Entity Linking to Knowledge Graphs to Infer Column Types and Properties

Avijit Thawani, Minda Hu, Erdong Hu, Husain Zafar, Naren Teja Divvala, Amandeep Singh, Ehsan Qasemi, Pedro Szekely, and Jay Pujara

About Us

Team ISI:

- Information Sciences Institute
- University of Southern California

Me:

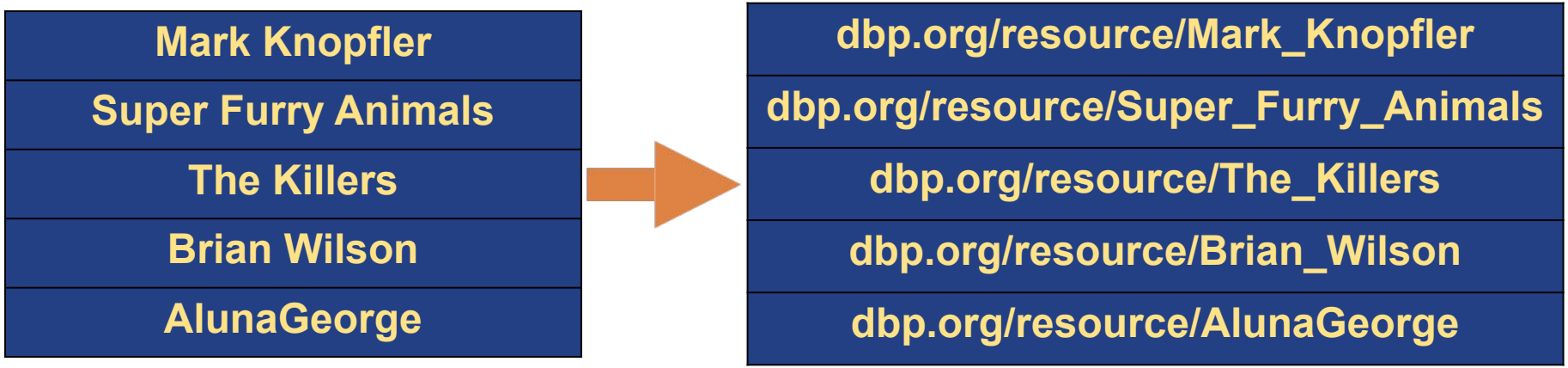
- PhD student, USC

Outline

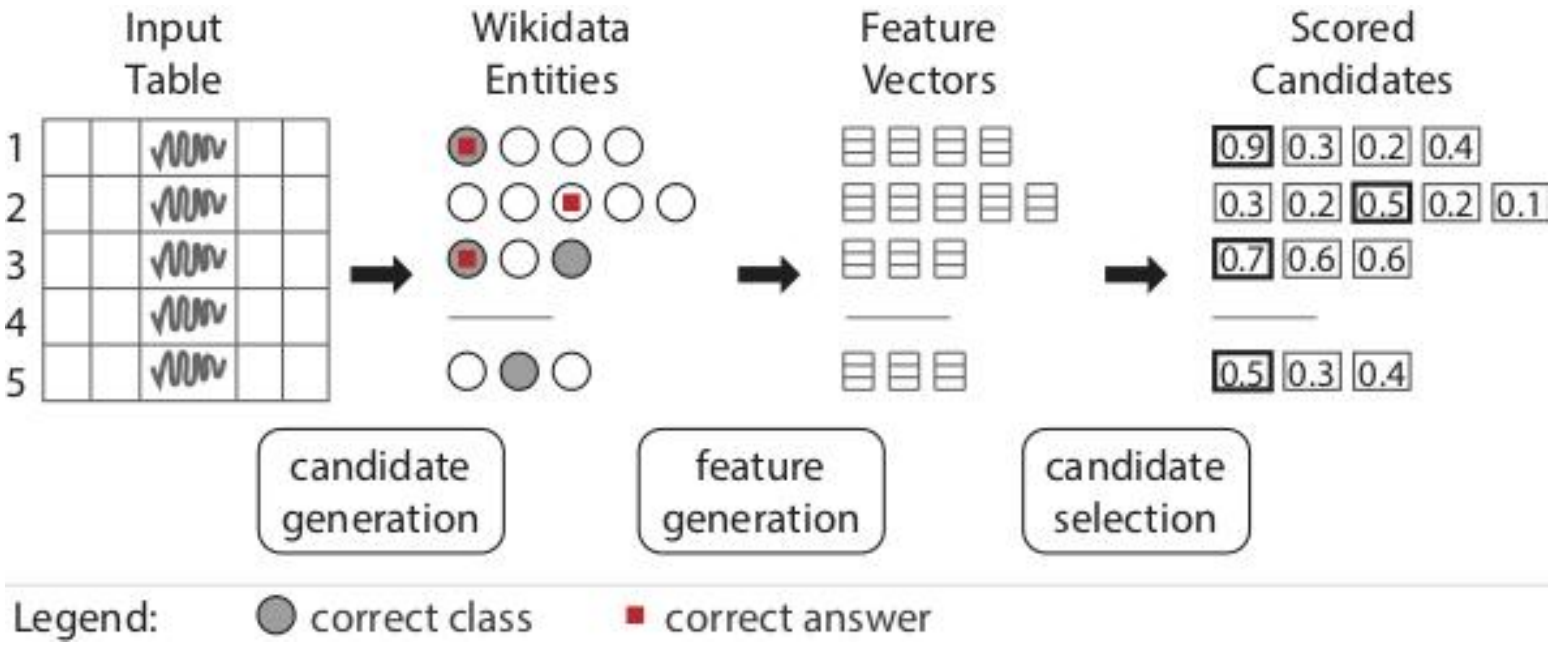
1. CEA
2. tf-idf
3. CTA and CPA
4. Shortcomings
5. Analysis
6. Appendix: PSL

