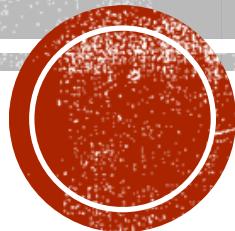
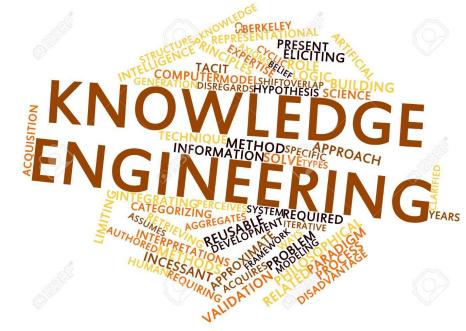
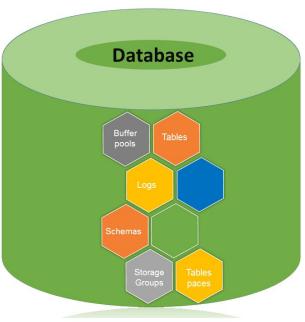


基于可视化的知识 图谱查询

linwang@SIF100 2019

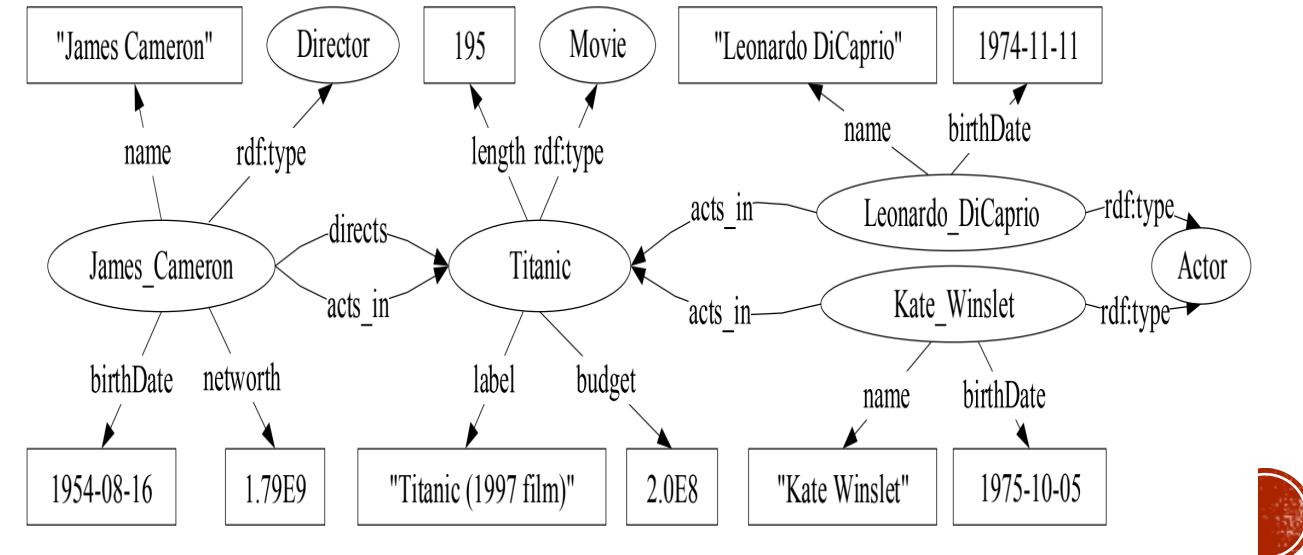
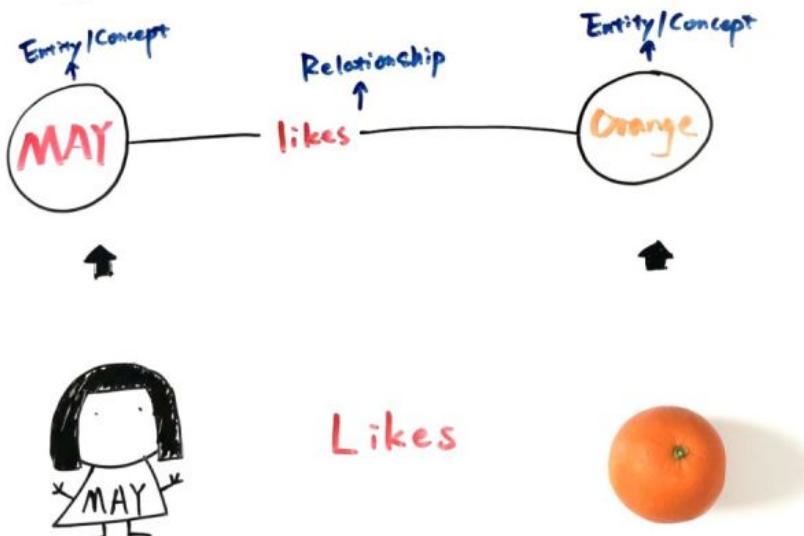


