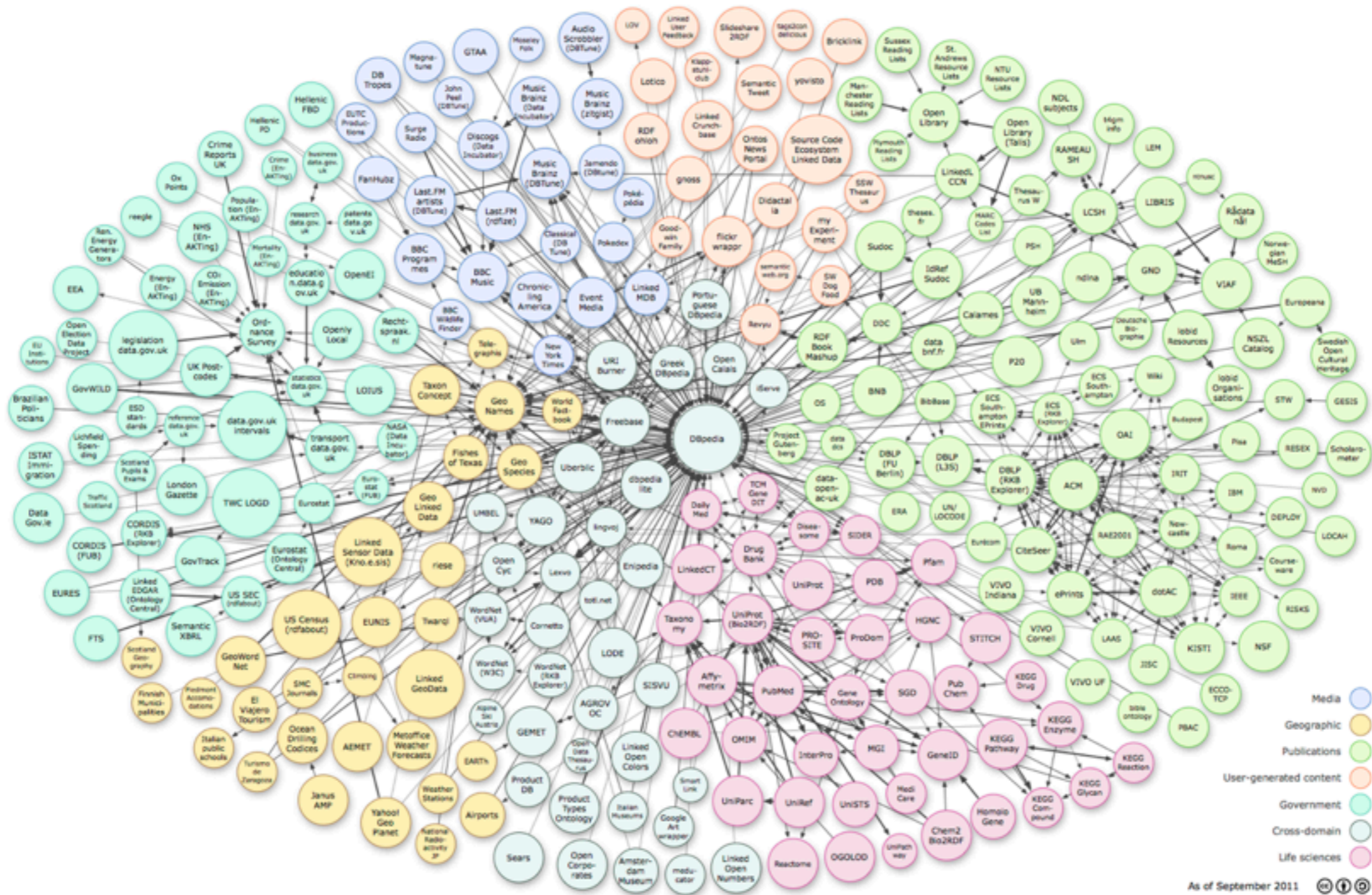


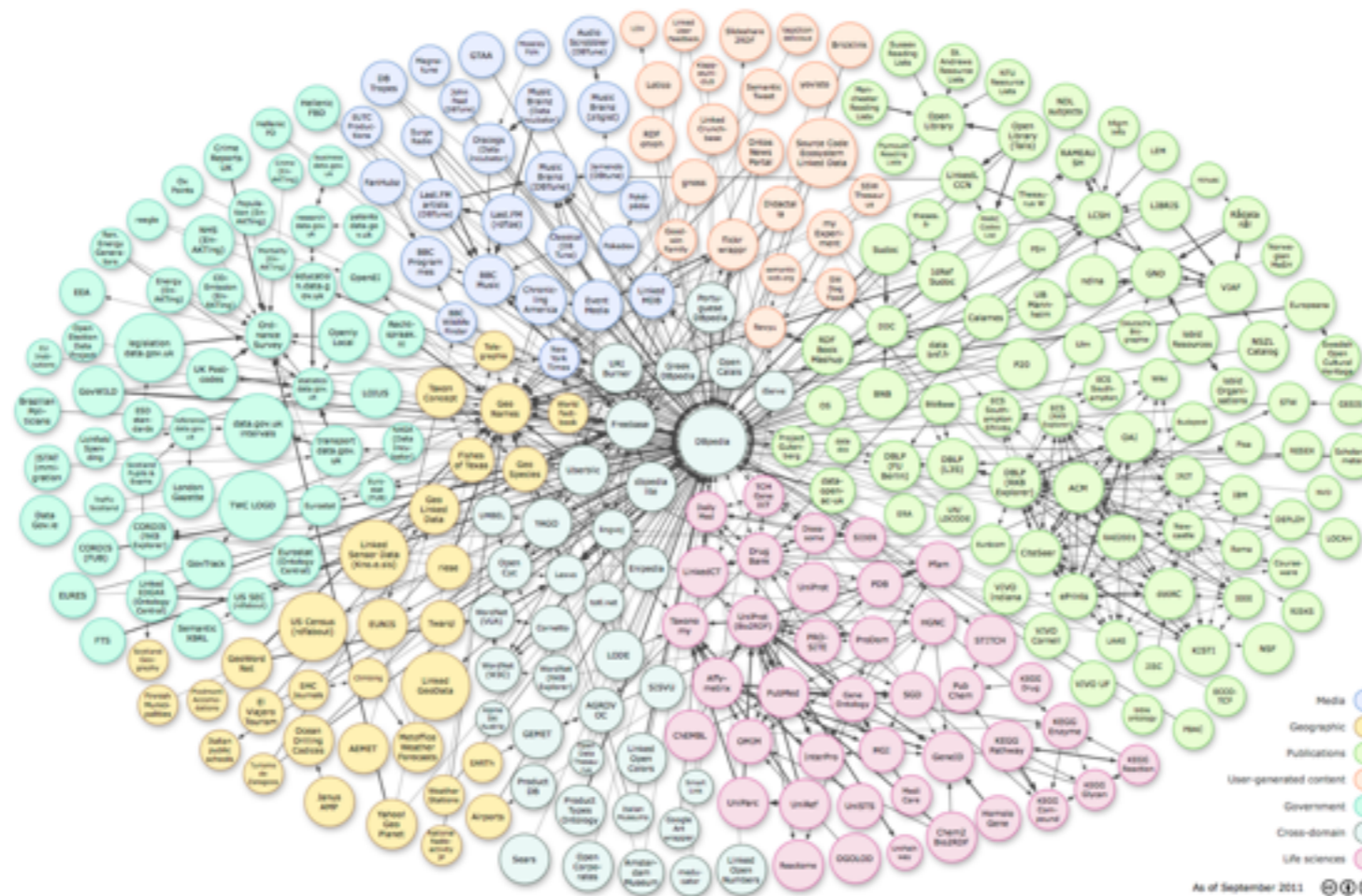
Matteo Collina  
Edgard Marx  
Hajar Ghaem Sigarchian  
Janez Starc  
Jon Tutchter

# TurtleComplete

Letting Users Search the Linked Open Data Cloud







The Linked Open  
Data Cloud is for us.

Unfortunately,  
us means  
researchers.



# Unreached Communities

- Web Developers
- Small Businesses
- General Public

} Why?

# “Goalkeeper playing for Real Madrid”

The image shows a search engine interface with a search bar containing the text "Goalkeeper playing for Real Madrid". Below the search bar, there are navigation tabs for "Web", "Immagini", "Maps", "Shopping", "Altro", and "Strumenti di ricerca". The search results are displayed in a list format, including a suggestion for "Cerca ri Preferenze" and several Wikipedia links related to "Iker Casillas" and "List of Real Madrid".

Overlaid on the search results is a window for "SIG.MA SEMANTIC INFORMATION MASHUP". This window has a search bar with the same query "Goalkeeper playing for Real Ma" and buttons for "Add More Info" and "Start New". Below the search bar, there is a section for "Carlos Manuel Brito" with a picture and a list of names: "given name: Jordi [2], Carlos [4], Hugo [8], Raul [13], Joaquin E".

