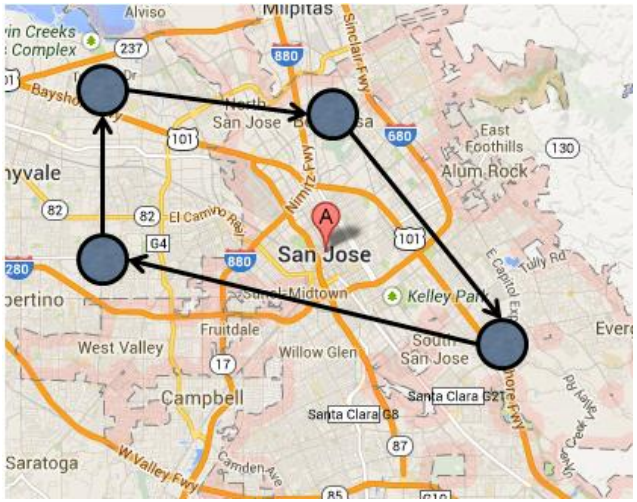
Prolog, Rules, Reasoning and SPARQLing Magic in the real world

Franz Inc

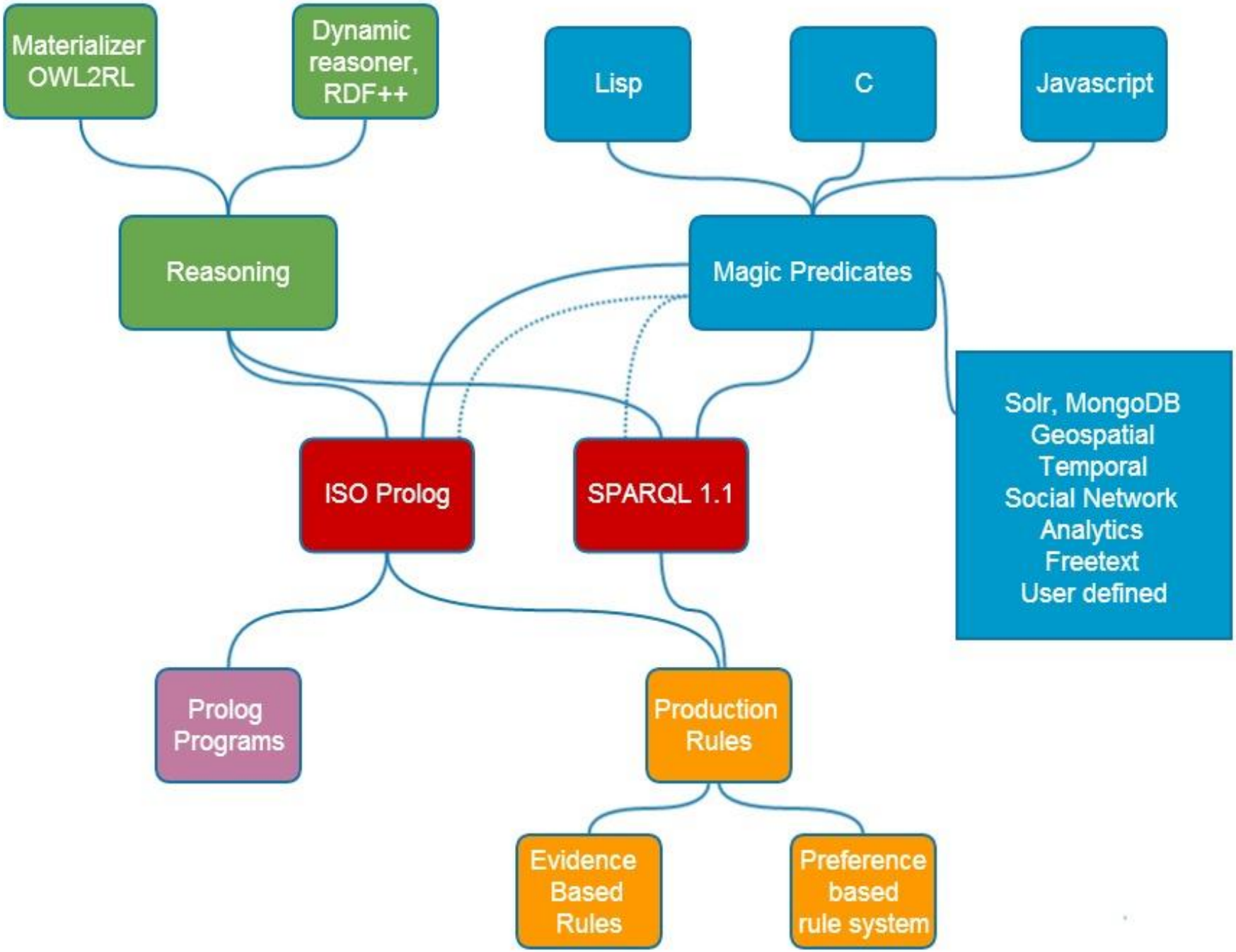
Contents

- How do we fit it all together: rules and prolog and reasoning and magic predicates and SPARQL

- Use case: BigBank
  - Event view of the world
  - Using Social Network Analysis, Geo and Temporal reasoning (embedded in Prolog and SPARQL 1.1)
  - Fraud detection



Franz Inc

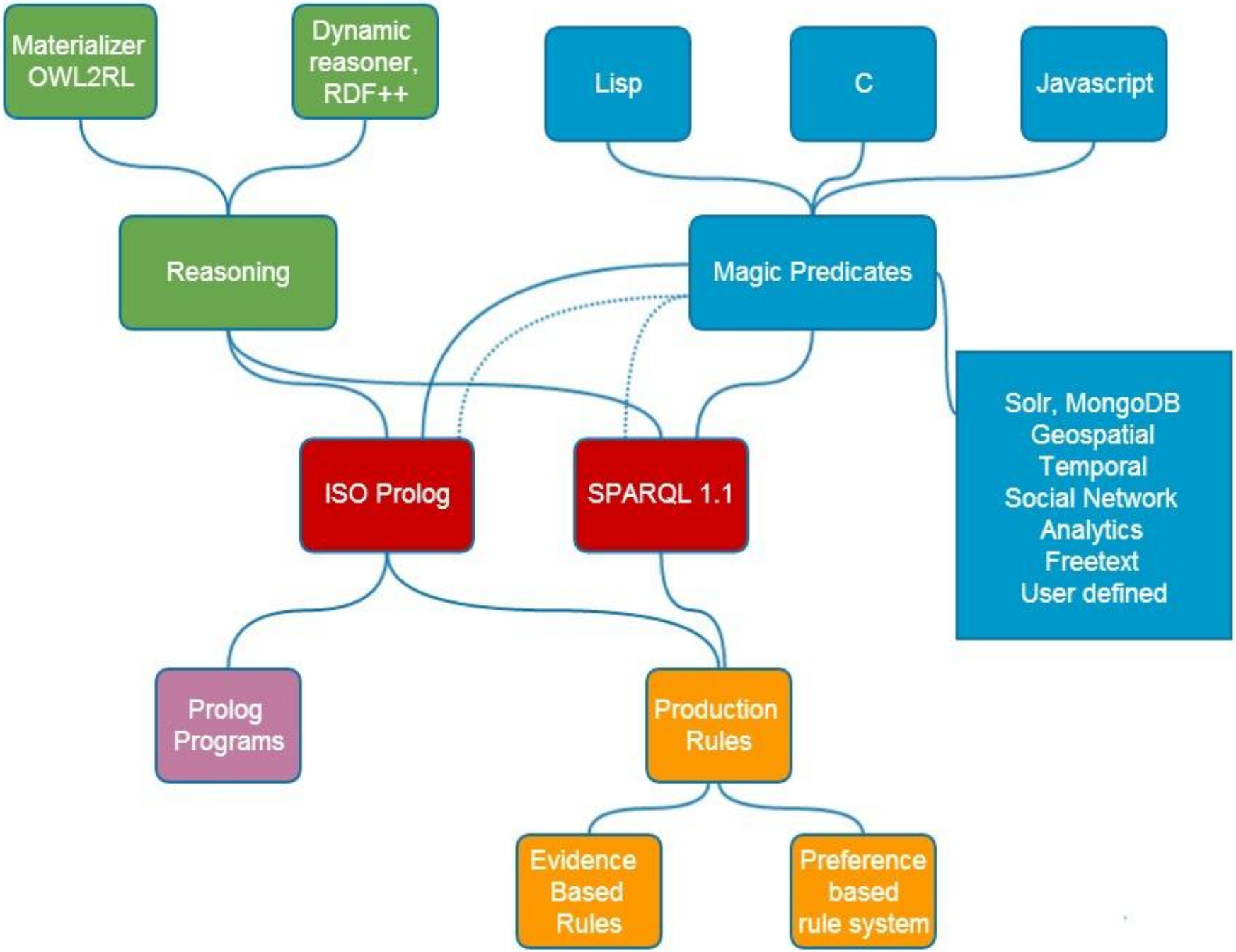

Materializer OWL2RL
Dynamic reasoner, RDF++
Lisp
C
Javascript
Reasoning
Magic Predicates
Solr, MongoDB
Geospatial
Temporal
Social Network
Analytics
Freetext
User defined
ISO Prolog
SPARQL 1.1
Prolog Programs
Production Rules
Evidence Based Rules
Preference based rule system

Reasoning

- For Static data sets we recommend the Materializer
  - OWL2RL, make it as complex as you want
- For dynamic applications we recommend the dynamic, RDFS++ reasoner
  - Mostly good for class hierarchies
  - Don’t go overboard with sub property chains
  - Keep your transitive relations simple
  - Avoid sameas if you can
  - Avoid inverse if you are in control of your application
- You can turn it on or off
- Both SPARQL and Prolog can use it.