KNOWLEDGE GRAPH HISTORY

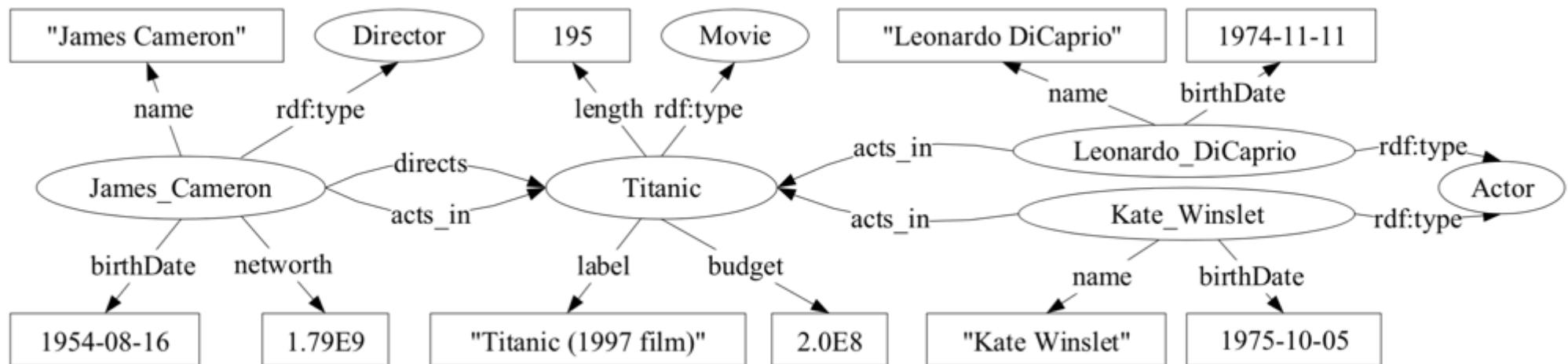


KNOWLEDGE GRAPH

- Knowledge Graph is a knowledge base used by Google to enhance its search engine's search results with semantic-search information gathered from a wide variety of sources.
- 知识图谱 2012 年 5 月由 Google 正式提出，其初衷是为了提高搜索引擎的能力，改善用户的搜索质量及体验。
- 知识图谱可看作是一张巨大的图，节点表示实体或概念，边则由属性或关系构成

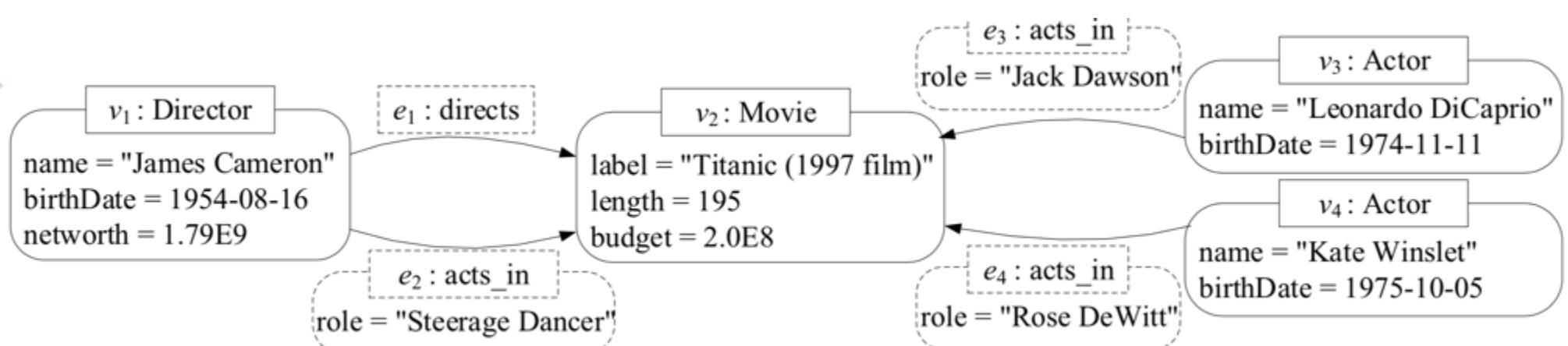


知识图谱数据模型



RDF
图模
型

属性
图模
型



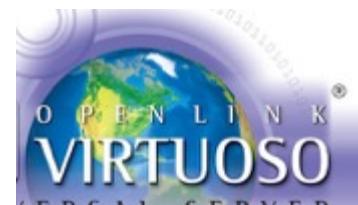
RDF DATABASE



rdf4j



gStore



Virtuoso



AllegroGraph
Franz Inc.

AllegroGraph



GraphDB



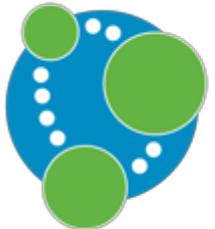
BlazeGraph



Stardog



ORIGINAL GRAPH DATABASE



neo4j

Neo4j



JanusGraph

JanusGraph



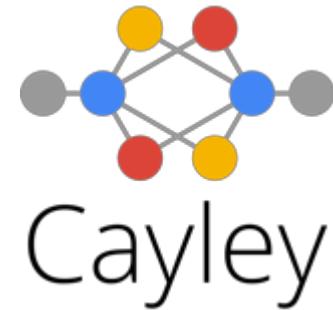
OrientDB®

OrientDB



NebulaGraph

NebulaGraph



Cayley

Cayley



KNOWLEDGE GRAPH QUERY

- 基于形式化语言的查询 (**SPARQL/Cypher**)
- 基于可视化交互的查询
- 基于自然语言的查询 (智能问答)

