Link Correlated Military Data for Better Decision Support

Science and Technology on Information System Engineering Laboratory,
Nanjing Research Institute of Electronics Engineering, Nanjing, China

June 2011

Problem: Locate required data from a “data sea”

—— To win the information superiority

- What a decider requires:
  - Data directly useful to solve decision problem
  - Data of interest or highly related to decision problem
- What a decider usually gets:
  - Huge amount of data while few have relevancy
- So the decider needs to:
  - Check each data’s relevancy till required data are found
  - Guess the location of required data by experience

Better solution?

One answer: Data navigation

- Data navigation:
  - Meaning: Lead the path to the required data step by step, based on some kind of guidance
- Three basic kinds of guidance:
  - Classification: guide through data taxonomy
  - Keyword: guide through keyword-based search results
  - Correlation: guide through correlations among data

Any difference?

Example decision problem – Who are robbers?

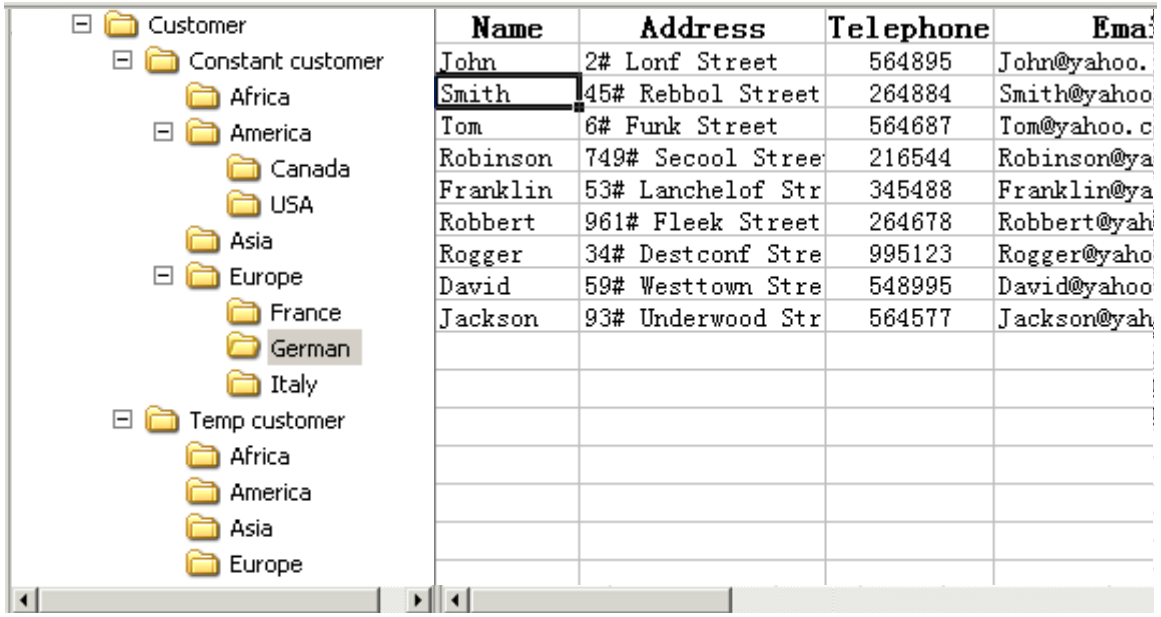
- "A bank was robbed last night. A man named Tom was arrested inside the bank, with a QSZ92 5.8mm gun found in his hand, made in company KGE. He refused to provide other robbers' names."
- To find other robbers, a possible way:
  1. Investigate company KGE, list its customers
  2. Gather intelligence, analyze each customer's recent activities
  3. Check features of each activity (time, place, weapons, etc.), compare with the bank robbery event, so as to find suspects

How to support this decision through data navigation?

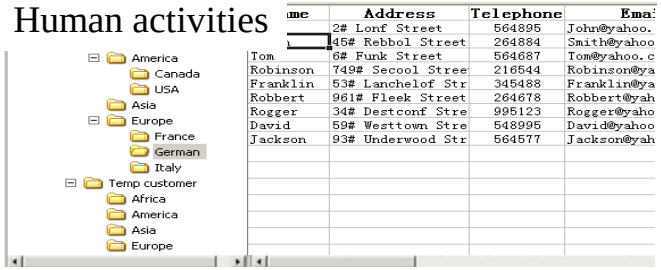
Data navigation method - Classification-based

- Users need to know exactly under which branch of the tree he can or may find the required data