1. CEA

Objective: CEA



Approach: CEA



Candidate Generation

string → **DBpedia, Wikidata entities**

The Killers →

Q220730 (American rock band)
Q205321 (1946 film by Robert Siodmak)
Q1197463 (1964 film by Don Siegel)
Q1213811 (book)
http://dbpedia.org/resource/Serial_killer

Lots of Cues

Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements		From related entities
discography	Brian Wilson discography	
instrument	bass guitar (electric or acoustic bass instrument)	
7 statements	guitar (fretted string instrument)	
	singing (act of producing musical sounds with the voice)	
work period (start)	1961	
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)	
topic's main template	Template:Brian Wilson (Wikimedia template)	
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)	
place of birth	Inglewood (city in California)	
birth name	Brian Douglas Wilson [en]	
country of citizenship	United States of America (federal republic in North America)	
given name	Brian (male given name)	

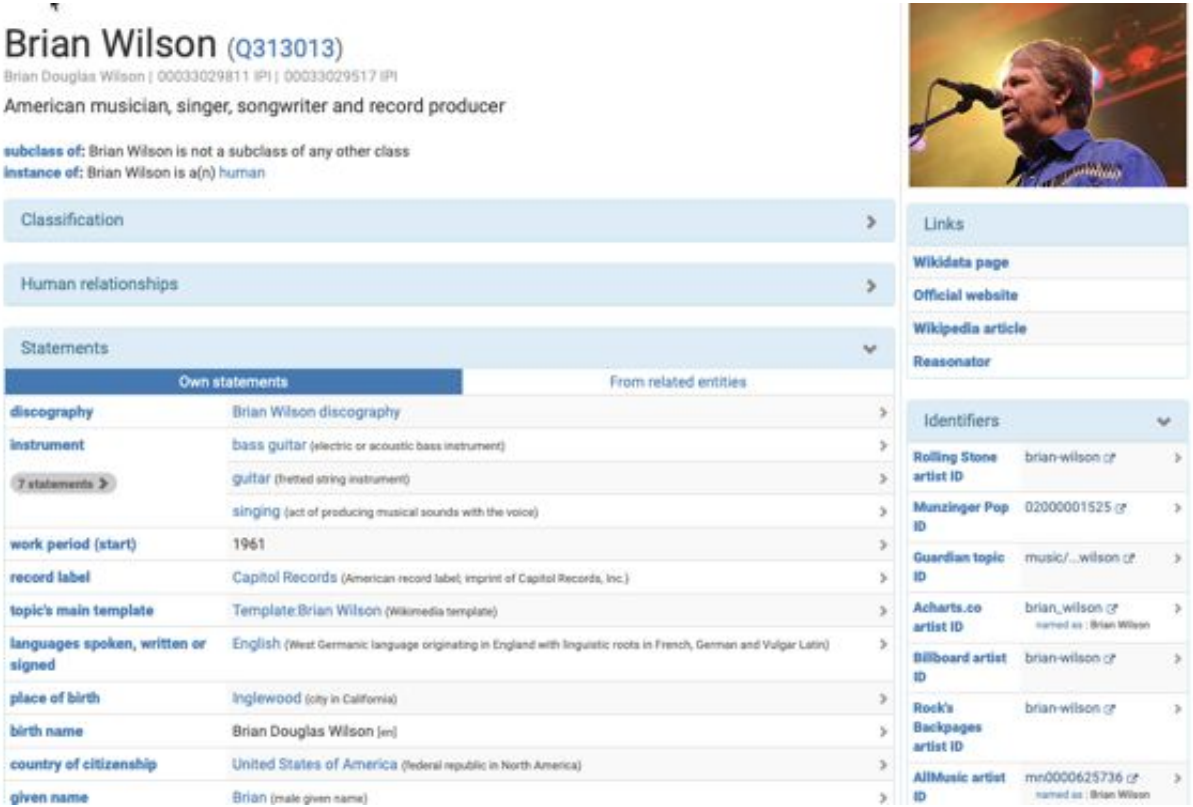
Links

- **Wikidata page**
- **Official website**
- **Wikipedia article**
- **Reasonator**

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson (named as : Brian Wilson)
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 (named as : Brian Wilson)

Lots of Cues



Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements | From related entities

discography	Brian Wilson discography
instrument	bass guitar (electric or acoustic bass instrument)
7 statements	guitar (fretted string instrument)
	singing (act of producing musical sounds with the voice)
work period (start)	1961
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)
topic's main template	Template:Brian Wilson (Wikimedia template)
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)
place of birth	Inglewood (city in California)
birth name	Brian Douglas Wilson [en]
country of citizenship	United States of America (federal republic in North America)
given name	Brian (male given name)

Links

Wikidata page
Official website
Wikipedia article
Reasonator

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson named as : Brian Wilson
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 named as : Brian Wilson

- Class

Lots of Cues

Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements		From related entities
discography	Brian Wilson discography	
instrument	bass guitar (electric or acoustic bass instrument)	
7 statements	guitar (fretted string instrument)	
	singing (act of producing musical sounds with the voice)	
work period (start)	1961	
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)	
topic's main template	Template:Brian Wilson (Wikimedia template)	
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)	
place of birth	Inglewood (city in California)	
birth name	Brian Douglas Wilson [en]	
country of citizenship	United States of America (federal republic in North America)	
given name	Brian (male given name)	

Links

- **Wikidata page**
- **Official website**
- **Wikipedia article**
- **Reasonator**

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson (named as : Brian Wilson)
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 (named as : Brian Wilson)

- Class
- Properties

Lots of Cues

Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements		From related entities
discography	Brian Wilson discography	
instrument	bass guitar (electric or acoustic bass instrument)	
7 statements	guitar (fretted string instrument)	
	singing (act of producing musical sounds with the voice)	
work period (start)	1961	
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)	
topic's main template	Template:Brian Wilson (Wikimedia template)	
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)	
place of birth	Inglewood (city in California)	
birth name	Brian Douglas Wilson [en]	
country of citizenship	United States of America (federal republic in North America)	
given name	Brian (male given name)	