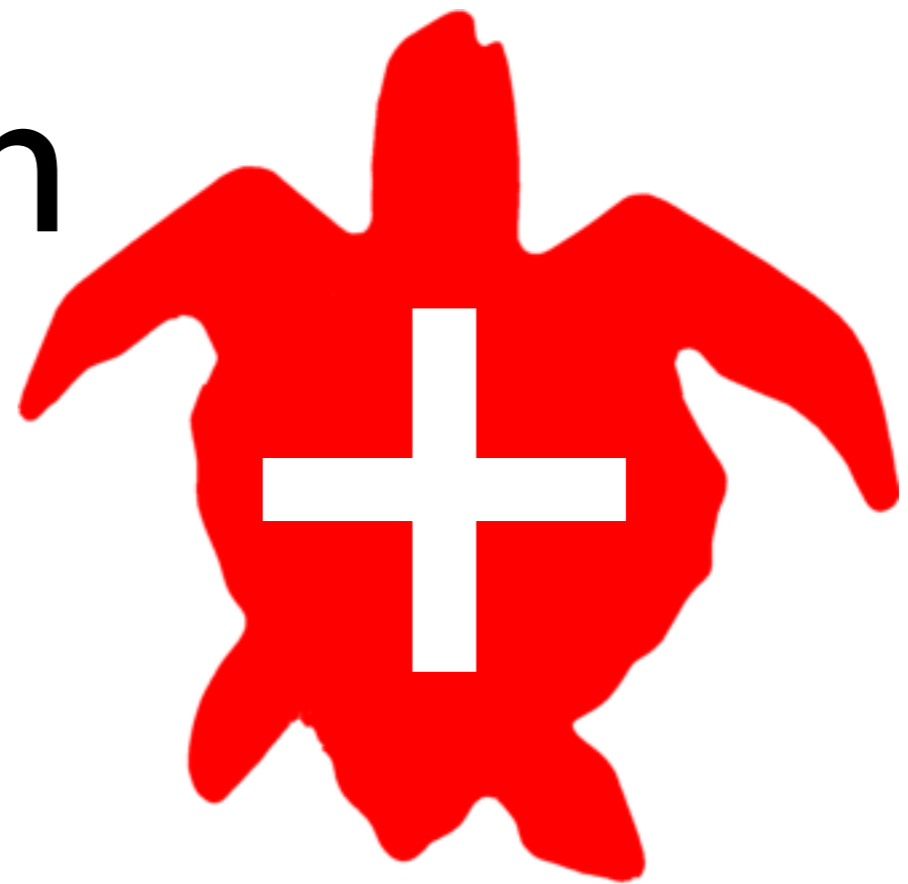
Another window, "indice THE SEMANTIC WEB INDEX", is overlaid on the bottom right. It shows a search interface with the same query "Goalkeeper playing for Real Madrid" and a "SEARCH" button. The results are sorted by "relevance" and show a list of documents found, including "Cricket stars bat in campaign against AIDS, goalkeeper score...", "Jordi Codina Rodriguez, ジョルディ・コディーナ, Кодина, Хорди, Jordi Co...", and "Carlos Manuel Brito Leal Queiroz, Carlos Queiroz, 卡路士·昆洛斯, C...".

However...  
we have the data!

dbpedia.org/page/Iker\_Casillas

dbpprop:caps	<ul style="list-style-type: none"><li>4 (xsd:integer)</li><li>26 (xsd:integer)</li><li>457 (xsd:integer)</li></ul>
dbpprop:clubnumber	<ul style="list-style-type: none"><li>1 (xsd:integer)</li></ul>
dbpprop:clubs	<ul style="list-style-type: none"><li>dbpedia:Real_Madrid_C.F.</li><li>dbpedia:Real_Madrid_Castilla</li><li>dbpedia:Real_Madrid_C</li></ul>
dbpprop:currentclub	<ul style="list-style-type: none"><li>dbpedia:Real_Madrid_C.F.</li></ul>

