Everything is Connected™

GeoWeb 2008

geoweb updates: subscribe Send

- home
- about GeoWeb 2008
- registration
- location
  - about Vancouver
  - conference venue
  - accommodations
  - travel partner
  - going green
  - visitor information
- call for papers & workshops
- students & academia
- program
- sponsors & vendors
- press & media
- podcasts
- photos
- youtube

about GeoWeb 2008

The term "GeoWeb" or Geospatial Web refers to the ability to locally/globally integrate and share geospatial information via the Internet.

The GeoWeb 2008 conference reflects the breadth, evolution and growing maturity of the GeoWeb and is one of the only conferences focusing exclusively on the convergence of GIS and the Internet, and the economic potential associated with the convergence of XML, web services and GIS. This exciting annual conference will take place in Vancouver, Canada from July 21-25th, 2008 at the Morris J. Wosk Centre for Dialogue. The GeoWeb 2008 conference welcomes both public and private organizations to meet, discuss and learn about today's most innovative geospatial technologies.

GeoWeb 2008 reflects a growing awareness of the need to interpret geospatial information in the broadest possible terms, especially in the context of the built environment, where information sharing and collaboration is key to the increased productivity, efficiency and vastly improved decision making. We hi-light this in GeoWeb 2008 with the theme on CAD/BIM/GIS Integration as we build toward GeoWeb 2009 Cityscapes.

The theme for the 2008 conference is **"Infrastructure: Local to Global"**, which implies the GeoWeb has a local community dimension as well as a global dimension. The integration of global aggregators will drive the creation of local infrastructures and will give rise to a global infrastructure. Additional points of discussion will include:

- Global Aggregators and Data Communities

CSBE | SCGAB

platinum sponsors

ESRI

Google

Microsoft Virtual Earth

GeoTemporal Reasoning in a Web 3.0 World

(or the joy of having a spatial database in an RDF Triple Store)

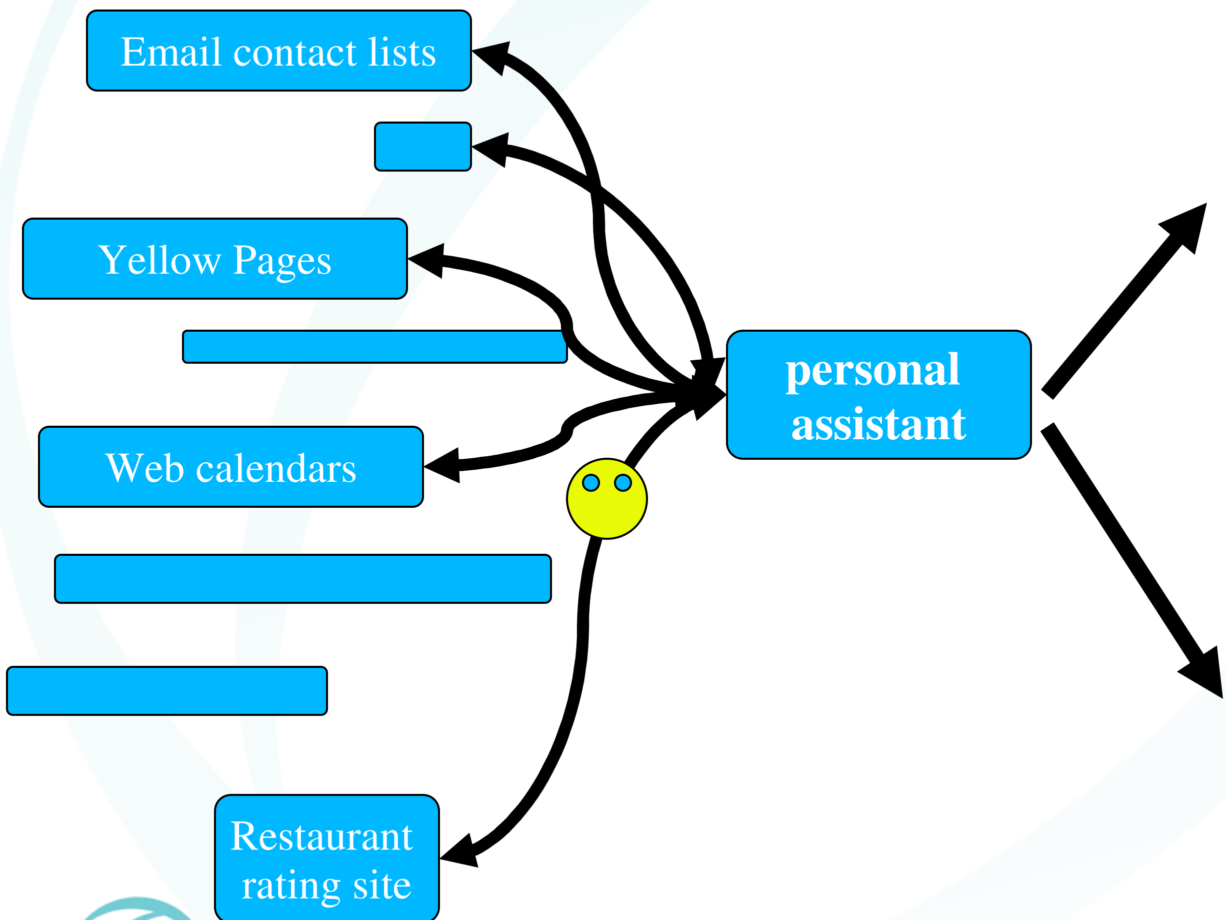
Jans Aasman
Franz Inc.
www.franz.com

This talk

- What do people do with an RDF Database
- How to combine
  - Geospatial
  - Temporal
  - Social Network Analysis