Both Prolog and SPARQL can use the reasoner

```
with reasoning OFF

(select ?type (q !fr:Jans !rdf:type ?type))

select ?type {q fr:Jans a ?type }

{Human}

with reasoning ON

{Human}
{Mammal}
{Animal}
{Thing}
```


Materializer OWL2RL
Dynamic reasoner, RDF++
Lisp
C
Javascript
Reasoning
Magic Predicates
Solr, MongoDB
Geospatial
Temporal
Social Network
Analytics
Freetext
User defined
ISO Prolog
SPARQL 1.1
Prolog Programs
Production Rules
Evidence Based Rules
Preference based rule system

Magic Properties

- Predicates as functions (instead of real triples)
- Here is a SPARQL query with three magic properties from the Social Network Analysis Library

AllegroGraph WebView 4.12 repository big-bankfraud4.db

centrality

Query language: SPARQL Query planner: default Result limit: 100

```
1 prefix sna:  <http://franz.com/ns/allegrograph/4.11/sna/>
2 select ?email ?centrality {
3   ?who bb:email 'Sonia.Madrid@gmail.com' .
4   ?group sna:egoGroup ( bb:paid ?who 3 ) .
5   (?member ?centrality) sna:actorDegreeCentrality (bb:paid ?group) .
6   FILTER(?member != ?who)
7   ?member bb:email ?email
8 } order by desc(?centrality)
9 limit 10
```

Execute Save as centrality Add to repository

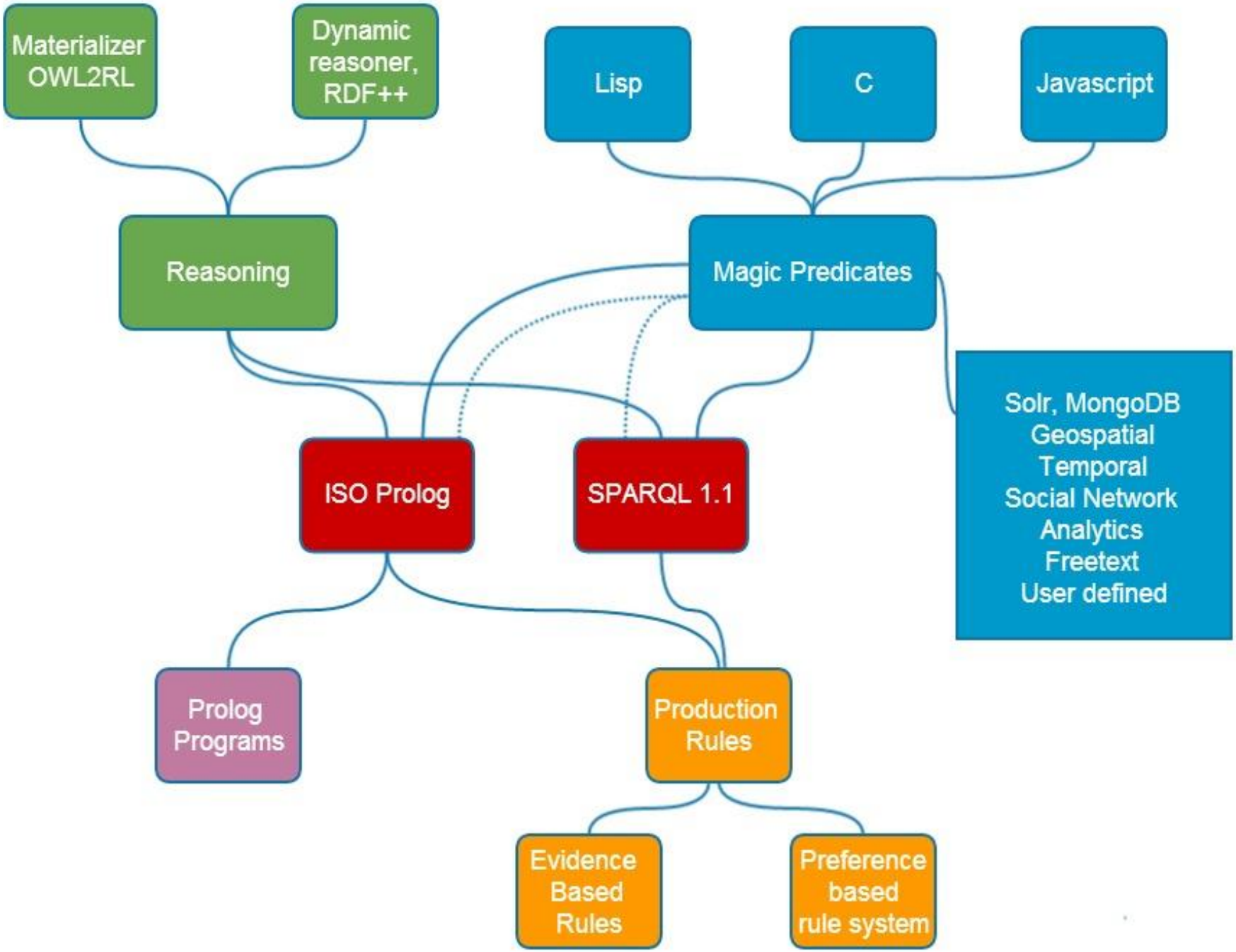
Result

Download as SPARQL JSON

email	centrality

SPARQL Magic Properties

Table of Contents


Materializer OWL2RL
Dynamic reasoner, RDF++
Lisp
C
Javascript
Reasoning
Magic Predicates
Solr, MongoDB
Geospatial
Temporal
Social Network
Analytics
Freetext
User defined
ISO Prolog
SPARQL 1.1
Prolog Programs
Production Rules
Evidence Based Rules
Preference based rule system

Prolog

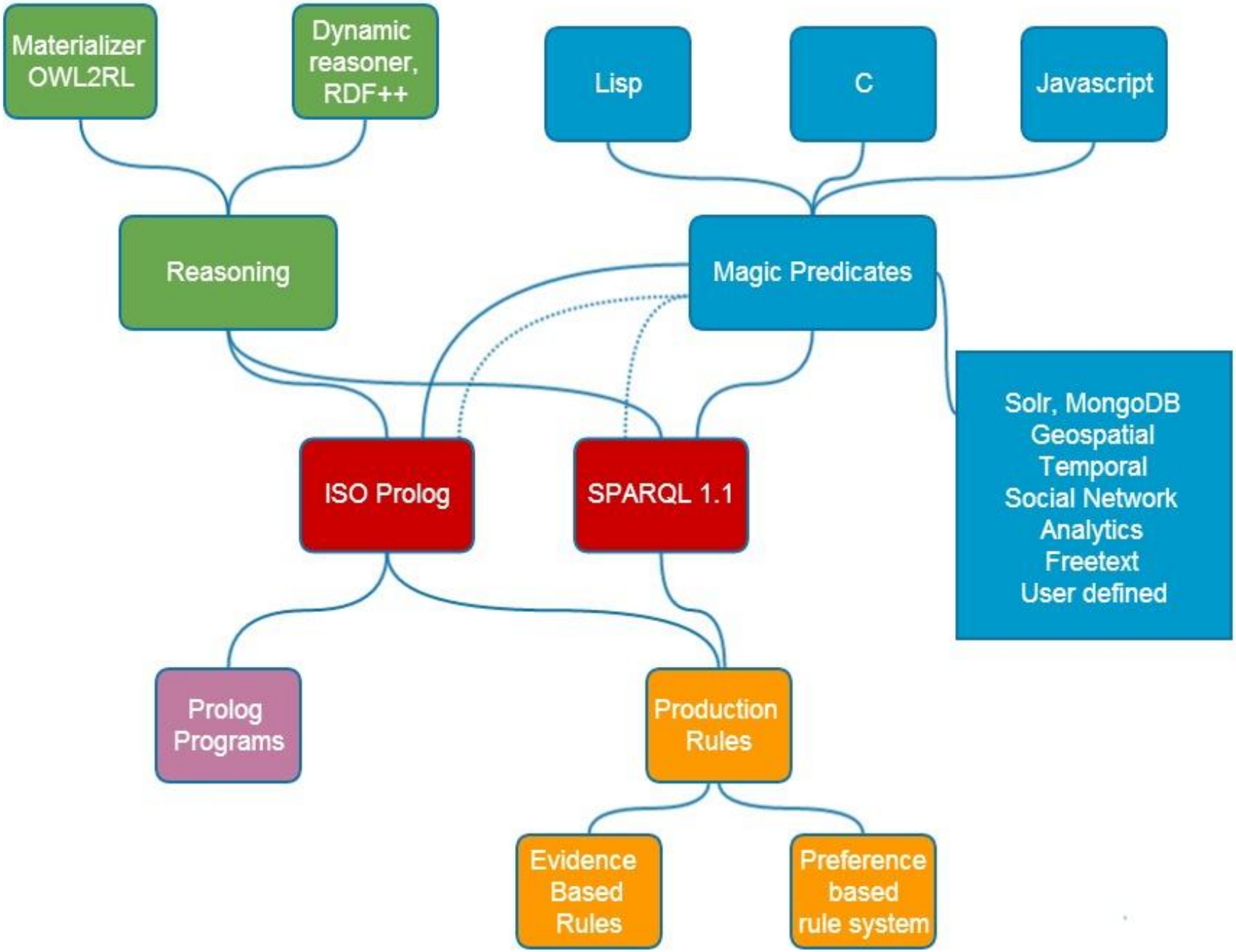
- ISO Standard
- A full Query and Programming Language
  - Where SPARQL is only an access language
- How do customers use it:
  - Create libraries and domain specific languages
  - Build up hierarchies of functors
- So that the final queries and rules are simpler to read and maintain.