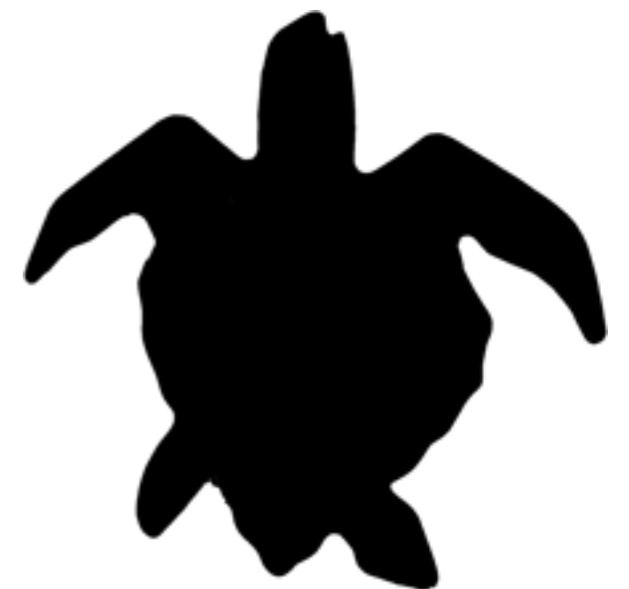
**TurtleComplete  
is the Solution**



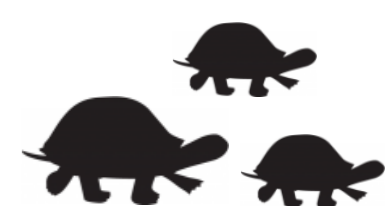
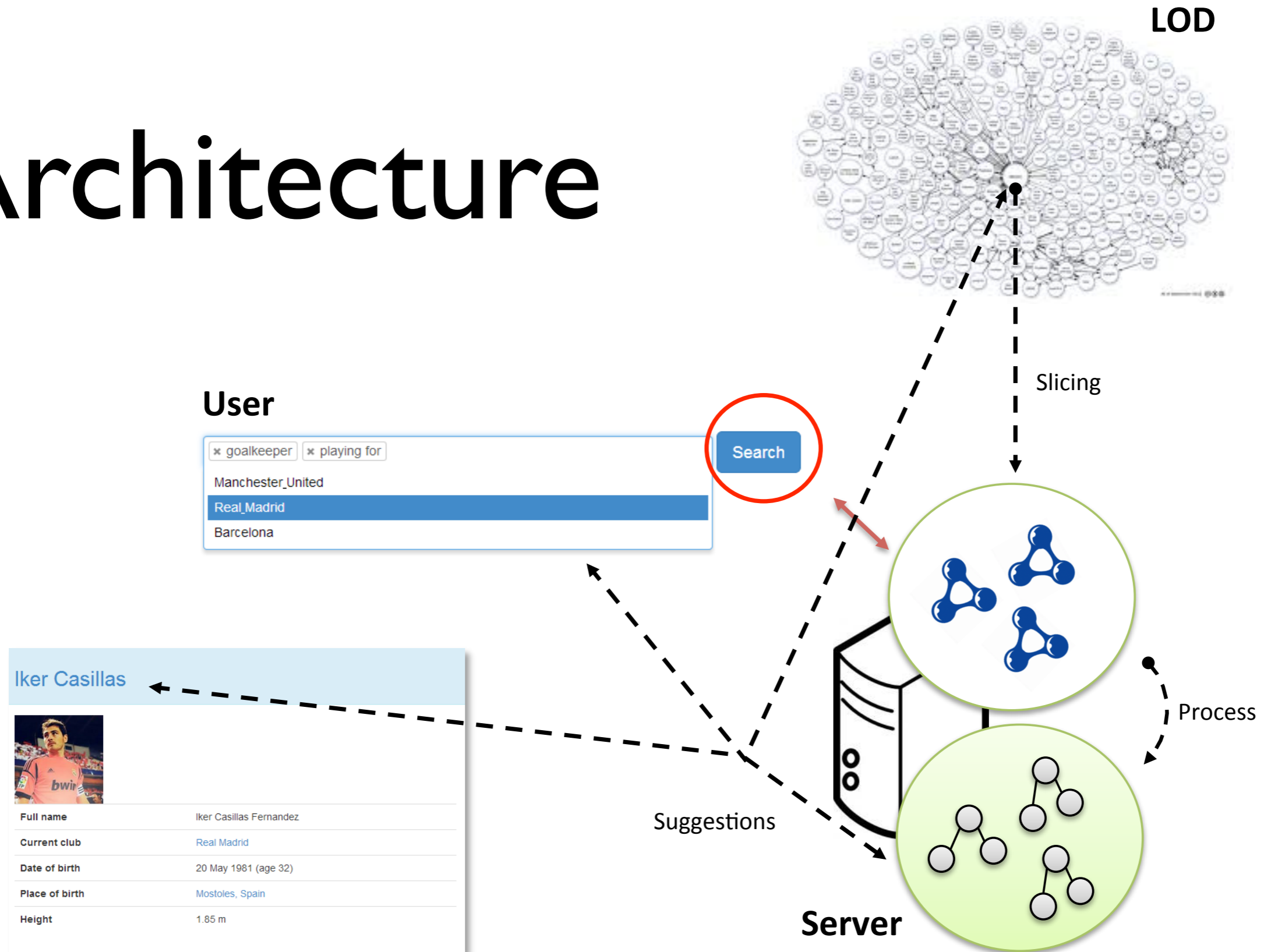