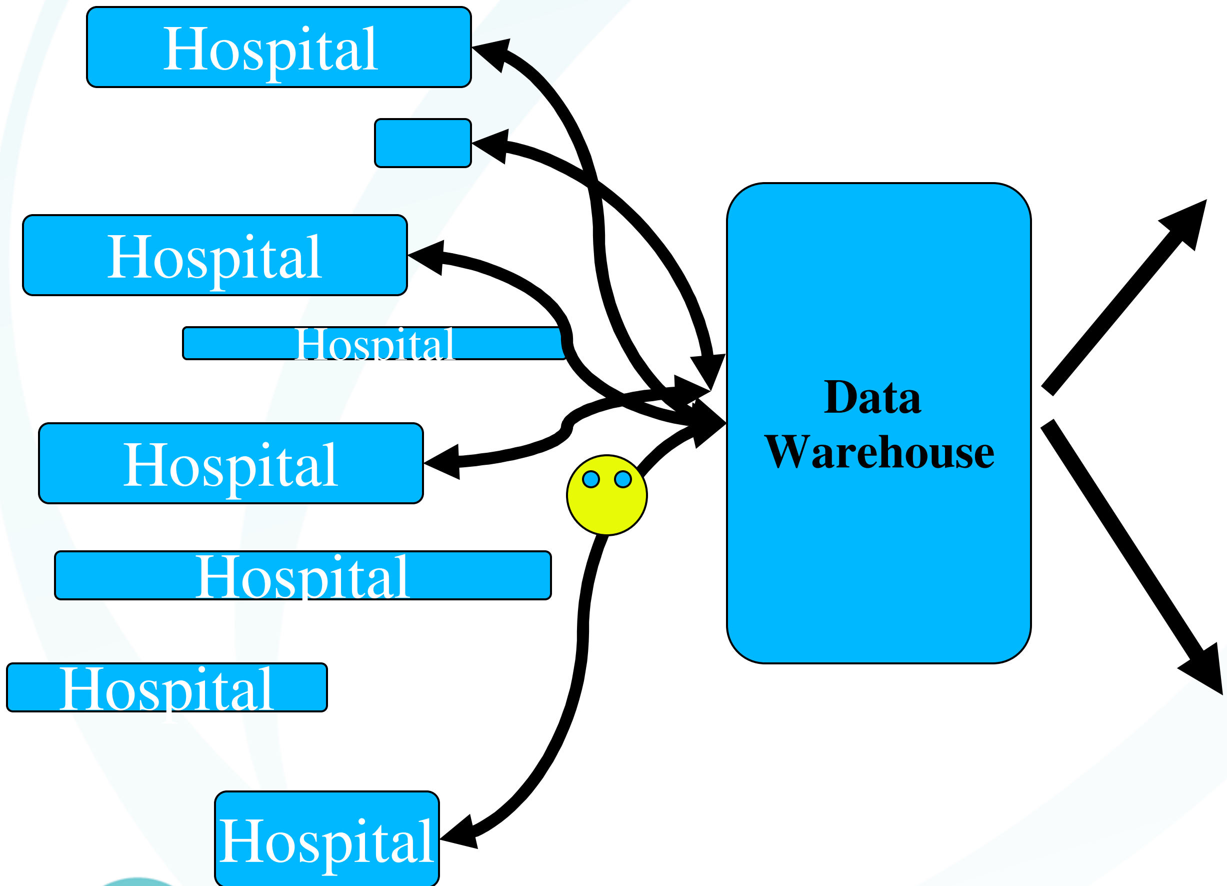
What users do with RDF?

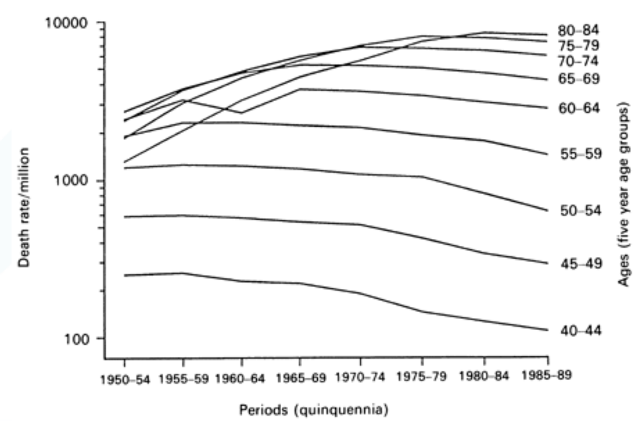
Semantic Web View: Your personal assistant



Web 3.0's Database

Enterprise View: UT Houston unifies data from many hospitals for advanced data analytics





What is AllegroGraph [1]

- Scalable and persistent quadstore
- Federated
  - Create an abstract store that is a collection of other triple stores. Prolog and SPARQL and Reasoning work transparently against abstract store
- Compliant with standards
  - RDF, RDFS, OWL, SPARQL, Named Graphs, ISO Prolog, OWL-lite reasoning

What is AllegroGraph [2]

- Relational database efficiency for range queries
  - We support most xml schema types (dates, times, longitudes, latitudes, durations, telephone numbers, etc)
- Spatial database efficiency for geospatial primitives
  - Find elements in bounding boxes as fast as in spatial databases
- Temporal reasoning
  - Reasoning about times and intervals (Allen Logic)
- Social Network Analytics library
  - Find actor degrees and centrality, cliques, group centrality and cohesiveness

The difference with a relational database?

```
<triple 32: "person2" "type" "person">
<triple 33: "person2" "first-name" "Rose">
<triple 34: "person2" "middle-initial" "Elizabeth">
<triple 35: "person2" "last-name" "Fitzgerald">
<triple 36: "person2" "suffix" "none">
<triple 37: "person2" "alma-mater" "Sacred-Heart-Convent">
<triple 38: "person2" "birth-year" "1890">
<triple 39: "person2" "death-year" "1995">
<triple 40: "person2" "sex" "female">
<triple 41: "person2" "spouse" "person1">
<triple 58: "person2" "has-child" "person17">
<triple 56: "person2" "has-child" "person15">
<triple 54: "person2" "has-child" "person13">
<triple 52: "person2" "has-child" "person11">
<triple 50: "person2" "has-child" "person9">
<triple 48: "person2" "has-child" "person7">
<triple 60: "person2" "profession" "home-maker">
<triple 46: "person2" “birth-place“ “gn2988834”>
<triple 44: " gn2988834 " “name" “boston">
<triple 42: " gn2988834 " “centerlocation" “42.3,-714">
```

Franz Inc.
Web 3.0's Database

An artist's impression of the same info in a RDBM

Table Person

ID	First-Name	Last-Name	Middle-In.	DOB	DOD	PlaceOB	Sex
2	Rose	Fitzgerald	Elizabeth	1890	1995	1	F

Table Spouses

ID1	ID2
2	1

Table to-schools

ID1	SchoolID
2	3

Table Schools

ID	Name
3	Sacred-Heart

Table has-profession

ID1	ProfID
2	3

Table Professions

ID	Name
3	Home-maker

Table Has-Child

ID1	ID2
2	17
2	15
2	14
2	13

Table Place

ID	Name	State	Longitude	Latitude
1	Boston	MA	42.3	-71.4

With a Graph Database – RDF Database

- you add new predicates without changing any schema
- one-to-many relations are directly encoded without the indirection of tables
- You never think about what to index because everything is indexed

Events and Activity recognition

- Creating an event database with social network analysis and geospatial and temporal reasoning

`Find all meetings that happened in July within 5 miles of Berkeley that was attended by the most important person in Jans' friends and friends of friends.`

