

Querying Wikidata: Comparing SPARQL, Relational and Graph Databases

Daniel Hernández

Aidan Hogan

Cristian Riveros

Carlos Rojas

Enzo Zerega

Center for Semantic Web Research
Chile



What is Wikidata?



- Main page
- Community portal
- Project chat
- Create a new item
- Item by title
- Recent changes
- Random item
- Query Service
- Nearby
- Help
- Donate

- Print/export
 - Create a book
 - Download as PDF
 - Printable version

- In other projects
 - Wikimedia Commons
 - MediaWiki
 - Meta-Wiki
 - Wikispecies
 - Wikibooks
 - Wikinews
 - Wikipedia

English Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

[Main Page](#)


[Discussion](#)

[Read](#)

[View source](#)


[View history](#)

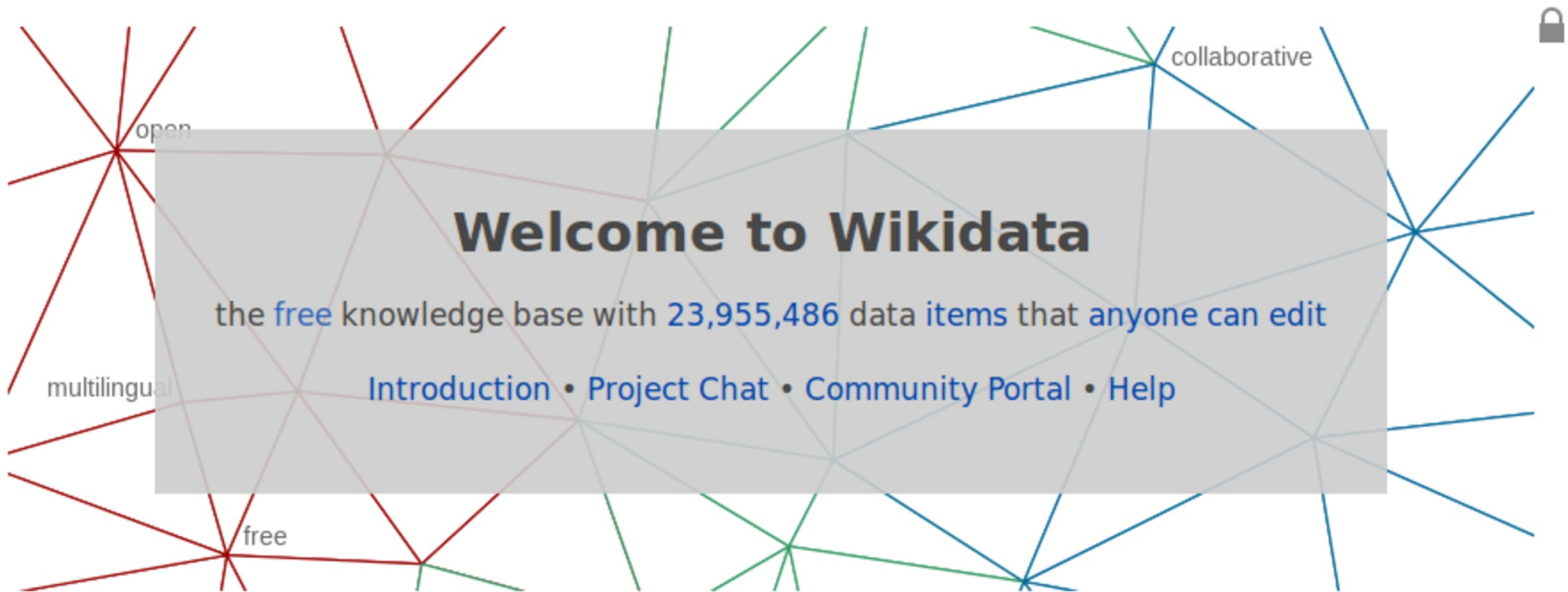




Should Wikimedia content be licensed under the latest Creative Commons license?

Participate in the discussion between October 5 and November 8!





Welcome to Wikidata

the free knowledge base with 23,955,486 data items that anyone can edit

[Introduction](#) • [Project Chat](#) • [Community Portal](#) • [Help](#)

collaborative

open

multilingual

free



Welcome!

Wikidata is a free and open knowledge base that can be



Learn about data

New to the wonderful world of data? [Develop and](#)

PROBLEMS WITH WIKIPEDIA

Problem 1: Inconsistencies between different languages

Alexis Sánchez



Sánchez with Chile in 2013

Personal information	
Full name	Alexis Alejandro Sánchez Sánchez ^[1]
Date of birth	19 December 1988 (age 26) ^{[1][2]}
Place of birth	Tocopilla, Chile ^{[3][2]}

Alexis Sánchez



Datos personales

Nombre completo Alexis Alejandro Sánchez Sánchez

Alexis



Alexis Sánchez im Dezember 2011

Spielerinformationen	
Voller Name	Alexis Alejandro Sánchez Sánchez
Geburtstag	19. Dezember 1988
Geburtsort	Tocopilla, Chile

International goals: 32

International goals: 34

International goals: 33

Problem 2: Complex lists of things manually edited by users



WIKIPEDIA
The Free Encyclopedia

[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)
[Wikipedia store](#)

Interaction

[Help](#)
[About Wikipedia](#)
[Community portal](#)
[Recent changes](#)
[Contact page](#)

Tools

[What links here](#)
[Related changes](#)
[Upload file](#)
[Special pages](#)
[Permanent link](#)
[Page information](#)
[Wikidata item](#)

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Article [Talk](#)

[Read](#) [Edit](#) [View history](#)

List of international goals scored by Alexis Sánchez

From Wikipedia, the free encyclopedia

Alexis Sánchez is the joint second-highest goalscorer in the history of the [Chile national football team](#), with 34 goals in 99 matches since his debut in 2006.

Goals [[edit](#)]

Score and Result lists Chile's goal tally first^[1]

#	Date	Venue	Opponent	Score	Result	Competition
1.	7 September 2007	Ernst-Happel-Stadion, Vienna, Austria	 Switzerland	1–1	1–2	Friendly
2.	4 June 2008	Estadio El Teniente, Rancagua, Chile	 Guatemala	1–0	2–0	Friendly
3.				2–0		
4.	11 February 2009	Peter Mokaba Stadium, Polokwane, South Africa	 South Africa	2–0	2–0	Friendly
5.	28 March 2009	Estadio Monumental "U", Lima, Peru	 Peru	1–0	3–1	2010 FIFA World Cup qualification



Alexis Sánchez with Chile in 2013

Solution: Wikidata

Collaboratively edit structured data in one place, with multi-lingual labels.



- Main page
- Community portal
- Project chat
- Create a new item
- Item by title
- Recent changes
- Random item
- Query Service
- Nearby
- Help
- Donate

- Print/export
 - Create a book
 - Download as PDF
 - Printable version

- In other projects
 - Wikimedia Commons
 - MediaWiki
 - Meta-Wiki
 - Wikispecies

English Not logged in Talk Contributions Create account Log in

Main Page

Discussion

Read

View source

View history

Search



Should Wikimedia content be licensed under the latest Creative Commons license?
Participate in the discussion between October 5 and November 8!

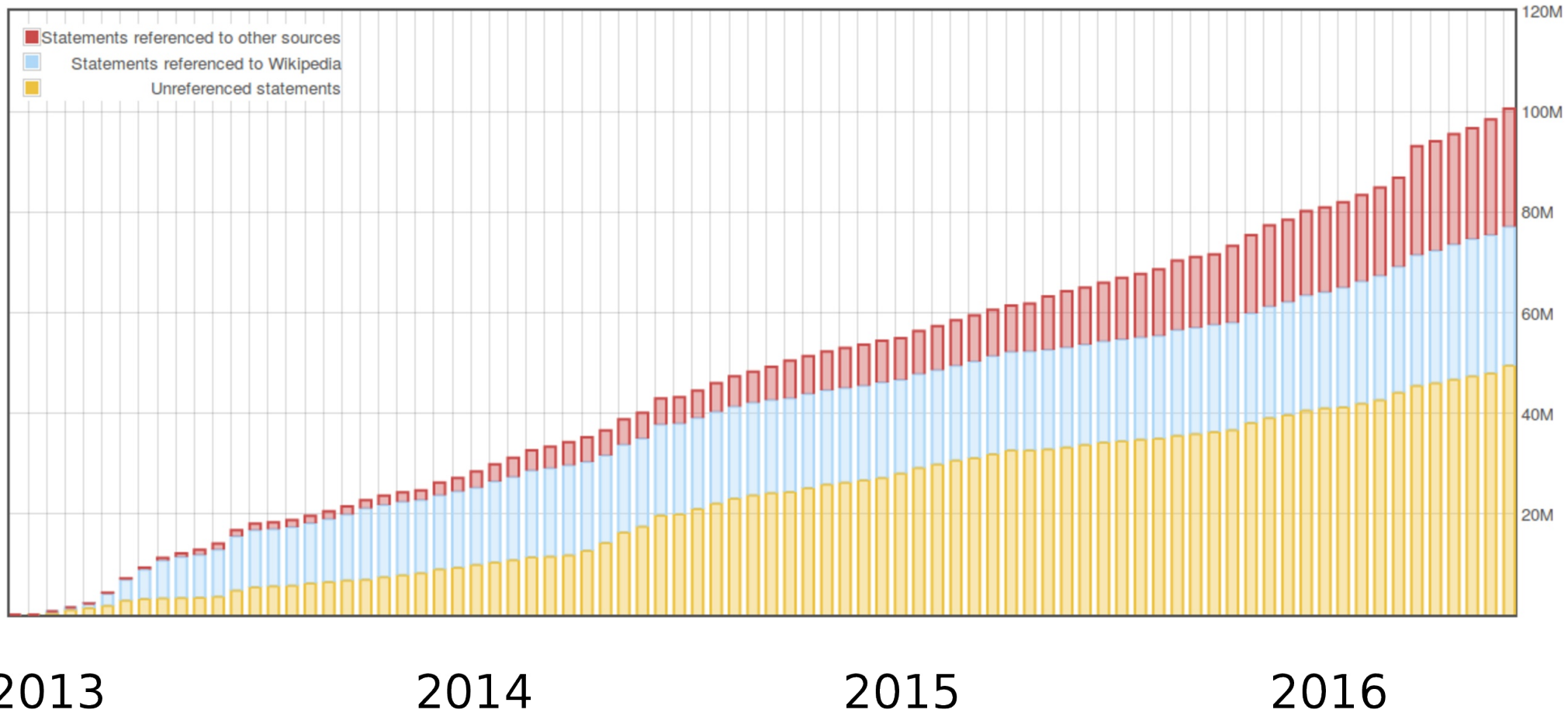
The Creative Commons logo, consisting of the letters "CC" inside a circle, with a small "X" in the top right corner.

