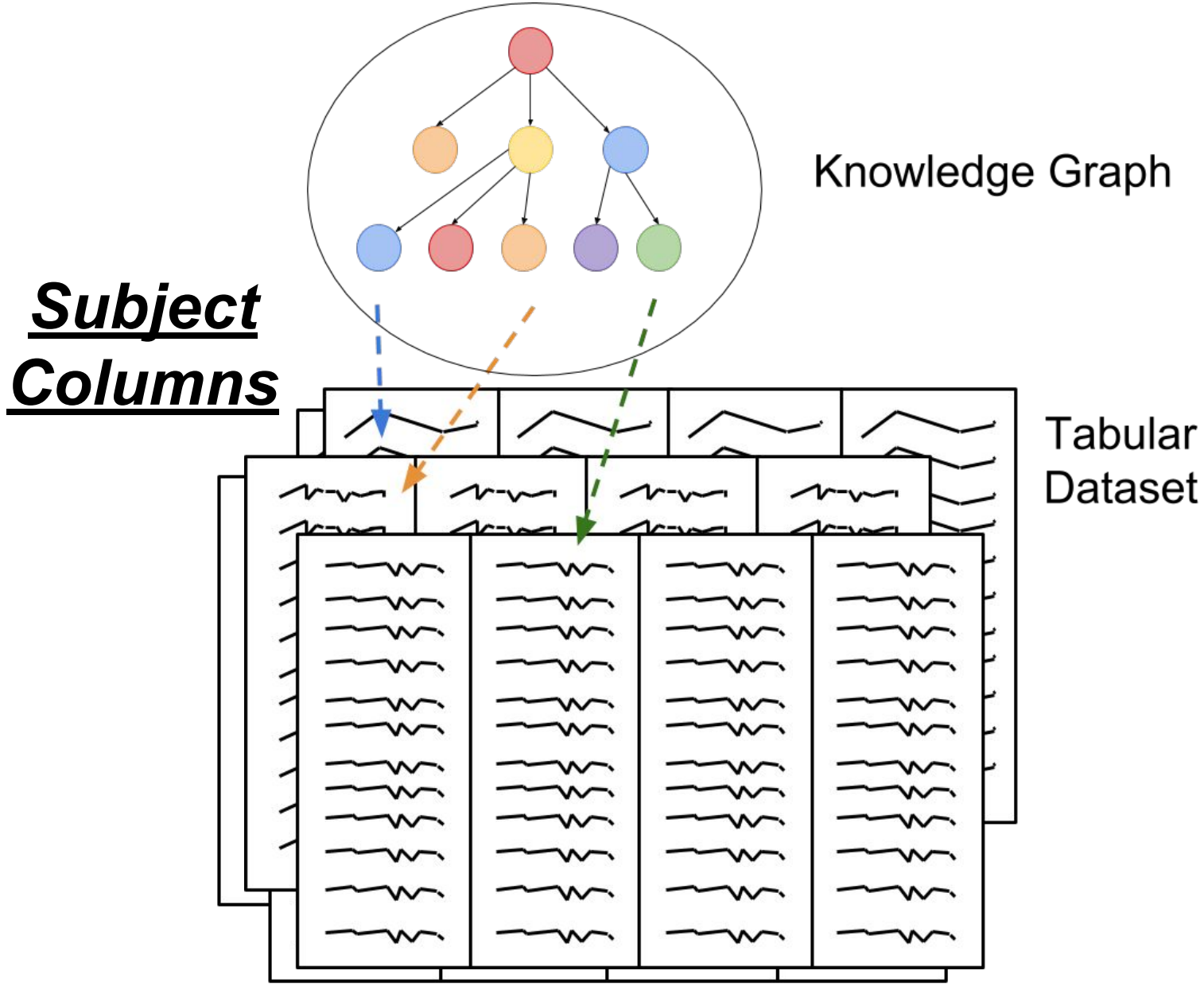
Knowledge-Graph-Based Semantic Labeling: Balancing Coverage and Specificity

Ahmad Alobaid and Oscar Corcho
Ontology Engineering Group
Universidad Politécnica de Madrid, Spain

aalobaid@fi.upm.es
ahmad88csc

11-4-2019
1003, block 1, Montegancedo

CC BY SA


Knowledge Graph
Subject Columns
Tabular Dataset

Subject Column

creator	{|creation}	{|age (years)
louis braille	invented a system of printing and writir	15
blaise pascal	formulated pascal,Äôs theorem	16
galileo galilei	discovered the laws of pendulum motio	17
edwin land	patented his first polarizing light filter	19
george westinghouse	first patent for a rotary steam engineir	1922
guglielmo marconi	invented a system of radio telegraphy	21
joshua lederberg	discovered bacterial conjugation	21
thomas edison	invented automated relaying communi	22
carl gauss	proved the theorem of complex coeffici	22
john nash	published his theory of non-cooperative	22
brian josephson	predicted the josephson effect	22
james hillier	developed the electron microscope	22
isaac newton	calculus principles of optics elements of	23
louis parker	invented a low frequency receiver for ra	23
srinivasa ramanujan	published his first mathematical papers	24
satyendra bose	published is first statistical mechanics p	24
paul dirac	quantum mechanics for motion of atom	24
richard feynman	published his theory of electromagnetic	24