The common elements of an event

- **A type**
  - Meetings, communications event, financial transactions, visit, attack/truce, an insurance claim, a purchase order
  - Reasoning over types of events requires **RDFS++ reasoning**
- **List of actors**
  - Reasoning over relationships between actors requires **Social Network Analysis**
- **Location**
  - Reasoning over where something happened or how far away something happened requires **GeoSpatial Reasoning**
- **Start-time and possible an end-time**
  - Reasoning over when or in what order something happened requires **Temporal Reasoning**
- **Anything else that describes the event**
  - Goods that changed hands

Events at the core of many Business Processes

- **Health care**
  - A hospital visit, a visit to a drugstore, a medical procedure
- **Communications Industry**
  - A telephone call (and they store your location now too)
  - An Email or SMS
- **Financial industry**
  - Every financial transaction is an event
- **The Insurance industry**
  - Track behavior of customers & find fraud
- **The Enterprise**
  - Combine your ERS system, your email archive and your HR data
- **The Government**
  - Homeland Security is interested in every type of imaginable event

Integrated in query language.

Find all meetings that happened in July within 5 miles of Berkeley that was attended by the most important person in Jans' friends and friends of friends.

```
(select (?x)
   (ego-group !person:jans knows ?group 2)
   (actor-centrality-members ?group knows ?x ?num)
   (q ?event !fr:actor ?x)
   (qs ?event !rdf:type !fr:Meeting)
   (interval-during ?event "2008-07-01" "2008-07-25")
   (geo-box-around !geoname:Berkeley ?event 5 miles)
   !)
```

Integrated in select language.

Find a meetings that happened in July within 5 miles of Berkeley that was attended by the most important person in Jans' friends and friends of friends.

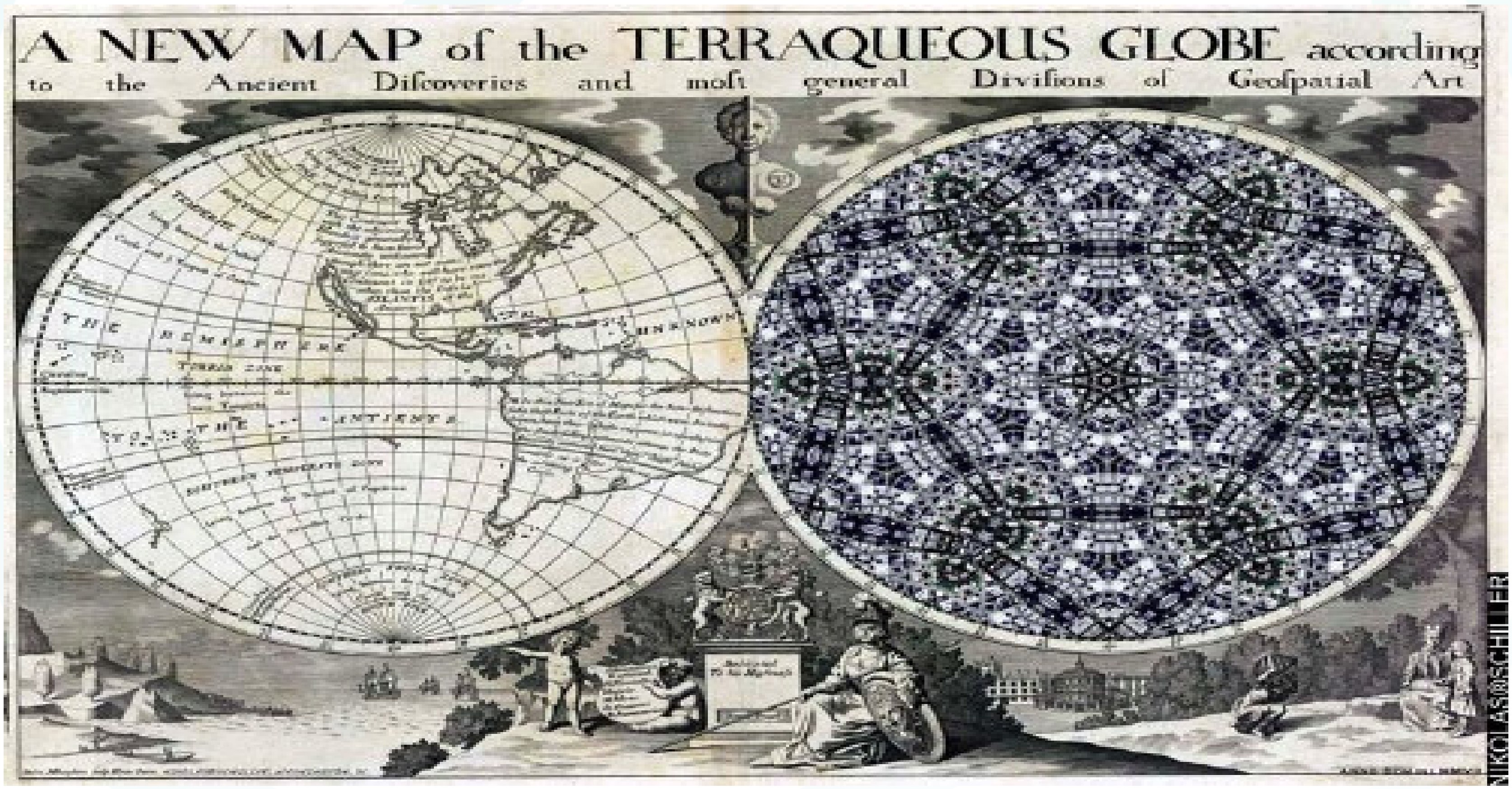
```
(select (?x)
   (ego-group !person:jans knows ?group 2)
   (actor-centrality-members ?group knows ?x ?num)
   (q ?event !fr:actor ?x)
   (qs ?event !rdf:type !fr:Meeting)
   (interval-during ?event "2008-07-01" "2008-07-25")
   (geo-box-around !geoname:Berkeley ?event 5 miles)
   !)
```

Integrated in select language.