Customers of company KGE



Human activities

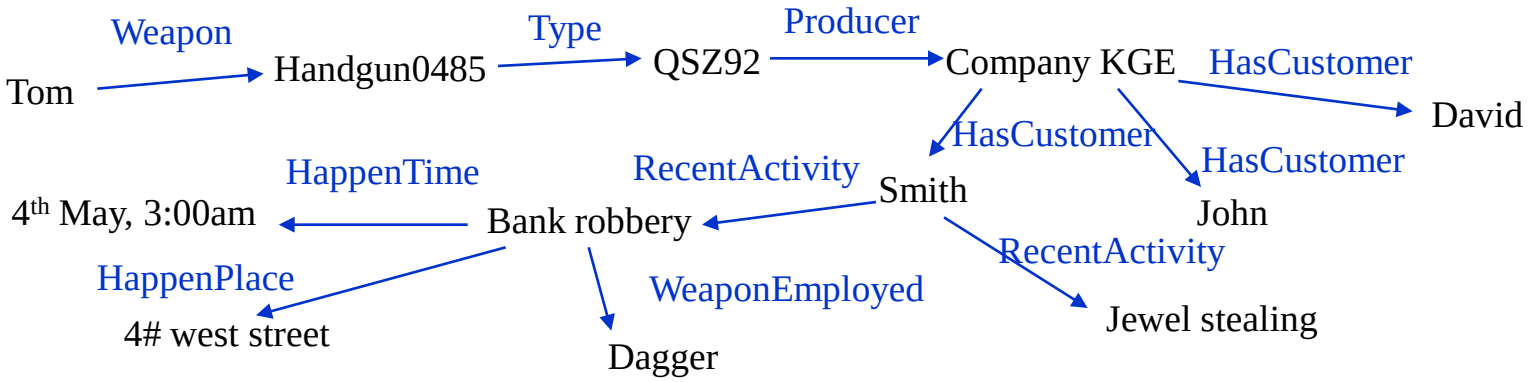


• • •

Data navigation method - Keyword-based

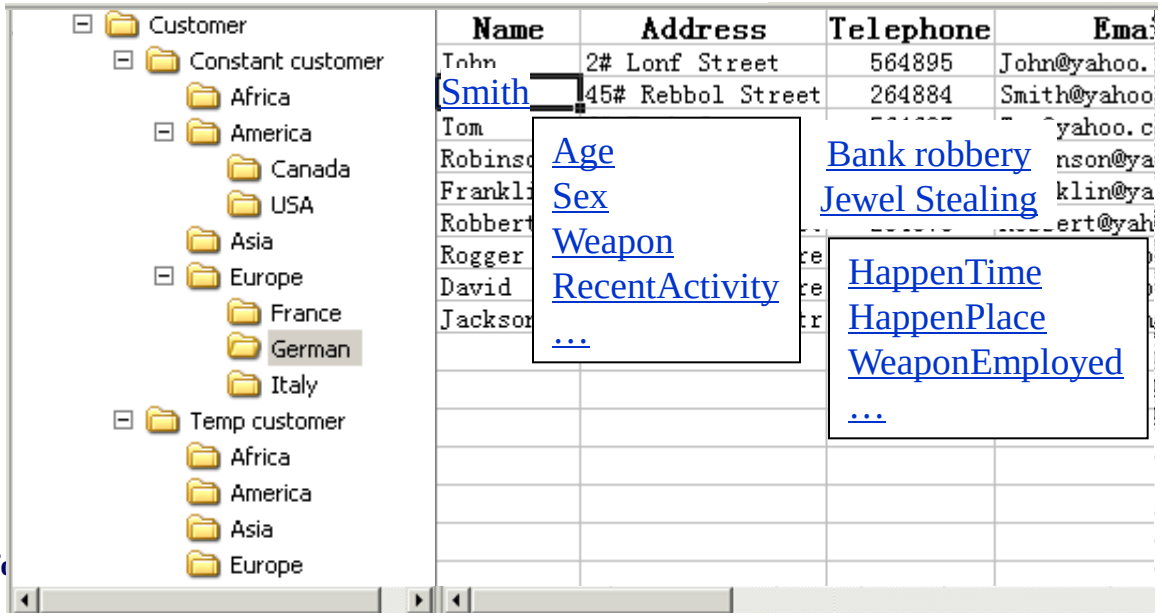
- Users need to master precise keywords to both extract required data, and filter out irrelevant ones