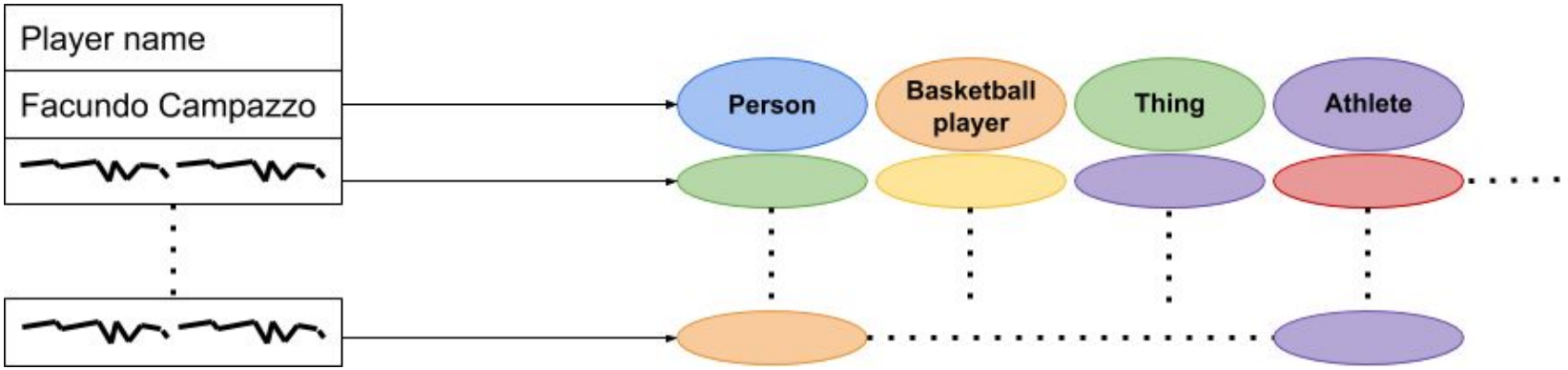
How? link cells to entities?

1. Get entities for each cell

```
select distinct ?subject
where{select distinct ?subject
where{
?subject ?property "Facundo Campazzo"@en}
```

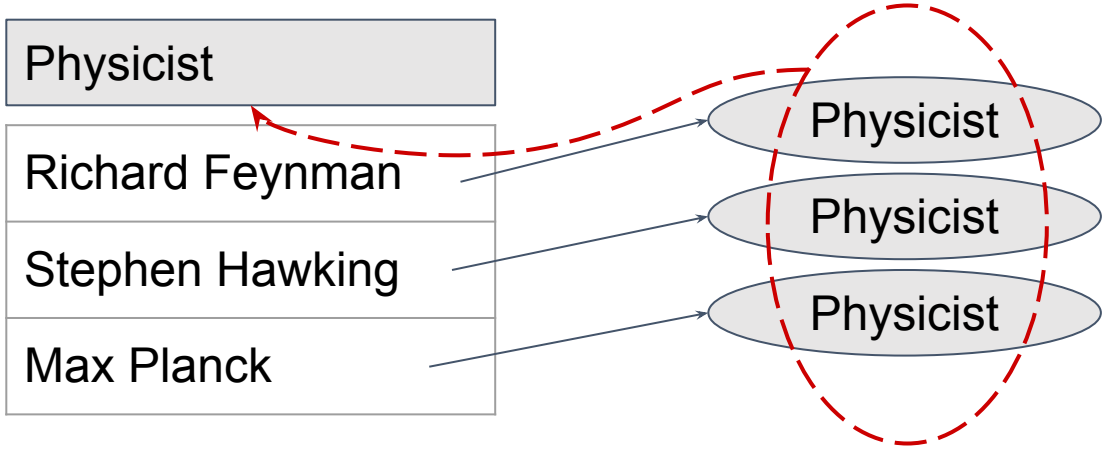
2. Get types for each entity

```
select distinct ?class where{
<http://dbpedia.org/resource/Facundo_Campazzo> a ?class}
```

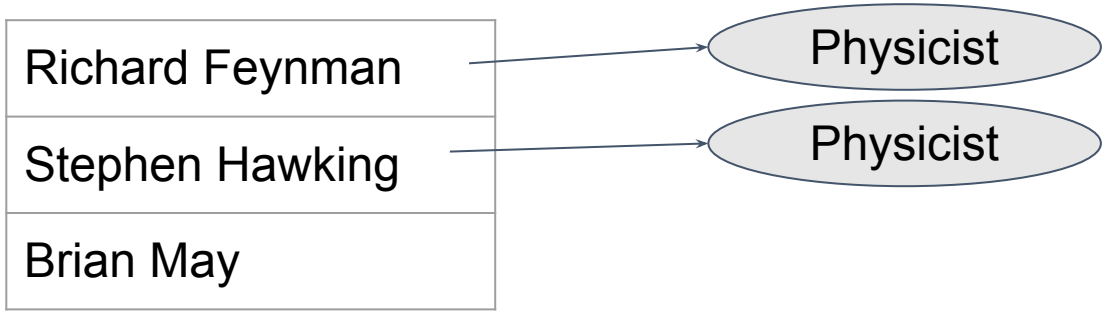


Now what?