Find a meetings that happened in December within 5 miles of Berkeley that was attended by the most important person in Jans' friends and friends of friends.

```
(select (?x)
   (ego-group !person:jans knows ?group 2)          ← SNA
   (actor-centrality-members ?group knows ?x ?num) ← SNA
   (q ?event !fr:actor ?x)                          ← DB Lookup
   (qs ?event !rdf:type !fr:Meeting)                ← RDFS Reasoning
   (interval-during ?event "2007-12-01" "2007-12-31") ← Temporal
   (geo-box-around !geoname:Berkeley ?event 5 miles) ← Spatial
   !)
```

Geospatial Reasoning

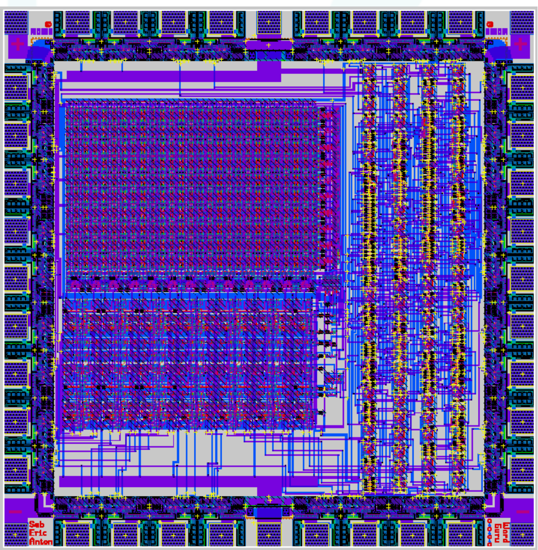


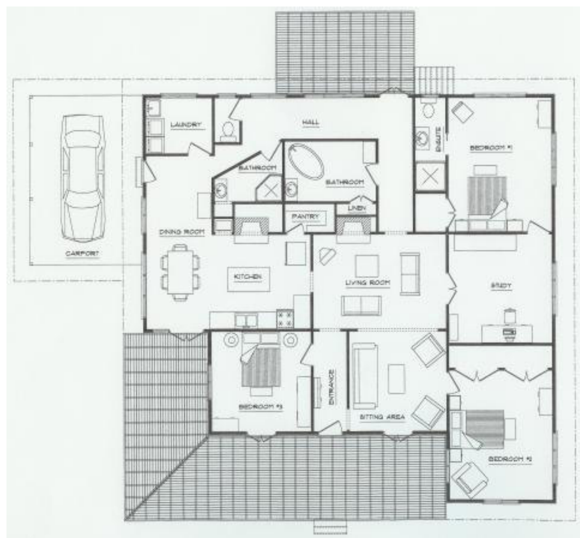
Geospatial Challenge

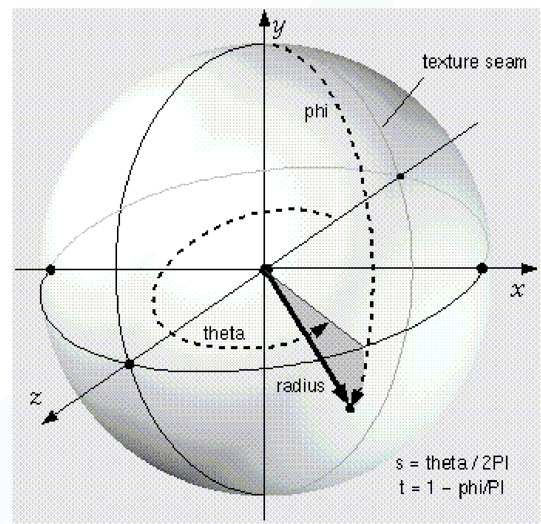
- Make the following super efficient
  - Where did something happen?
  - How far was event1 from event2?
  - Find all the events that occurred in a bounding box or radius of M miles?
  - Do these two shapes overlap?
  - Find all the objects in the intersection of two shapes
- On a very large scale
  - when things don't fit in memory
  - millions of events and polygons

Scale

- We work both in a flat plane (x, y) and on a sphere (lat, lon)
- We work on different scales : From ångström to miles







The law of haversines

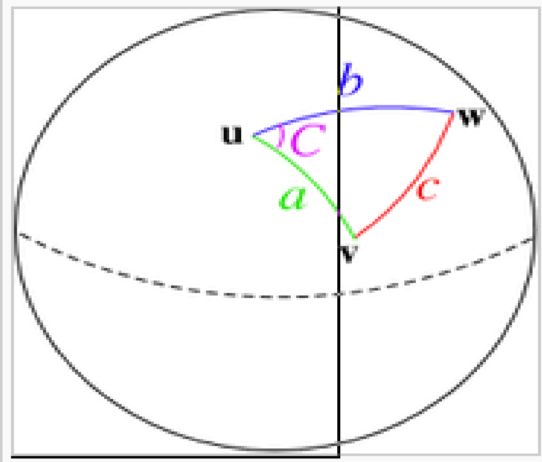
[edit]

Given a unit sphere, a "triangle" on the surface of the sphere is defined by the great circles connecting three points **u**, **v**, and **w** on the sphere. If the lengths of these three sides are *a* (from **u** to **v**), *b* (from **u** to **w**), and *c* (from **v** to **w**), and the angle of the corner opposite *c* is *C*, then the law of haversines states:

(the law of haversines)

$$\operatorname{haversin}(c) = \operatorname{haversin}(a-b) + \sin(a)\sin(b)\operatorname{haversin}(C)$$

Since this is a unit sphere, the lengths *a*, *b*, and *c* are simply equal to the angles (in radians) subtended by those sides from the center of the sphere (for a non-unit sphere, each of these arc lengths is equal to its central angle multiplied by the radius of the sphere).



Spherical triangle solved by the law of haversines.

In order to obtain the haversine formula of the previous section from this law, one simply considers the special case where **u** is the ~~north pole~~, while **v** and **w** are the two points whose separation *d* is to be determined. In that case, *a* and *b* are $\pi/2 - \varphi_{1,2}$ (i.e., 90° − latitude), *C* is the longitude separation $\Delta\lambda$, and *c* is the desired *d*/*R*. Noting that $\sin(\pi/2 - \varphi) = \cos(\varphi)$, the haversine formula immediately follows.

To derive the law of haversines, one starts with the spherical law of cosines:

$$\cos(c) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(C)$$

(spherical law of cosines)

As mentioned above, this formula is an ill-conditioned way of solving for *c* when *c* is small. Instead, we substitute the identity that $\cos(\theta) = 1 - 2\operatorname{haversin}(\theta)$, and also employ the addition identity $\cos(a - b) = \cos(a)\cos(b) + \sin(a)\sin(b)$, to obtain the law of haversines, above.