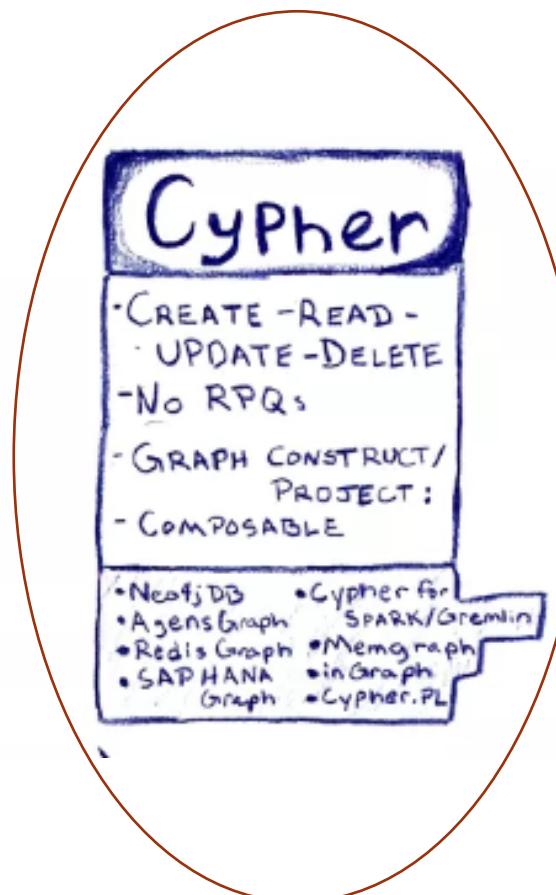
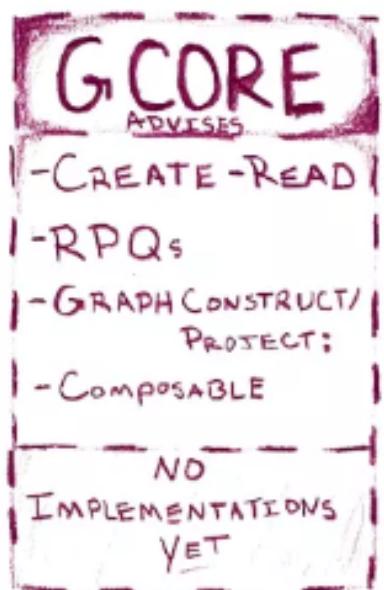
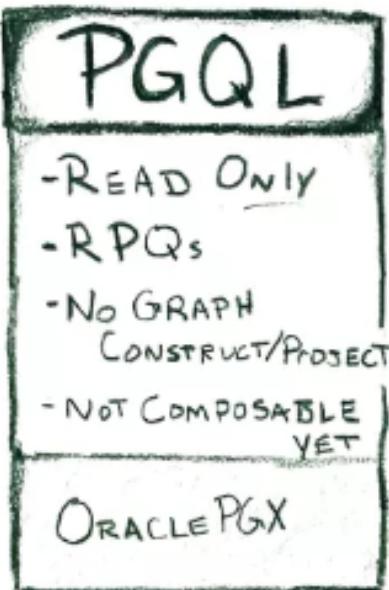


KNOWLEDGE GRAPH QUERY

- 基于形式化语言的查询 (**SPARQL/Cypher**)
- 基于可视化交互的查询
- 基于自然语言的查询 (智能问答)



QUERY LANGUAGE



Gremlin
 $G = (V, E)$



LARGE-SCALE KNOWLEDGE GRAPH

名称	规模	
DBpedia	英文:4百万个实体, 5亿个RDF三元组 685个概念, 2,795个属性 125种语言	
YAGO	1千万个实体, 1.2亿个RDF三元组 35万个概念, 100个属性	
Freebase	4千万个实体, 10亿个RDF三元组 1.5万个概念, 4,000个属性	
Google	6亿个实体, 35亿条RDF三元组	
	Wikidata, Wolfram Alpha, CMU NELL, Probase	



QUERY ON DBPEDIA

Virtuoso SPARQL Query Editor

[About](#) | [Namespace Prefixes](#) | [Inference rules](#) | [RDF views](#) | [iSPARQL](#)

Default Data Set Name (Graph IRI)

Query Text

```
select ?obj
where{
<http://dbpedia.org/resource/Karl_Marx> <http://dbpedia.org/ontology/influencedBy> ?obj.
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see [details](#).)

Results Format:

Execution timeout: milliseconds (values less than 1000 are ignored)

Options:

- Strict checking of void variables
- Log debug info at the end of output (has no effect on some queries and output formats)
- Generate SPARQL compilation report (instead of executing the query)

(The result can only be sent back to browser, not saved on the server, see [details](#))



SPARQL-CASE STUDY

SPARQL ENDPOINT CONTENT TYPE (SELECT) CONTENT TYPE (GRAPH)

/DBPedia_Person/query JSON Turtle

```
1 ▼ PREFIX res: <http://dbpedia.org/resource/>
2 prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 prefix owl: <http://www.w3.org/2002/07/owl#>
4
5 SELECT ?subject
6
7 ▼ WHERE {
8
9   ?subject <http://xmlns.com/foaf/0.1/gender> "female"@en
10
11 }
12
13 limit 25
14
```

◀ ▶ ⏪ ⏩

QUERY RESULTS

Table Raw Response