Welcome to Wikidata
the free knowledge base with 23,955,486 data items that anyone can edit

[Introduction](#) • [Project Chat](#) • [Community Portal](#) • [Help](#)

A network diagram with nodes and connecting lines. The nodes are labeled with words: "open", "multilingual", "free", "collaborative", and "wikidata". The lines are colored in red, green, and blue.

Wikidata expansion in four years



WIKIDATA QUERY SERVICE (WDQS)

Example query

What are the teams where Alexis has played?

When did he play in each team?

Example query



```
1 SELECT ?teamLabel ?startDate ?endDate
2 WHERE {
3   ?id rdfs:label "Alexis Sánchez"@es ;
4     schema:description "futbolista chileno"@es ;
5     p:P54 ?statement .
6   ?statement ps:P54 ?team ;
7             pq:P580 ?startDate ;
8             pq:P582 ?endDate .
9   ?team rdfs:label ?teamLabel .
10  FILTER (lang(?teamLabel)="en")
11 }
```



Find anything

Show member of sports team , start time , end time

Limit

Press [CTRL-SPACE] to activate auto completion.

Data updated **a few seconds** ago

Run

Clear

6 Results in 2703 ms

Display

Data updated 30 seconds ago.

Results

```
7      pq:P580 ?startDate ;  
8      pq:P582 ?endDate .  
9      ?team rdfs:label ?teamLabel .  
10     FILTER (lang(?teamLabel)="en")  
11 }  
12 LIMIT 10
```

time

Limit 10

Press [CTRL-SPACE] to activate auto completion.

Data updated **a few seconds** ago

▶ Run

Clear

6 Results in 2703 ms

👁 Display ▾

📄 Download ▾

🔗 Link ▾

teamLabel	startDate	endDate
Moñitos clup	Jan 1, 2006	Jan 1, 2011
FC Barcelona	Jan 1, 2011	Jan 1, 2014
Club Atlético River Plate	Jan 1, 2007	Jan 1, 2008
Colo-Colo	Jan 1, 2006	Jan 1, 2007
Cobreloa	Jan 1, 2005	Jan 1, 2006
Chile national under-20 football team	Jan 1, 2006	Jan 1, 2008

Thousands queries per minute



Wikidata Query Service

Deployments

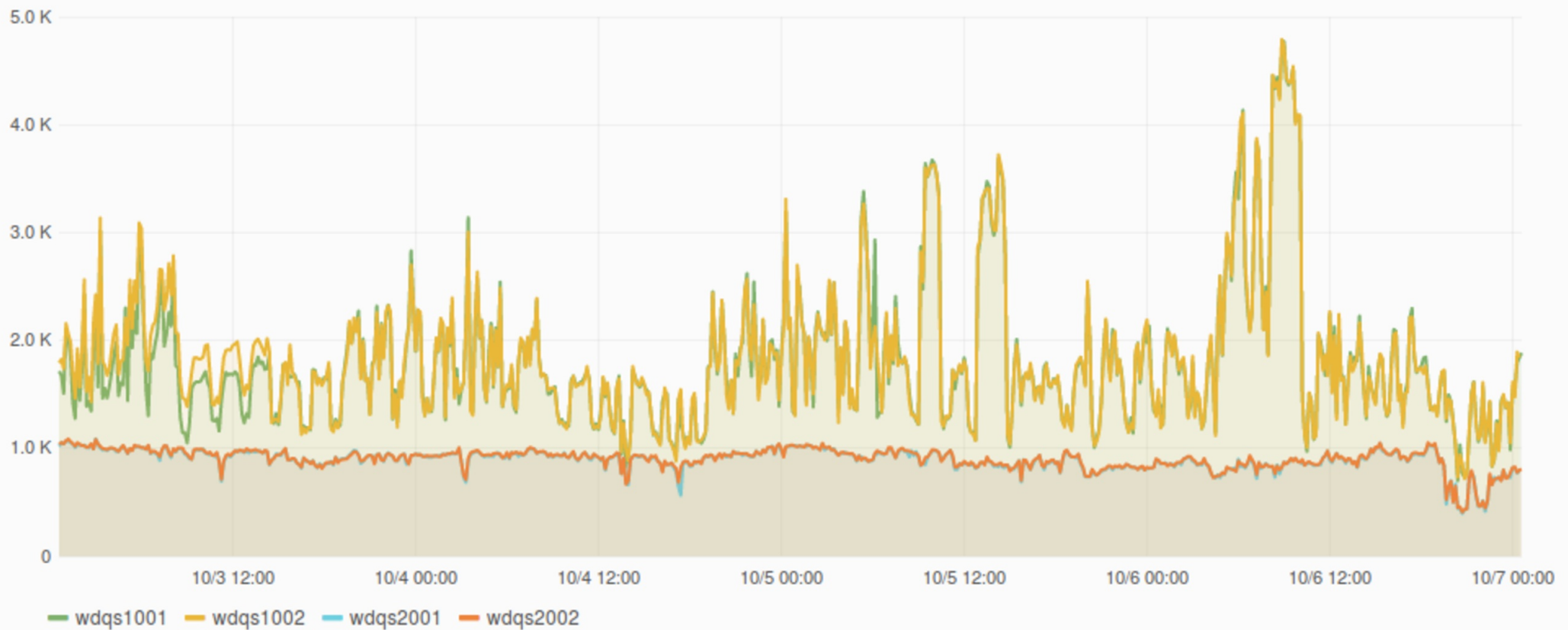
WDQS

Discovery dashboard

Wikidata dashboards

WDQS dashboards

Done rate



Our main question

Why RDF+SPARQL+Blazegraph?

We focus on performance

A WIKIDATA EXAMPLE

Item [Discussion](#)

Read [View history](#)


Search



Alexis Sánchez (Q180553)

Chilean footballer

[edit](#)

Statements

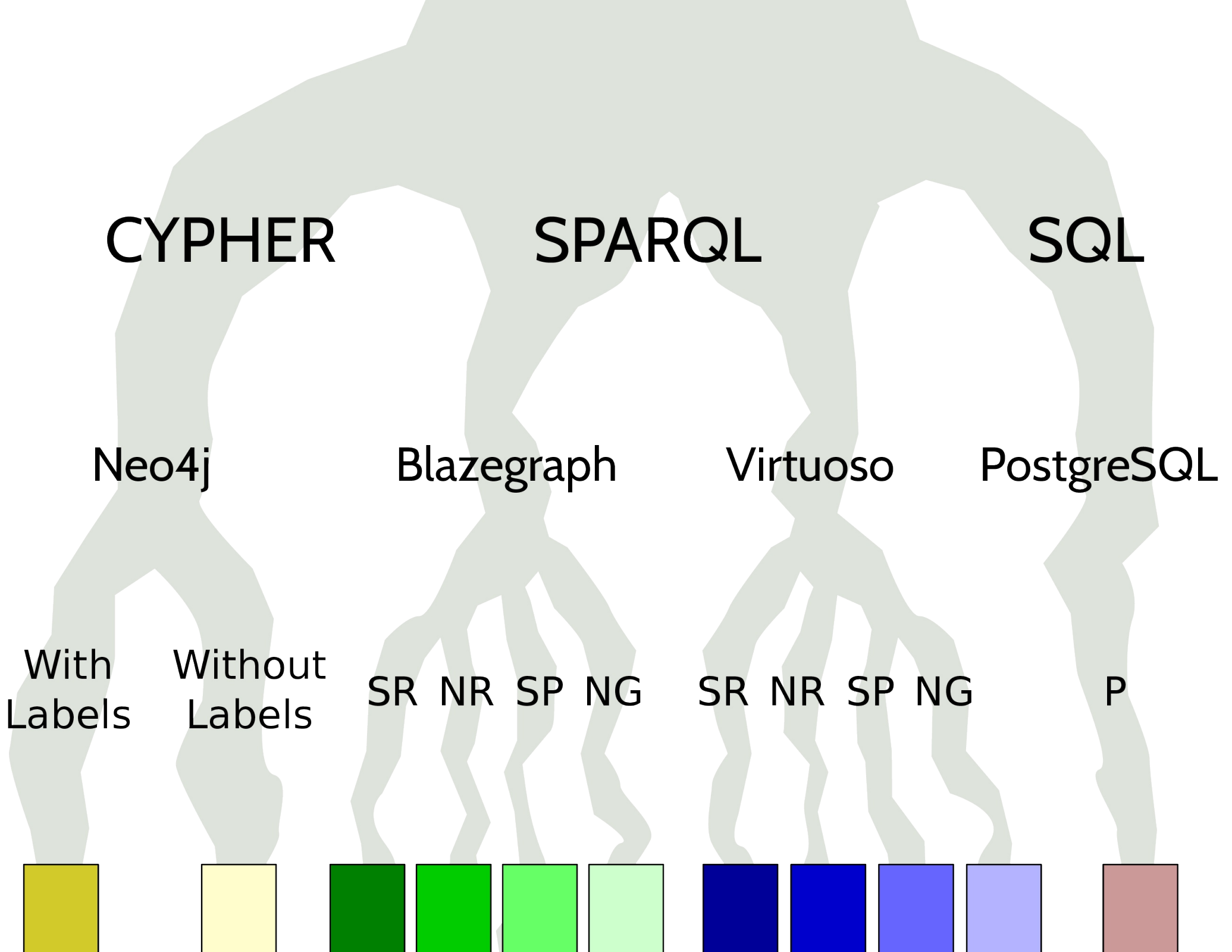
sex or gender	 male edit
	▶ 1 reference
+ add	

member of sports team	 Moñitos clup edit								
	<table border="0"> <tr><td>start time</td><td>2006</td></tr> <tr><td>end time</td><td>2011</td></tr> <tr><td>number of matches played</td><td>95</td></tr> <tr><td>number of points/goals scored</td><td>20</td></tr> </table>	start time	2006	end time	2011	number of matches played	95	number of points/goals scored	20
	start time	2006							
	end time	2011							
number of matches played	95								
number of points/goals scored	20								
▶ 1 reference									
+ add									
member of sports team	 Arsenal F.C. edit								
	<table border="0"> <tr><td>start time</td><td>2014</td></tr> </table>	start time	2014						
	start time	2014							
▶ 1 reference									
+ add									
member of sports team	 Club Atlético River Plate edit								
	<table border="0"> <tr><td>start time</td><td>2007</td></tr> <tr><td>end time</td><td>2008</td></tr> <tr><td>acquisition transaction</td><td>loan</td></tr> </table>	start time	2007	end time	2008	acquisition transaction	loan		
	start time	2007							
	end time	2008							
	acquisition transaction	loan							
<table border="0"> <tr><td>number of matches played</td><td>23</td></tr> <tr><td>number of points/goals scored</td><td>5±1</td></tr> </table>	number of matches played	23	number of points/goals scored	5±1					
number of matches played	23								
number of points/goals scored	5±1								
▶ 1 reference									
+ add									

- [Main page](#)
- [Community portal](#)
- [Project chat](#)
- [Create a new item](#)
- [Item by title](#)
- [Recent changes](#)
- [Random item](#)
- [Query Service](#)
- [Nearby](#)
- [Help](#)
- [Donate](#)

