce upon the validity of any transaction or agreement entered into by entities with their customers.

2. SBI MF gets RBI nod to buy 9.99% stake in Karur Vysya Bank

- The Reserve Bank has given its approval to SBI Mutual Fund (SBI MF) for acquiring up to 9.99% of the paid-up share capital or voting rights of Karur Vysya Bank Ltd (KVB).
- If SBI MF fails to acquire major shareholding within a period of one year from the date of RBI's letter, this approval shall stand cancelled.
- Besides, SBI MF must ensure that the aggregate holding in the bank does not not exceed 9.99% at all times.
- Further, if the aggregate holding falls below 5%, prior approval of RBI will be required to increase it to 5% or more of the paid-up share capital or voting rights of KVB.
- KVB, meanwhile, announced the opening of four branches, of which three are in Tamil Nadu and one in Bengaluru. With this, the total number of branches has increased to 831.

3. Bank lending to pvt corporate rises 14.9 pc in Sep: RBI data

- Bank lending to the private corporate sector grew 14.9 per cent in September 2023 from 14.7 per cent a year ago, according to the Reserve Bank data.
- Another set of banking data released by the central bank showed the share of term deposits bearing a 6 to 8 per cent interest rate rose to 78.6 per cent in September 2023 against 12.5 per cent in March 2022.
- According to data on outstanding credit of banks at September-end 2023, loans to industry accounted for nearly one-fourth of total bank credit; they increased by 8.6 per cent (year-on-year) in September 2023 (against 12.3 per cent a year ago); working capital loans growth has remained in double digits for the last six quarters.

Experimentation Results

LLM	File name	Page number	Node count	Relation count
OpenAI-GPT-3.5-turbo-16k	../data/December_2023_Bolt.pdf	3	28	29
Diffbot	../data/December_2023_Bolt.pdf	3	0	0
OpenAI-GPT-4	../data/December_2023_Bolt.pdf	3	27	15
Human	../data/December_2023_Bolt.pdf	3	22	25

Human Interpretation

Triplets:

[RBI → imposed penalty on → Bank of America,

RBI → imposed penalty on → HDFC bank Ltd,

Bank of America -> penalty of -> Rs 10000,

Bank of America -> violation of -> Liberalised Remittance Scheme of FEMA,

HDFC bank Ltd -> penalty of -> Rs 10000,

Brian Thomas → is Ceo/Chairman of → Bank of America,

RBI → imposed penalty on → The Patliputra Central Cooperative Bank,

RBI -> imposed penalty on -> The Balasore Bhadrak Central Co-operative Bank,

RBI -> imposed penalty on -> The Dhrangadhra People's Cooperative Bank,

RBI -> imposed penalty on -> Patan Nagarik Sahakari Bank Ltd,

RBI -> imposed penalty on -> The Mandal Nagarik Sahakari Bank,

The Balasore Bhadrak Central Co-operative Bank->Located in ->Odisha,

Patliputra Central Cooperative Bank-> Located in -> Bihar,

The Dhrangadhra People's Cooperative Bank -> Located in -> Gujarat,

Patan Nagarik Sahakari Bank Ltd -> Located in -> Gujarat,

The Mandal Nagarik Sahakari Bank -> Located in ->Gujarat,

The Balasore Bhadrak Central Co-operative Bank->violation of ->regulatory compliance,

Patliputra Central Cooperative Bank-> violation of -> regulatory compliance,

The Dhrangadhra People's Cooperative Bank -> violation of ->regulatory compliance,

Patan Nagarik Sahakari Bank Ltd -> violation of -> regulatory compliance,

The Mandal Nagarik Sahakari Bank -> violation of ->regulatory compliance,

RBI → Approves → SBI Mutual Fund,

SBI mutual fund → Intends to acquire -> KVB,

KVB → announced branch in → Bengaluru,

KVB → announced branch in → Tamil Nadu]

Relation:

[imposed penalty on, imposed penalty on, imposed penalty on, imposed penalty on, imposed penalty on, imposed penalty on, imposed penalty on, penalty of,penalty of,is Ceo/Chairman of,Approves,announced branch in, announced branch in, Intends to acquire,violation of,violation of,violation of,violation of,violation of,violation of,violation of,located in,located in,located in,located in,located in

]

Nodes:

[RBI, SBI Mutual Fund, non-residents, Karur Vysya Bank, Hdfc Bank, Bank of America, Central Bank, National stock exchange of India, Interglobe Aviation, Patliputra Central Cooperative Bank, The Balasore Bhadrak Central Cooperative Bank, The Dhrangadhra Cooperative Bank, Patan Nagarik Sahakari Bank, The Mondal Nagarik Sahakari Bank, Liberalised Remittance Scheme of FEMA, regulatory compliance, Brian Thomas Moynihan, Bihar, Orissa, Gujarat, Tamil Nadu, Bengalure, Website]

KG creation Open AI GPT 3.5 analysis
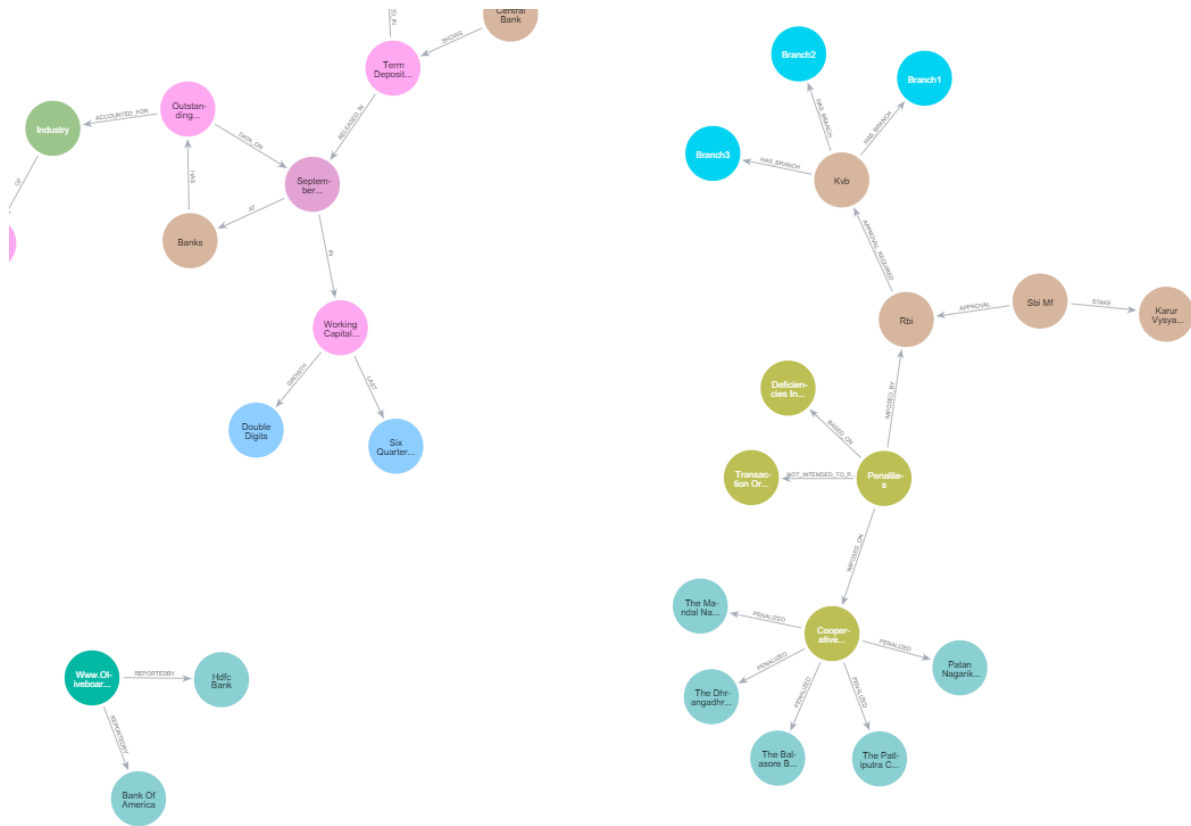
Nodes:

['Central Bank', 'The Balasore Bhadrak Central Co-Operative Bank, Odisha', 'Branch3', 'September 2023', 'Term Deposits', 'Working Capital Loans', '""The Dhrangadhra People'S Co-Operative Bank, Gujarat""', 'Rbi', 'The Mandal Nagarik Sahakari Bank, Gujarat', 'Hdfc Bank', 'Banks', 'Non-Residents', 'Kvb', 'Branch1', 'Karur Vysya Bank', 'Bank Of America', 'Deposits', 'March 2022', 'Industry', 'Www.Oliveboard.In', 'Patan Nagarik Sahakari Bank Ltd., Patan, Gujarat', 'Outstanding Credit', 'Penalties', 'The Patliputra Central Cooperative Bank, Bihar', 'Total Bank Credit', 'Branch2', 'Cooperative Banks', 'Sbi Mf']
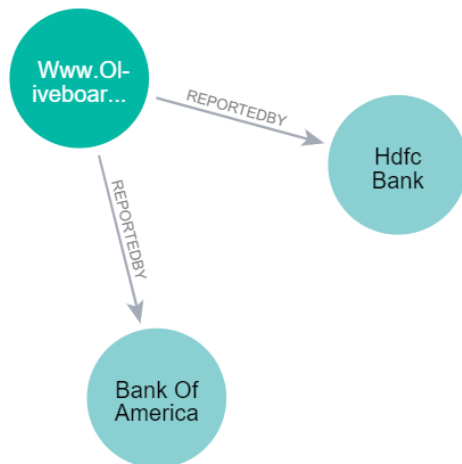
Relationships:

['reportedBy', 'reportedBy', 'from', 'imposed on', 'penalized', 'penalized', 'penalized', 'penalized', 'penalized', 'imposed by', 'based on', 'not intended to pronounce upon', 'approval', 'stake', 'approval_required', 'has_branch', 'has_branch', 'has_branch', 'shows', 'released in', 'released in', 'has', 'data on', 'at', 'accounted for', 'of', 'in', 'growth', 'last']

Observations:

1. Gpt-3.5-turbo-16k was unable to understand that the location of banks should be a separate node.

For example: 'The Mandal Nagarik Sahakari Bank is a financial institution so this should be a separate entity and 'Gujarat' is the location of the bank. So, two separate nodes should have been created here. But instead it created one node 'The Mandal Nagarik Sahakari Bank, Gujarat'.

2. Some nodes and relationships generated were unnecessary. For example:

Also, It missed the actual relationship for the nodes 'HDFC bank' and 'Bank of America'

RBI → imposed penalty on → Bank of America,

RBI → imposed penalty on → HDFC bank Ltd

3. Contrary to my expectations the LLM did a good job in understanding the short forms. It did not create separate nodes for SBI MF and SBI mutual funds, same for RBI which is Reserve bank of india.

KG creation Open AI GPT 4 analysis
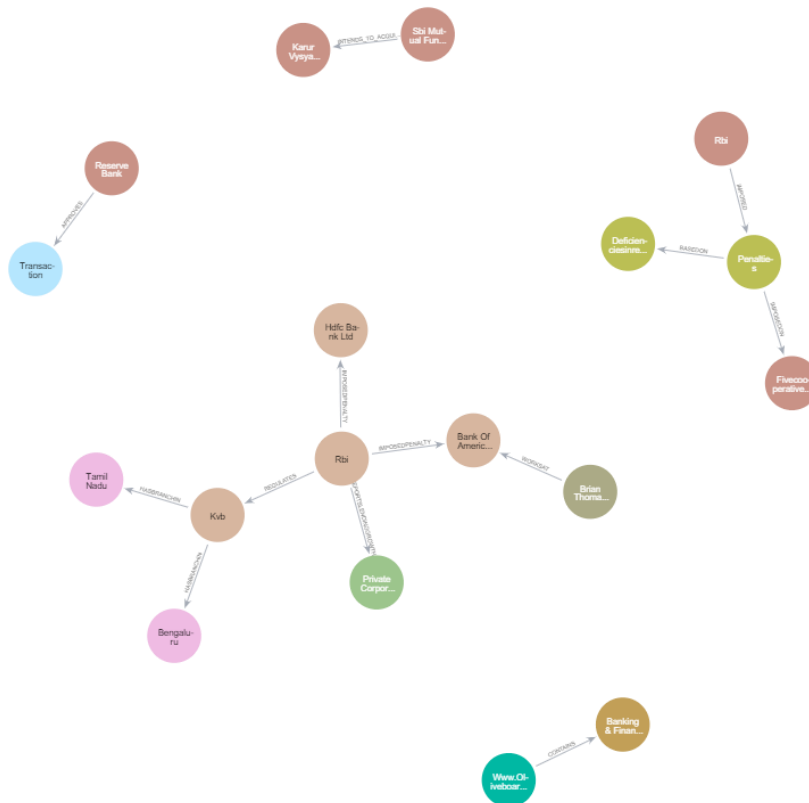
Nodes:

['Transaction', 'Thepatliputracentralcooperativebankbihar', 'Themandalnagariksahakaribankgujarat', 'Reserve Bank', 'Www.Oliveboard.In', 'Karur Vysya Bank Ltd', 'Nonresidents', 'Tamil Nadu', 'Bengaluru', 'Thedhrangadhrapeoplescooperativebankgujarat', 'Patannagariksahakaribankltdpatangujarat', 'Centralbank', 'Kvb', 'Rbi', 'Private Corporate Sector', 'Hdfc Bank Ltd', 'Brian Thomas Moynihan', 'Bank Of America, N.A.', 'Sbi Mutual Fund', 'Termdeposits', 'Banking & Finance', 'Penalties', 'Outstandingcredit', 'Deficienciesinregulatorycompliance', 'Fivecooperativebanks', 'Workingcapitalloans', 'Thebalasorebhadrakcentralcooperativebankodisha']"

Relationships:

['contains', 'imposedPenalty', 'imposedPenalty', 'worksAt', 'imposed', 'imposedOn', 'basedOn', 'intends to acquire stake', 'approves', 'regulates', 'hasBranchIn', 'hasBranchIn', 'reportsLendingGrowth', 'releasedData', 'releasedData']

Observations:

1. Gpt 4 is removing space in some of the nodes like 'Thepatliputracentralcooperativebank'. For others it isn't doing the same. So inconsistency can be observed.
2. Does a better job than gpt 3.5-turbo-16k at finding relationships. For example,
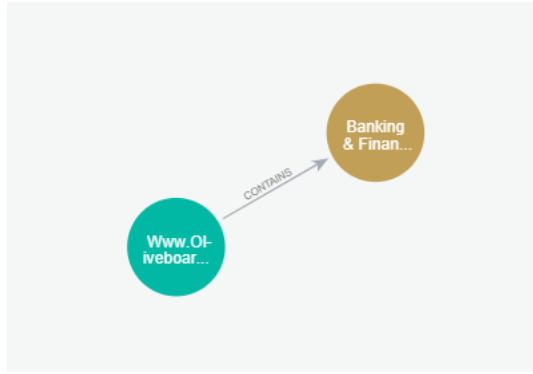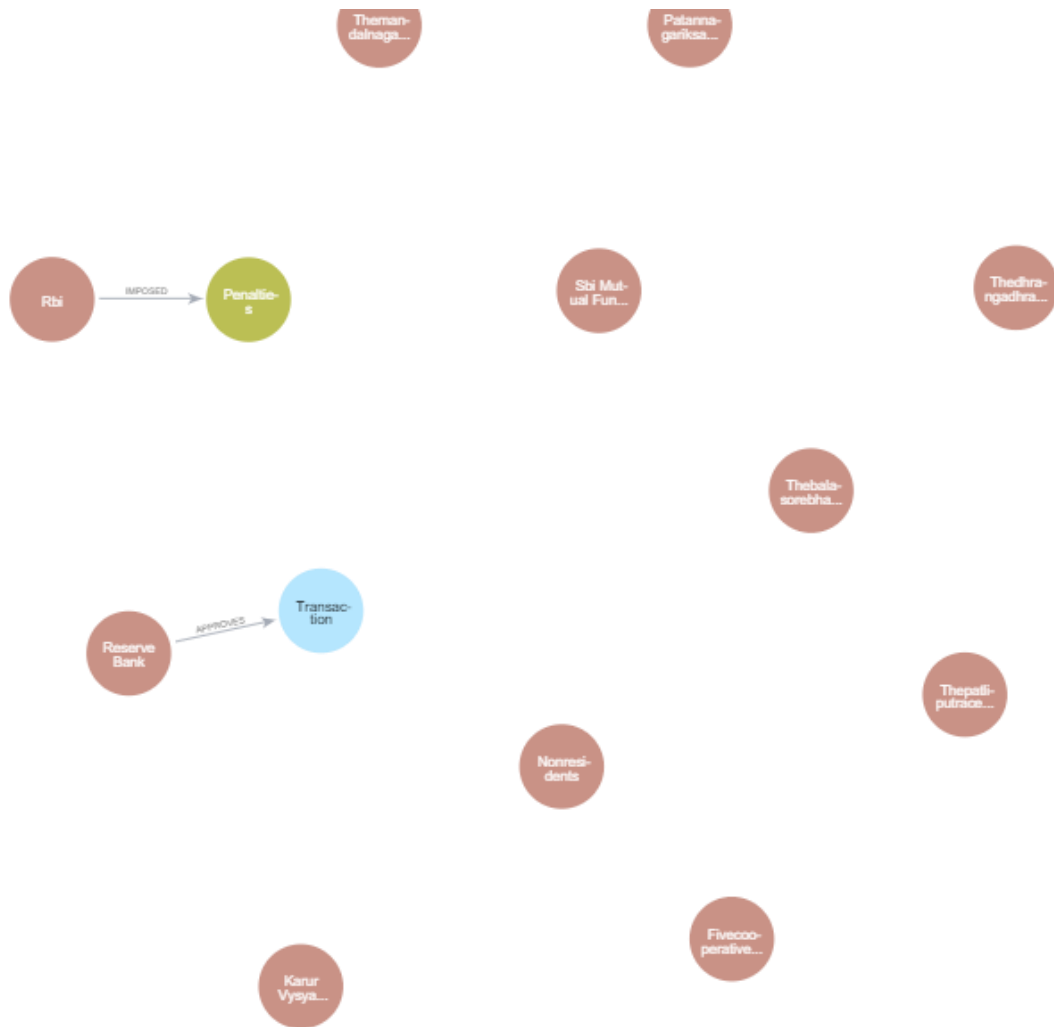
3. Gpt 4 did a better job at identifying the financial concepts. For example,