Links

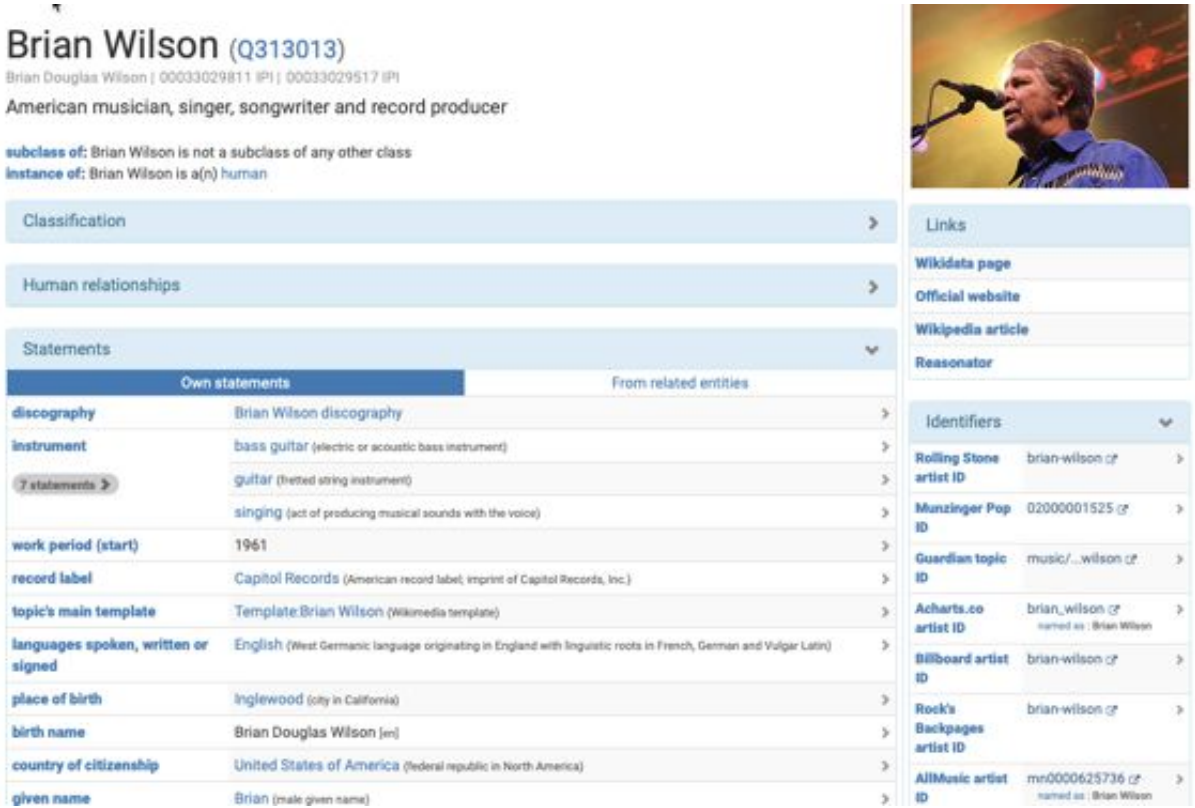
- **Wikidata page**
- **Official website**
- **Wikipedia article**
- **Reasonator**

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson (named as : Brian Wilson)
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 (named as : Brian Wilson)