```
1 ▼ {
2 ▼   "head": {
3 ▼     "vars": [ "subject" ]
4   },
5 ▼   "results": {
```

SPARQL-CASE STUDY

PREFIXES

rdf rdfs owl xsd +

SPARQL ENDPOINT /DBPedia_Person/query CONTENT TYPE (SELECT) JSON CONTENTS Turtle

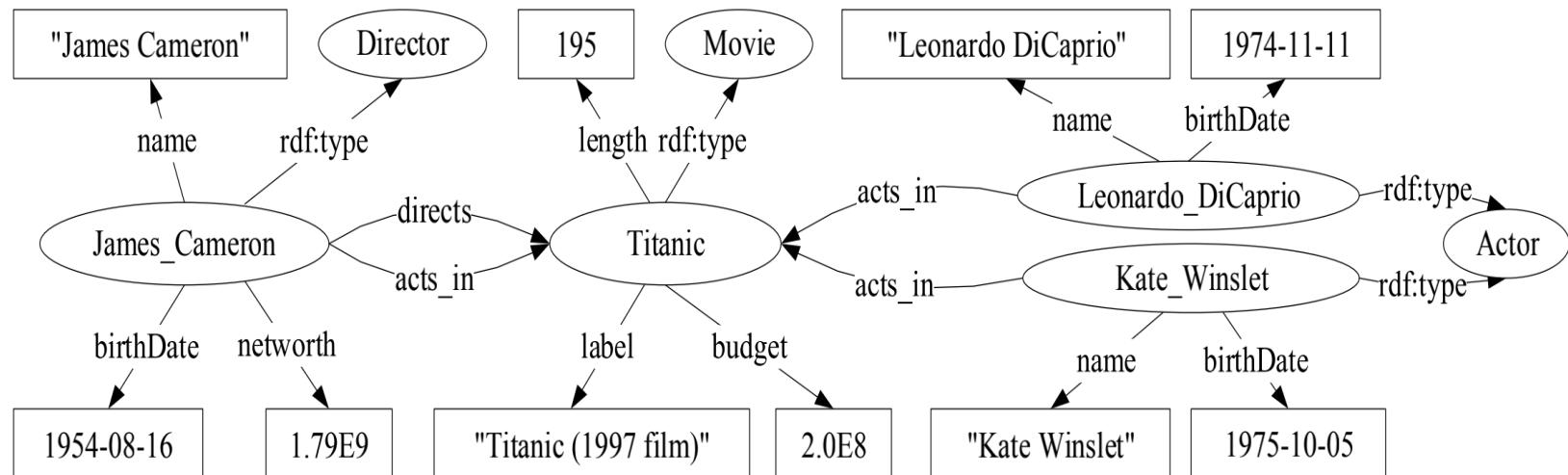
```
1 PREFIX res: <http://dbpedia.org/resource/>
2
3 SELECT ?object
4
5 WHERE {
6   res:Holly_Brooks <http://purl.org/dc/terms/description> ?object
7 }
8 LIMIT 25
9
```

QUERY RESULTS Table Raw Response

A photograph of Holly Brooks, a cross-country skier, standing on a snowy slope. She is wearing a blue ski suit with 'APU' printed on the leg, a white helmet with 'APU' and 'SWIX' logos, and is holding a pair of yellow and black skis and red ski poles. In the background, there are other people and a wooden building.

SPARQL-MORE

我想知道1950年之后出生的资产大于1.0E9的导演执导的电影的出演演员！