R-tree

From Wikipedia, the free encyclopedia

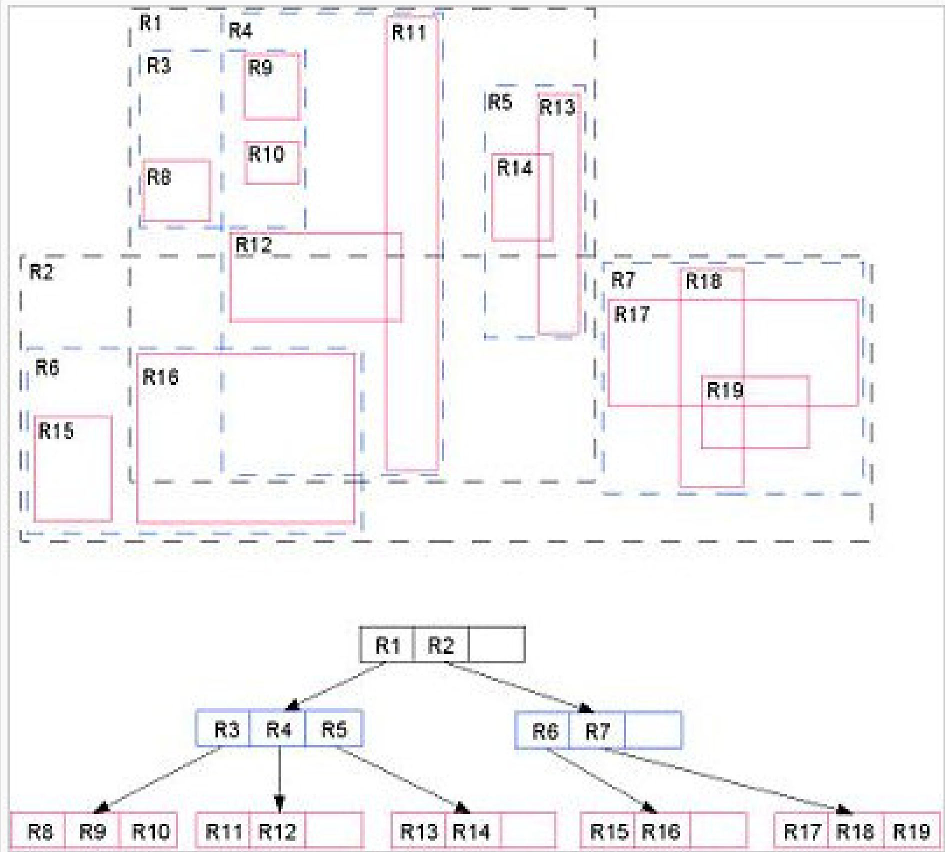
This article is about the data structure. For the type of metric space, see Real tree.

R-trees are tree data structures that are similar to B-trees, but are used for spatial access methods i.e., for indexing multi-dimensional information; for example, the (X, Y) coordinates of geographical data. A common real-world usage for an R-tree might be: "Find all museums within 2 miles of my current location".

The data structure splits space with hierarchically nested, and possibly overlapping, minimum bounding rectangles (otherwise known as bounding boxes).

Each node of an R-tree has a variable number of entries (up to some pre-defined maximum). Each entry within a non-leaf node stores two pieces of data: a way of identifying a child node, and the bounding box of all entries within this child node.

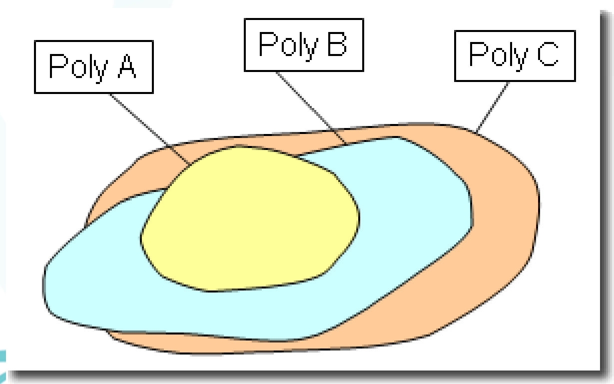
The insertion and deletion algorithms use the bounding boxes from the nodes to ensure that "nearby" elements are placed in the same leaf node (in particular, a new element will go into the leaf node that requires the least enlargement in its bounding box). Each entry within a leaf node stores two pieces of information; a way of identifying the actual data element (which, alternatively, may be placed directly in the node), and the bounding box of the data element.

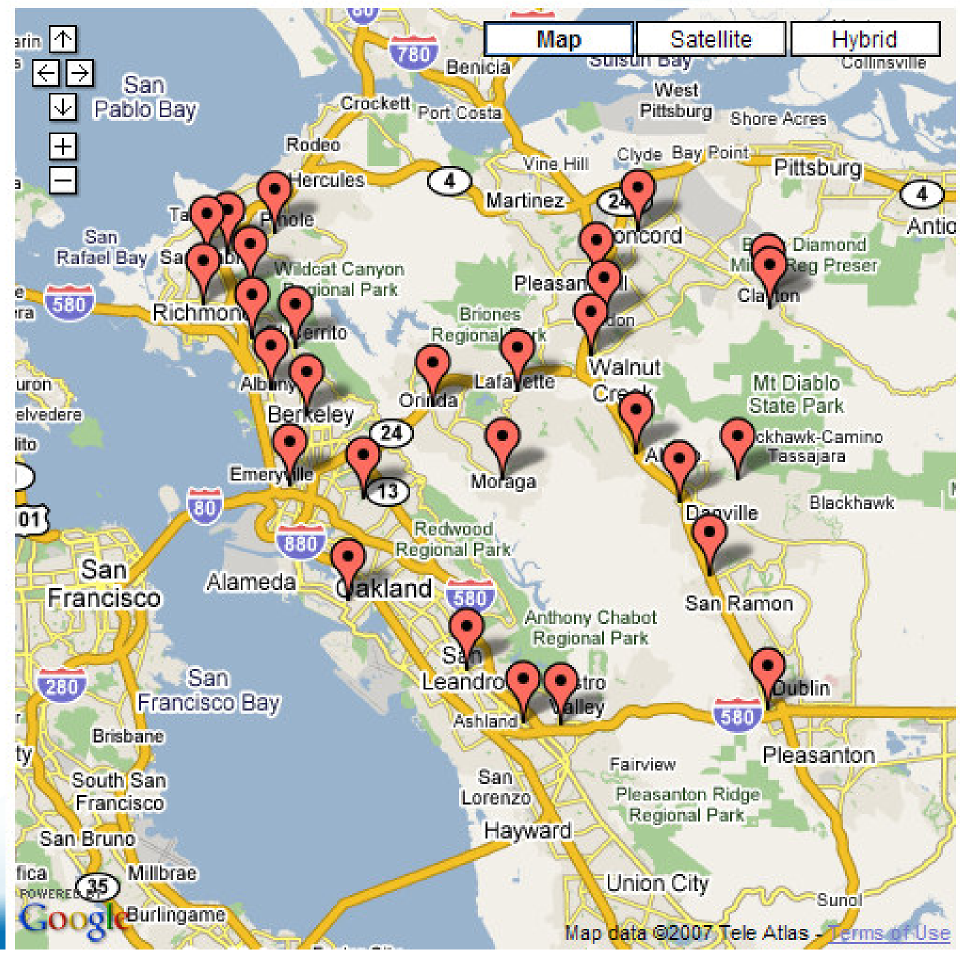


Simple example of R-tree for a 2D rectangles

Sample Geospatial Primitives

- (geo-bounding-box ?x +minlat +maxlat +minlon +maxlon)
- (geo-box-around +x ?y +miles)
- (geo-distance +x +y ?dist)
- (geo-radius-around +x ?y +miles)
- (polygon-in ?p1 ?p2)
- (polygon-touch ?p1 ?p2)
- (polygon-overlap ?p1 ?p2)
- Etc.