A silly example of building up hierarchy of functors

```
(<-- (male ?x) (q ?x !o:sex !o:male))
(<-- (female ?x) (q ?x !o:sex !o:female))

(<-- (father ?x ?y)
     (male ?x)
     (q ?x !o:has-child ?y))

(<-- (mother ?x ?y)
     (female ?x)
     (q ?x !o:has-child ?y))

(<-- (parent ?x ?y)
     (or (father ?x ?y)
         (mother ?x ?y)))

(<-- (ancestor ?x ?y)
     (parent ?x ?y))

(<-  (ancestor ?x ?y)
     (parent ?x ?z)
     (ancestor ?z ?y))
```

Real World Example..

```
;; Rule 6

(defrule poi-has-partner-through-closeness (poi)
  (select ()
    (most-recent-event (?? poi) ?event)
    (q ?event !f:location ?loc)
    (nearby ?other-event ?loc !f:location 3)
    (not (= ?event ?other-event))
    (most-recent-event-p ?other-event)
    (q ?other-event !f:actor ?poi2)
    (suggest (?? poi) 1)))

;; Rule 7

(defrule closeness-from-poi-to-target (poi)
  (select ()
    (most-recent-event (?? poi) ?event1)
    (member ?target (?? *targets*))
    (most-recent-event ?target ?event2)
    (distance-between-events ?event1 ?event2 ?dist)
    (lisp ?value (if* (< ?dist 1) then 10
                  elseif (< ?dist 2) then 5
                  elseif (< ?dist 3) then 4
                  elseif (< ?dist 10) then 1 else 0))
    (suggest (?? poi) ?value)))
```


Materializer OWL2RL
Dynamic reasoner, RDF++
Lisp
C
Javascript
Reasoning
Magic Predicates
Solr, MongoDB
Geospatial
Temporal
Social Network
Analytics
Freetext
User defined
ISO Prolog
SPARQL 1.1
Prolog Programs
Production Rules
Evidence Based Rules
Preference based rule system

Production Rules

- Can be written in Prolog or Sparql

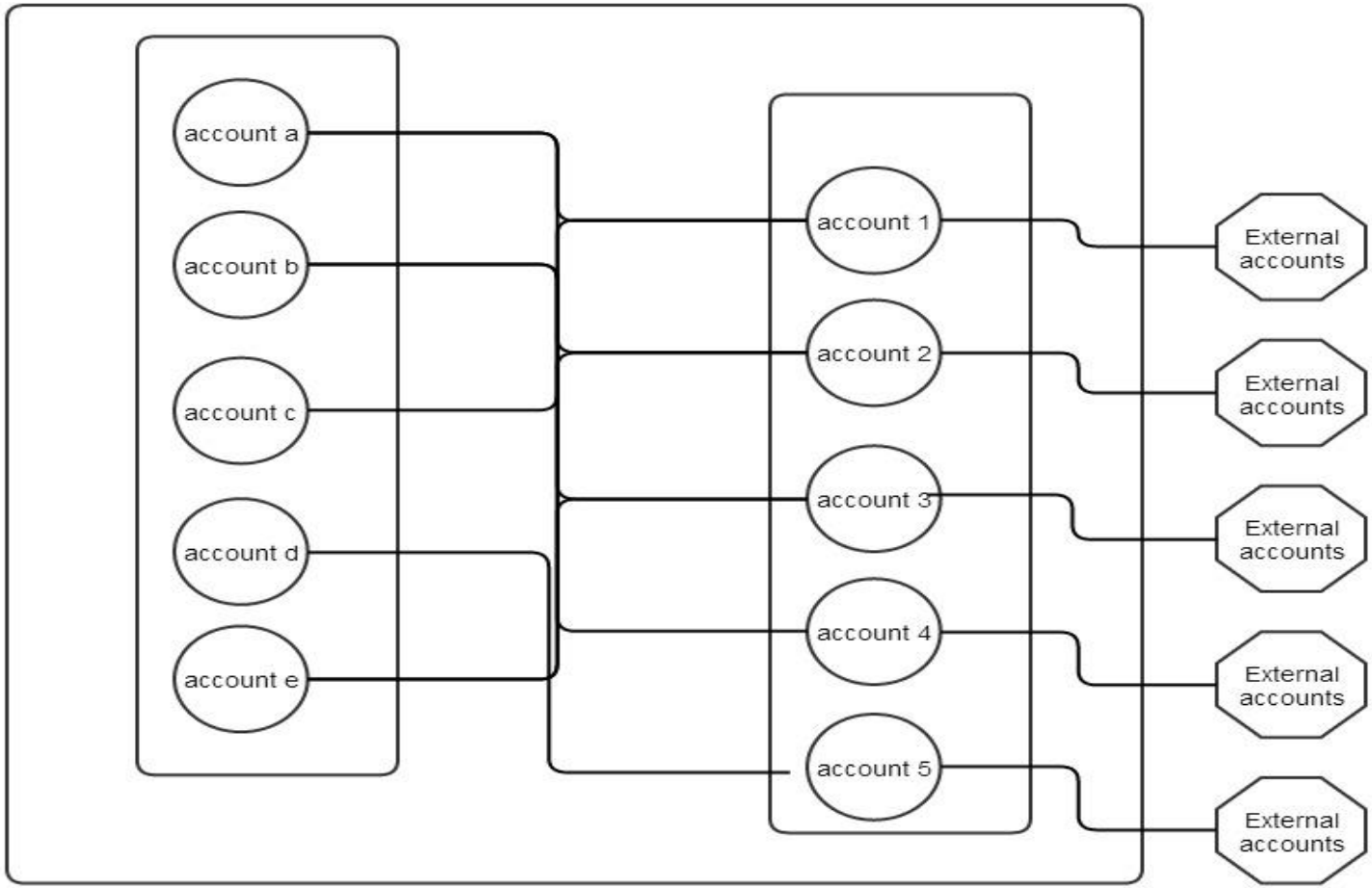
- Evidence based rule systems
  - Applied in a complex event handling situation
  - Apply 100’s of rules after every few seconds or after X new events
  - Add or subtract points from hypotheses, or the danger level of an entity or the threat level of a potential target.

- Also mix with our Bayesian Belief Network interface to Netica

Real World Example..

```
;; Rule 6

(defrule poi-has-partner-through-closeness (poi)
  (select ()
    (most-recent-event (?? poi) ?event)
    (q ?event !f:location ?loc)
    (nearby ?other-event ?loc !f:location 3)
    (not (= ?event ?other-event))
    (most-recent-event-p ?other-event)
    (q ?other-event !f:actor ?poi2)
    (suggest (?? poi) 1)))

;; Rule 7

(defrule closeness-from-poi-to-target (poi)
  (select ()
    (most-recent-event (?? poi) ?event1)
    (member ?target (?? *targets*))
    (most-recent-event ?target ?event2)
    (distance-between-events ?event1 ?event2 ?dist)
    (lisp ?value (if* (< ?dist 1) then 10
                  elseif (< ?dist 2) then 5
                  elseif (< ?dist 3) then 4
                  elseif (< ?dist 10) then 1 else 0))
    (suggest (?? poi) ?value)))
```