```
SELECT ?x4 ?x5
WHERE { ?x1 rdf:type Director . ?x1 birthDate ?x2 . FILTER (?x2 >= 1950-01-01) .
?x1 networth ?x3 . FILTER (?x3 > 1.0E9) .
?x1 directs ?x4 . ?x5 acts_in ?x4 . }
```

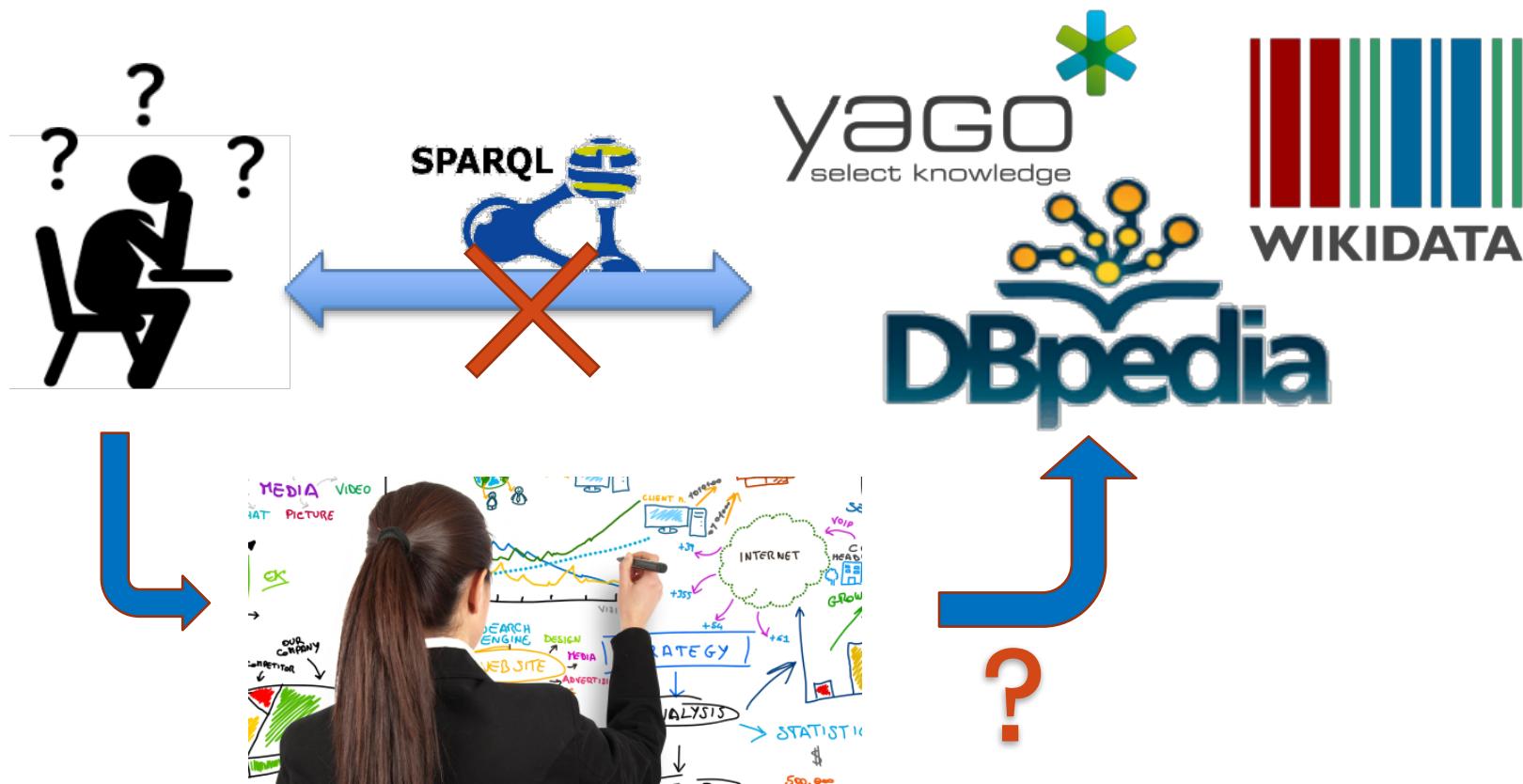


KNOWLEDGE GRAPH QUERY

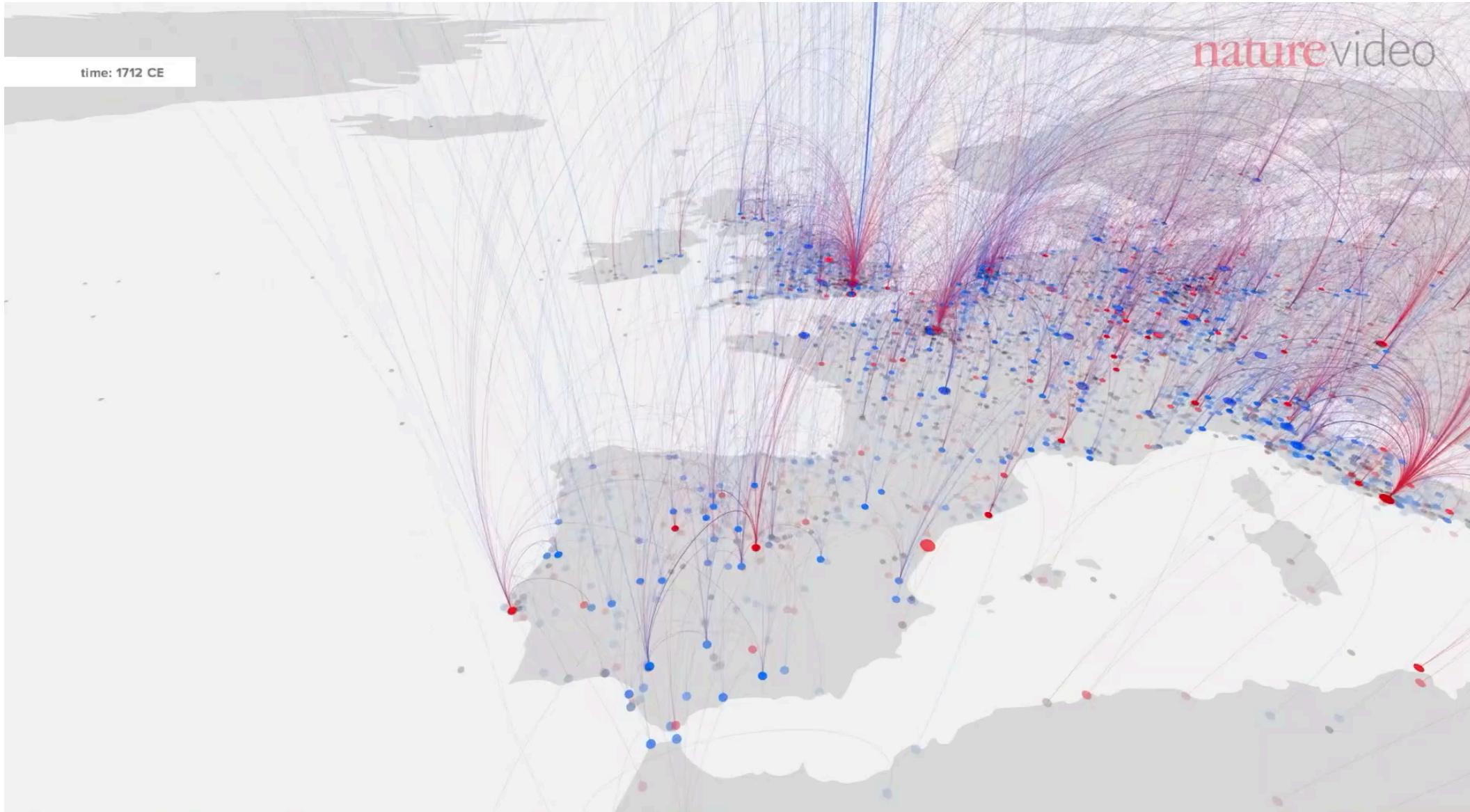
- 基于形式化语言的查询 (**SPARQL/Cypher**)