1. Obvious case:



2. Not-so-obvious case:



https://cdn.britannica.com/s:300x300/73/20973-004-F71E20CB.jpg
https://upload.wikimedia.org/wikipedia/en/thumb/4/42/Richard_Feynman_Nobel.jpg/220px-Richard_Feynman_Nobel.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685

1. Obvious case:

Physicist
Richard Feyr
Stephen Hav
Max Planck

2. Not-so-c

Richard Feyr
Stephen Hawking
Brian May

https://cdn.britannica.com/s:300x300/73/20973-004-F71E20CB.jpg
https://upload.wikimedia.org/wikipedia/en/thumb/4/42/Richard_Feynman_Nobel.jpg/220px-Richard_Feynman_Nobel.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685

1. Obvious case:

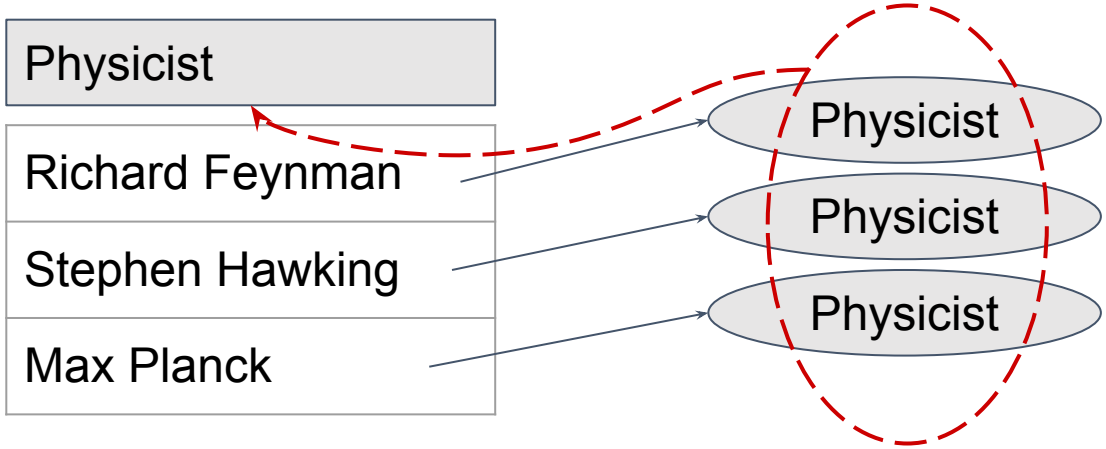
Physicist
Richard Feyr
Stephen Hav
Max Planck

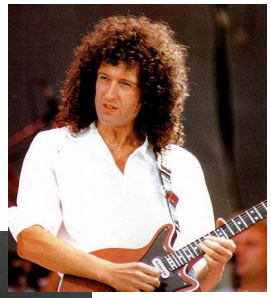
2. Not-so-c

Richard Feyr
Stephen Hawking
Brian May

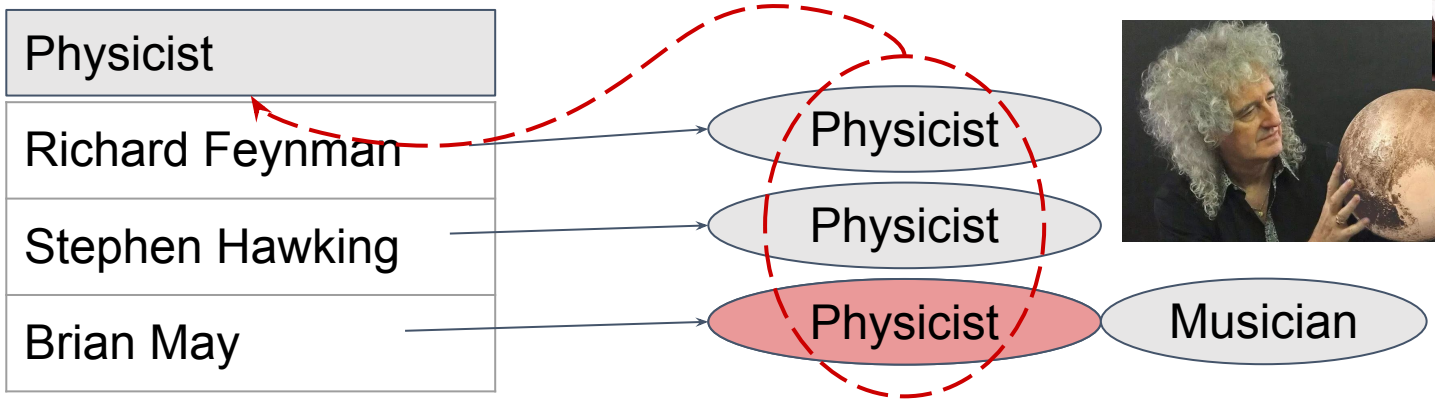
https://cdn.britannica.com/s:300x300/73/20973-004-F71E20CB.jpg
https://upload.wikimedia.org/wikipedia/en/thumb/4/42/Richard_Feynman_Nobel.jpg/220px-Richard_Feynman_Nobel.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685