Data navigation method - Correlation-based



- Users just follow links of interest
- No restriction on content of links – by all possible data correlations

Customers of company KGE



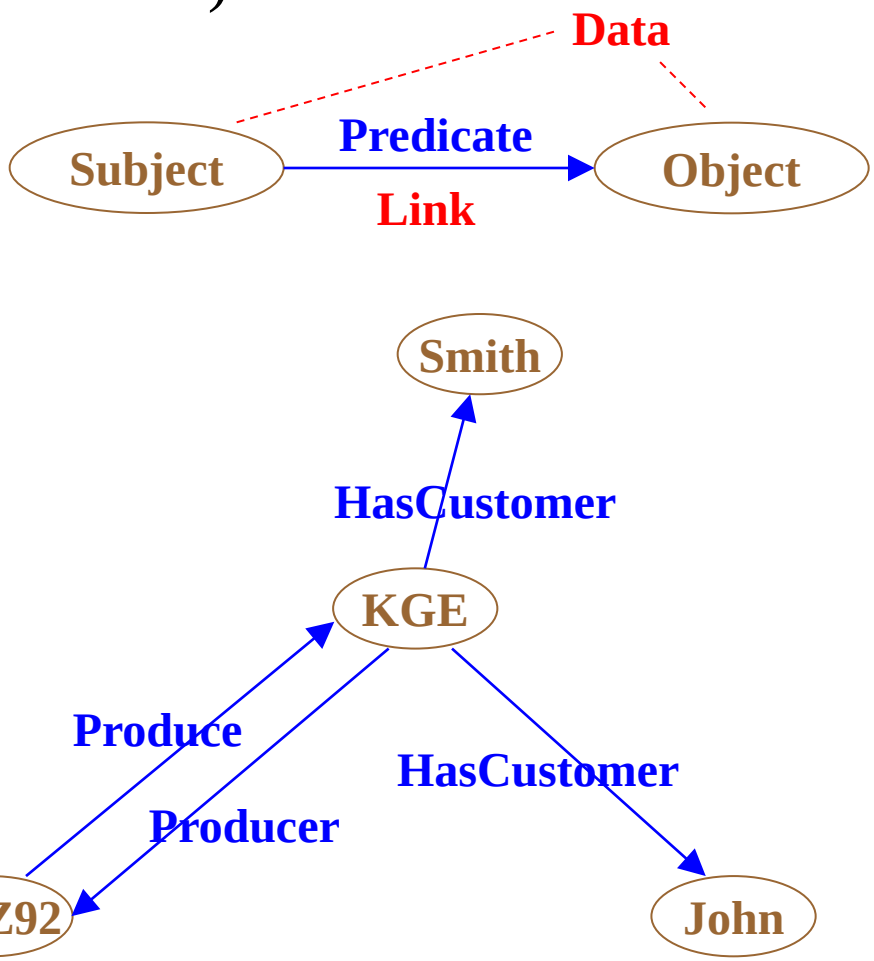
Merits of correlation-based method

- By links, one can jump from data to data directly.
- Navigation by links is as easy as surfing on Internet.
- Link construction is based on data correlations. Link selection is upon user interest.
- What user may associate, there is a link to support him, given links rich enough.

A method suited to human association habit?