- 基于可视化交互的查询
- 基于自然语言的查询 (智能问答)



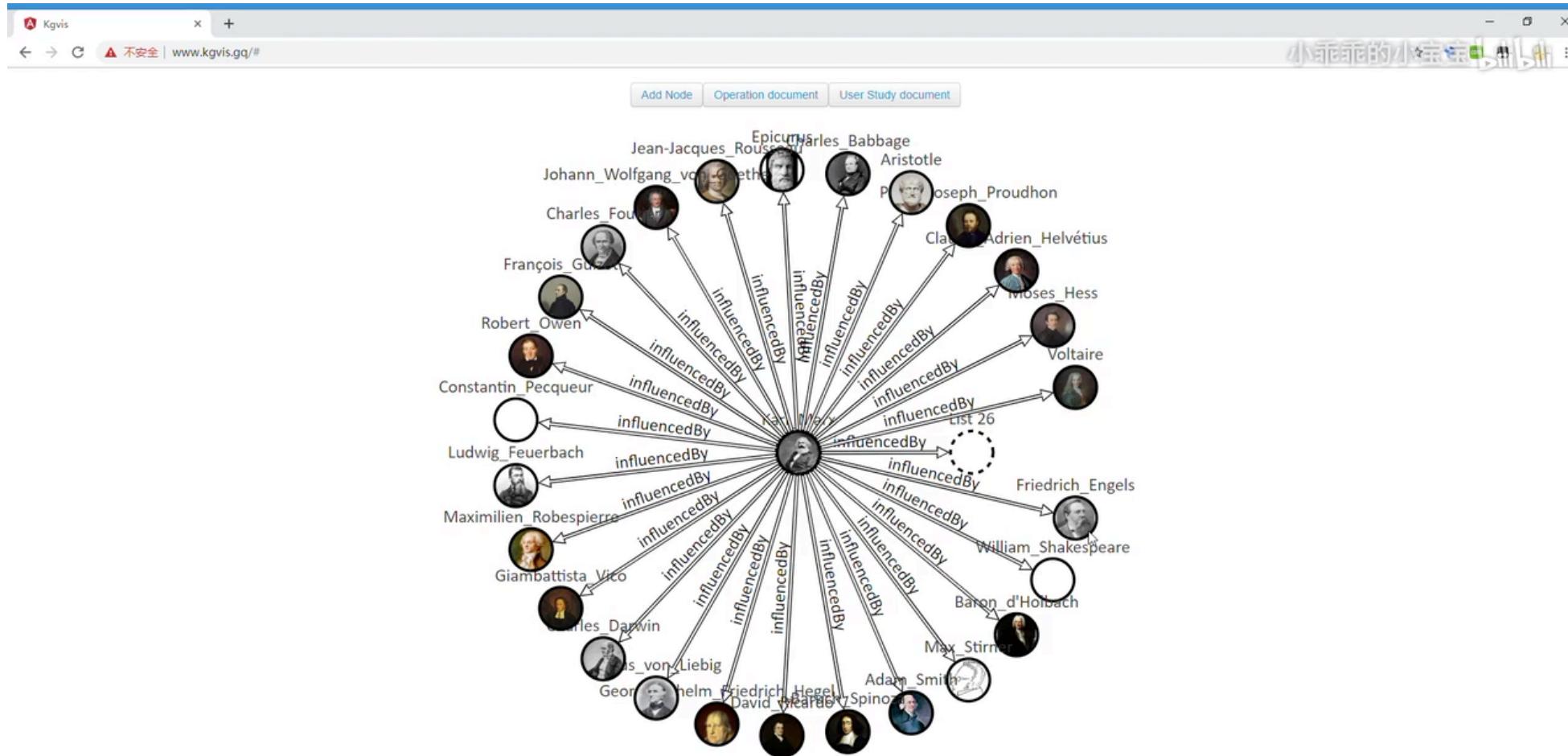
VISUALIZATION QUERY



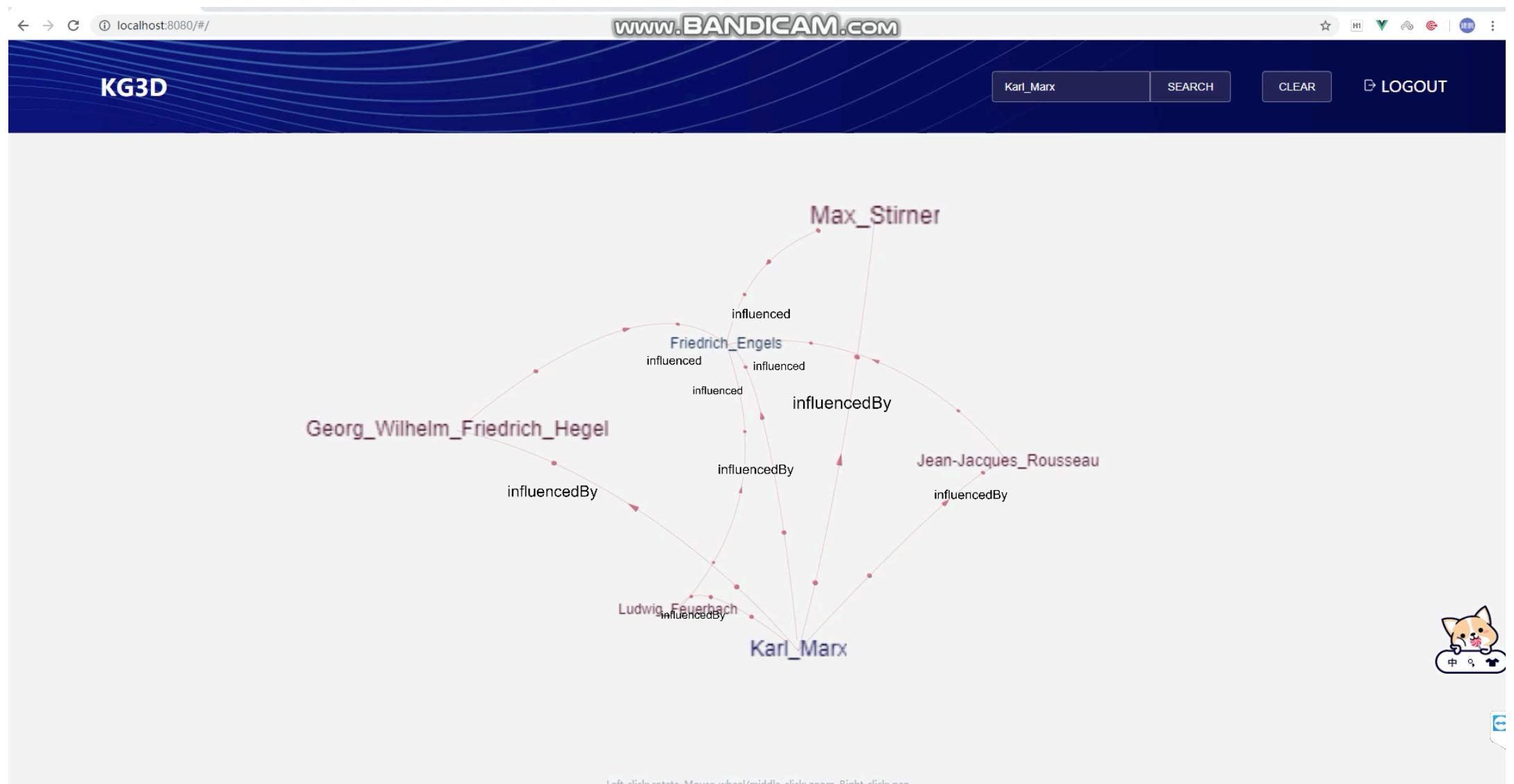
DATA VISUALIZATION



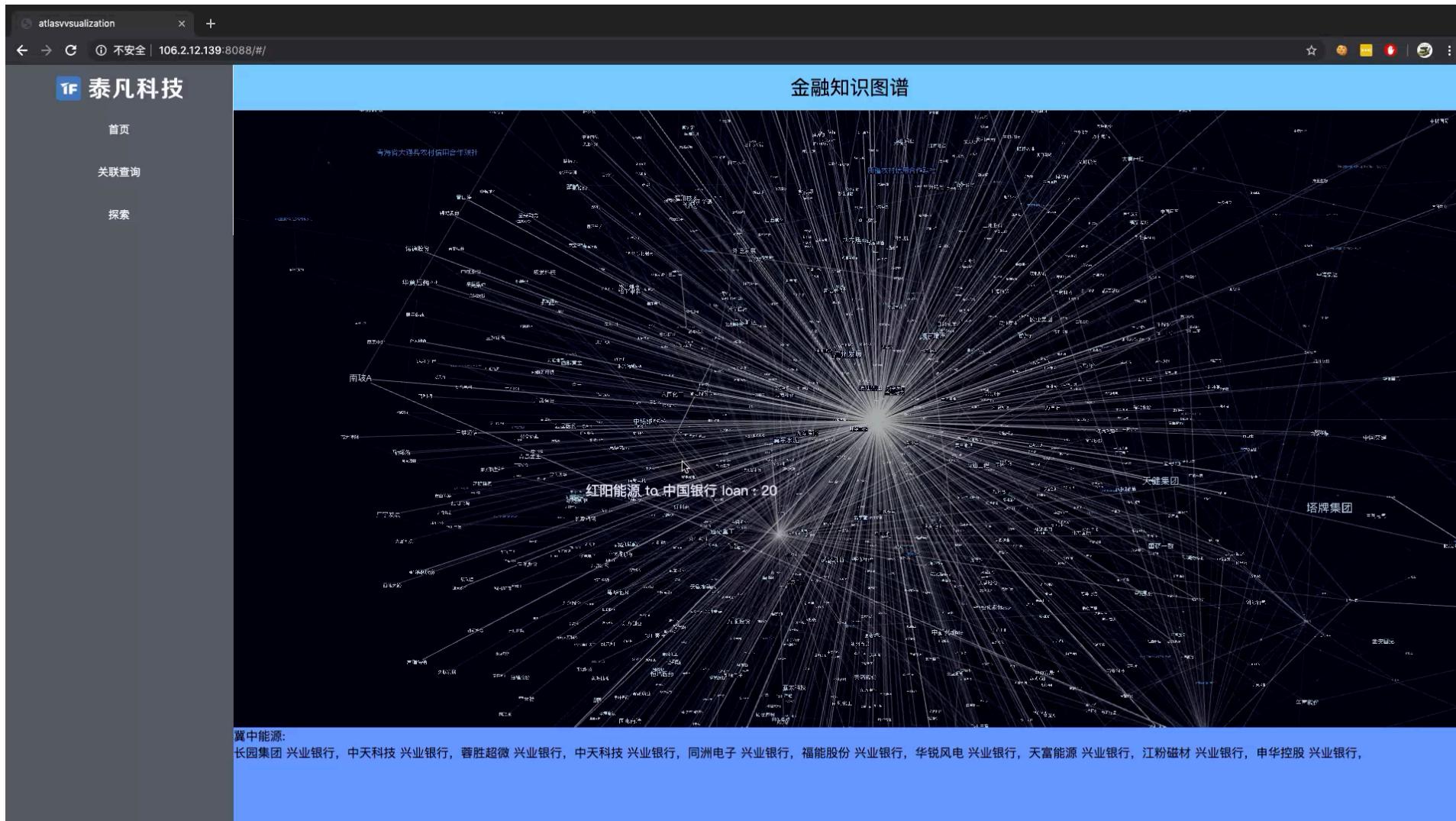
KNOWLEDGE GRAPH VIS LANGUAGE



KNOWLEDGE GRAPH 3D VISUALIZATION



KNOWLEDGE GRAPH 3D VISUALIZATION-2