1. Obvious case:





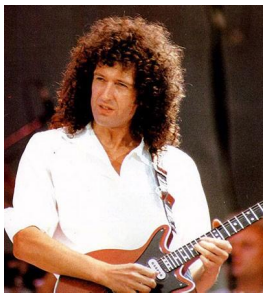
2. Not-so-obvious case:



https://cdn.britannica.com/s:300x300/73/20973-004-F71E20CB.jpg
https://upload.wikimedia.org/wikipedia/en/thumb/4/42/Richard_Feynman_Nobel.jpg/220px-Richard_Feynman_Nobel.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685
https://i1.wp.com/blog.eil.com/wp-content/uploads/2018/12/brian-may-astrophysicist-07212016.jpg?fit=970%2C545&ssl=1

3. Tricky cases

?
Brian May
Leonardo da Vinci
Pharrell Williams



Physicist, Musician

Physicist, Musician, Mathematician, Botanist

https://cdn.britannica.com/s:300x300/73/20973-004-F71E20CB.jpg
https://upload.wikimedia.org/wikipedia/en/thumb/4/42/Richard_Feynman_Nobel.jpg/220px-Richard_Feynman_Nobel.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685
https://i1.wp.com/blog.eil.com/wp-content/uploads/2018/12/brian-may-astrophysicist-07212016.jpg?fit=970%2C545&ssl=1

3. Tricky cases

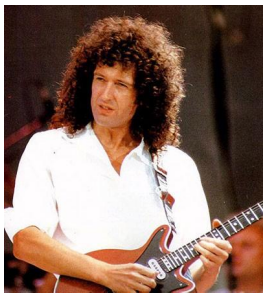
Brian Ma
Leonard
Pharrell

an

Botanist

https://cdn.britannica.com/s:300x300/73/20973-004-F71E20CB.jpg
https://upload.wikimedia.org/wikipedia/en/thumb/4/42/Richard_Feynman_Nobel.jpg/220px-Richard_Feynman_Nobel.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685
https://i1.wp.com/blog.eil.com/wp-content/uploads/2018/12/brian-may-astrophysicist-07212016.jpg?fit=970%2C545&ssl=1

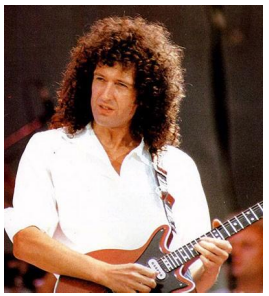
3. Tricky cases



Musician					
Brian May	→	Physicist	Musician		
Leonardo da Vinci	→	Physicist	Musician	Mathematician	Botanist
Pharrell Williams	→	Musician			

https://cdn.britannica.com/s:300x300/73/20973-004-F71E20CB.jpg
https://upload.wikimedia.org/wikipedia/en/thumb/4/42/Richard_Feynman_Nobel.jpg/220px-Richard_Feynman_Nobel.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685
https://i1.wp.com/blog.eil.com/wp-content/uploads/2018/12/brian-may-astrophysicist-07212016.jpg?fit=970%2C545&ssl=1

3. Tricky cases



?				
Brian May	Physicist	Musician		
Leonardo da Vinci	Physicist	Musician	Mathematician	Botanist
Pharrell Williams	Musician			
Bill Gates	Investor	Author	philanthropist	

https://cdn.britannica.com/s:300x300/73/20973-004-F71E20CB.jpg
https://upload.wikimedia.org/wikipedia/en/thumb/4/42/Richard_Feynman_Nobel.jpg/220px-Richard_Feynman_Nobel.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685
https://i1.wp.com/blog.eil.com/wp-content/uploads/2018/12/brian-may-astrophysicist-07212016.jpg?fit=970%2C545&ssl=1