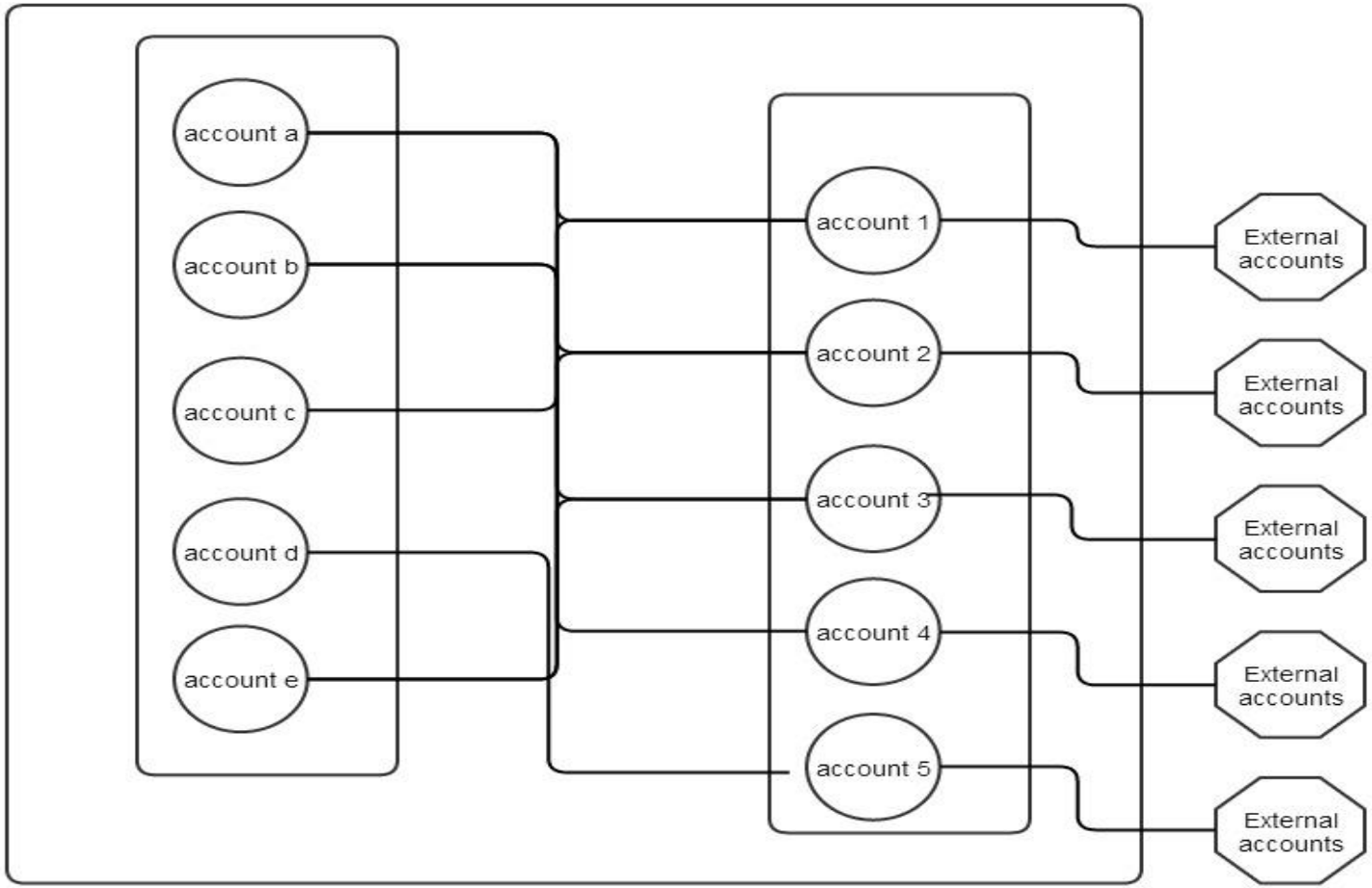
A Big Asian Bank: find fraud



Franz Inc

A Big Asian Bank: find fraud

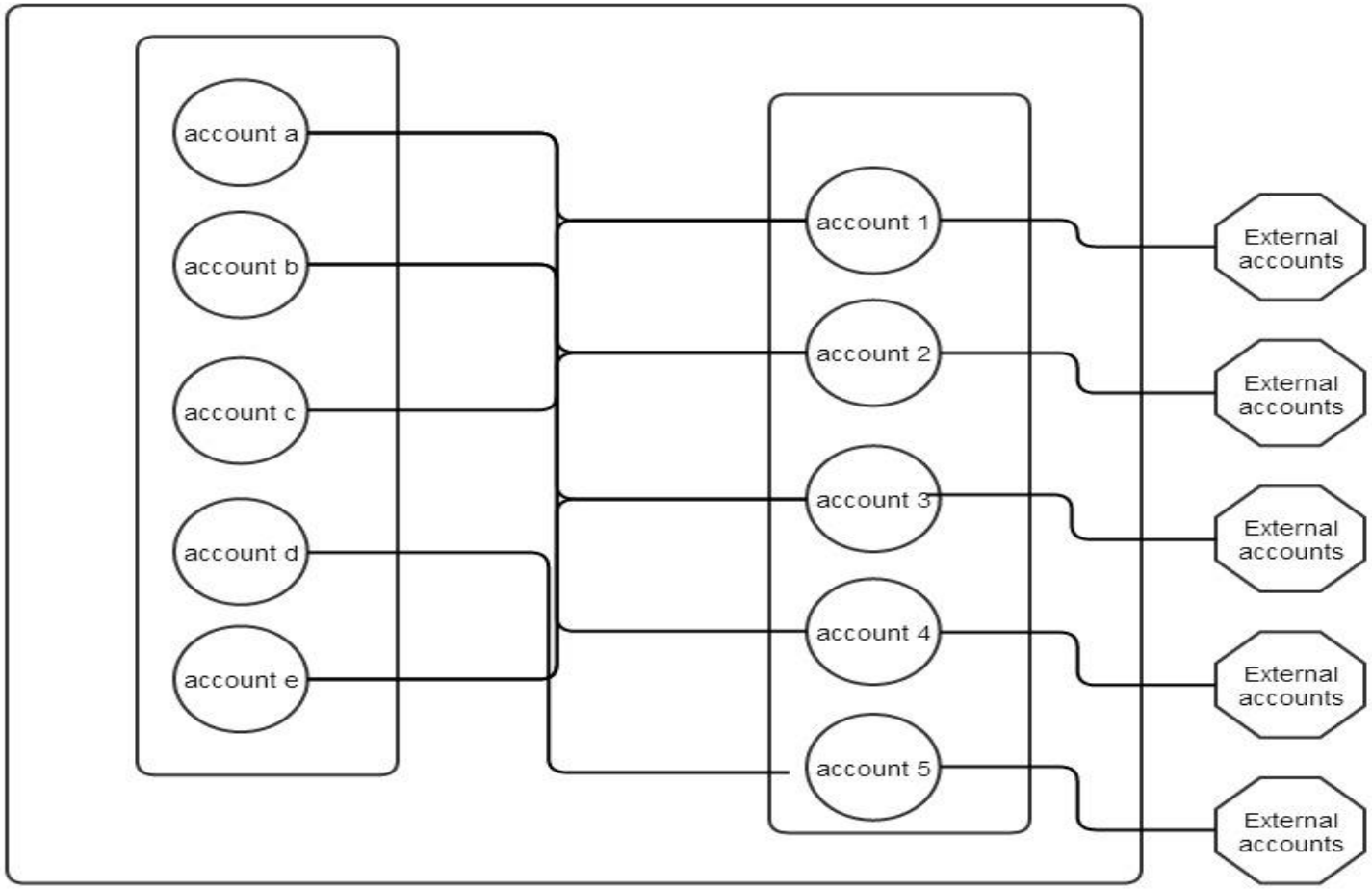
1. > 5 accounts opened in 1 hour



Franz Inc

A Big Asian Bank: find fraud

1. > 5 accounts opened in 1 hour

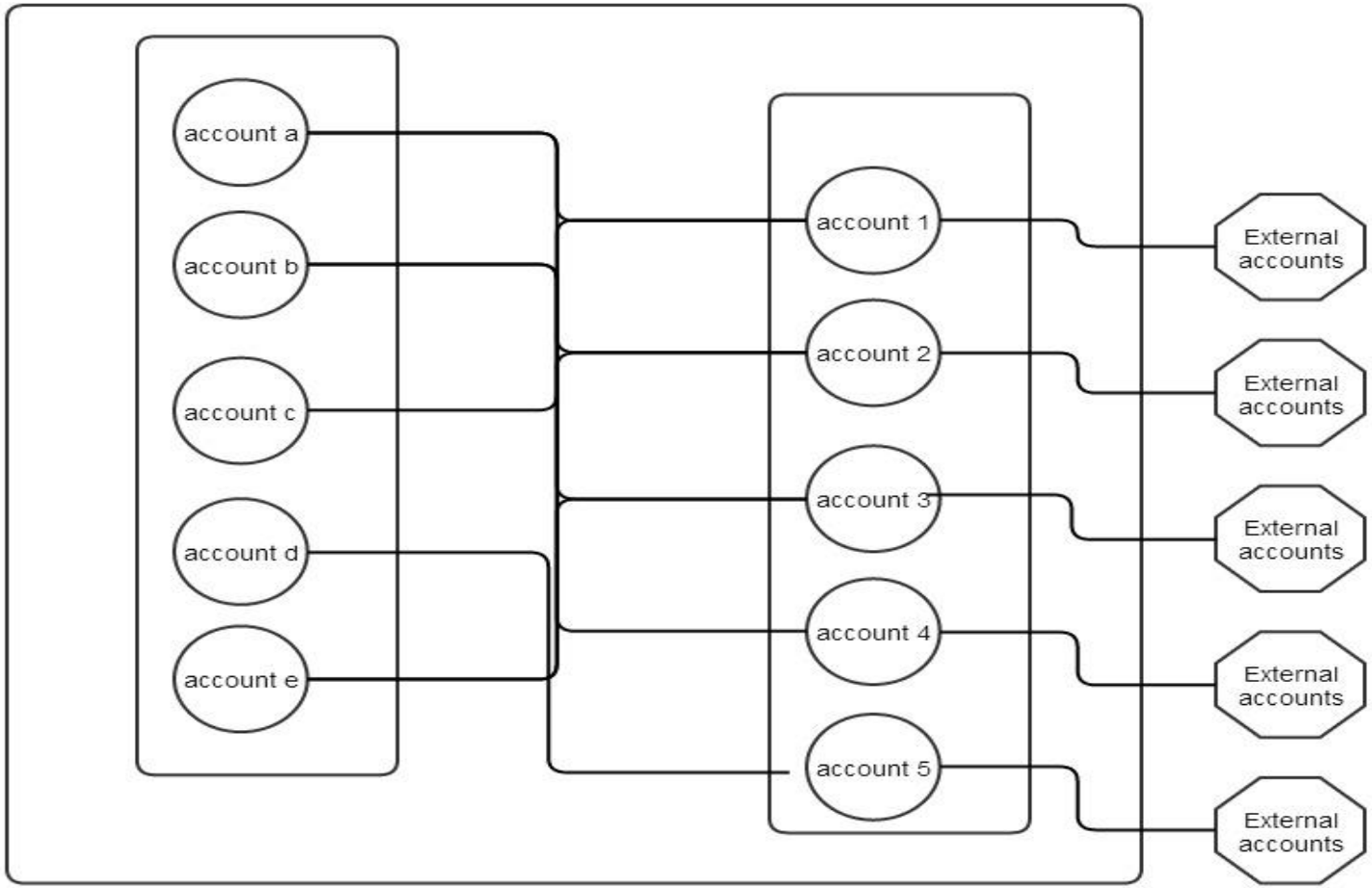


2. Within 30 minutes other accounts pay large money into new accounts

Franz Inc

A Big Asian Bank: find fraud

1. > 5 accounts opened in 1 hour



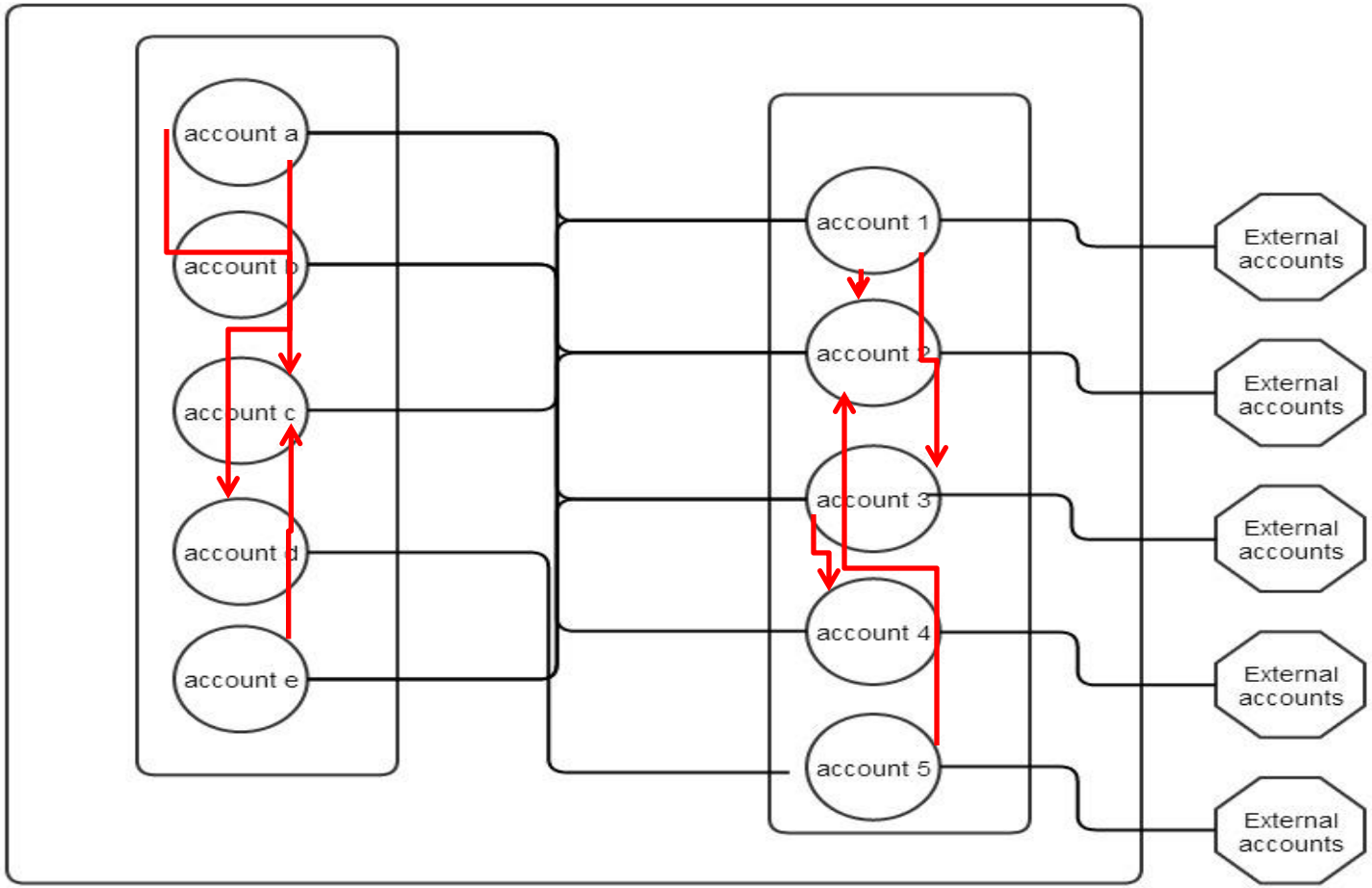
3. In the next hour money flows
Into external accounts and credit cards

2. Within 30 minutes other accounts
pay large money into new accounts

Franz Inc

A Big Asian Bank: find fraud

1. > 5 accounts opened in 1 hour



3. In the next hour money flows
Into external accounts and credit cards

2. Within 30 minutes other accounts
pay large money into new accounts

4. Evidence that accounts are
Somehow related

Franz Inc

A very simple event ontology

A type