Now with SPARQL Support [1]

```
PREFIX fr:        <http://franz.com/ns/allegrograph/3.0/geospatial/>
  PREFIX geo:         <http://www.geonames.org/ontology#>
  PREFIX country:     <http://www.geonames.org/Countries#>

  SELECT ?placename ?population WHERE {
    GEO OBJECT
    HAVERSINE ( ?londonpos, 50 MILES ) {
      ?place fr:pos ?pos ;
             geo:name ?placename ;
             geo:population ?population ;
             geo:countryCode ?cc
    }
    WHERE {
      # Select London, UK.
      ?london geo:name 'London' ;
              geo:countryCode 'GB' ;
              fr:pos ?londonpos .
    }
    FILTER (?population > 25000)
```

Now with SPARQL Support [2]

- ; cpu time (non-gc) 30 msec user, 10 msec system
- ; cpu time (total) 40 msec user, 10 msec system
- ; (#({Worthing} {99110}) #({Brighton} {139001}) #({Hove} {75174})
- #({Burgess Hill} {31183}) #({Haywards Heath} {29660}) #({Horsham} {50680})
- #({Crawley} {107061}) #({East Grinstead} {26523})
- #({Royal Tunbridge Wells} {61075}) #({Tonbridge} {36894})
- #({Basingstoke} {96348}) #({Farnham} {36971}) #({Aldershot} {61339})
- #({Farnborough} {59902}) #({Guildford} {71873}) #({Reigate} {52123})
- #({Redhill} {51559}) #({Sevenoaks} {27871}) #({Maidstone} {90894})
- #({Fleet} {34218}) ...)

GeoNames - Geospatial in RDF

Using GeoNames

- coordinates for 6,445,201 places
- 109,568,417 RDF triples

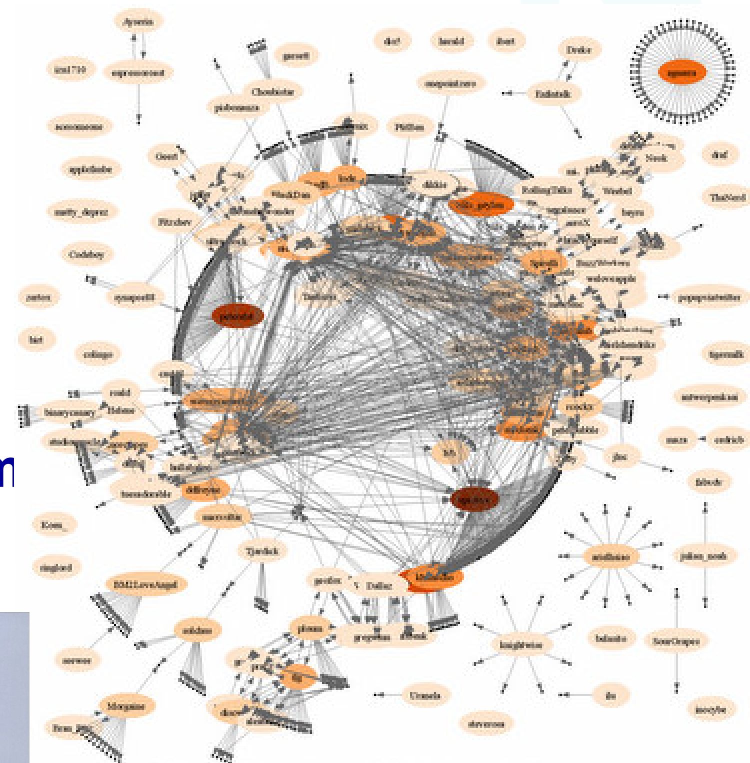
AllegroGraph requires only 95 msec real time to return the 502 entries within a 3 mile radius of the Franz Inc offices.

Questions in SNA (1) -

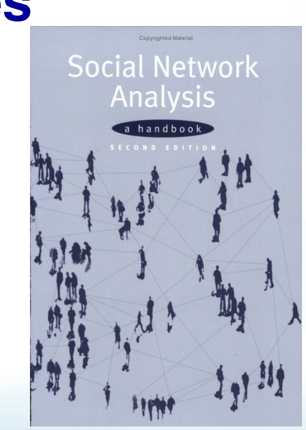
How far is Actor1 from Actor2?

- **Degrees of separation**
  - **How far is P1 from P2**

- **Connection strength**
  - **How many shortest paths from P1 to P2 through a series of predicates and rules**