3. Tricky cases

Person
Brian May
Leonardo da Vinci
Pharrell Williams
Bill Gates

Person

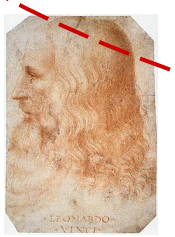
Musician

Person

philanthropist

Investor

Author



Person

Physicist

Botanist

Mathematician

Musician

Person

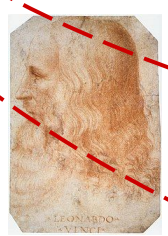
Physicist

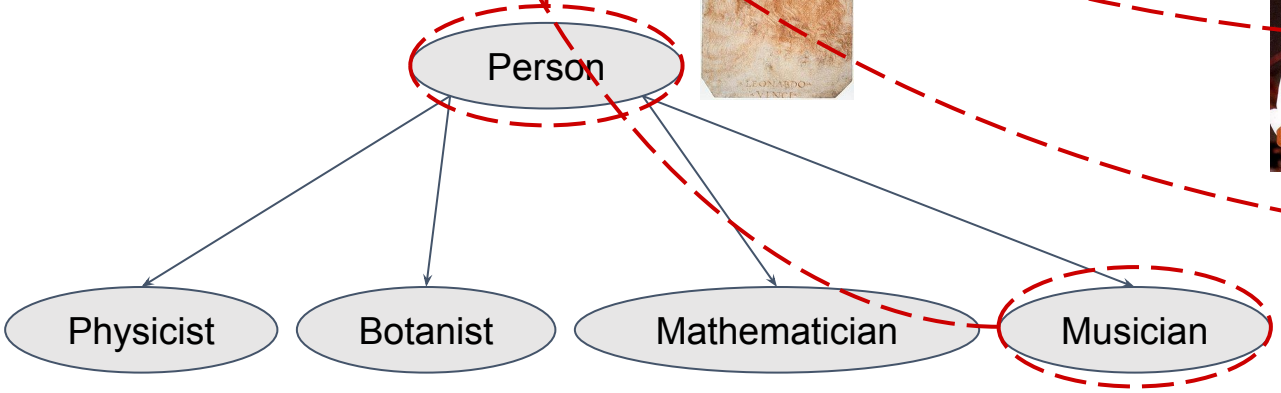
Musician

http://cdn01.cdn.justjared.com/wp-content/uploads/2014/03/williams-shorts/pharrell-williams-wear-shorts-on-oscars-2014-red-carpet-03.jpg
https://pbs.twimg.com/profile_images/988775660163252226/XpgonN0X.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685
https://i1.wp.com/blog.eil.com/wp-content/uploads/2018/12/brian-may-astrophysicist-07212016.jpg?fit=970%2C545&ssl=1

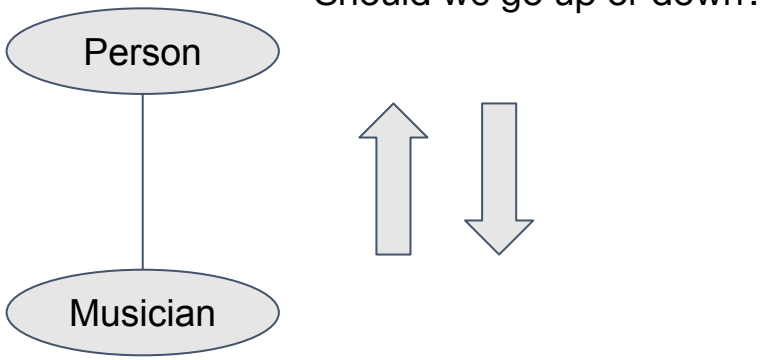
3. Tricky cases

Person
Brian May
Leonardo da Vinci
Pharrell Williams





http://cdn01.cdn.justjared.com/wp-content/uploads/2014/03/williams-shorts/pharrell-williams-wear-shorts-on-oscars-2014-red-carpet-03.jpg
https://pbs.twimg.com/profile_images/988775660163252226/XpgonN0X.jpg
https://www.queenie.cz/storage/temp/8f2b24db44d09bd8e3d563d0bb099fc2-400x800x1.jpg
https://www.thetimes.co.uk/imageserver/image/methode%2Ftimes%2Fprod%2Fweb%2Fbin%2F3f44abc8-2774-11e8-acc5-262aff1ca7a6.jpg?crop=988%2C556%2C910%2C214&resize=685
https://i1.wp.com/blog.eil.com/wp-content/uploads/2018/12/brian-may-astrophysicist-07212016.jpg?fit=970%2C545&ssl=1