- Meetings, communications event, financial transactions, visits, attack/truce, insurance claims, purchase orders
- RDFS++ reasoning

A list of actors

- Social Network Analysis

A place

- GeoSpatial Reasoning

A Start-time and possible an end-time

- Temporal Reasoning

Anything else that describes the event

Events are everywhere

- Communication events
  - Telephone, SMS, Email
- Social media
  - Tweets: sender, receiver, topics about people..
- Payments
  - Online, credit card,
- Insurance
  - Claims, payments
- E-commerce
  - Purchase orders
  - Website visits

For our analysis we use

- Social Network Analysis and Graph Analytics
- Temporal reasoning
- Geospatial reasoning

Embedded in Prolog and SPARQL 1.1

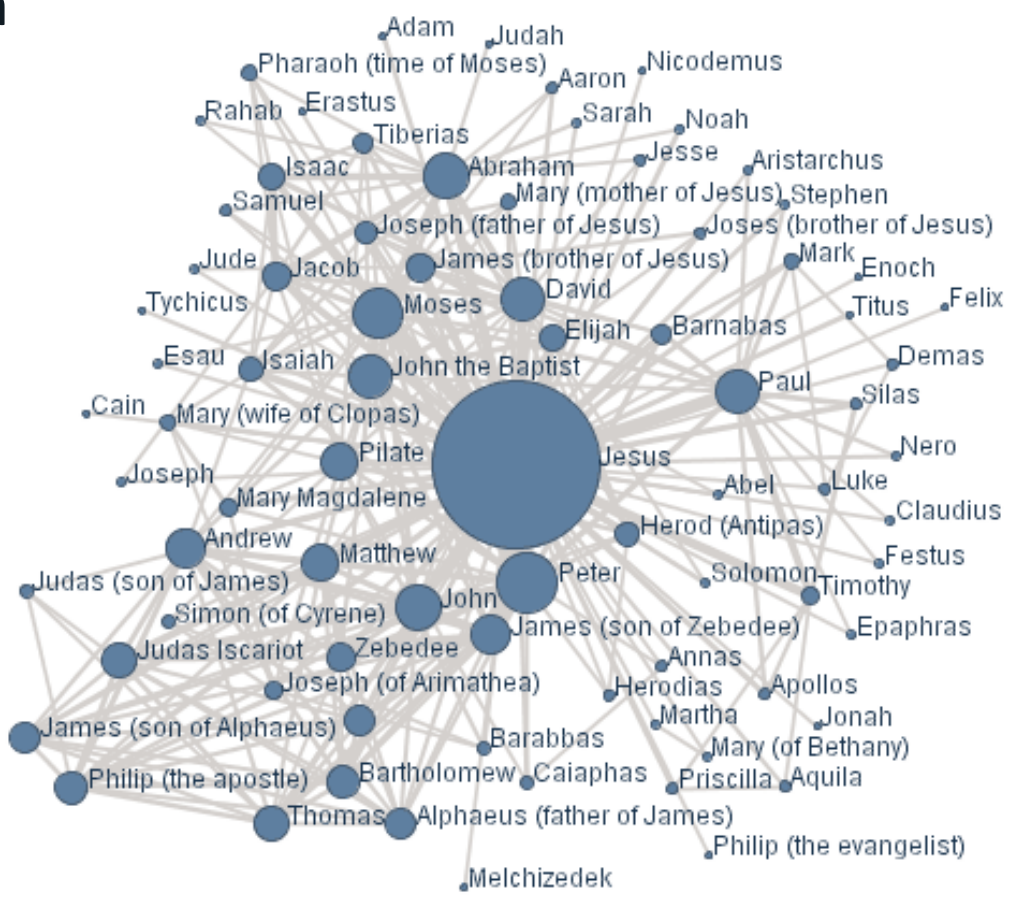
Social Network Analysis answers 4 questions

How far is P1 from P2
and how strong is the relation

To what groups does this person belong (ego groups, cliques?)

How important is this person in the group?

Does this group have a leader, how cohesive are they?