- Tools
- [What links here](#)
 - [Related changes](#)
 - [Special pages](#)
 - [Permanent link](#)
 - [Page information](#)
 - [Concept URI](#)
 - [Cite this page](#)

MODELS, LANGUAGES, ENGINES



MODELING WIKIDATA AS RDF

Item [Discussion](#)

Read [View history](#)


Search



Alexis Sánchez (Q180553)

Chilean footballer

[edit](#)

Statements

sex or gender	 male edit
	▶ 1 reference
+ add	

member of sports team	 Moñitos clup edit								
	<table border="0"> <tr><td>start time</td><td>2006</td></tr> <tr><td>end time</td><td>2011</td></tr> <tr><td>number of matches played</td><td>95</td></tr> <tr><td>number of points/goals scored</td><td>20</td></tr> </table>	start time	2006	end time	2011	number of matches played	95	number of points/goals scored	20
	start time	2006							
	end time	2011							
number of matches played	95								
number of points/goals scored	20								
▶ 1 reference									
+ add									
member of sports team	 Arsenal F.C. edit								
	<table border="0"> <tr><td>start time</td><td>2014</td></tr> </table>	start time	2014						
	start time	2014							
▶ 1 reference									
member of sports team	 Club Atlético River Plate edit								
	<table border="0"> <tr><td>start time</td><td>2007</td></tr> <tr><td>end time</td><td>2008</td></tr> <tr><td>acquisition transaction</td><td>loan</td></tr> </table>	start time	2007	end time	2008	acquisition transaction	loan		
	start time	2007							
	end time	2008							
	acquisition transaction	loan							
<table border="0"> <tr><td>number of matches played</td><td>23</td></tr> <tr><td>number of points/goals scored</td><td>5±1</td></tr> </table>	number of matches played	23	number of points/goals scored	5±1					
number of matches played	23								
number of points/goals scored	5±1								
▶ 1 reference									
+ add									

- [Main page](#)
- [Community portal](#)
- [Project chat](#)
- [Create a new item](#)
- [Item by title](#)
- [Recent changes](#)
- [Random item](#)
- [Query Service](#)
- [Nearby](#)
- [Help](#)
- [Donate](#)

Tools

- [What links here](#)
- [Related changes](#)
- [Special pages](#)
- [Permanent link](#)
- [Page information](#)
- [Concept URI](#)
- [Cite this page](#)

Alexis Sánchez (Q180553)

Chilean footballer  edit

Statements

sex or gender

predicate

male

object

▶ 1 reference

 edit

+ add

member of sports team



Moñitos clup  edit

start time 2006
end time 2011
number of matches played 95
number of points/goals scored 20
▶ 1 reference



Arsenal F.C.  edit

start time 2014
▶ 1 reference



Club Atlético River Plate  edit

start time 2007
end time 2008
acquisition transaction loan
number of matches played 23
number of points/goals scored 5±1

- Main page
- Community portal
- Project chat
- Create a new item
- Item by title
- Recent changes
- Random item
- Query Service
- Nearby
- Help
- Donate

Tools

- What links here
- Related changes
- Special pages
- Permanent link
- Page information
- Concept URI
- Cite this page

Item **Discussion** subject
Read View history Search

Alexis Sánchez (Q180553)

Chilean footballer edit

- Main page
- Community portal
- Project chat
- Create a new item
- Item by title
- Recent changes
- Random item
- Query Service
- Nearby
- Help
- Donate

- Tools
- What links here
 - Related changes
 - Special pages
 - Permanent link
 - Page information
 - Concept URI
 - Cite this page

Statements

sex or gender predicate **male** object edit
▶ 1 reference

+ add

member of sports team predicate **Moñitos clup** object edit

start time	2006
end time	2011
number of matches played	95
number of points/goals scored	20

▶ 1 reference

Arsenal F.C. edit

start time	2014
------------	------

▶ 1 reference

Club Atlético River Plate edit

start time	2007
end time	2008
acquisition transaction	loan
number of matches played	23
number of points/goals	5±1

Item **Discussion** **subject**
Read View history Search

Alexis Sánchez (Q180553)
Chilean footballer edit

- Main page
- Community portal
- Project chat
- Create a new item
- Item by title
- Recent changes
- Random item
- Query Service
- Nearby
- Help
- Donate

- Tools
- What links here
 - Related changes
 - Special pages
 - Permanent link
 - Page information
 - Concept URI
 - Cite this page

Statements

sex or gender **predicate** **male** **object** edit
▶ 1 reference

+ add

member of sports team **predicate** **Moñitos clup** **object** edit

start time	2006
end time	2011
number of matches played	95
number of points/goals scored	20

▶ 1 reference **qualifiers?**

Arsenal F.C. edit

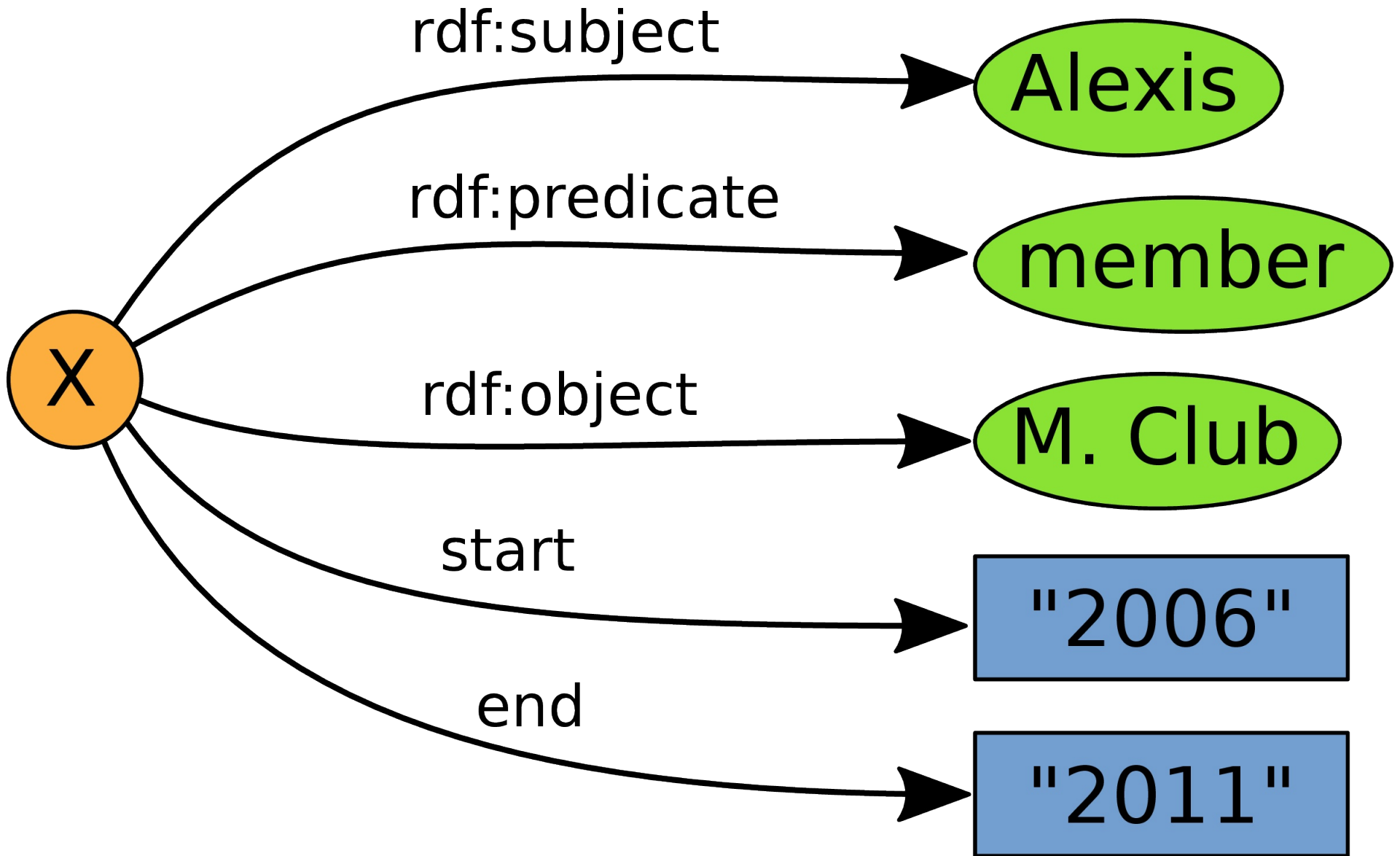
start time	2014
------------	------

▶ 1 reference

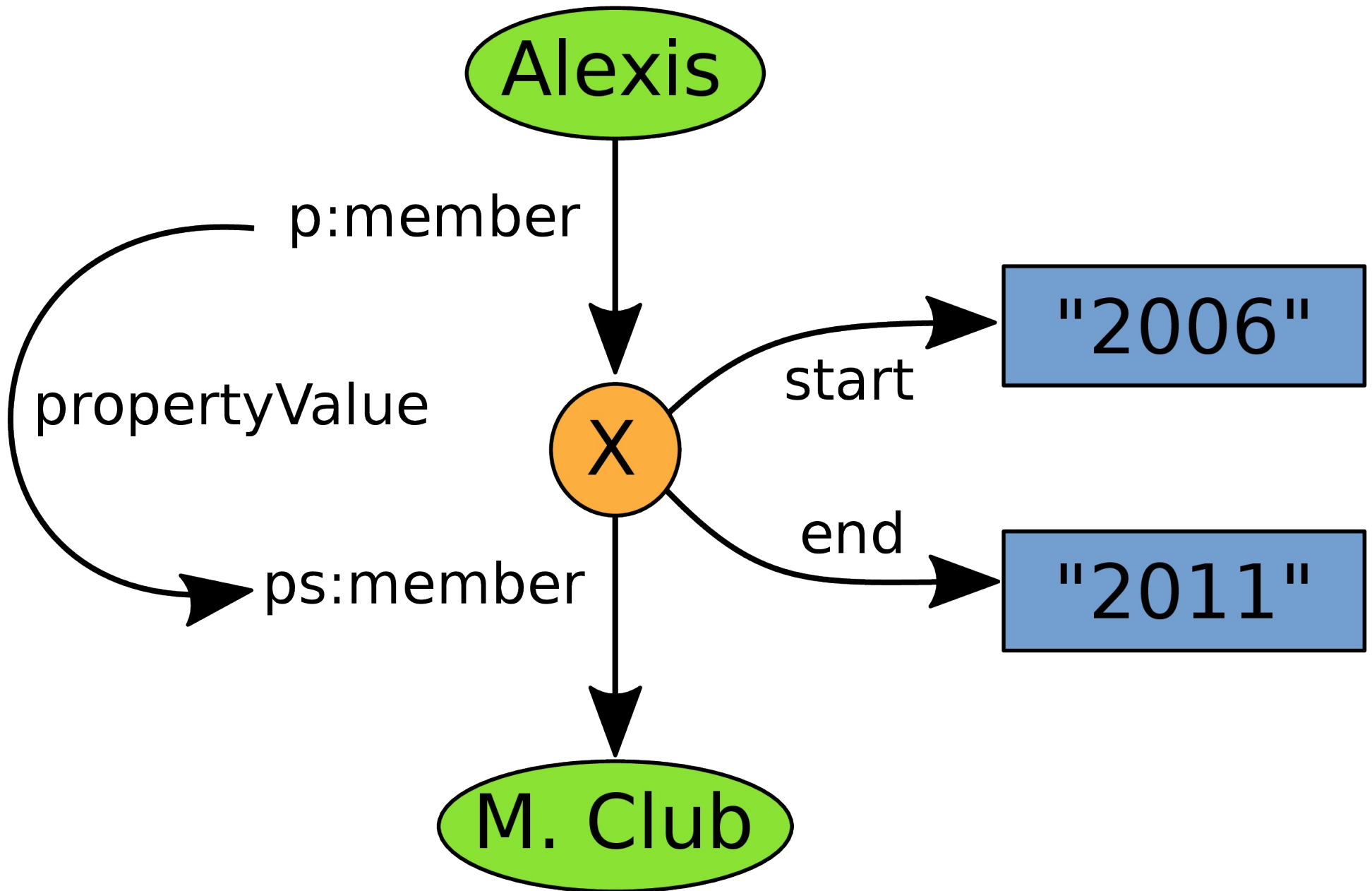
Club Atlético River Plate edit

start time	2007
end time	2008
acquisition transaction	loan
number of matches played	23
number of points/goals	5±1