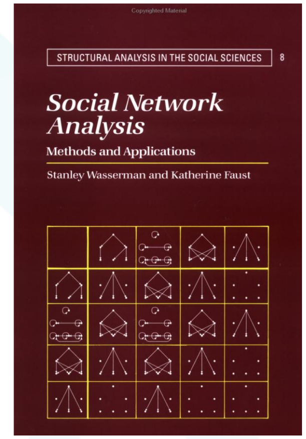


Twitter Friends van Belgische Twitteraars

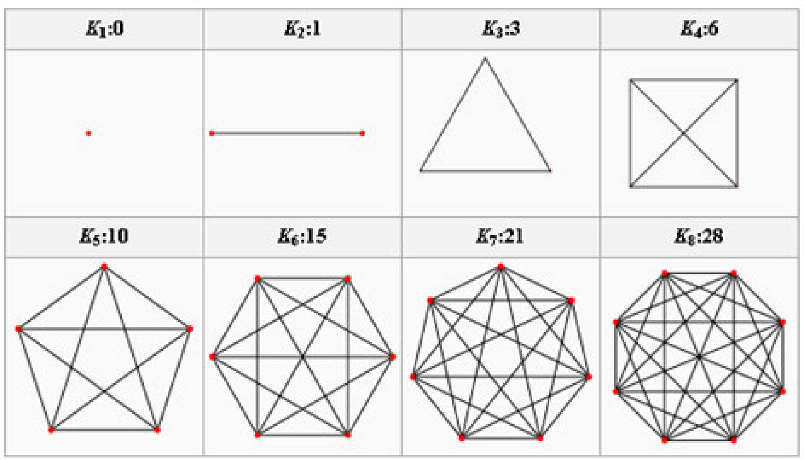


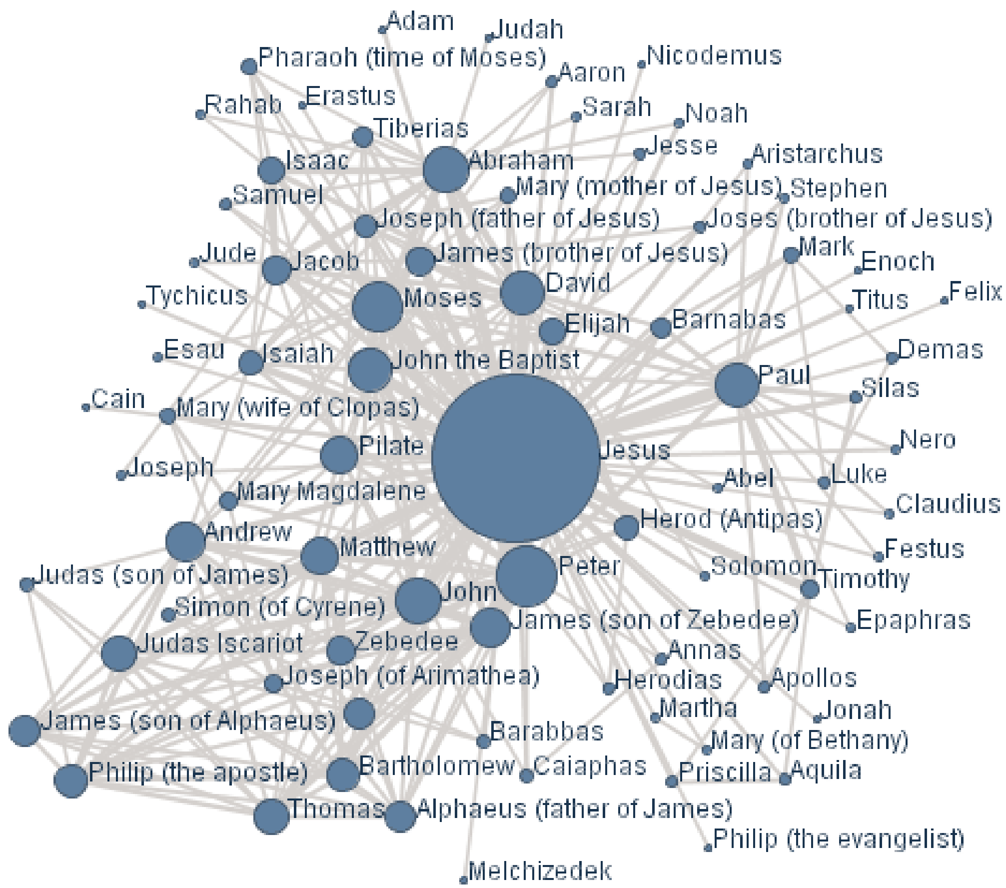
Questions in SNA (2)

In what groups is this actor?



- **Find the ego-network & cliques around a person**

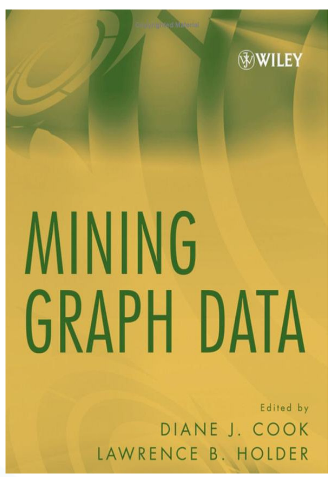


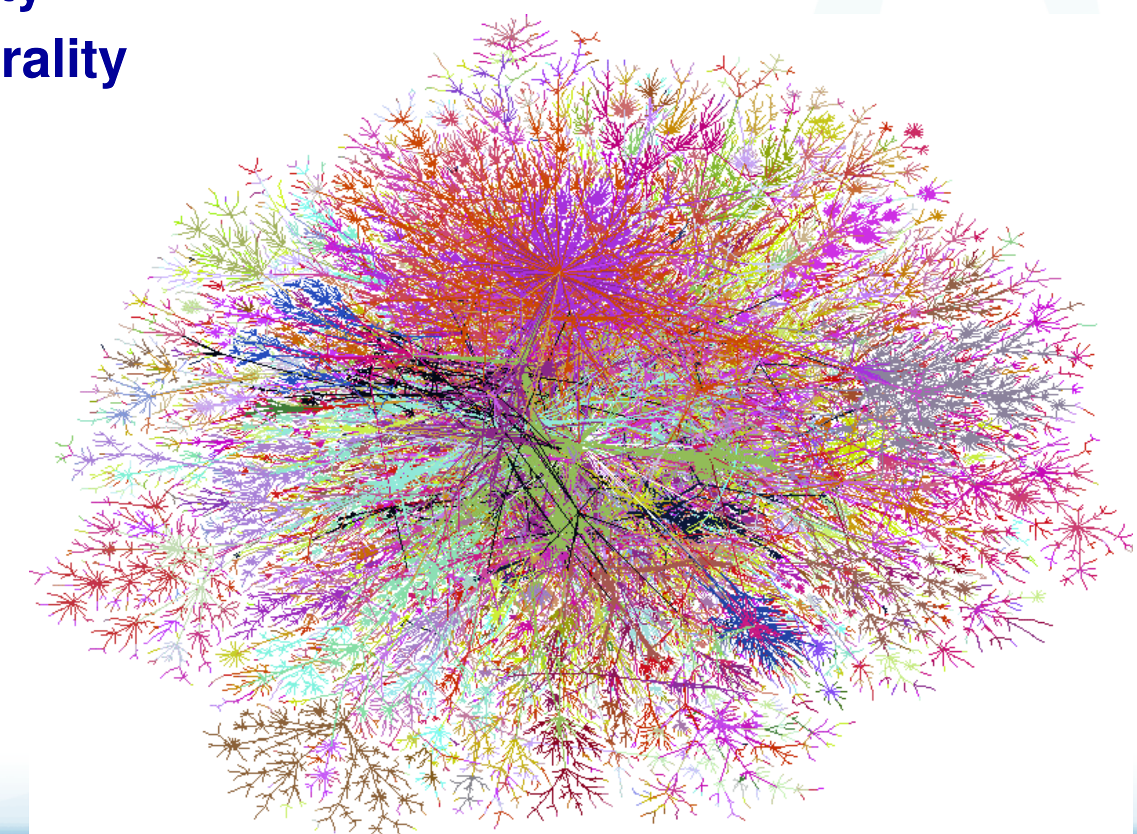


Questions in SNA (3)

Who are the key players in a network and how important is an actor?