KNOWLEDGE GRAPH 3D VISUALIZATION-3



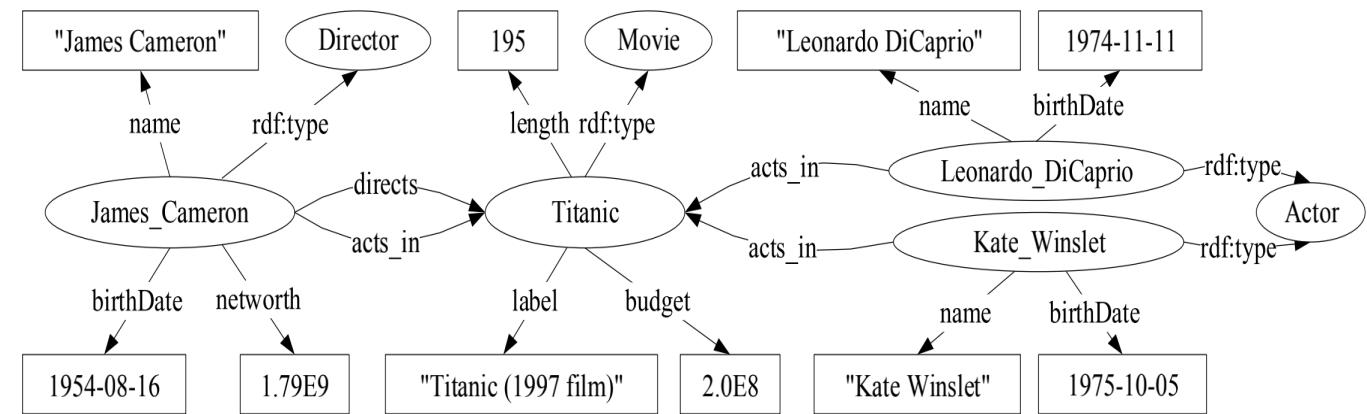
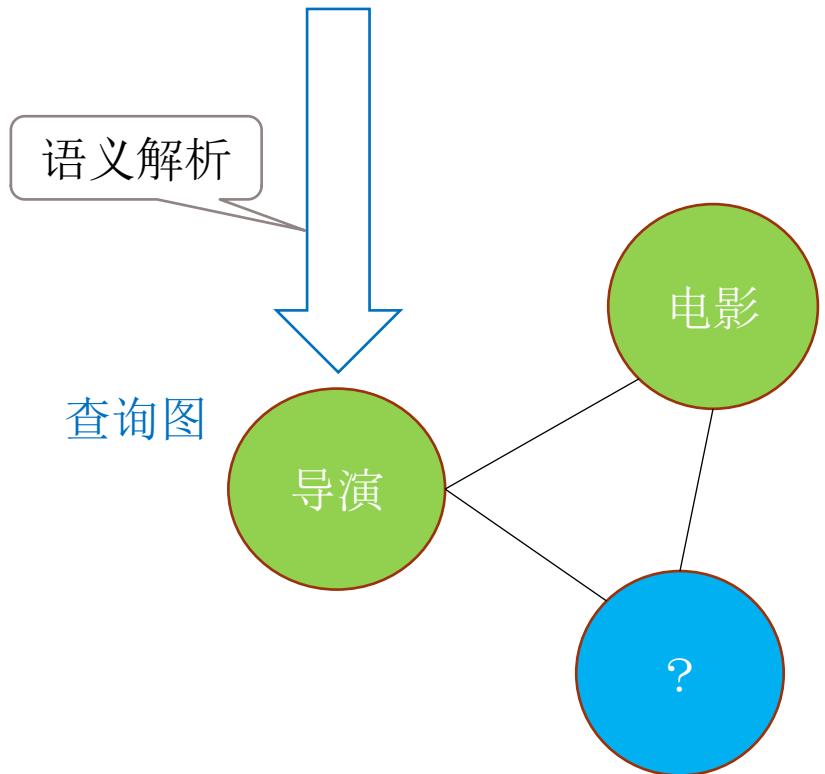
KNOWLEDGE GRAPH QUERY

- 基于形式化语言的查询 (**SPARQL/Cypher**)
- 基于可视化交互的查询
- 基于自然语言的查询 (智能问答)



KBQA

1950年之后出生的资产大于1.0E9的导演执导的电影的出演演员



将用户问题转换为查询语句，从知识图谱库中查询得到答案



泰凡KBQA ARCHITECTURE



知识构建与存储

结构化数据 半结构化数据 非结构化数据



知识抽取

知识处理平台

知识融合

语义解析与数据训练

查询模板库

匹配模型

排序模型

分类模型

标准问法生成

标准问法库

模型训练数据

数据标注平台

计算平台

Ai算法库



FURTHERMORE

- Visualization Interaction for Query Pattern code generation
- Intelligent Query code auto-completion
- Relationship detection for GKG



THANK YOU FOR YOUR ATTENTION !

2019.8 @ 杭州