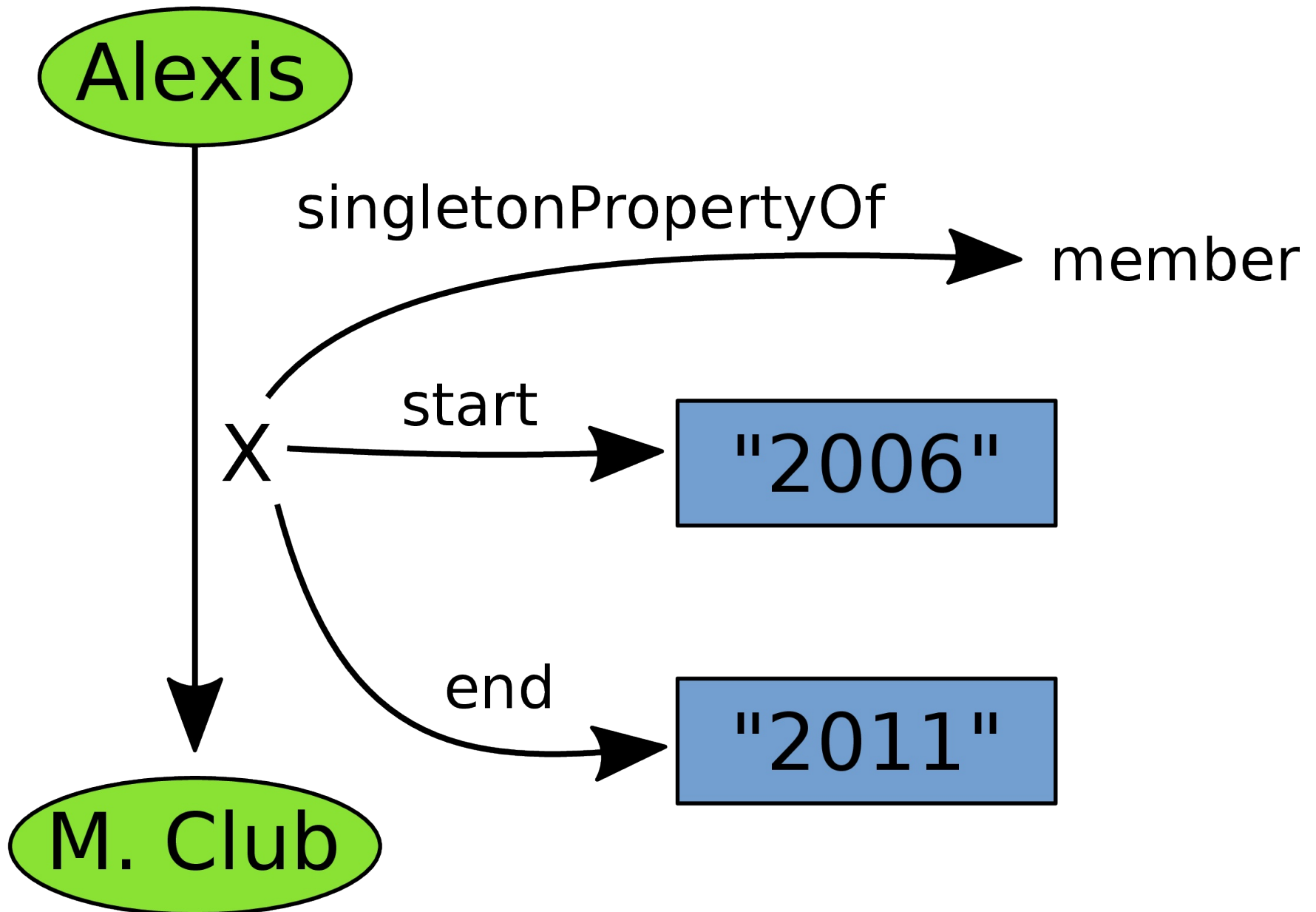
Standard Reification (SR)



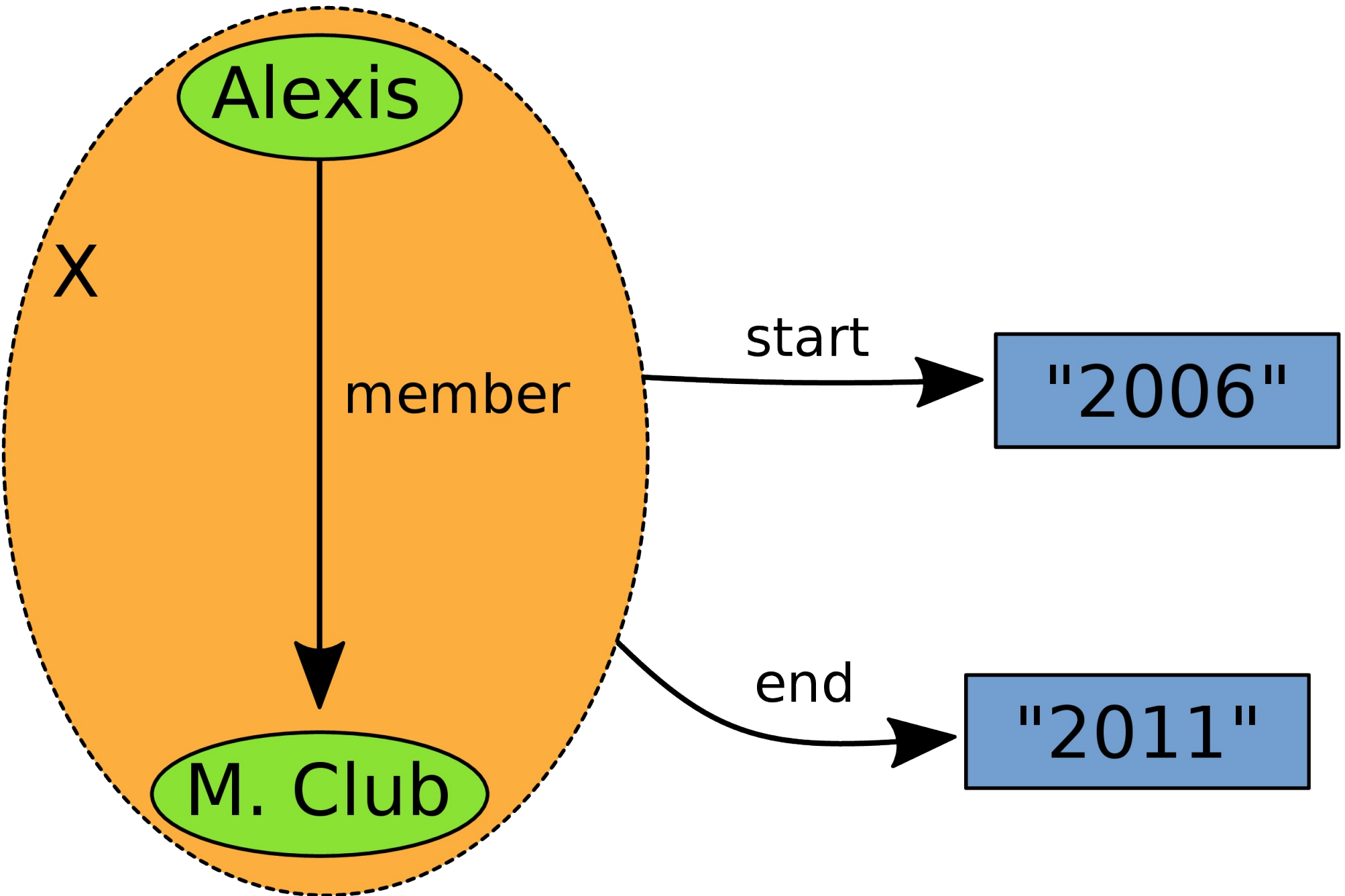
N-ary relations (NR)



Singleton Properties (SP)