# Underlying Principles

- No Dataset Naming
- Guided Search
- ML for Improved Experience

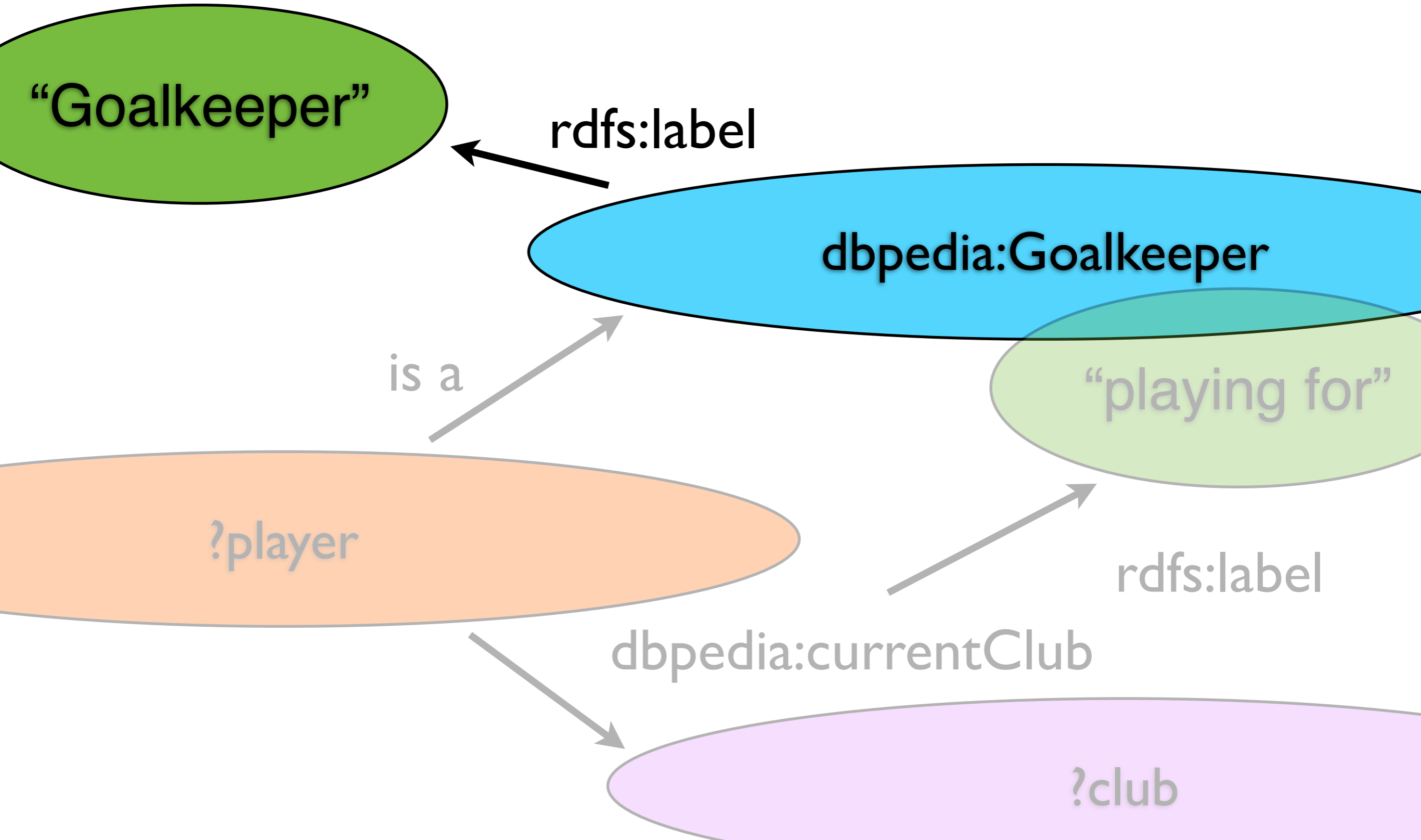


# Architecture



goal  
goalkeeper  
goal  
goat

Search



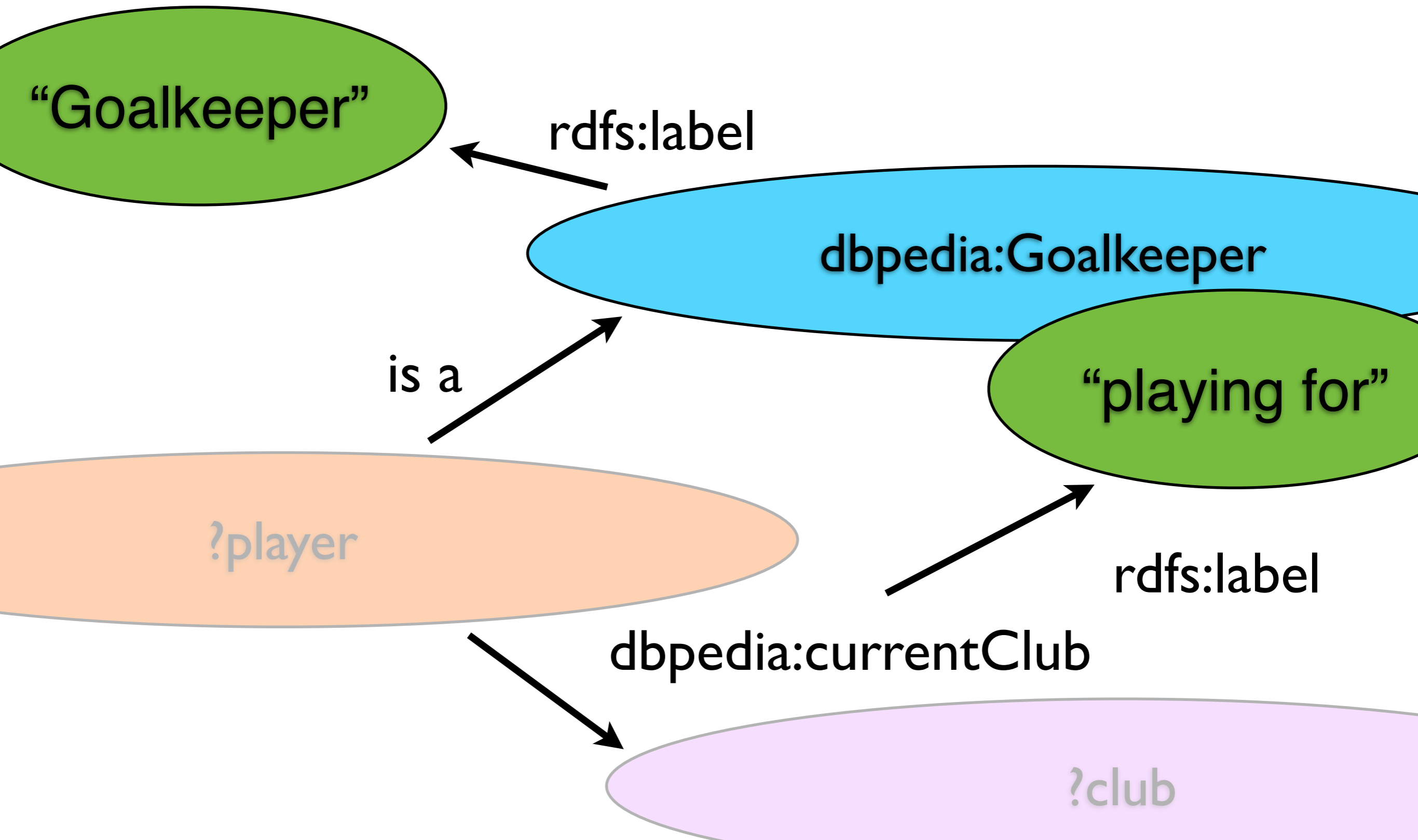
\* goalkeeper p|

playing for

place of birth

photo collection

Search



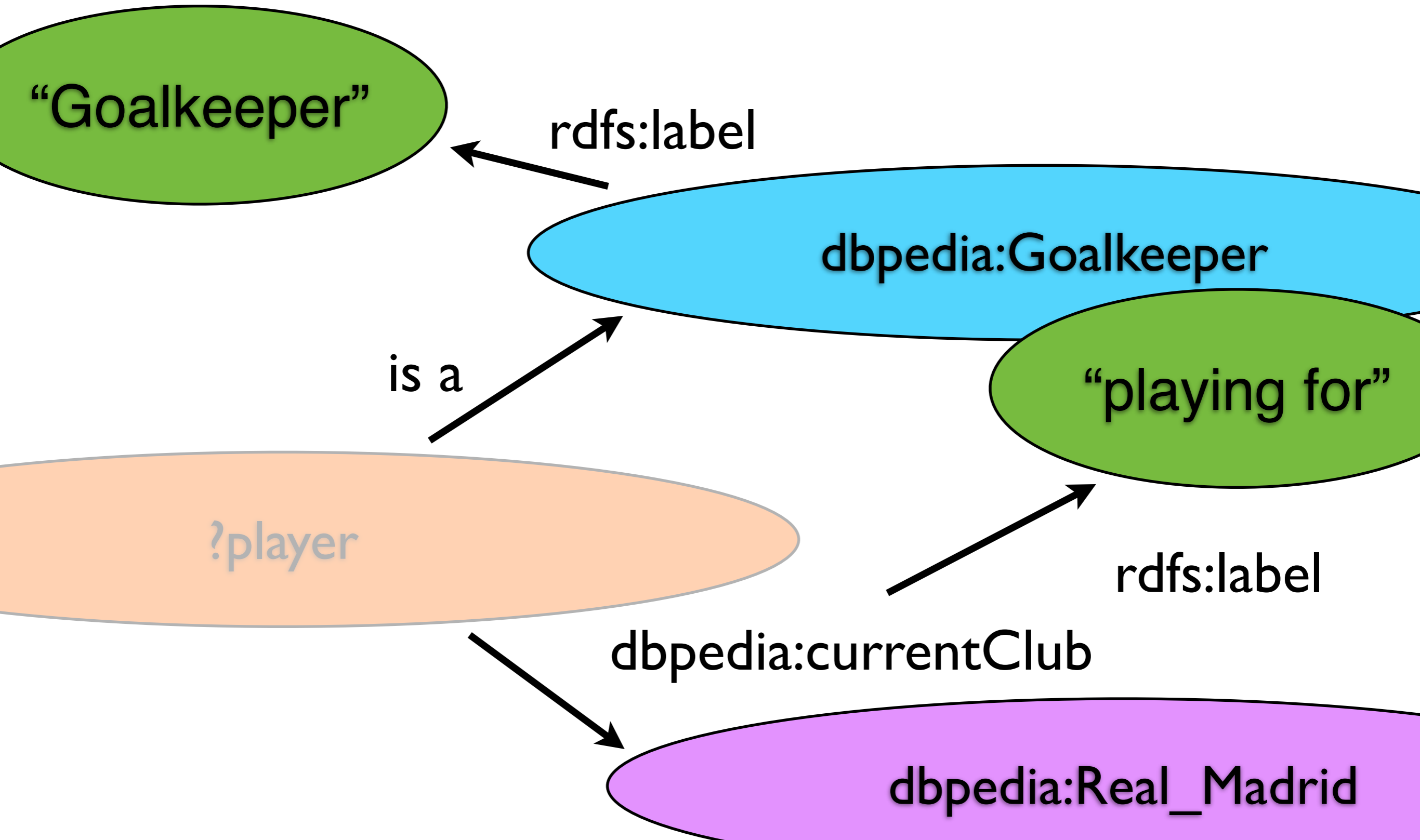
× goalkeeper × playing for

Manchester\_United

**Real\_Madrid**

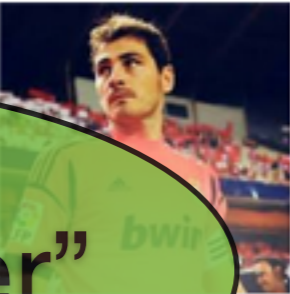
Barcelona

Search



× goalkeeper × playing for × Real Madrid Search

## Iker Casillas



Full name	Iker Casillas Fernández
Current club	Real Madrid
Date of birth	20 May 1981 (age 32)
Place of birth	Mostoles, Spain
Height	1.85 m