Franz Inc

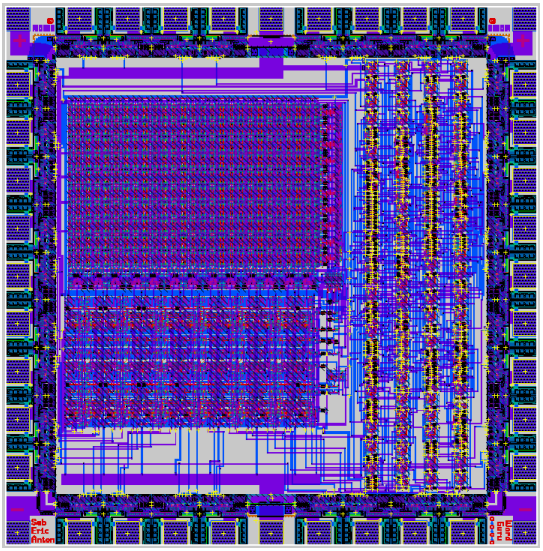
Geospatial Reasoning

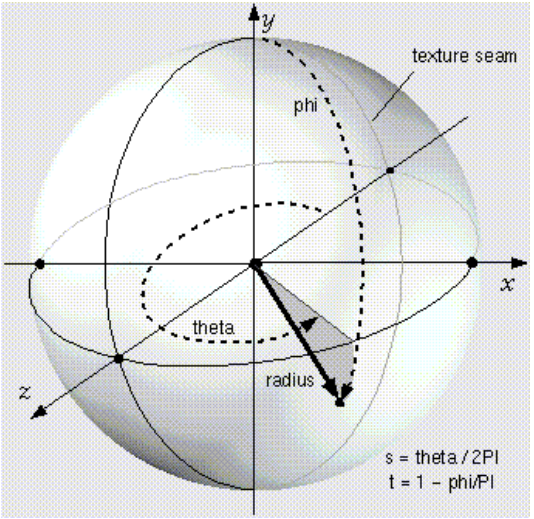
Make the following super efficient

- 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



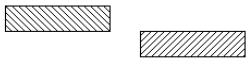
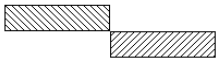
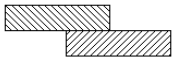
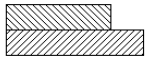
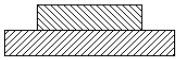
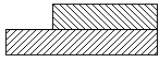
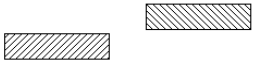
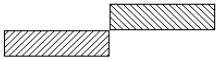
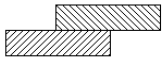
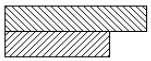
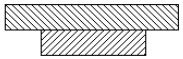
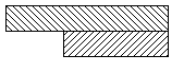
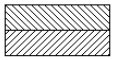


Franz Inc

Temporal Reasoning

Adhere to our convention
- to encode StartTimes and EndTimes and enjoy efficient temporal primitives

Implementation of
- Allen's interval logic 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)