Should we go up or down?
Person
Musician

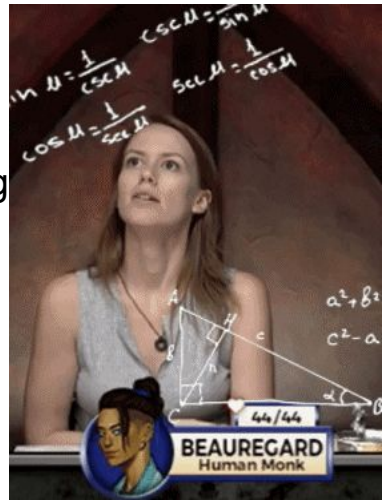
Coverage vs Specificity

$$\max_t f(t) = \alpha * f_c(t) + (1 - \alpha) * f_s(t)$$

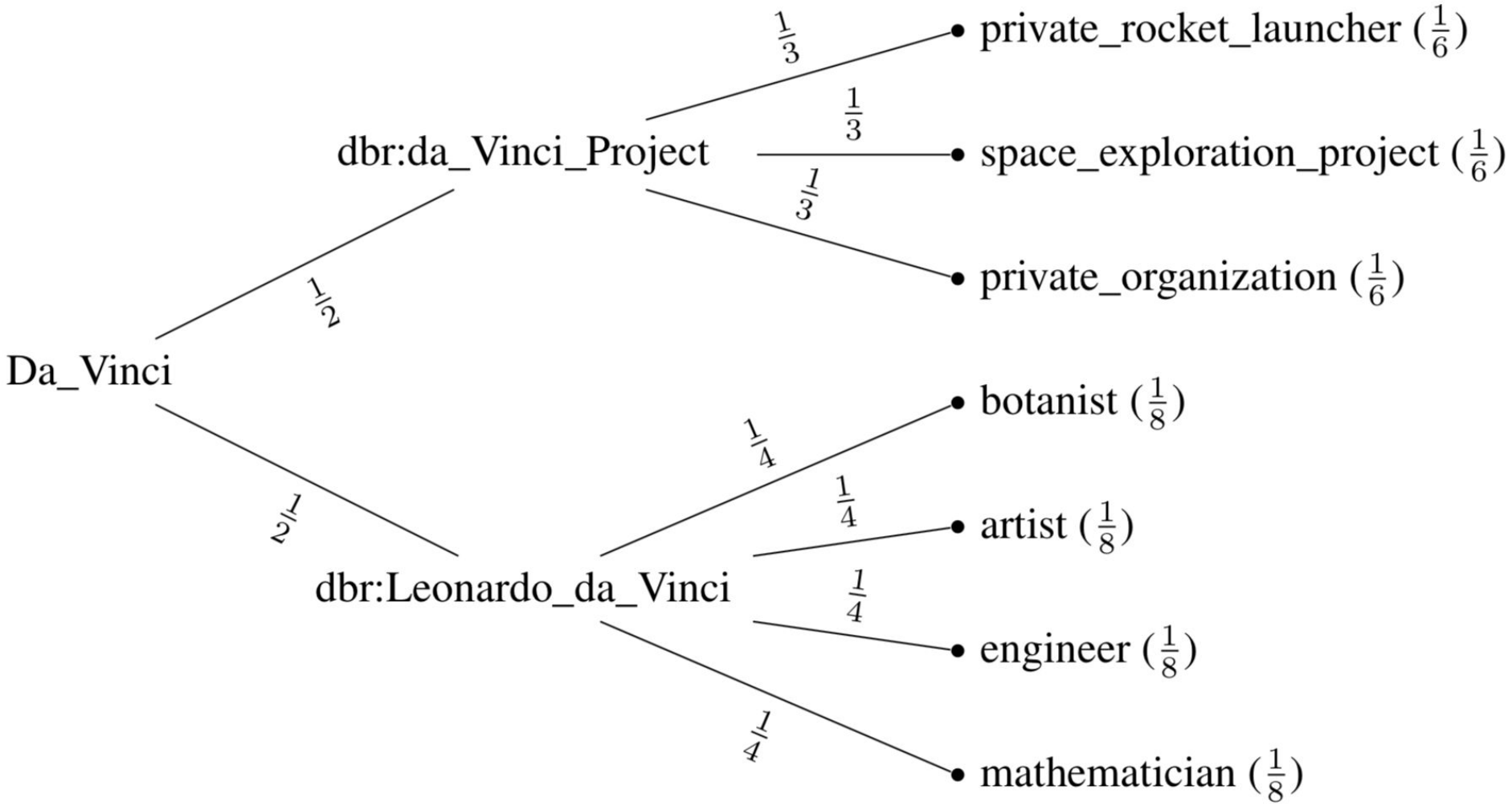
total score

types

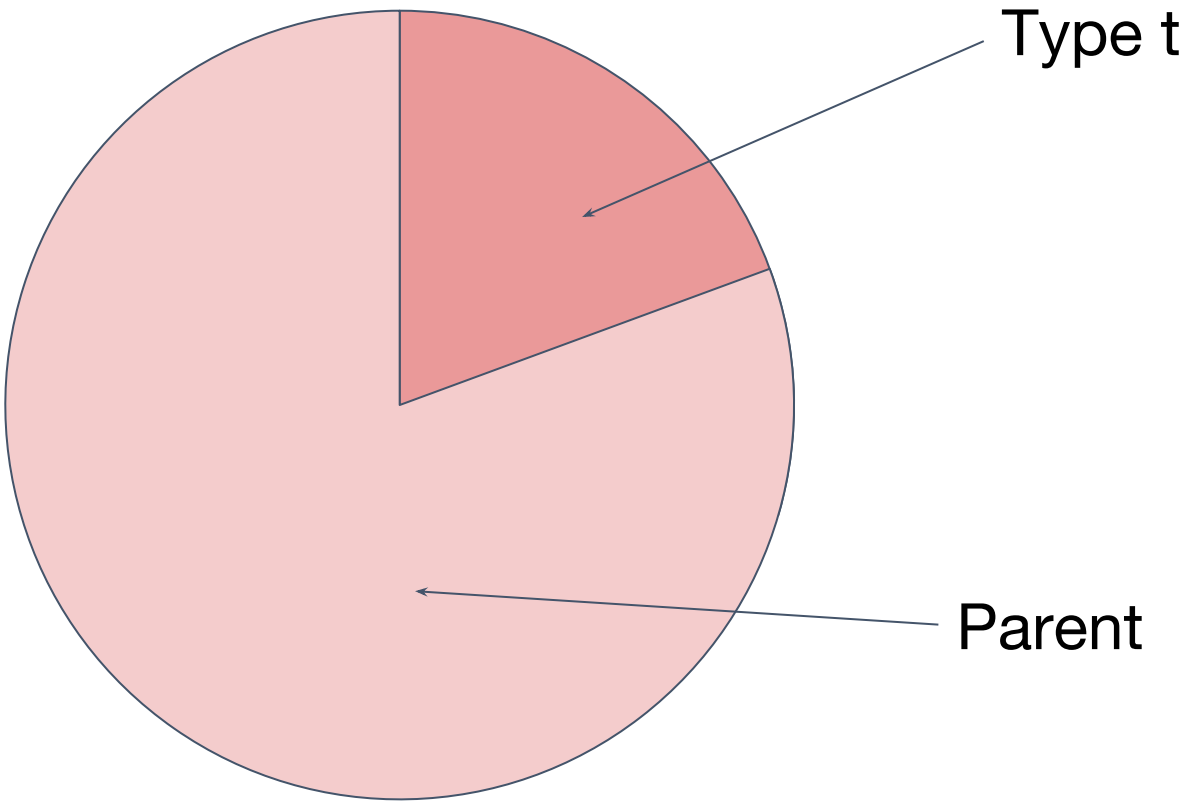
weig



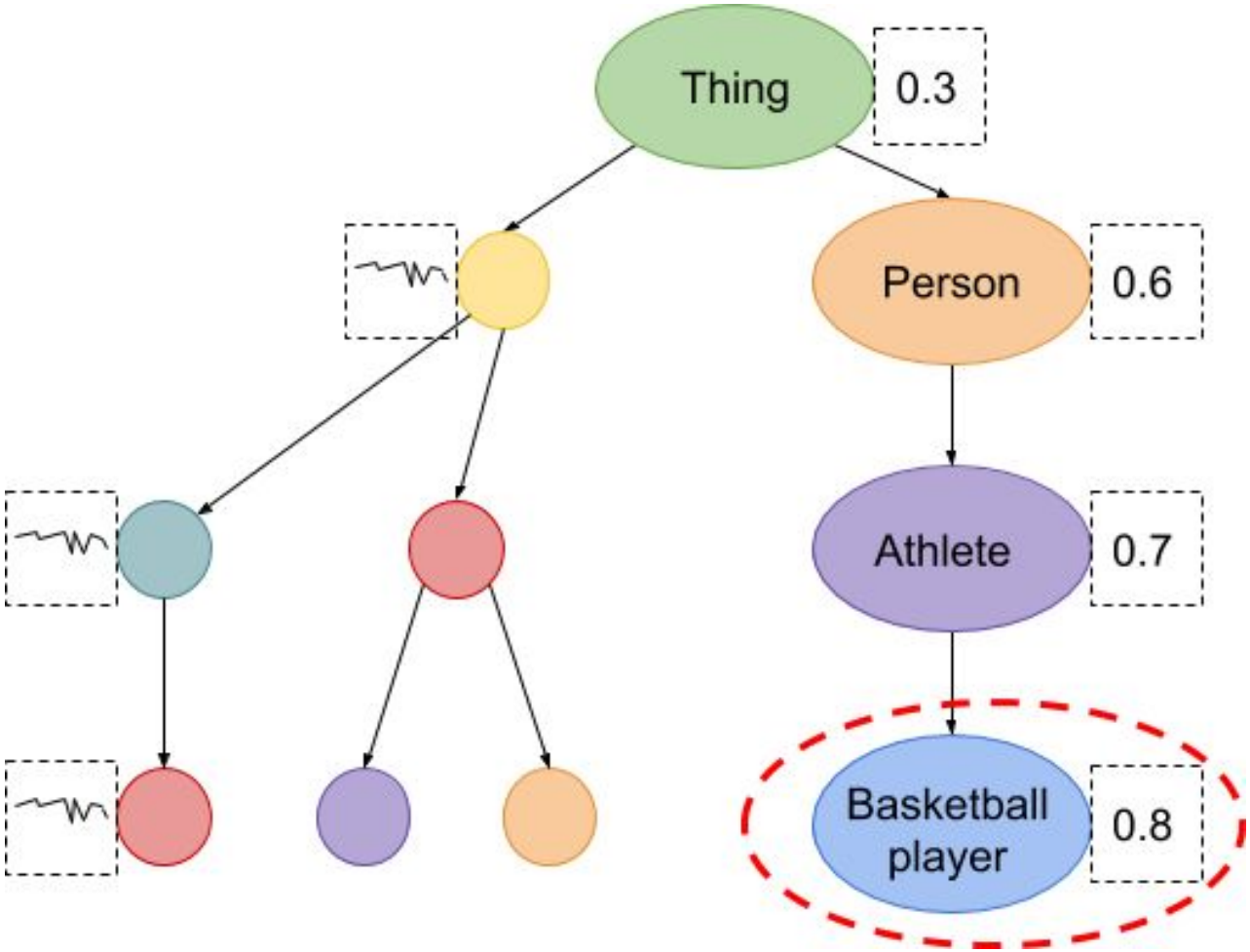
specificity score


Da_Vinci
1/2
dbr:da_Vinci_Project
1/3
private_rocket_launcher (1/6)
1/3
space_exploration_project (1/6)
1/3
private_organization (1/6)
1/2
dbr:Leonardo_da_Vinci
1/4
botanist (1/8)
1/4
artist (1/8)
1/4
engineer (1/8)
1/4
mathematician (1/8)

Specificity



Scored graph

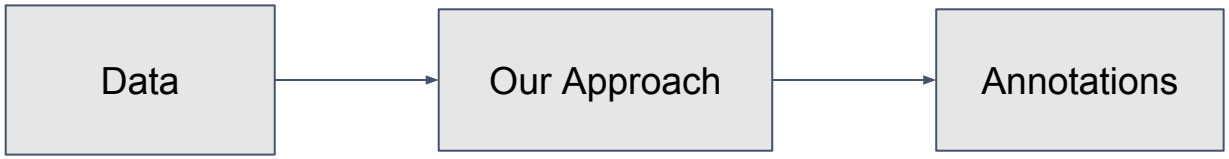


1. Show the potential of the scoring functions by proving their match to our intuitions.

2. Experiment with real data

Datasets:

Olympic Games
T2Dv1
T2Dv2



Olympic Games

f_s	Precision	Recall	F1
f_{s1}	1.0	1.0	1.0
f_{s2}	1.0	1.0	1.0
f_{s3}	1.0	1.0	1.0
f_{s4}	1.0	1.0	1.0
f_{s5}	1.0	1.0	1.0

T2Dv1

Approach	Precision	Recall	F1
T2K	0.94	0.94	**0.94**
TADA-Entity (f_{s1})	0.71	0.96	0.82
TADA-Entity (f_{s2})	0.78	0.96	0.86
TADA-Entity (f_{s3})	0.93	0.97	**0.95**
TADA-Entity (f_{s4})	0.88	0.97	0.92
TADA-Entity (f_{s5})	0.88	0.97	0.92

T2Dv2

Approach	Precision	Recall	F1
T2K (Majority)	0.47	0.51	0.49
T2K (Majority + Frequency)	0.87	0.90	0.89
T2K (Page attributes)	0.97	0.37	0.53
T2K (Text)	0.75	0.34	0.46
T2K (Majority + Frequency + Page attributes + Text)	0.90	0.86	0.88
T2K Extended	0.93	0.91	**0.92**
TADA-Entity (f_{s1})	0.68	0.96	0.79
TADA-Entity (f_{s2})	0.75	0.96	0.84
TADA-Entity (f_{s3})	0.91	0.97	**0.94**
TADA-Entity (f_{s4})	0.85	0.97	0.90
TADA-Entity (f_{s5})	0.84	0.97	0.90

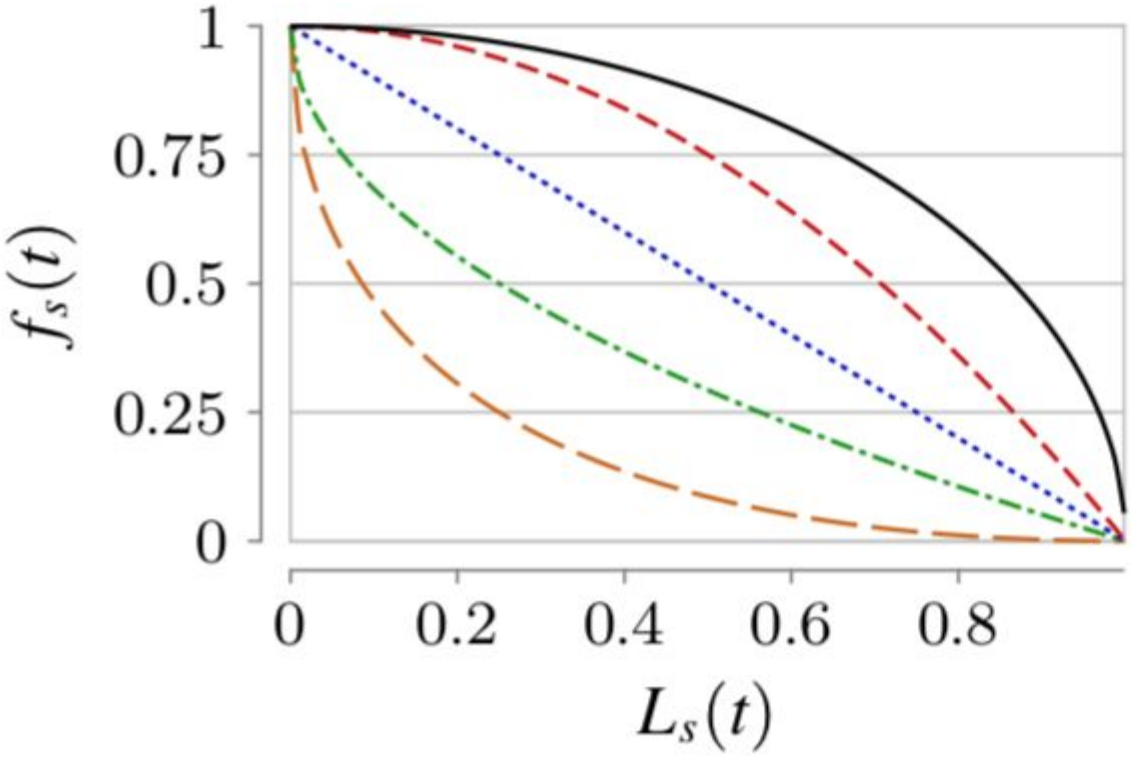
"We can perform semantic labeling without context and external linguistic resources, and yet reach high precision and recall."

Future work:

1. Find the optimal balancing factor
2. Use advanced entity-linking

Questions?

Specificity functions



$\sqrt{(1 - L_s(t)^2)}$

$-L_s(t)^2 + 1$

$-L_s(t) + 1$

$1 - \sqrt{L_s(t)}$

$(1 - \sqrt{L_s(t)})^2$