4.

The llm was not able to make relationships for some nodes like the bank names such as 'Thepatliputracentralcooperativebankbihar', 'Themandalnagariksahakaribankgujarat'.

Theman-
dalnaga...

Patanna-
gariksa...

Rbi —IMPOSED→ Penaltie-
s

Sbi Mut-
ual Fun...

Thedhra-
ngadhra...

Thebala-
sorebha...

Transac-
tion

Reserve
Bank —APPROVES→

Thepatti-
putrace...

Nonresi-
dents

Fivecoo-
perative...

Karur
Vysya...

5. Gpt 4 is able to identify the locations better than gpt-3.5-turbo. Gpt 4 has generalized it better. For example,
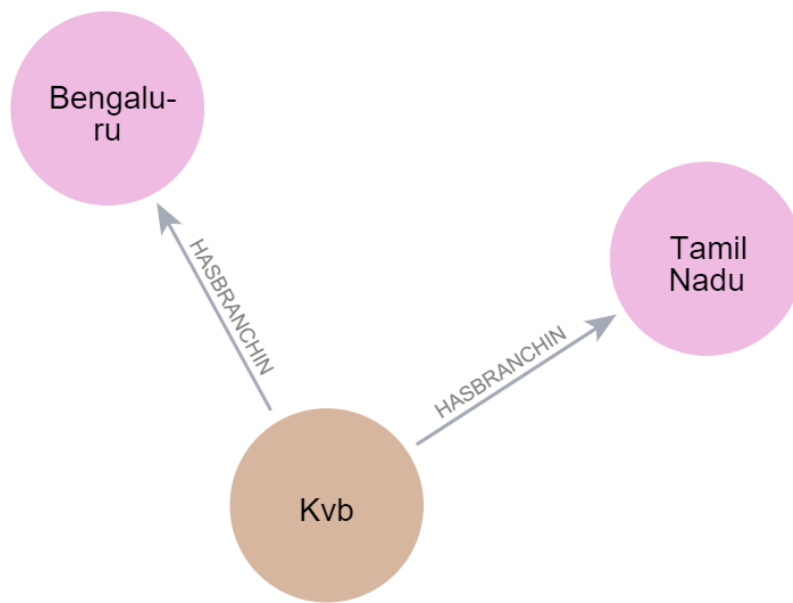
6. Gpt 4 was able to understand that RBI imposes penalties on 5 cooperative banks but it couldn't map the banks' names.

Conclusion

The purpose of this experimentation was to evaluate the performance of knowledge graph creation by LLMS like Diffbot, OpenAI GPT-3.5-turbo-16k and GPT-4.

While Diffbot by langchain couldn't extract any nodes and relationships in the sample pdf page, Open AI gpt 4 did a great job at identifying nodes and relationships. With better prompt engineering gpt-4 results can be improved upto an extent. While gpt-3.5 extracted more nodes and relationships in comparison to gpt-4, but gpt-4 did a better job at creating meaningful results.