Activity recognition

Find all 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
   (interval-during ?event "2013-12-01" "2013-12-18")        Temporal
   (geo-box-around geoname:Berkeley ?event 5 miles)          Spatial
   !)
```

Back to the Fraud example

Accounts

- Opened at some time, at an IP address, with email, in a particular location.

Events

- Payments between accounts from some IP address, at some location at some time. Sender, receivers, amounts

Locations with latitudes and longitudes

Bruce.Rosie@hitachi.com

Revisit Show All Triples

Property Left-click a property to collapse or expand its rows. Shift-left-click to add a property's triples to the graph	**Value** Left-click a value to visit it in the table view and add the triple to the graph view. Right-click a value or press M for a menu of navigation commands. J moves down a row, K moves up, and L moves to the other column.	
Account Date Time UT	3256525427	
Account IP	1.323.225.6	
Account Number	6673190296	
Account Place	Gun Barrel City	
Email	Bruce.Rosie@hitachi.com	
Label	Bruce.Rosie@hitachi.com	
Paid	Ursula.Bouillion@franz.com	00005976
	Ursula.Bouillion@franz.com	00005979
	Ginny.Blick@easychair.org	00005973
	Tracey.Spight@gmail.com	00005975
	Curt.Gallaher@franz.com	00005978
	Clarence.Baran@franz.com	00005977
	Kathryn.Klapp@gmail.com	00005974
Paid-medium	Ursula.Bouillion@franz.com	00005976
	Ursula.Bouillion@franz.com	00005979
Paid-medium-large	Curt.Gallaher@franz.com	00005978
	Clarence.Baran@franz.com	00005977
Paid-normal	Kathryn.Klapp@gmail.com	00005974
Paid-small	Ginny.Blick@easychair.org	00005973
Paid-very-small	Tracey.Spight@gmail.com	00005975
Time	2003-03-13T06:23:47Z	
Type	Buyer	
is Paid of	Jeremy.Whitesides@franz.com	00004501
	Kathryn.Errico@wellsfargo.com'	00004831
	Angel.Readnour@vulcan.com	00004479
	Dave.Eichhorn@sgi.com	00003315
	Nellie.Hahn@reply.facebook.com	00003252
is Paid-large of	Angel.Readnour@vulcan.com	00004479
is Paid-medium of	Kathryn.Errico@wellsfargo.com'	00004831
is Paid-medium-small of	Jeremy.Whitesides@franz.com	00004501
is Paid-normal of	Dave.Eichhorn@sgi.com	00003315
is Paid-very-small of	Nellie.Hahn@reply.facebook.com	00003252

Gun Barrel City

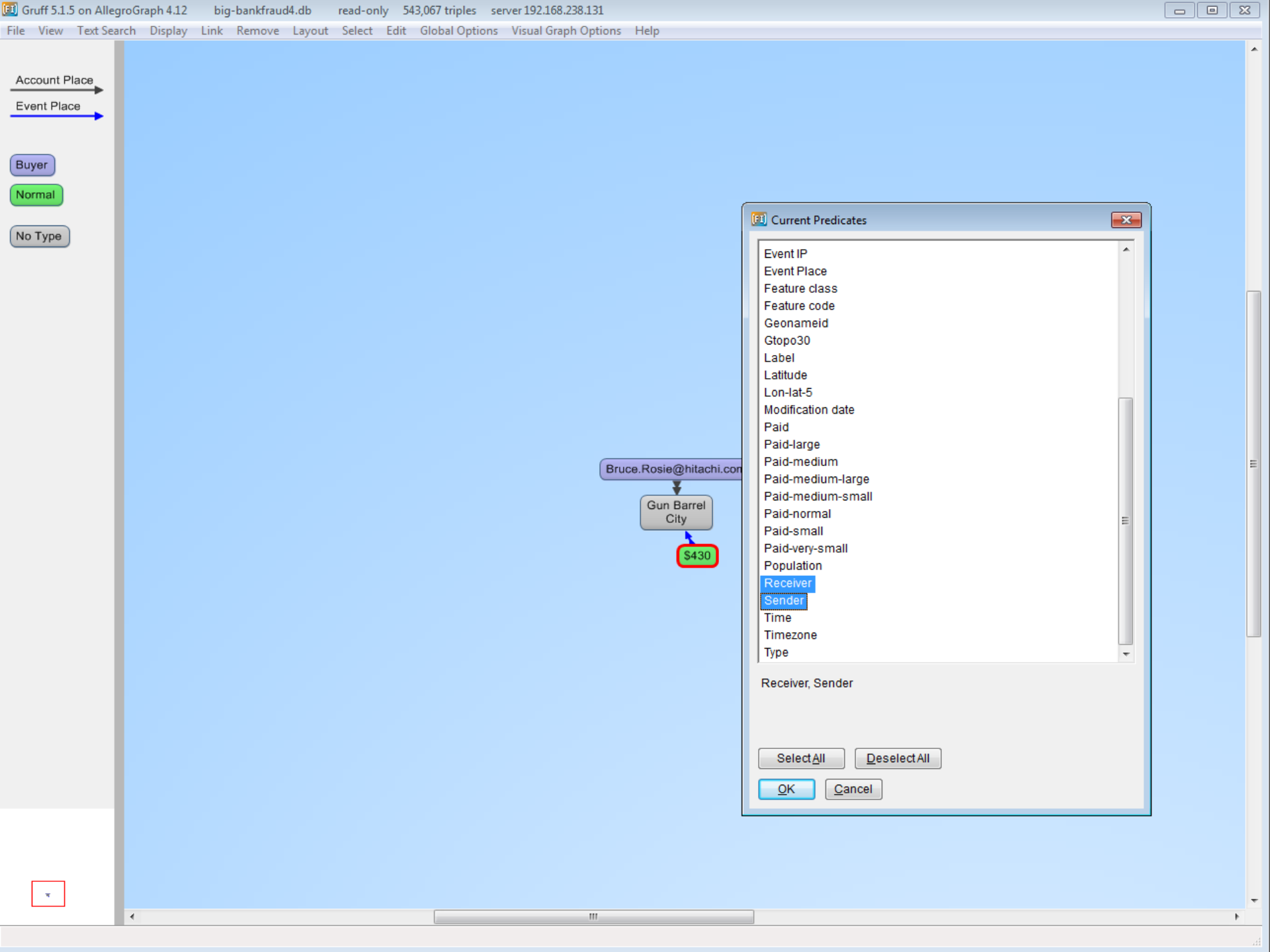
Revisit Show All Triples

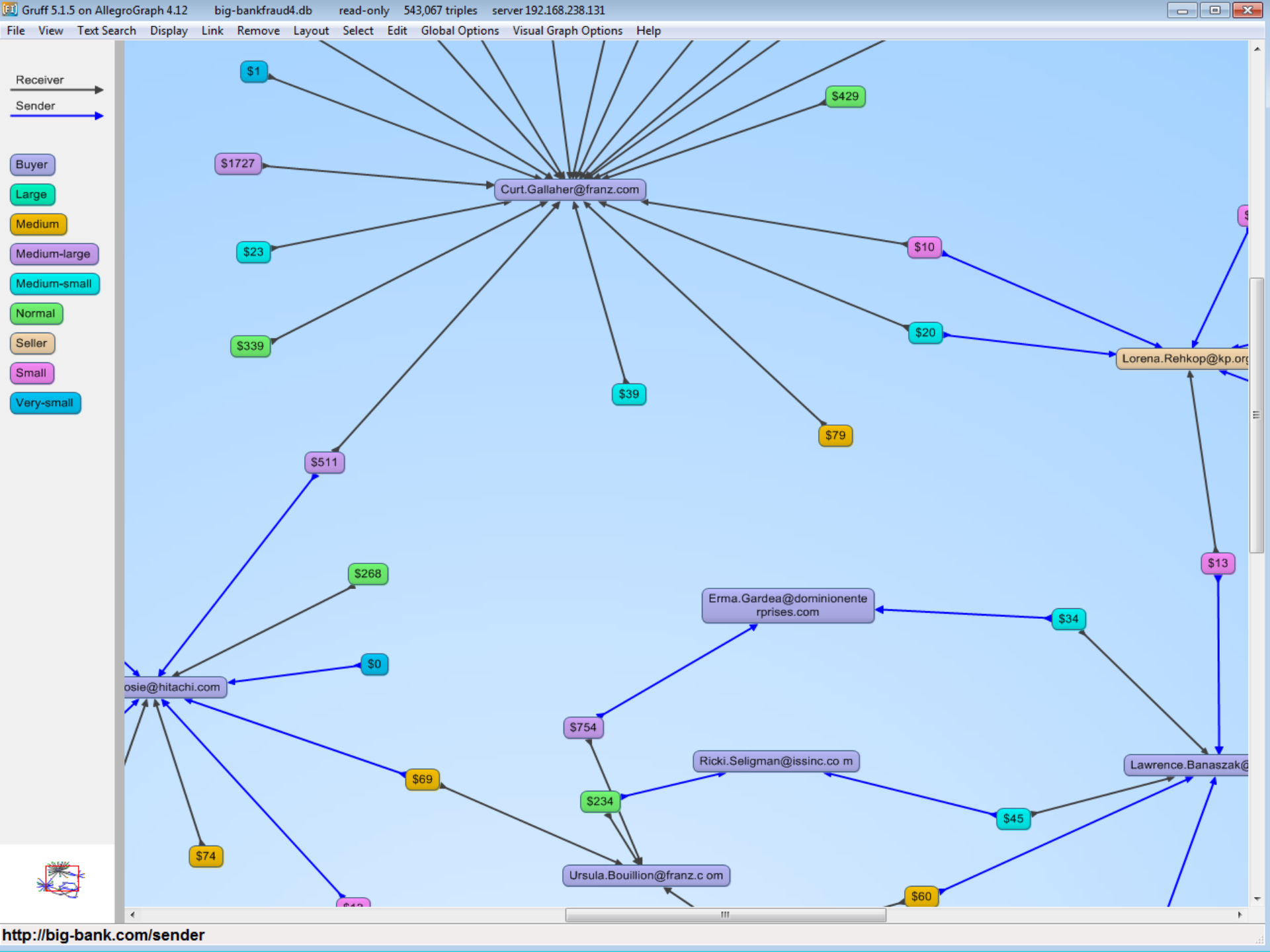
Property Left-click a property to collapse or expand its rows. Shift-left-click to add a property's triples to the graph	**Value** Left-click a value to visit it in the table view and add the triple to the graph view. Right-click a value or press M for a menu of navigation commands. J moves down a row, K moves up, and L moves to the other column.
Admin1 code	TX
Admin2 code	213
Alternatenames	
Asciiname	Gun Barrel City
Cc2	
Country code	US
Elevation	107
Feature class	P
Feature code	PPL
Geonameid	4695535
Gtopo30	106
Label	Gun Barrel City
Latitude	32.33458709716797d0 -96.15135955810547d0
Lon-lat-5	-96.1513595117845d0 32.33458707912458d0
Modification date	2006-01-17
Population	5856
Timezone	America/Denver
is Account Place of	Bruce.Rosie@hitachi.com
is Event Place of	$13 $430 $511 $106

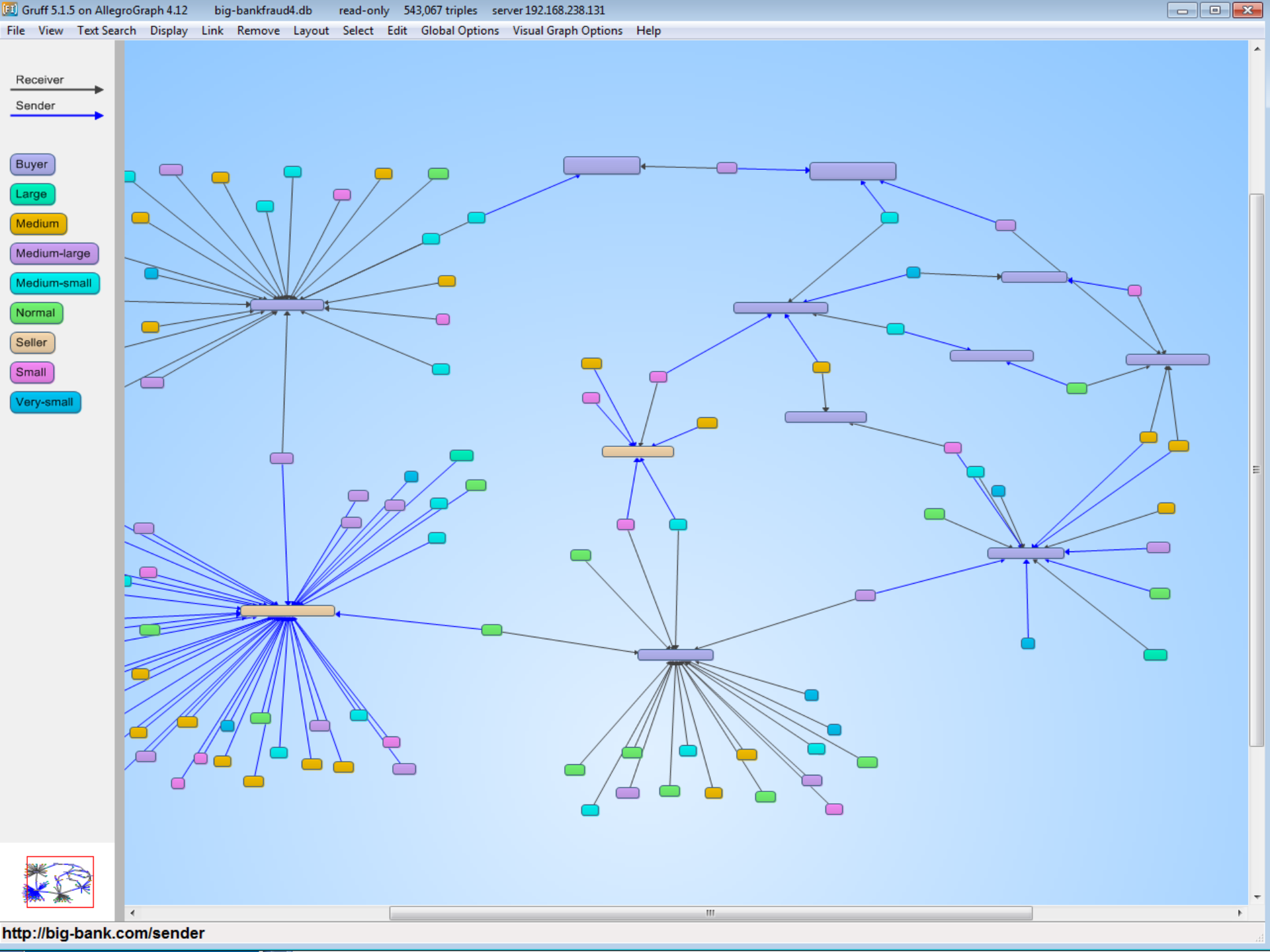
$430

Revisit Show All Triples

Property	Value
Left-click a property to collapse or expand its rows. Shift-left-click to add a property's triples to the graph	Left-click a value to visit it in the table view and add the triple to the graph view. Right-click a value or press M for a menu of navigation commands. J moves down a row, K moves up, and L moves to the other column.