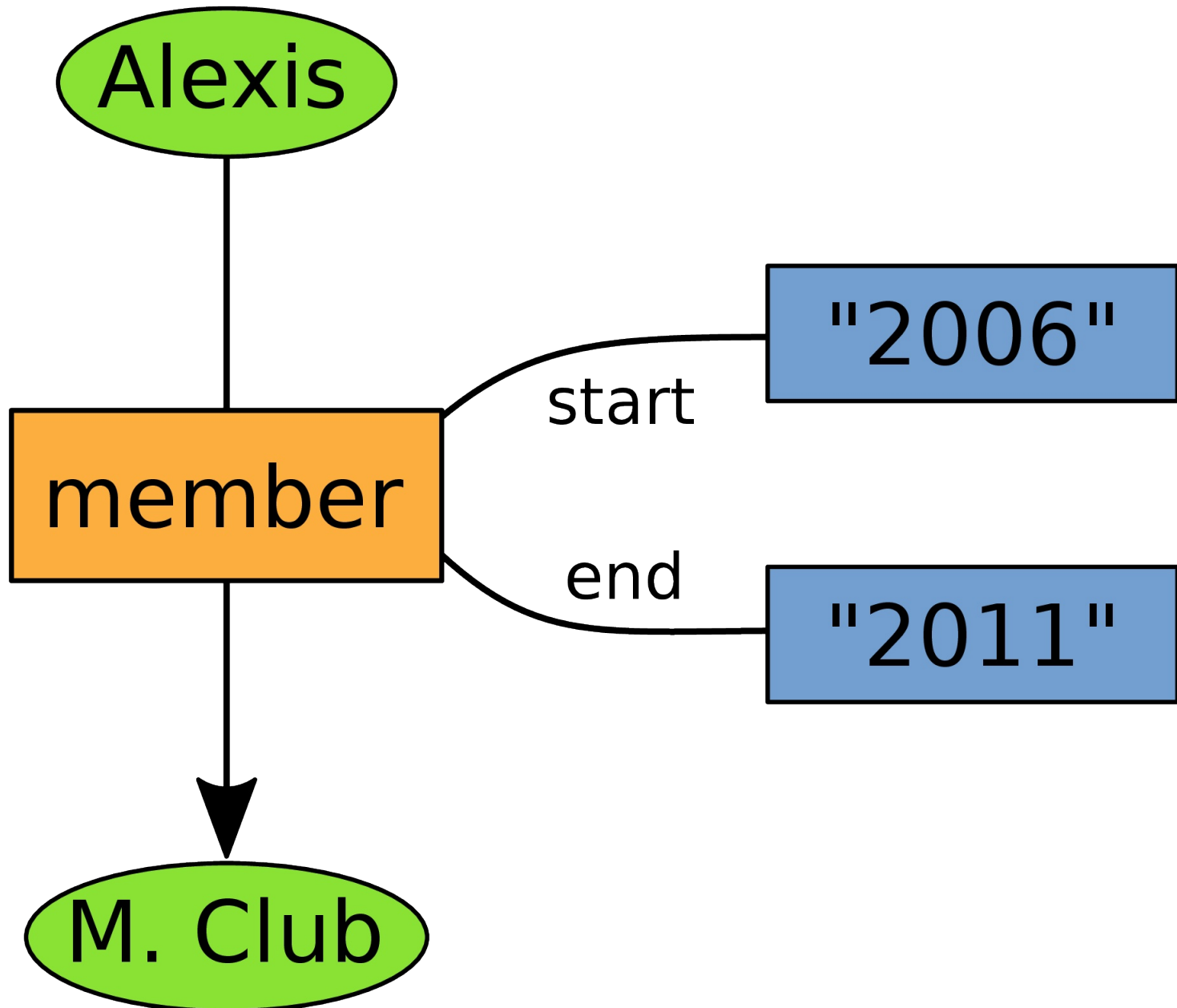


Named Graphs (NG)

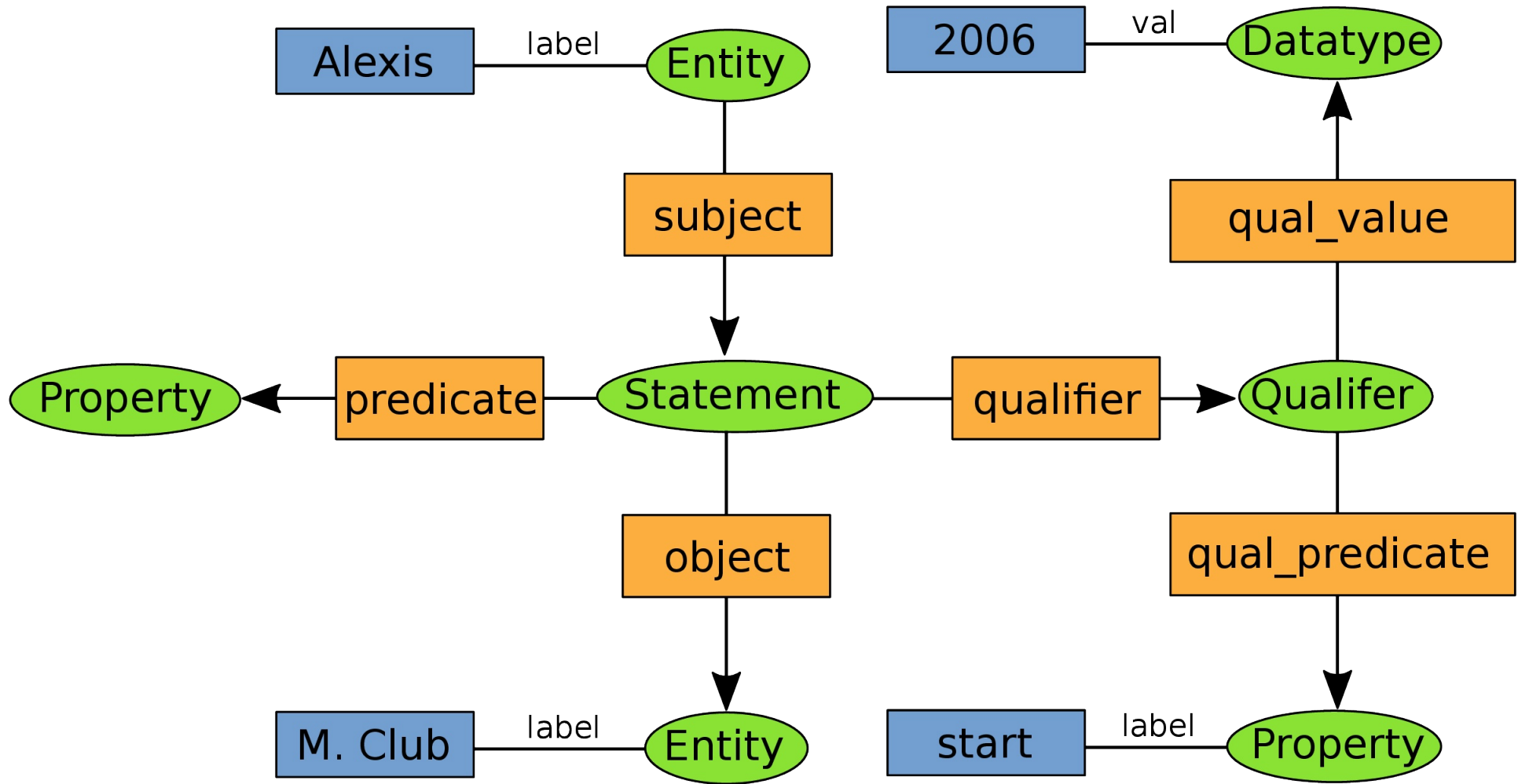


MODELING WIKIDATA AS PROPERTY GRAPHS

Property graph



Reified property graph



MODELING WIKIDATA AS A RELATIONAL DATABASE

Relational model (1/3)

statements

id	s	p	o	o_date
X1	Q91	Q31	Q11696	-

qualifiers

q	v	v_date	id
P155	Q12325	-	X1
P156	Q8612	-	X1
P580	-	1861-03-04	X1
P582	-	1865-04-15	X1

labels

e	v	lang
P39	position held	en
...
Q91	Abraham Linkcoln	en
...

Relational model (2/3)

statements

id	s	p	o	o_date
X1	Q91	Q31	Q11696	-

qualifiers_entity

q	v	id
P155	Q12325	X1
P156	Q8612	X1

labels

e	v	lang
P39	position held	en
...
Q91	Abraham Linkcoln	en
...

qualifiers_date

q	v	id
P580	1861-03-04	X1
P582	1865-04-15	X1

Relational model (3/3)

statements

id	s	p	o	o_date
X1	Q91	Q31	Q11696	-

qualifiers

q	v	type	id
P155	Q12325	entity	X1
P156	Q8612	entity	X1
P580	1861-03-04	date	X1
P582	1865-04-15	date	X1

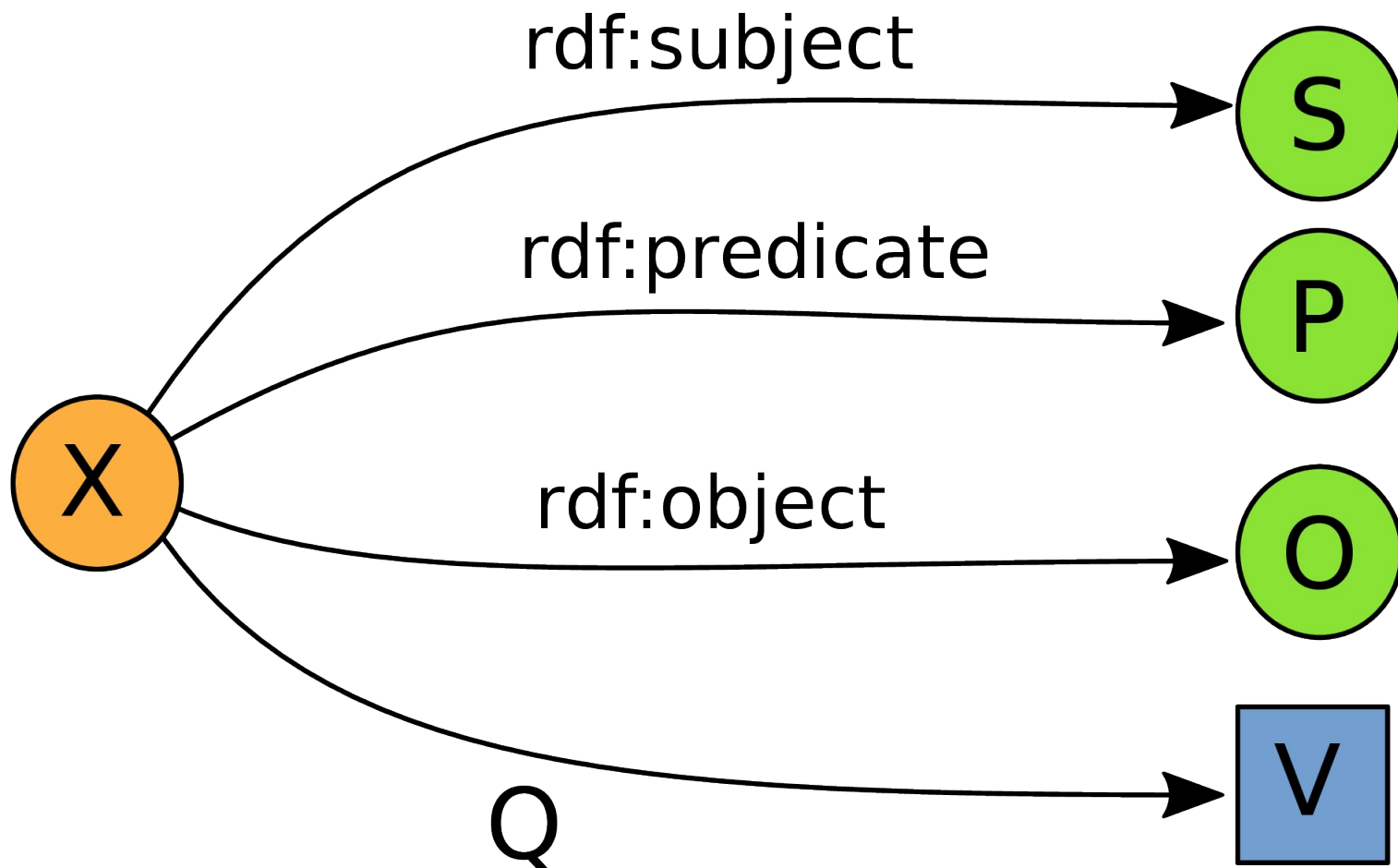
labels

e	v	lang
P39	position held	en
...
Q91	Abraham Linkcoln	en
...