“Goalkeeper”

rdfs:label

dbpedia:Goalkeeper

“playing for”

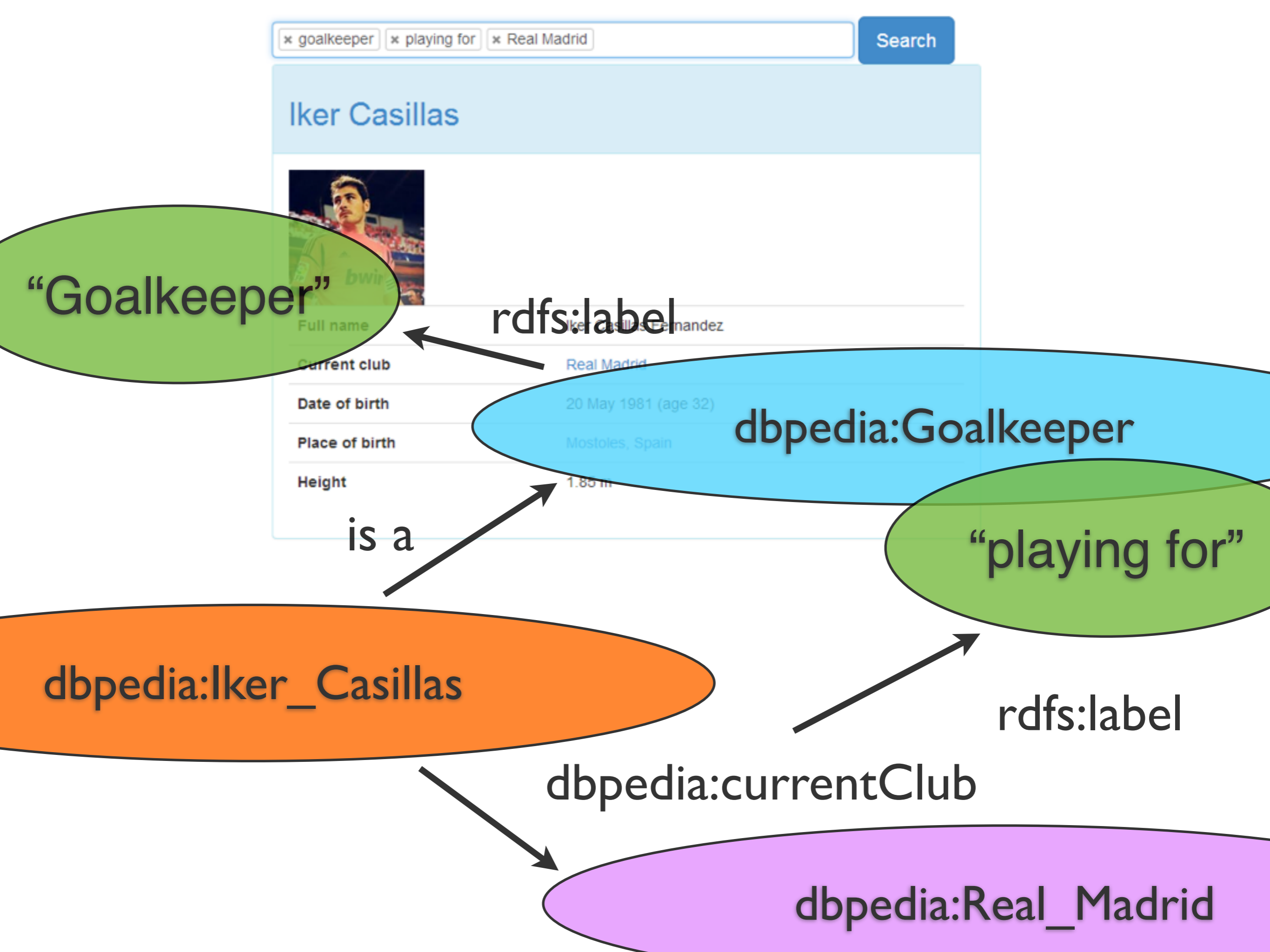
is a

dbpedia:Iker\_Casillas

rdfs:label

dbpedia:currentClub

dbpedia:Real\_Madrid

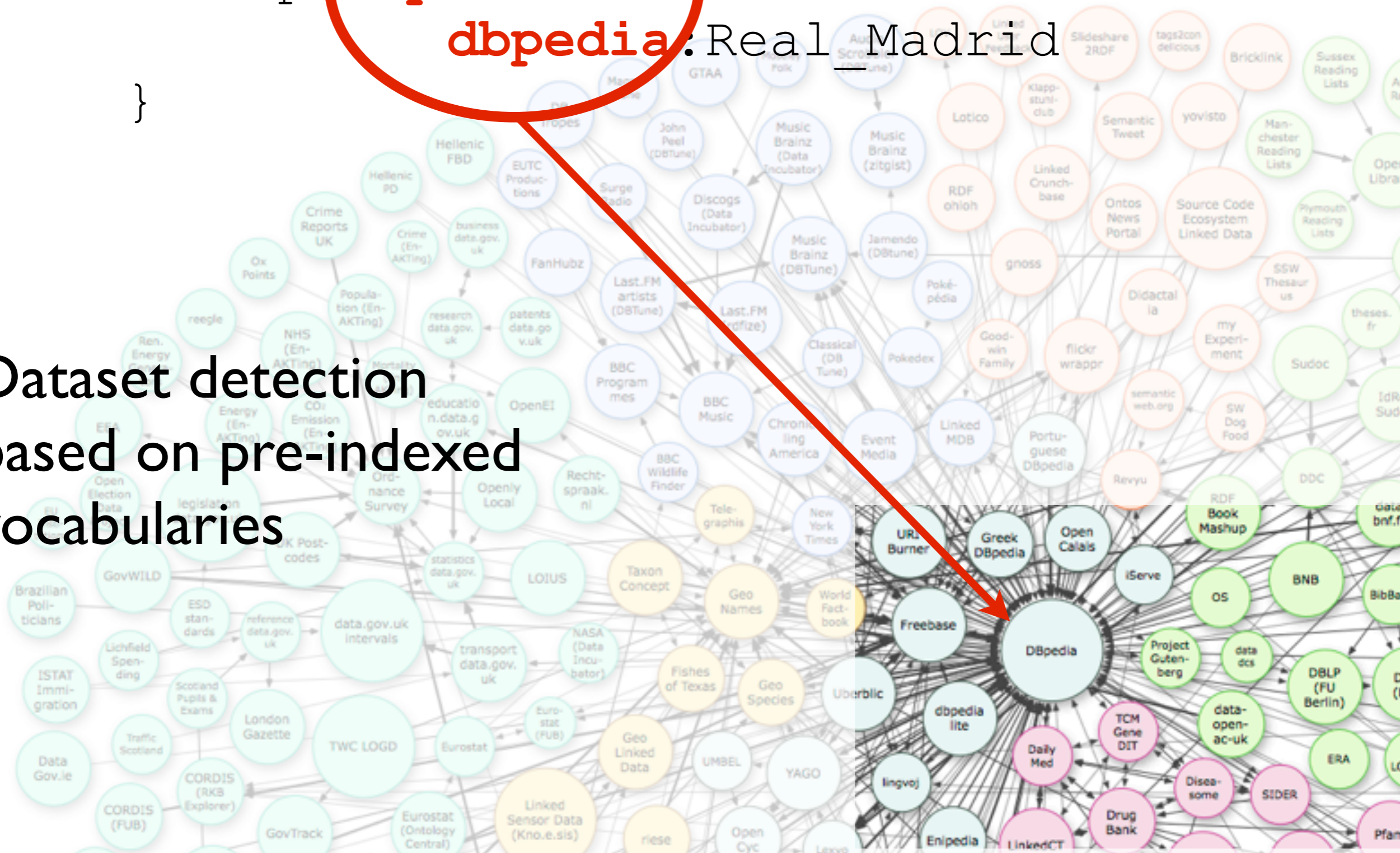


```

SELECT *
WHERE {
  ?p a dbpedia:Goalkeeper
  ?p dbpedia:currentClub
     dbpedia:Real_Madrid
}

```

Dataset detection  
based on pre-indexed  
vocabularies



# Machine Learning

- Previous successes refine autocomplete selection
- Ranking of datasets based on negative feedback



# DEMO!



[matteo.collina@unibo.it](mailto:matteo.collina@unibo.it)  
[marx@informatik.uni-leipzig.de](mailto:marx@informatik.uni-leipzig.de)  
[hajar.ghaemsigarchian@ugent.be](mailto:hajar.ghaemsigarchian@ugent.be)  
[janez.starc@ijs.si](mailto:janez.starc@ijs.si)  
[jxt548@bham.ac.uk](mailto:jxt548@bham.ac.uk)

Matteo Collina  
Edgard Marx  
Hajar Ghaem  
Sigarchian  
Janez Starc  
Jon Tutcher

Thanks!



# References



- Papadakis I., and Apostolatos I., “A LOD-based, query construction/refinement service for web search engines”. In proceedings to Journal of Semantic Web, Greece, 2013.
- Cammarano M. et. al., “Visualization of Heterogeneous Data”, IEEE Trans. Vis. Comput. Graph. 13(6):1200-1207 (2007)
- Simperl, E., Norton, B., and Vrandečić, D. (2011), “Crowdsourcing tasks in Linked Data management”. In Proceedings of the 2nd Workshop on Consuming Linked Data COLD2011 co-located with the 10th International Semantic Web Conference ISWC 2011.
- Joachims, T, “Optimizing search engines using clickthrough data”. In Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 133-142). ACM (2002, July).