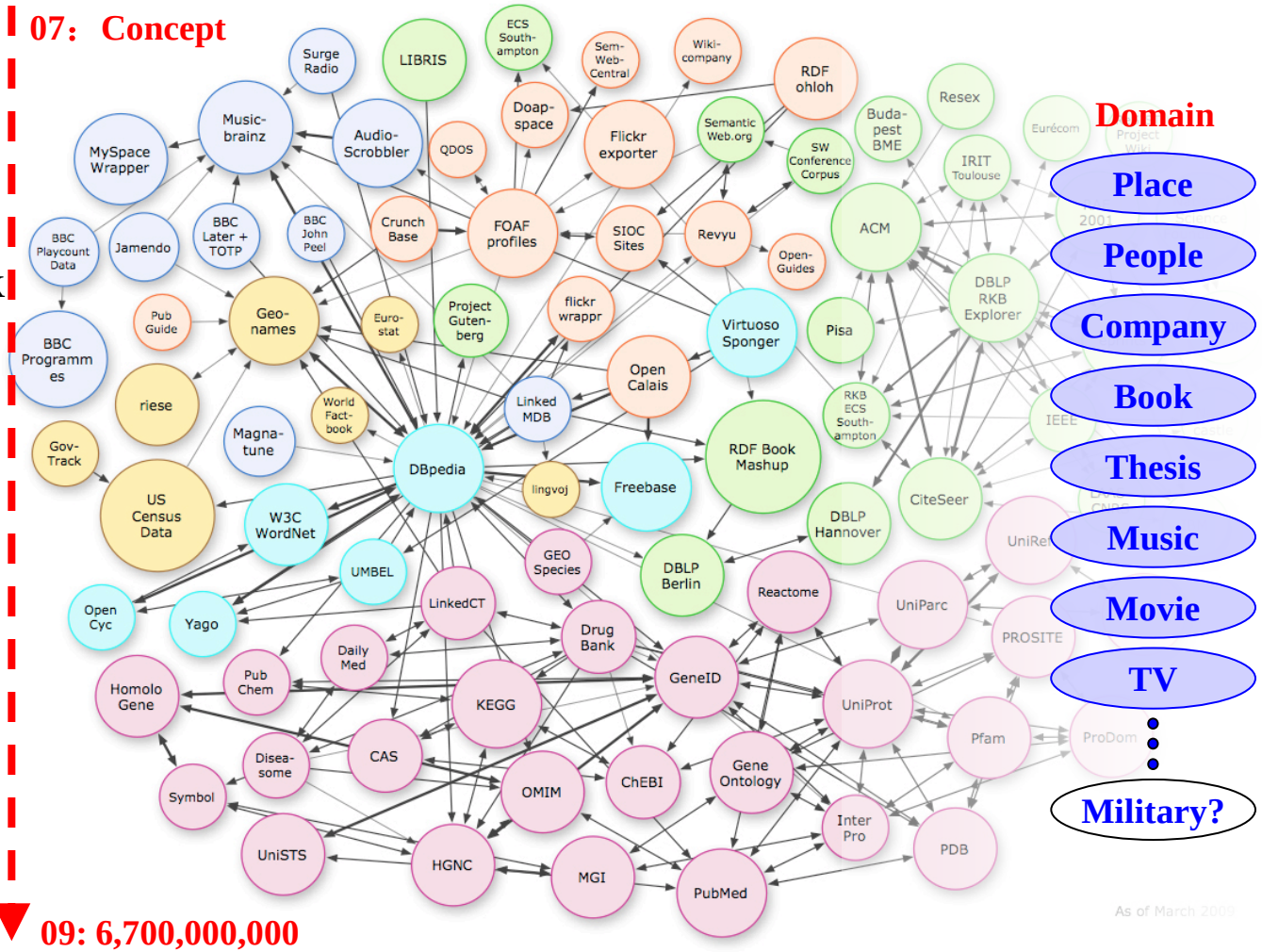
Method: Unified correlation description

- RDF (Resource Description Framework)
  - <Subject, Predicate, Object>
- Example：
  - <http://company.org/KGE, http://company.org/HasCustomer, http://people.org/Tom>
  - < http://company.org/KGE, http://company.org/HasCustomer, http://people.org/John>
  - < http://company.org/KGE, http://company.org/Produce, http://weapon.org/QSZ92>
  - < http://weapon.org/QSZ92, http://company.org/Producer, http://company.org/KGE >



Method: Unified correlation description

- Linked Data
  - Use URIs as names for things;
  - Use HTTP URIs so that people can look up those names;
  - When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL);
  - Include links to other URIs, so that they can discover more things.