EXPERIMENT 1: ATOMIC LOOKUPS

Statement pattern

(S,P,O,Q,V)



Atomic lookups

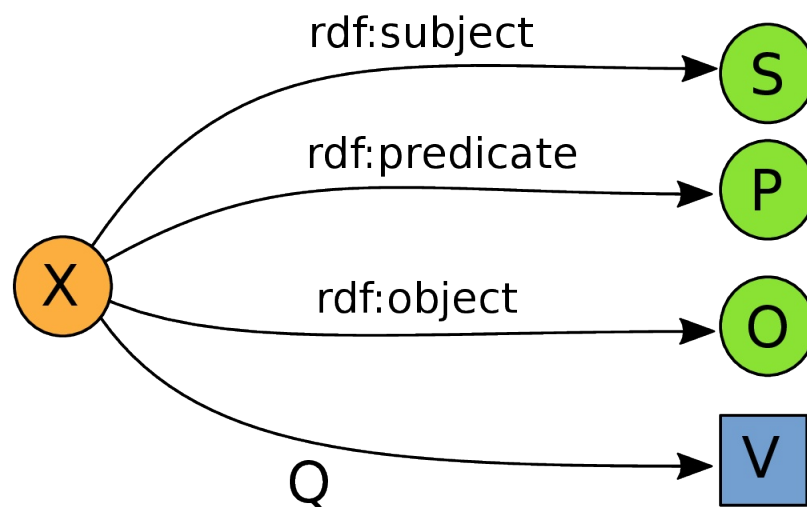
(S,?,?,?,?)

(?,P,?,?,?)

(S,P,?,?,?)

...

(S,P,O,Q,V)



Limit 10,000 solutions

Timeout 60 seconds

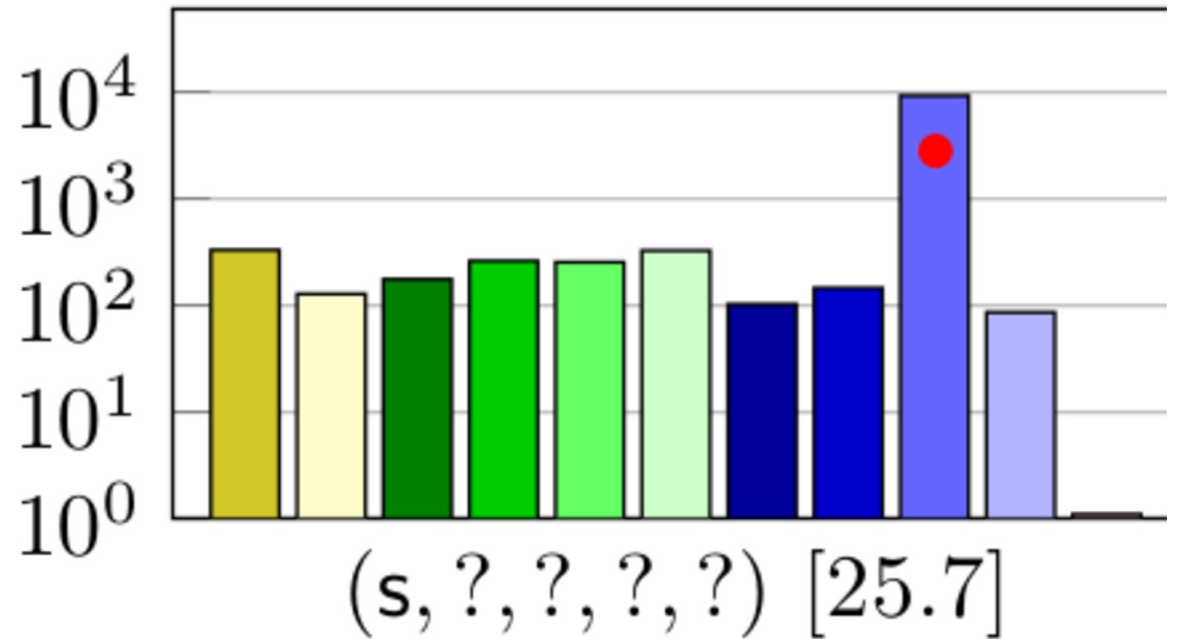
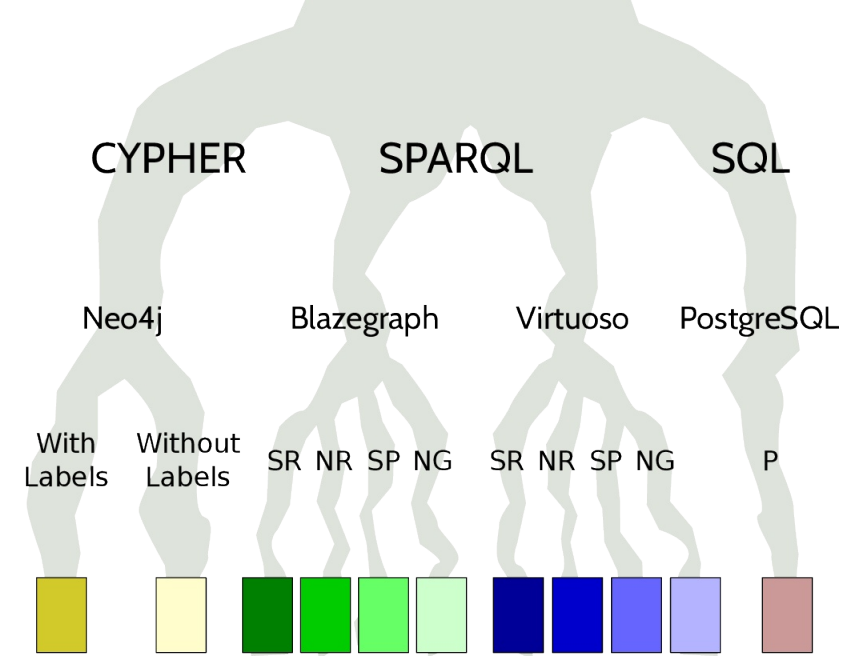
300 random queries per pattern

We compare result sizes to verify that we get the same

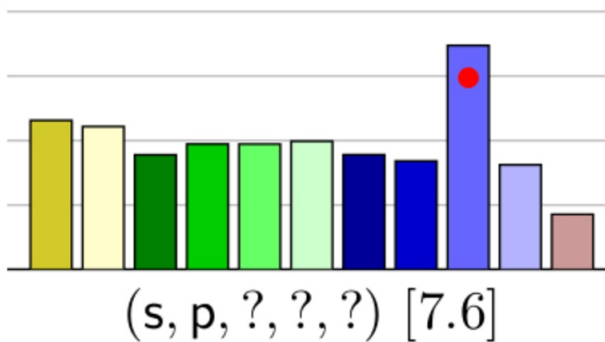
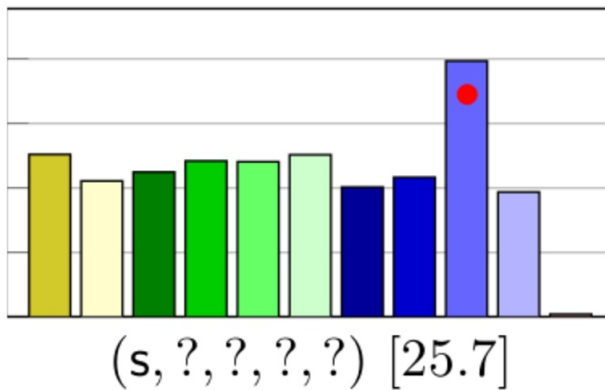
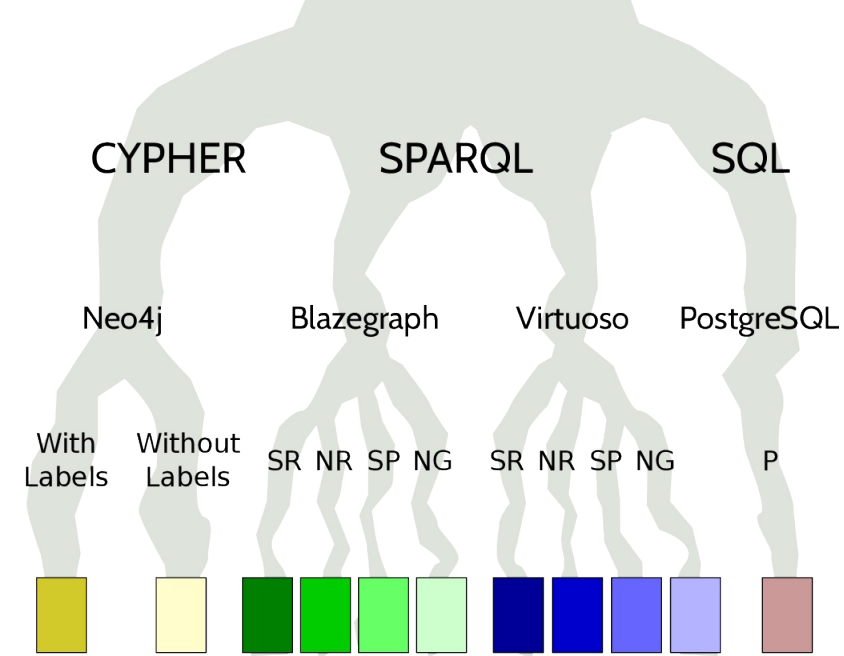
High selectivity

runtime (ms)

