

TA

- Xiuwen Zheng

- Office Hour: 6:00 - 7:00 pm Monday and Wednesday
- Office: SDSC E-309
- Email: xiz675@eng.ucsd.edu
- Research Interest: heterogeneous data management, polystore system and query optimization.
- Hobbies: Ping-pong, badminton, tennis and hiking



Different data stores

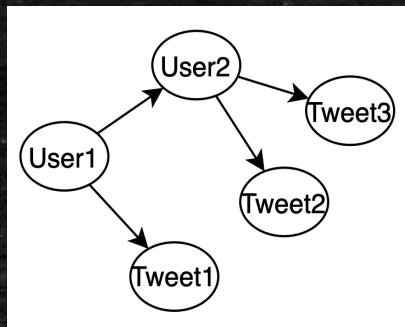
- Relational Database: Present the data to the user as relations (a presentation in tabular form consisting of a set of rows and columns);



	name	twittername	state
1	Lisa Murkowski	lisamurkowski	AK
2	Dan Sullivan	SenDanSullivan	AK
3	Doug Jones	SenDougJones	AL
4	Richard C. Shelby	SenShelby	AL
5	John Boozman	JohnBoozman	AR

select twittername from twitterhandle
where name='Lisa Murkowski'

- Graph Database:



Match (n:{name:'Lisa Murkowski'}) - [:friendwith] ->(m)-[:friendwith] -> (p) - [:friendwith] -> (k) return m,p,k

- Textual Database:



```
{
  "text-field": "Ailing Jed Lowrie, Todd Frazier racing the clock for Mets opener PORT ST. LUCIE - As mu
  "id": "199523:news",
  "information-class": "publications:news",
  "_version_": 1661344795584364544},
{
  "text-field": "Russell Westbrook dominates as Nets fade away in second half OKLAHOMA CITY - The last t
  "id": "199524:news",
  "information-class": "publications:news",
  "_version_": 1661344795585413120},
```

q= text-field: covid or corona

A Simple Example Use Case

- **Dataset:**

- Textual: A collection of political news articles published recently
 - Graph: A tweet social network contains tweets nodes, users nodes, users-write-tweets relationships and users-mention-users relationships.
 - Relational: A Twitter handler table contains politicians' real name and twitter account name

A Simple Example Use Case

- Task:
 - Given a list of keywords:
 - Task1: Find out politicians' Twitter accounts where these politicians are mentioned in the recent news that contains the keywords.
 - Task2: Find out other Twitter accounts that mention these politicians' accounts.
 - Task3: Find out tweets that contain these

Written in ADIL

```
/*specify configuration file*/
USE newsDB;
/* main code block */
create analysis politician as (
keywords := \"corona\", \"covid\", \"pandemic\", \"vaccine\"\;
temp := keywords.map(i => stringReplace(\"text-field: \$\", i));
t := stringJoin(\" OR \", temp);
doc<text-field:String> := executeSOLR(\"allnews\", \"q= \$t & rows=50\");
namedentity := NER(doc.text-field);
user := executeSQL(\"News\", \"select distinct t.name as name, t.twittername as twittername
from twitterhandle t, \$namedentity e where LOWER(e.name)=LOWER(t.name)\");
userNameList := toList(user.name);
userNameP := userNameList.map(i => stringReplace(\"t.text contains '\$\" \$\", i));
predicate := stringJoin(\" OR \", userNameP);
users<name:String> := executeCypher(\"tweetG\", \"match (u:User)-[:mention-](n:User)
where n.userName in \$user.twittername return u.userName as name\");
tweet<t:String> := executeCypher(\"tweetG\", \"match (t:Tweet) where \$predicate return t.text as t\");
);
```