Method: Automatic link construction

	Name	Product
1	KGE	QSZ92
2	KGE	QSS05

fk

pk

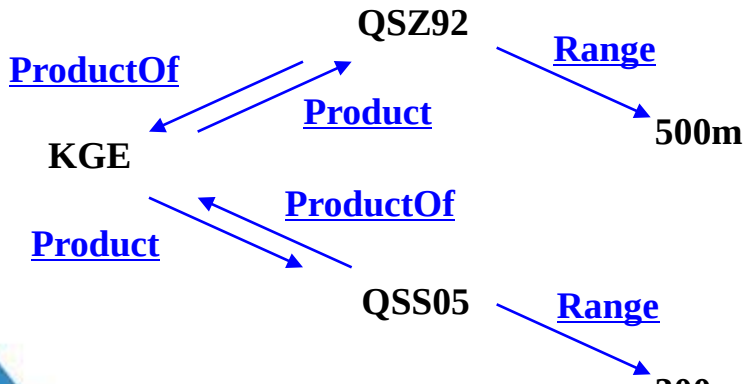
	Type	Range
1	QSZ92	500m
2	QSS05	300m

(a) Data in relational database

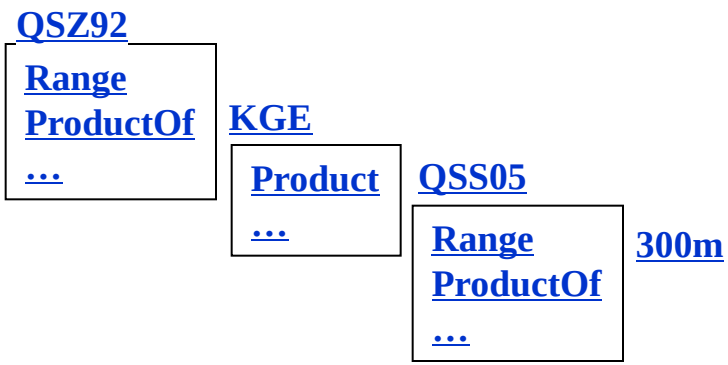
Company Info
Name VARCHAR(200) <pk>
Product VARCHAR(200) <fk>
…