- [0,33%] timeouts
- [33%,66%] timeouts
- [66%,100%] timeouts

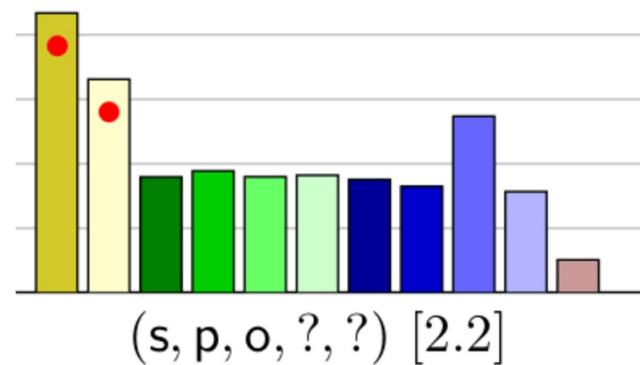
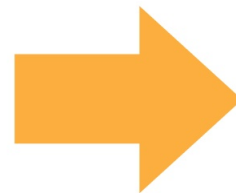
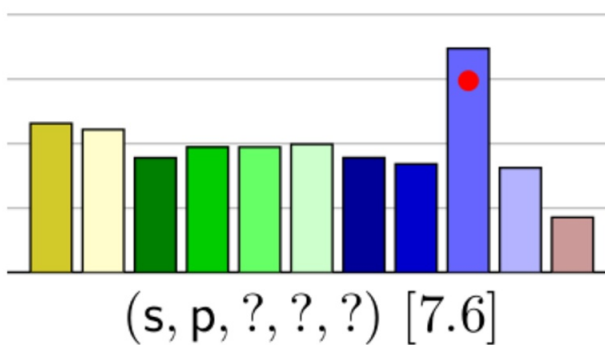
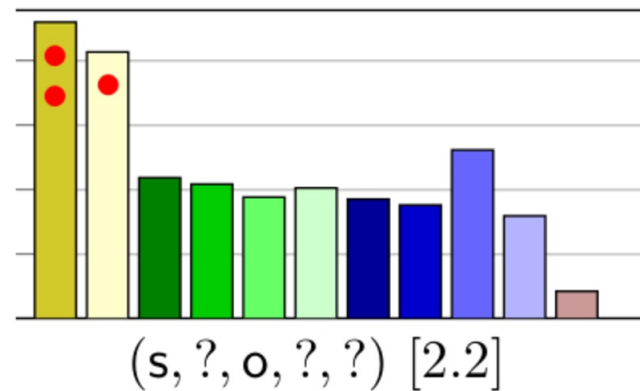
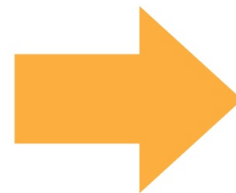
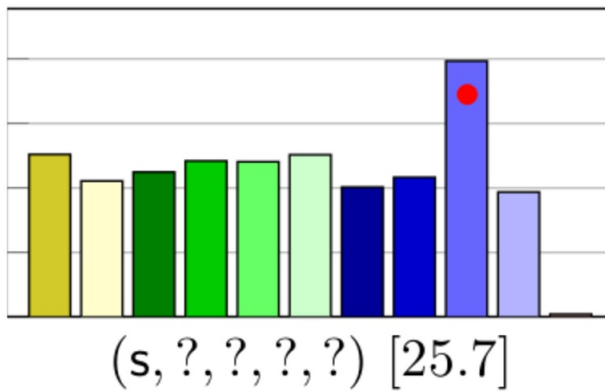
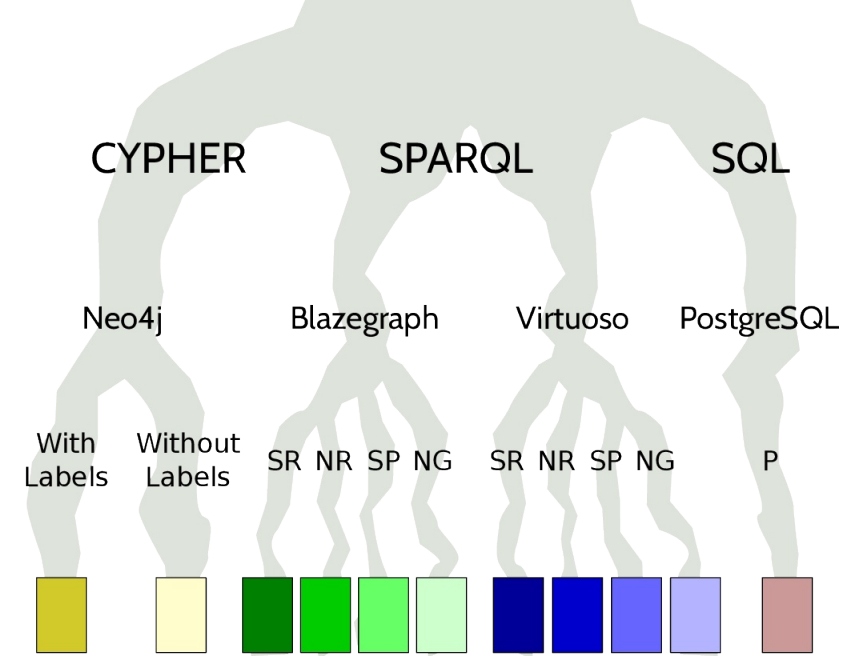


High selectivity



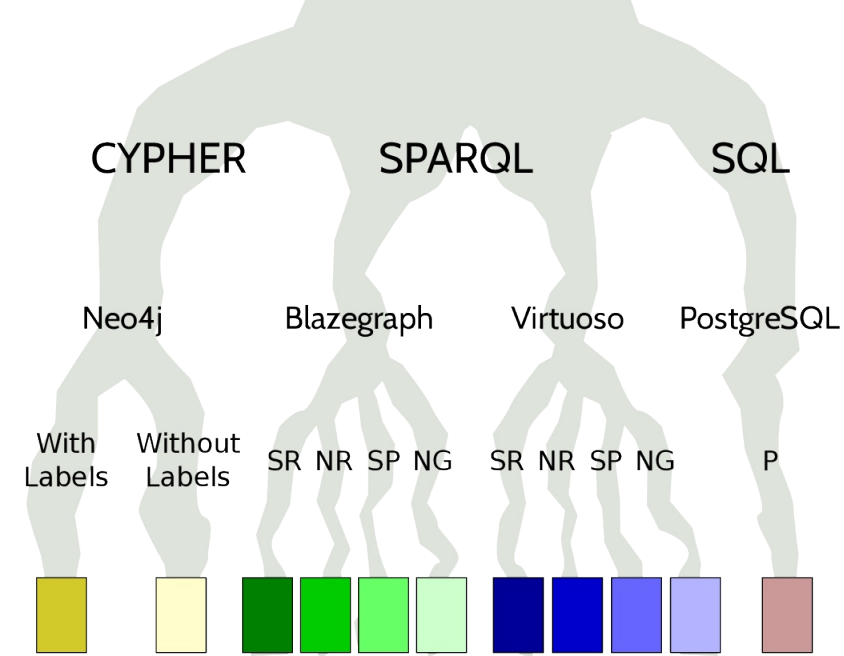
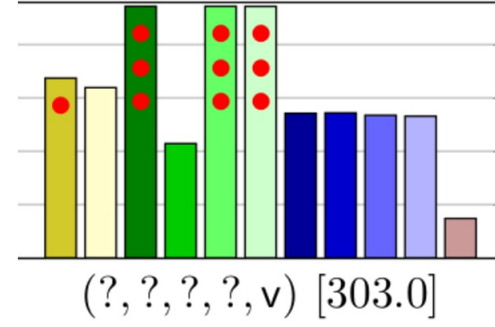
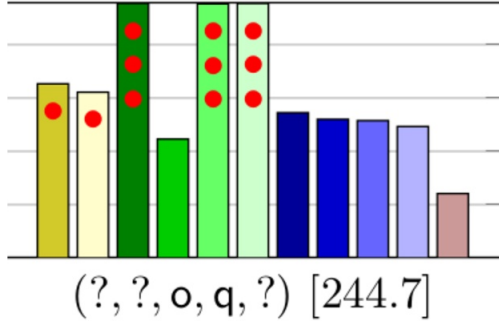
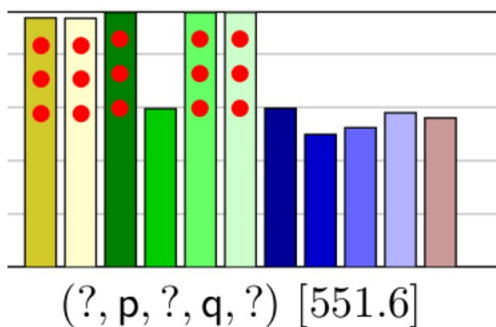
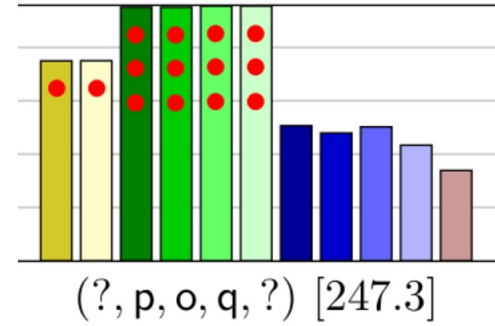
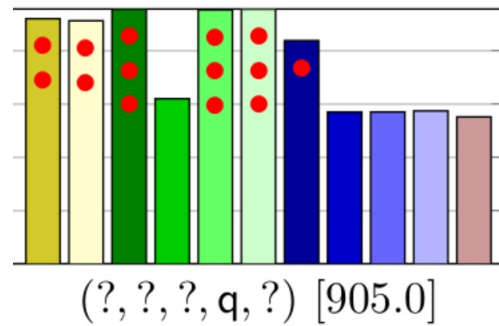
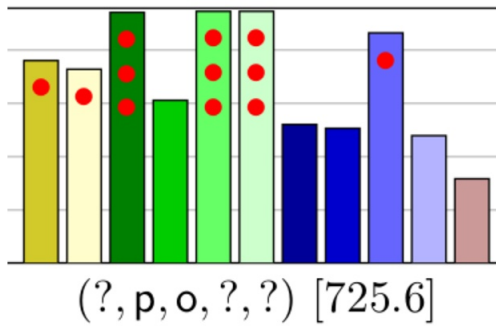
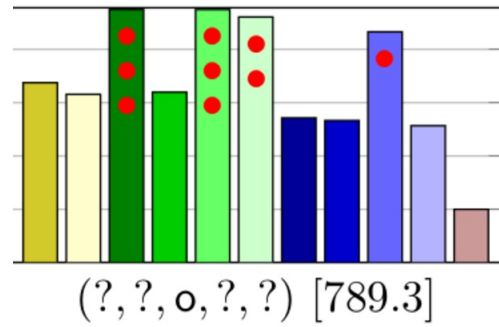
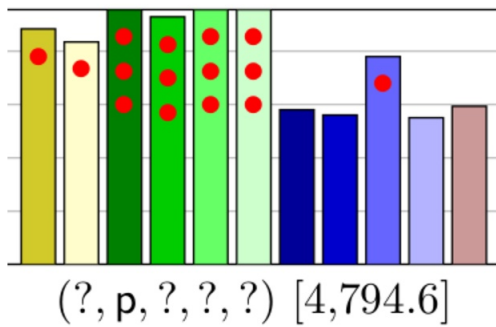
High selectivity

Note that Neo4j gets worst runtime when o is constant.