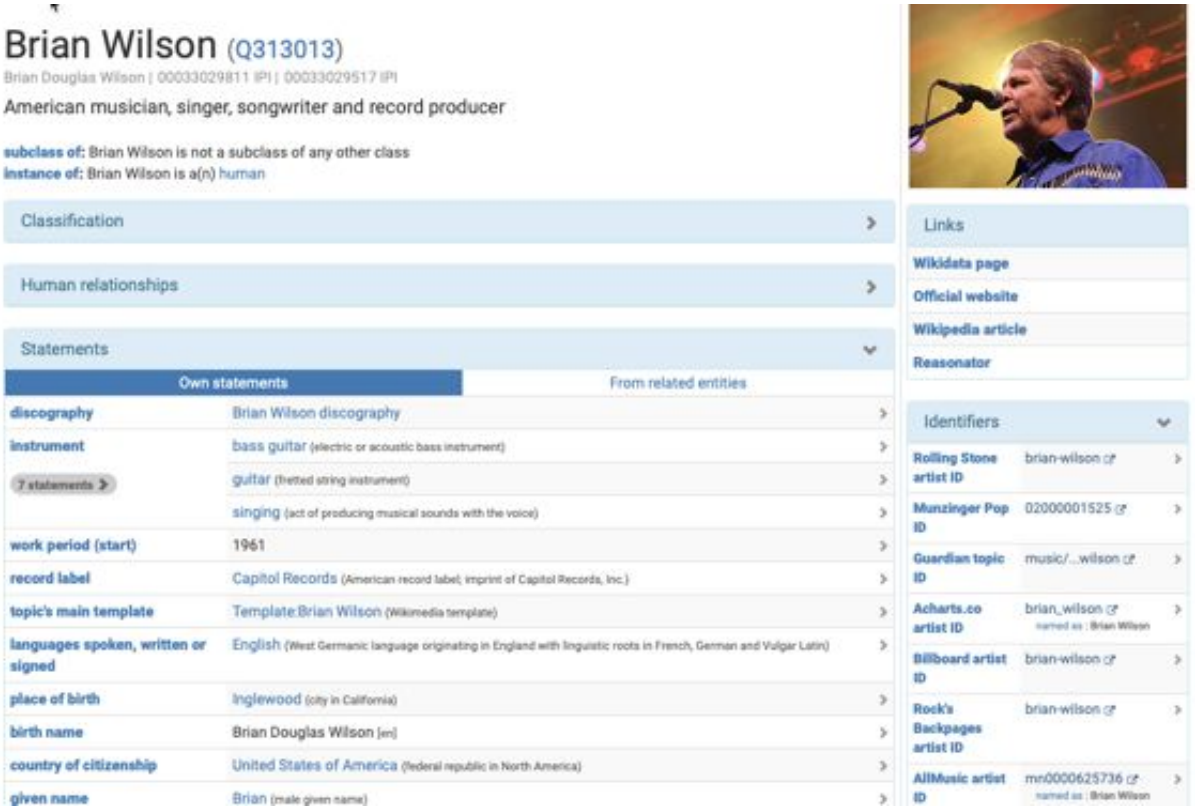
- Class
- Properties
- Values

Lots of Cues



- Class
- Properties
- Values

Lots of Cues



Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements | From related entities

discography	Brian Wilson discography
instrument	bass guitar (electric or acoustic bass instrument)
7 statements	guitar (fretted string instrument)
	singing (act of producing musical sounds with the voice)
work period (start)	1961
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)
topic's main template	Template:Brian Wilson (Wikimedia template)