Weapon Info
Type VARCHAR(200) <pk>
Range DOUBLE
…

(b) Table structure



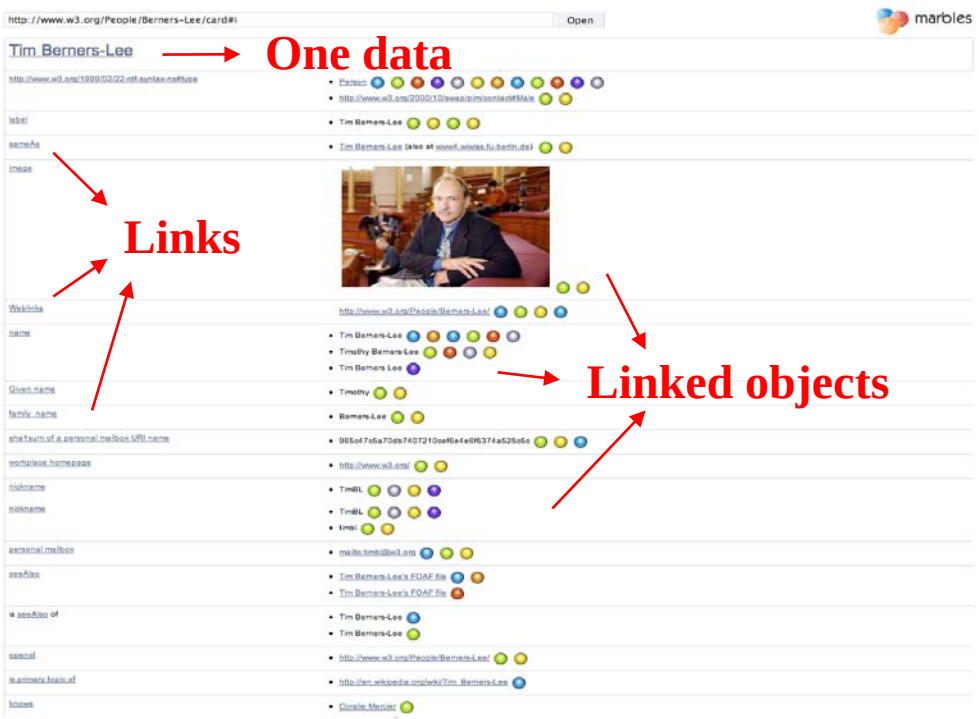
(c) Correlated data network

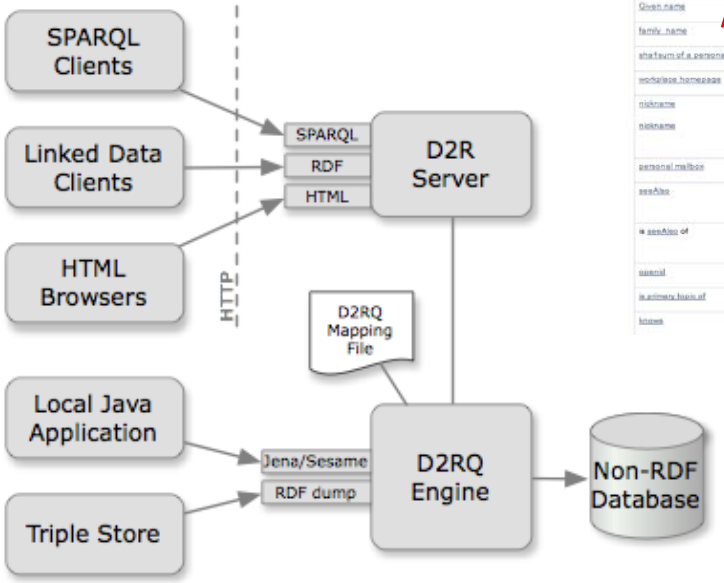


(d) Hyperlinks on user interfaces

Method: Automatic link construction

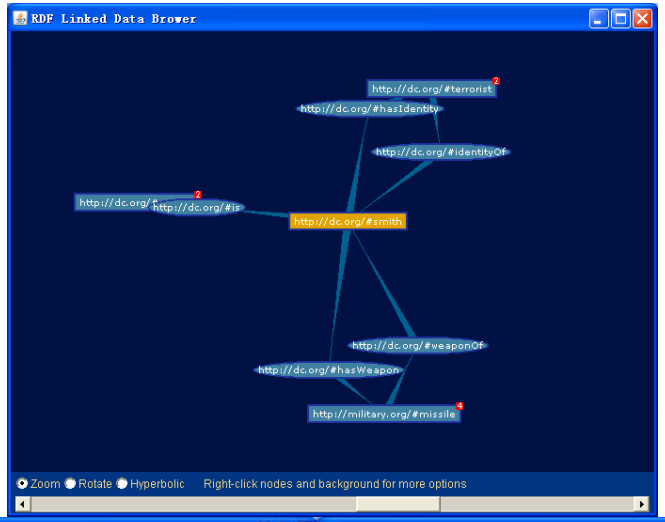
- D2R tool

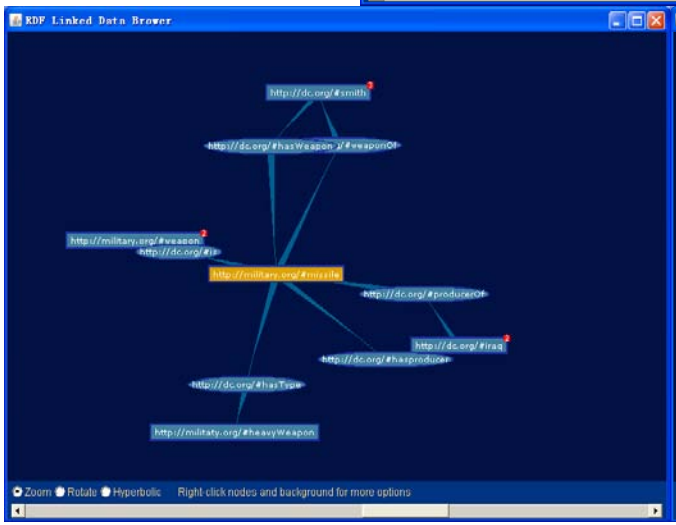


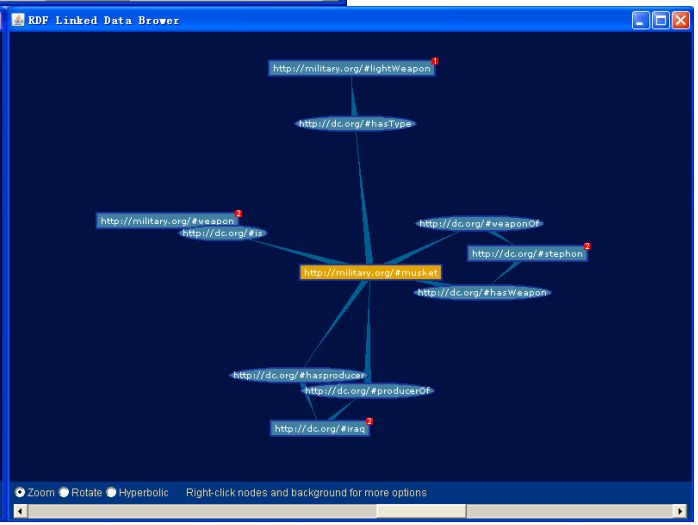


Implementation

- Correlations generated from relational database







Implementation

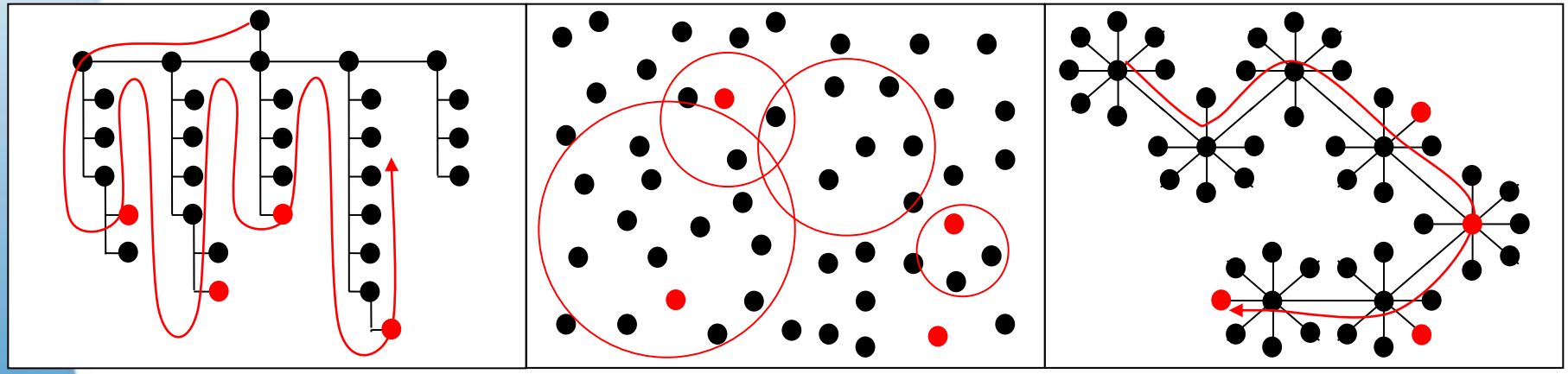
- Description script of data correlations

```
<rdf:RDF
  xmlns:dc="http://dc.org/#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:military="http://military.org/#" >
 <rdf:Description rdf:about="http://dc.org/#tom">
  <dc:hasWeapon rdf:resource="http://military.org/#gun"/>
  <dc:hasIdentity rdf:resource="http://dc.org/#police"/>
  <dc:is rdf:resource="http://dc.org/#people"/>
 </rdf:Description>
 <rdf:Description rdf:about="http://dc.org/#terrorist">
  <dc:identityOf rdf:resource="http://dc.org/#smith"/>
 </rdf:Description>