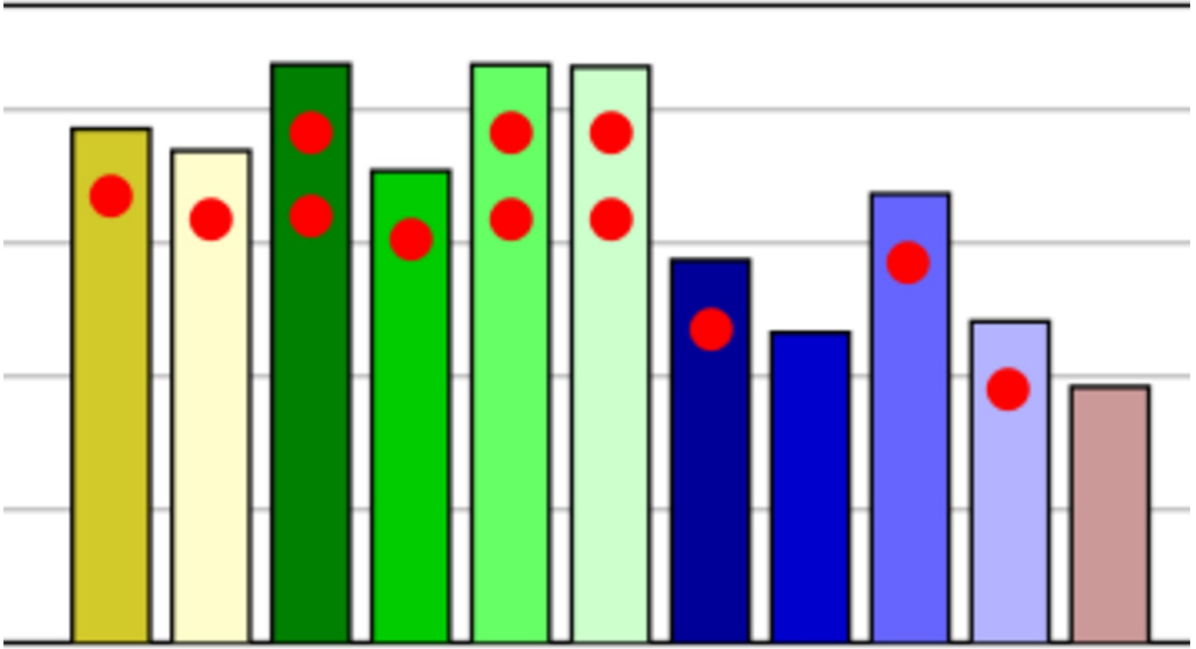
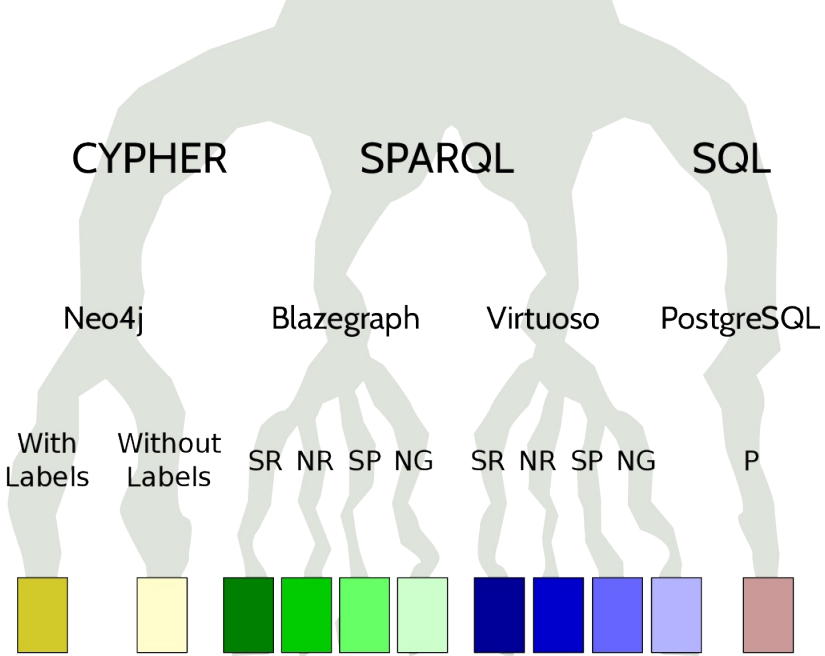


Low selectivity

Many timeouts in Neo4j and Blazegraph



Average

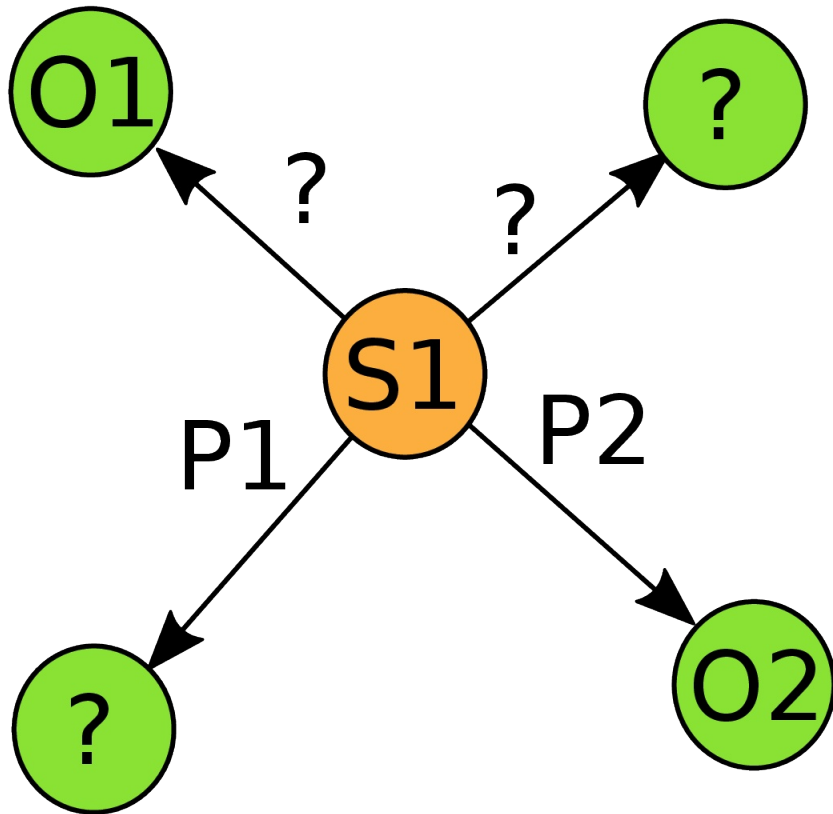


ALL [312.1]

EXPERIMENT 2:

SNOWFLAKE / STAR QUERIES

Snowflake queries depth 1



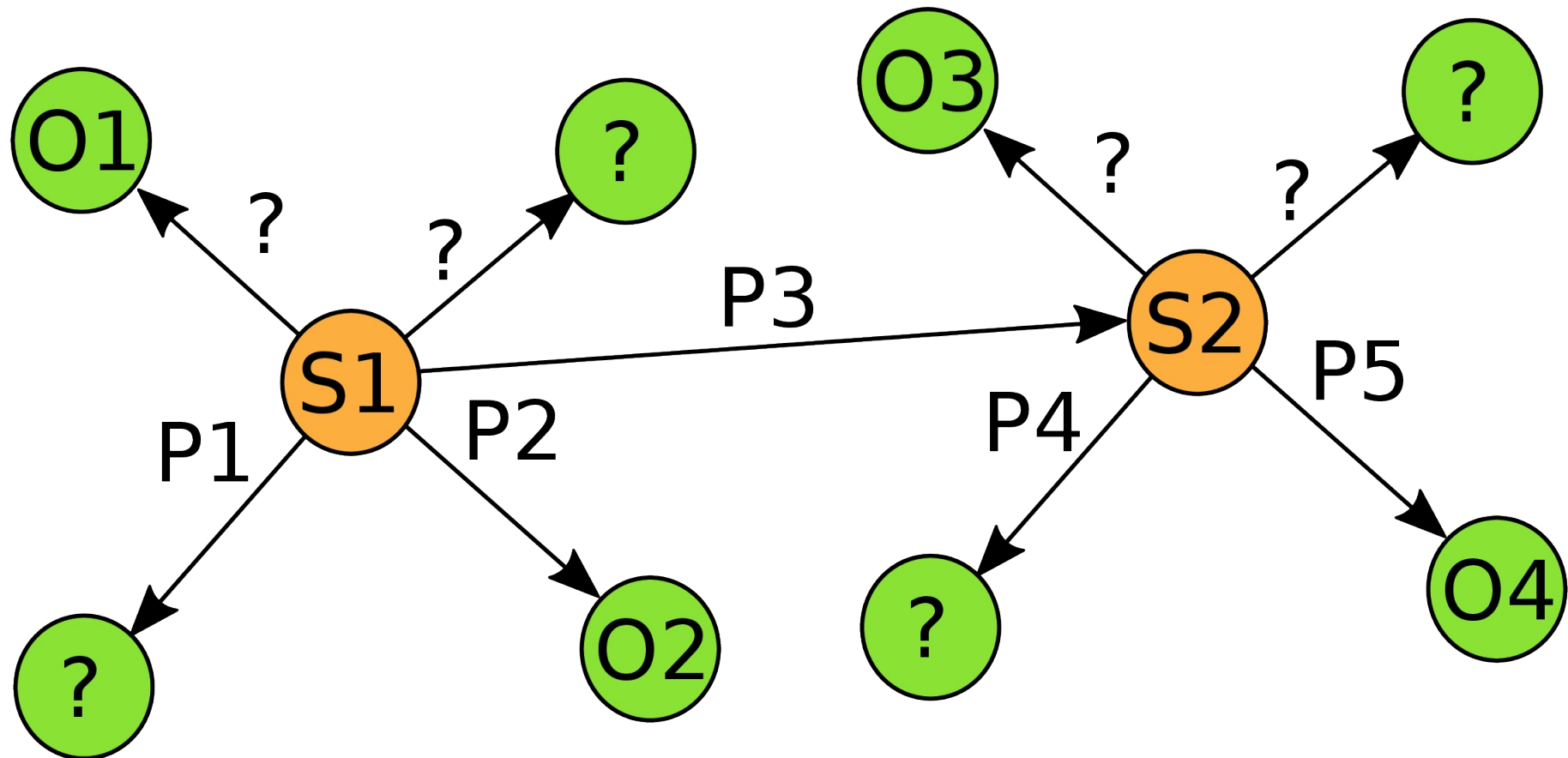
Limit 10,000 solutions

Timeout 60 seconds

300 random queries per pattern

We compare result sizes to verify that we get the same

Snowflake queries depth 2



Limit 10,000 solutions

Timeout 60 seconds

300 random queries per pattern

We compare result sizes to verify that we get the same

Conclusions

1. Neo4j and Blazegraph could be improved isolating queries to avoid a domino effect.
2. Neo4j has a naive planer for BGP, it is more navigational.
3. Neo4j could be improved allowing indices in edge properties. The current data model does not allow describing Wikidata statements without some sort of reification.
4. Virtuoso falls behind PostgreSQL in atomic lookups, but performs better in more complex graph patterns (snowflake queries).

Other considerations

1. Representing data types in the relational model is not straightforward.
2. Regarding Blazegraph performs, developers indicated that the engine was developed for SSD, where we used a standard SATA HD.
3. Path queries are only possible in nary relations and named graphs. Not possible in Neo4j, singleton properties or standard reification. Possible in PostgreSQL with recursion, but complex.
4. Different query semantics in Cypher, SPARQL and SQL.
5. Standarization, usability, licence, replication, distribution, interoperability, full text search, ...