- In-degree, out-degree
- Actor degree centrality
- Actor closeness centrality
- Actor betweenness centrality

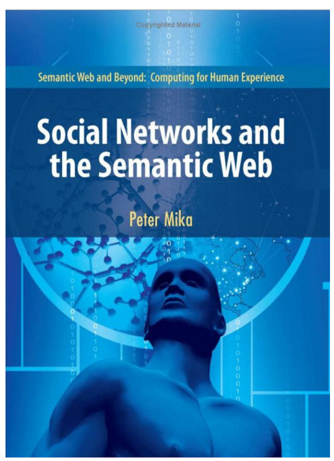


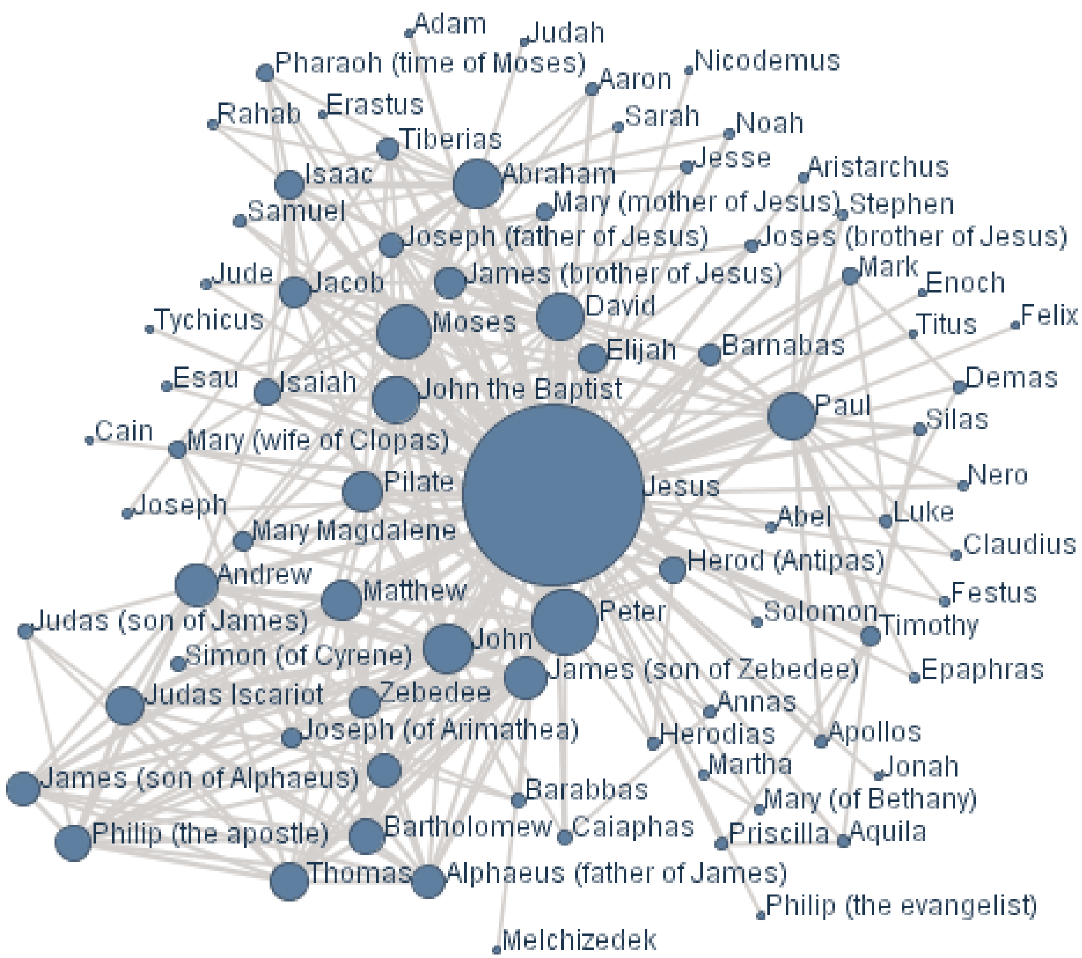


Franz Inc.
Web 3.0's Database

Questions in SNA (4) : Has the group a leader, is the group cohesive?

- **Group centralization**
- **Group cohesiveness**





GeoWeb 2008

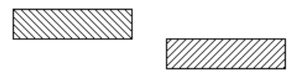
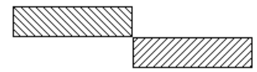
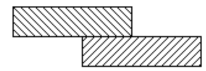
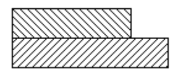
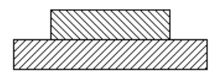
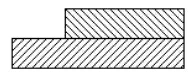
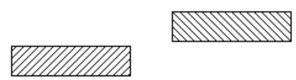
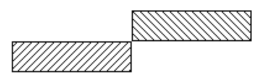
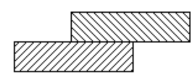
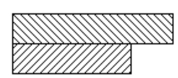
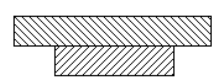
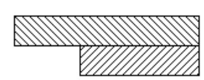
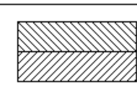
Temporal Primitives

- Based on Allen's temporal logic: see
  `http://www.ai.sri.com/snark/tutorial/tutorial.html#htoc46`
- Adhere to the convention to encode startimes and endtimes with franz:start and franz:end and enjoy efficient temporal primitives

```
add-triple (event1, franz:start, "Thu Sep 27 14:21:43 2007")
add-triple (event1, franz:end,  "Thu Sep 27 14:21:43 2007")
```

- The primitives make temporal reasoning user friendly

Allen's 13 temporal Interval primitives

	(interval-before ?e1 ?e2)
	(interval-meets ?e1 ?e2)
	(interval-overlaps ?e1 ?e2)
	(interval-starts ?e1 ?e2)
	(interval-during ?e1 ?e2)
	(interval-finishes ?e1 ?e2)
	(interval-after ?e1 ?e2)
	(interval-met-by ?e1 ?e2)
	(interval-overlapped-by ?e1 ?e2)
	(interval-started-by ?e1 ?e2)
	(interval-contains ?e1 ?e2)
	(interval-finished-by ?e1 ?e2)
	(interval-cotemporal ?e1 ?e2)

Thank you

www.franz.com

AllegroGraph – Free Edition
Learning Center

And finally:

- We are looking for testers for our new
  - .Net C# interface
  - Python interface
  - Jena interface