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)
place of birth	Inglewood (city in California)
birth name	Brian Douglas Wilson [en]
country of citizenship	United States of America (federal republic in North America)
given name	Brian (male given name)

Links

Wikidata page
Official website
Wikipedia article
Reasonator

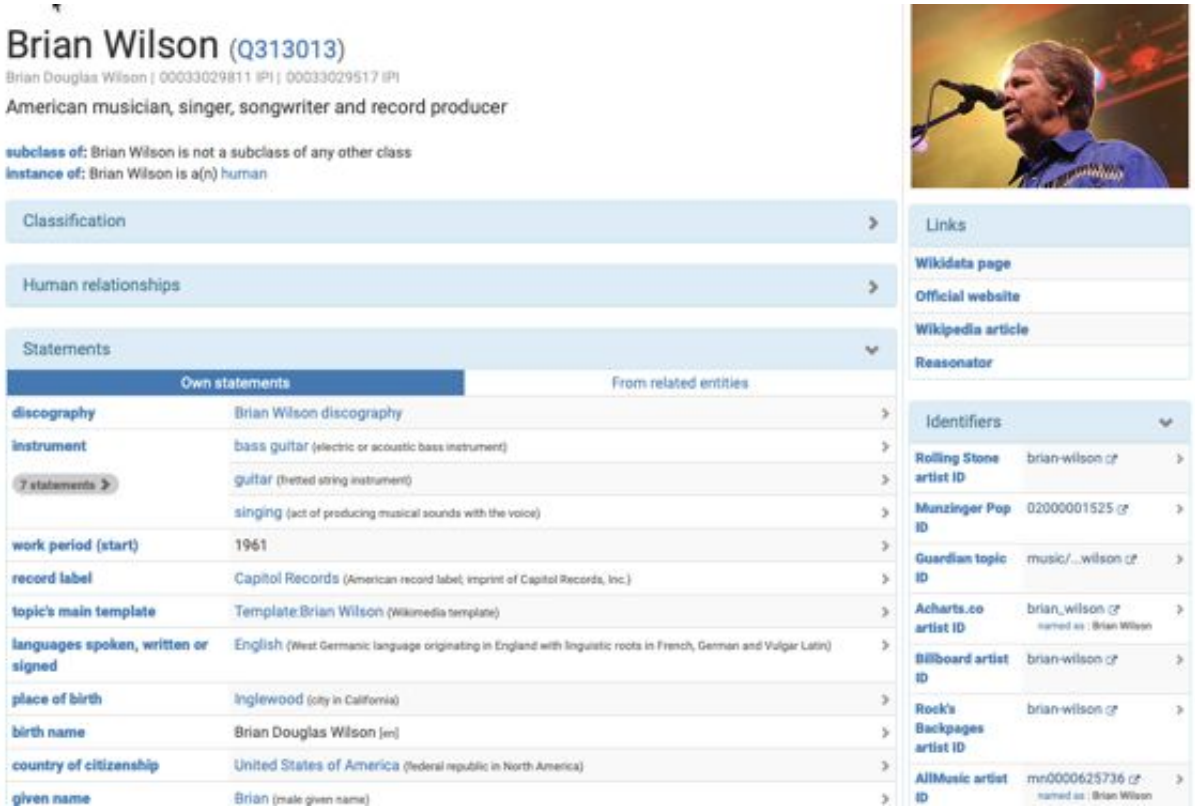
Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson named as : Brian Wilson
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 named as : Brian Wilson