Amount	430
Event Date Time UT	3539822738
Event IP	1.323.225.6
Event Place	Gun Barrel City
Label	$430
Receiver	Kathryn.Klapp@gmail.com
Sender	Bruce.Rosie@hitachi.com
Time	2012-03-04T04:05:38Z
Type	Normal


Gruff 5.1.5 on AllegroGraph 4.12
big-bankfraud4.db
read-only
543,067 triples
server 192.168.238.131
File
View
Text Search
Display
Link
Remove
Layout
Select
Edit
Global Options
Visual Graph Options
Help
Account Place
Event Place
Buyer
Normal
No Type
Bruce.Rosie@hitachi.com
Gun Barrel City
$430
Current Predicates
Event IP
Event Place
Feature class
Feature code
Geonameid
Gtopo30
Label
Latitude
Lon-lat-5
Modification date
Paid
Paid-large
Paid-medium
Paid-medium-large
Paid-medium-small
Paid-normal
Paid-small
Paid-very-small
Population
Receiver
Sender
Time
Timezone
Type
Receiver, Sender
Select All
Deselect All
OK
Cancel


Gruff 5.1.5 on AllegroGraph 4.12
big-bankfraud4.db
read-only
543,067 triples
server 192.168.238.131
File
View
Text Search
Display
Link
Remove
Layout
Select
Edit
Global Options
Visual Graph Options
Help
Receiver
Sender
Buyer
Large
Medium
Medium-large
Medium-small
Normal
Seller
Small
Very-small
$1
$429
$1727
Curt.Gallaher@franz.com
$23
$10
$20
$339
$39
$79
$511
$268
$0
$13
Erma.Gardea@dominionenterprises.com
$34
osie@hitachi.com
$754
Ricki.Seligman@issinc.co m
Lawrence.Banaszak@
$69
$234
$45
$74
Ursula.Bouillion@franz.c om
$60
http://big-bank.com/sender


Gruff 5.1.5 on AllegroGraph 4.12
big-bankfraud4.db
read-only
543,067 triples
server 192.168.238.131
File
View
Text Search
Display
Link
Remove
Layout
Select
Edit
Global Options
Visual Graph Options
Help
Receiver
Sender
Buyer
Large
Medium
Medium-large
Medium-small
Normal
Seller
Small
Very-small
http://big-bank.com/sender

AllegroGraph WebView 4.12 repository big-bankfraud4.db

« | Overview | Queries | Scripts | Namespaces | Admin | User test

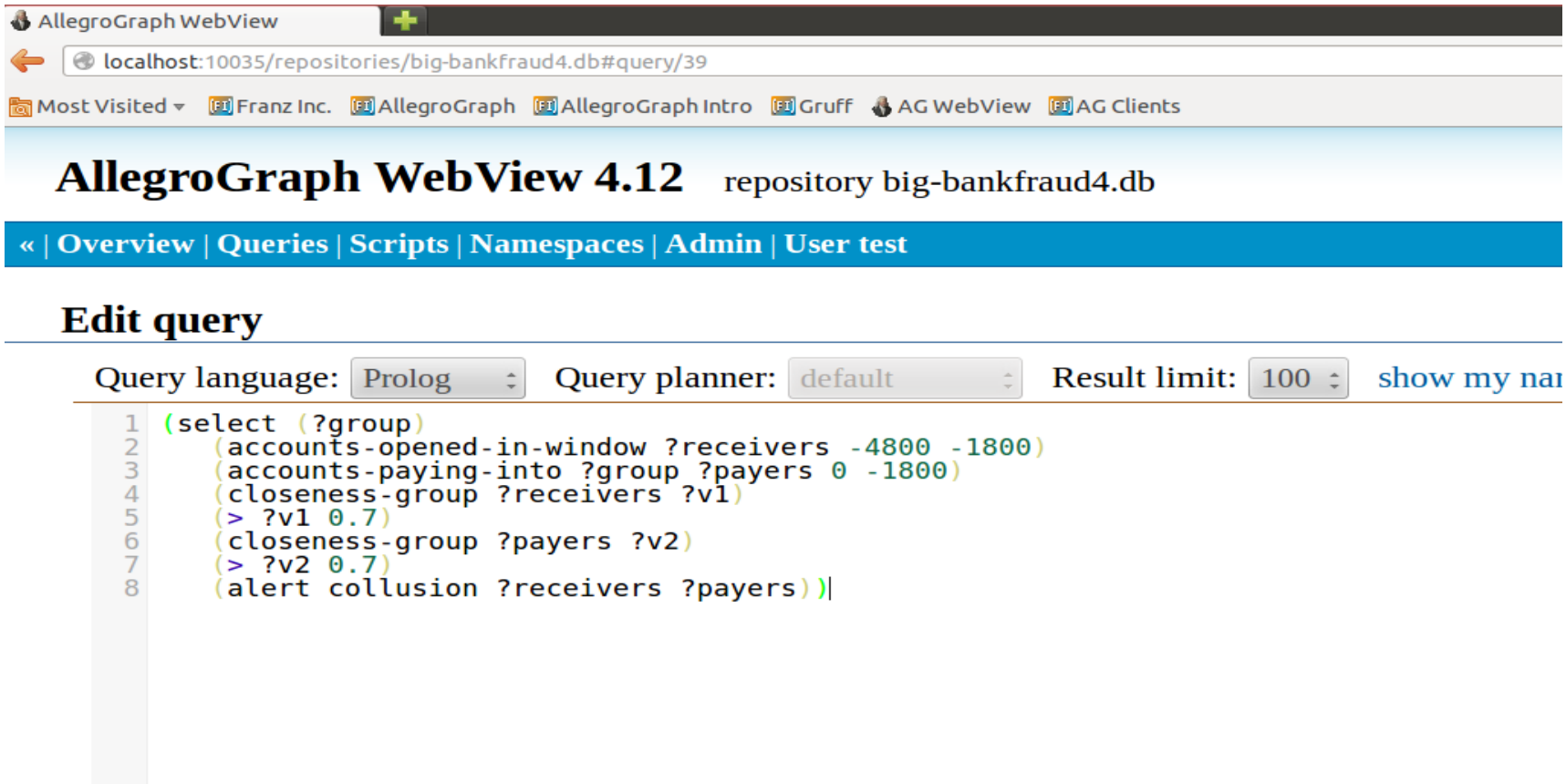
Edit query

Query language: SPARQL Query planner: default Result limit: 100 show

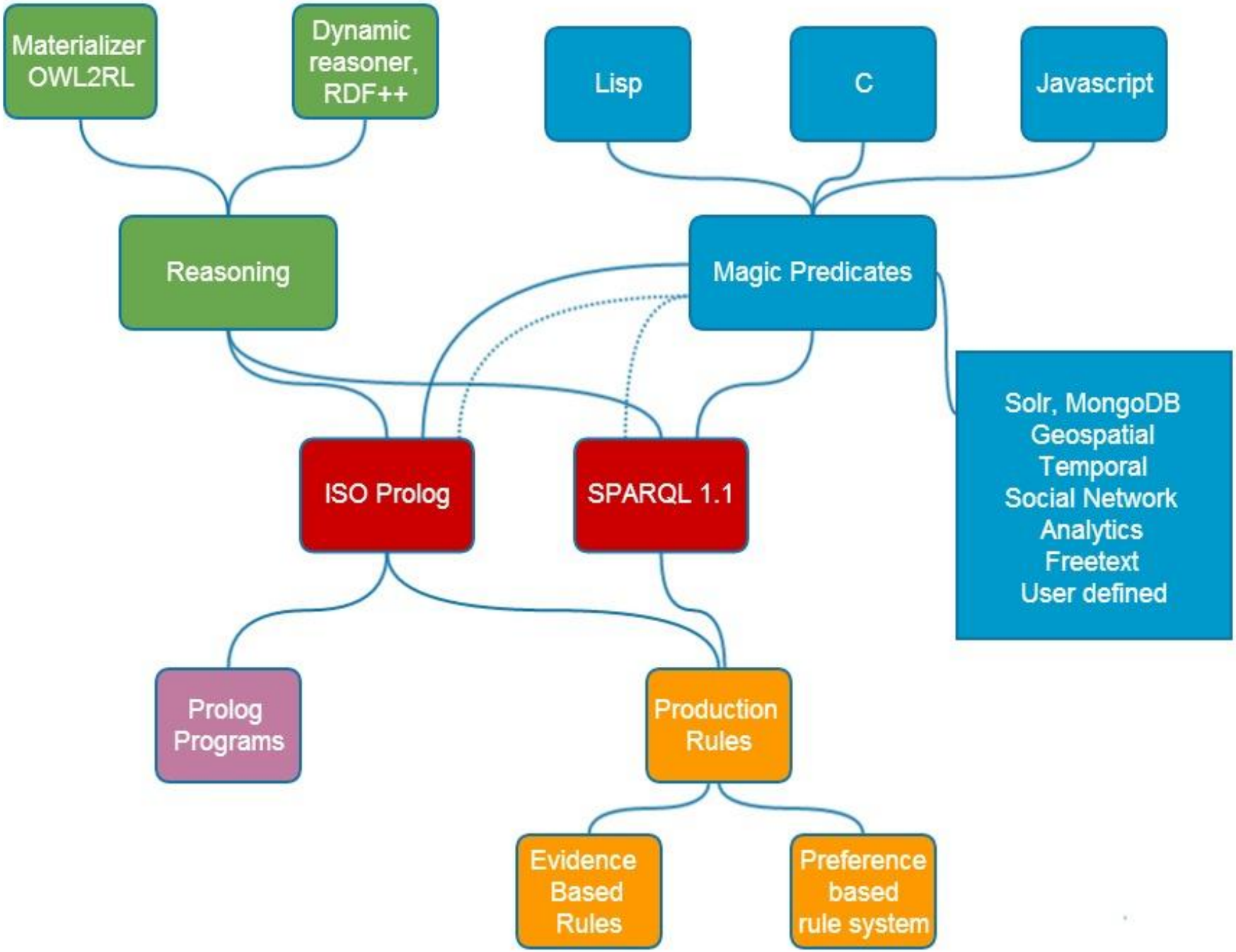
```
1  # Did the most important friends of Sonia make a payment
2  # within 100 miles of Rotterdam (NY)
3  # in the last 10 years?
4  prefix sna:  <http://franz.com/ns/allegrograph/4.11/sna/>
5  select ?email ?centrality ?amount {
6    ?place rdfs:label 'Rotterdam' ; gn:admin1_code 'NY' ; gn:lon-lat-5 ?location
7
8    ?who bb:email 'Sonia.Madrid@gmail.com' .
9    ?group sna:egoGroup ( bb:paid ?who 2 ) .
10   (?member ?centrality) sna:actorDegreeCentrality (bb:paid ?group) .
11   FILTER(?member != ?who)
12   # filter (?centrality > 0.1)
13   ?event bb:sender ?member ; franzTime:time ?time ; bb:amount ?amount ; franz:
14   filter( ?time >= '2002-01-01'^^xsd:dateTime && ?time <= '2013-01-01'^^xsd:da
15   ?member bb:email ?email .
16   ?otherPlace franzGeo:inCircleMiles (gn:lon-lat-5 ?location 100) .
17   ?otherPlace rdfs:label ?otherPlaceName .
18 } order by desc(?centrality)
19
20
```

Execute Save as Add to repository

Analysts have a chance of reading and writing this



Franz Inc


Materializer OWL2RL
Dynamic reasoner, RDF++
Lisp
C
Javascript
Reasoning
Magic Predicates
Solr, MongoDB
Geospatial
Temporal
Social Network
Analytics
Freetext
User defined
ISO Prolog
SPARQL 1.1
Prolog Programs
Production Rules
Evidence Based Rules
Preference based rule system

The End