 <rdf:Description rdf:about="http://dc.org/#USA">
  <dc:producerOf rdf:resource="http://military.org/#gun"/>
 </rdf:Description>
 <rdf:Description rdf:about="http://military.org/#missile">
  <dc:hasproducer rdf:resource="http://dc.org/#iraq"/>
  <dc:weaponOf rdf:resource="http://dc.org/#smith"/>
 <dc:hasType rdf:resource="http://militaty.org/#heavyWeapon"/>
  <dc:is rdf:resource="http://military.org/#weapon"/>
 </rdf:Description>
 <rdf:Description rdf:about="http://dc.org/#police">
  <dc:identityOf rdf:resource="http://dc.org/#tom"/>
 </rdf:Description>
 <rdf:Description rdf:about="http://military.org/#musket">
  <dc:hasproducer rdf:resource="http://dc.org/#iraq"/>
  <dc:weaponOf rdf:resource="http://dc.org/#stephon"/>
  <dc:hasType df:resource="http://military.org/#lightWeapon"/>
  <dc:is rdf:resource="http://military.org/#weapon"/>
 </rdf:Description>
 <rdf:Description rdf:about="http://dc.org/#iraq">
  <dc:producerOf rdf:resource="http://military.org/#musket"/>
  <dc:producerOf rdf:resource="http://military.org/#missile"/>
</rdf:Description>
…
</rdf:RDF>
```

Methods comparison

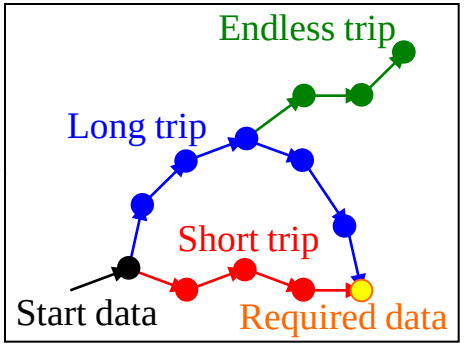
- User's view on operation mode

(a) Classification-based (b) Keyword-based (c) Correlation-based



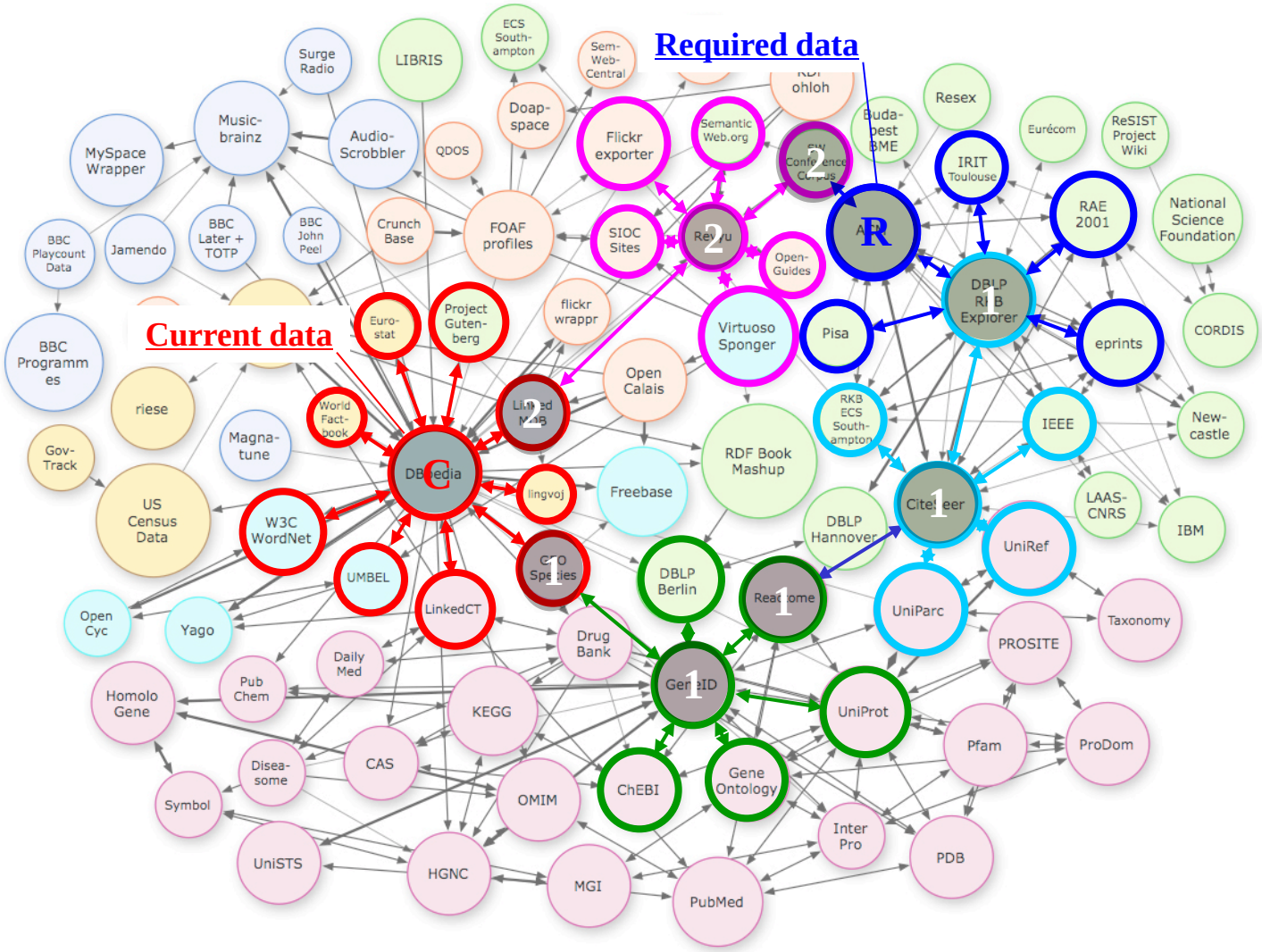
- Method performance

	Problem Relativity	*Operation Efficiency*	*Navigation Convergence*	*User Skill Requirement*
Classification-based	Low	Dependent	Assured	High
Keyword-based	Medium	Dependent	Not assured	High
Correlation-based	High	Dependent	Not assured	Low



Tree depth Information entropy Correlation richness

Navigation within correlated data network



Summary

- Correlation-based method is a good complement to traditional methods, not a substitute.
- A method is suited to human association habit.
- With more data networked through correlations, and statistical-analytic tools to support network mining, existing data will be more interesting.

Discussion

- Is it possible to link all military data by correlations?
- How to link data in different formats, text, media…?
- How to make better use of data correlations?