- Class
- Properties
- Values

instanceOf: Human

Lots of Cues



Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements | From related entities

discography	Brian Wilson discography
instrument	bass guitar (electric or acoustic bass instrument)
7 statements	guitar (fretted string instrument)
	singing (act of producing musical sounds with the voice)
work period (start)	1961
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)
topic's main template	Template:Brian Wilson (Wikimedia template)
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)
place of birth	Inglewood (city in California)
birth name	Brian Douglas Wilson [en]
country of citizenship	United States of America (federal republic in North America)
given name	Brian (male given name)

Links

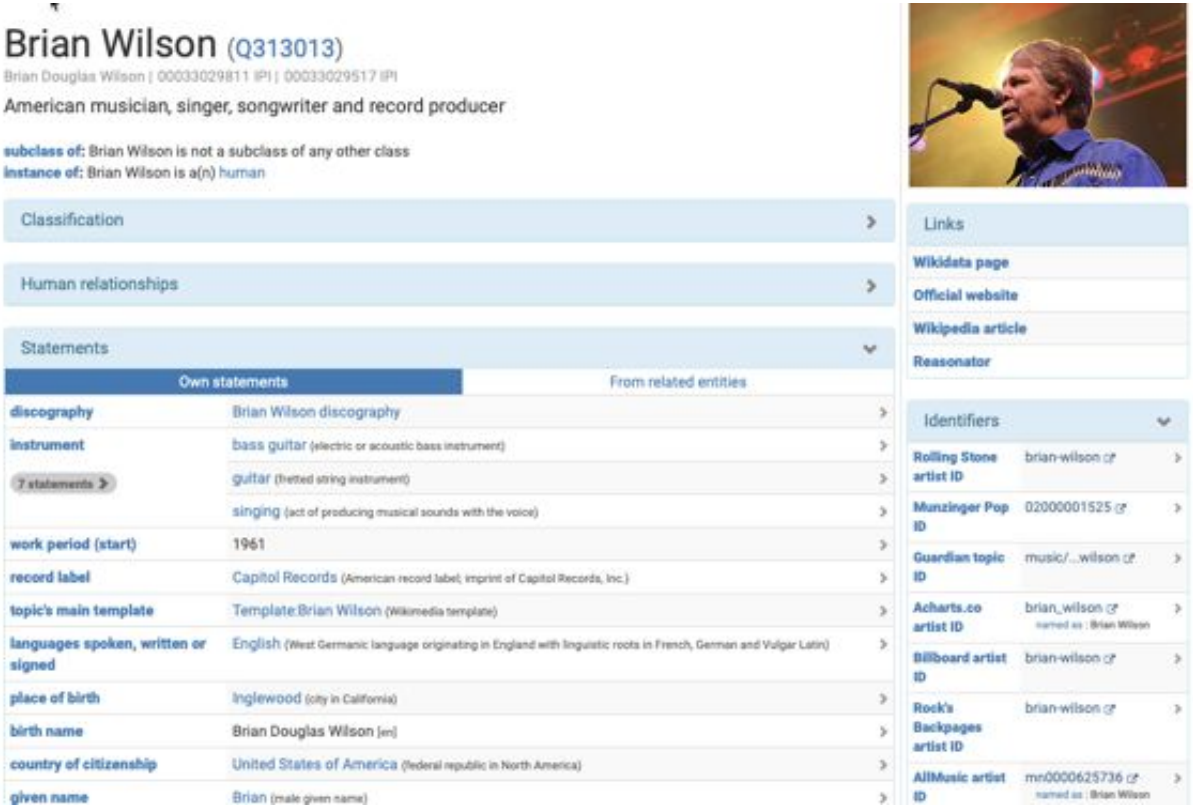
Wikidata page
Official website
Wikipedia article
Reasonator

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson named as : Brian Wilson
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 named as : Brian Wilson

- Class
- Properties
- Values

Lots of Cues



Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements | From related entities

discography	Brian Wilson discography
instrument	bass guitar (electric or acoustic bass instrument)
7 statements	guitar (fretted string instrument)
	singing (act of producing musical sounds with the voice)
work period (start)	1961
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)
topic's main template	Template:Brian Wilson (Wikimedia template)
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)
place of birth	Inglewood (city in California)
birth name	Brian Douglas Wilson [en]
country of citizenship	United States of America (federal republic in North America)
given name	Brian (male given name)

Links

- Wikidata page
- Official website
- Wikipedia article
- Reasonator

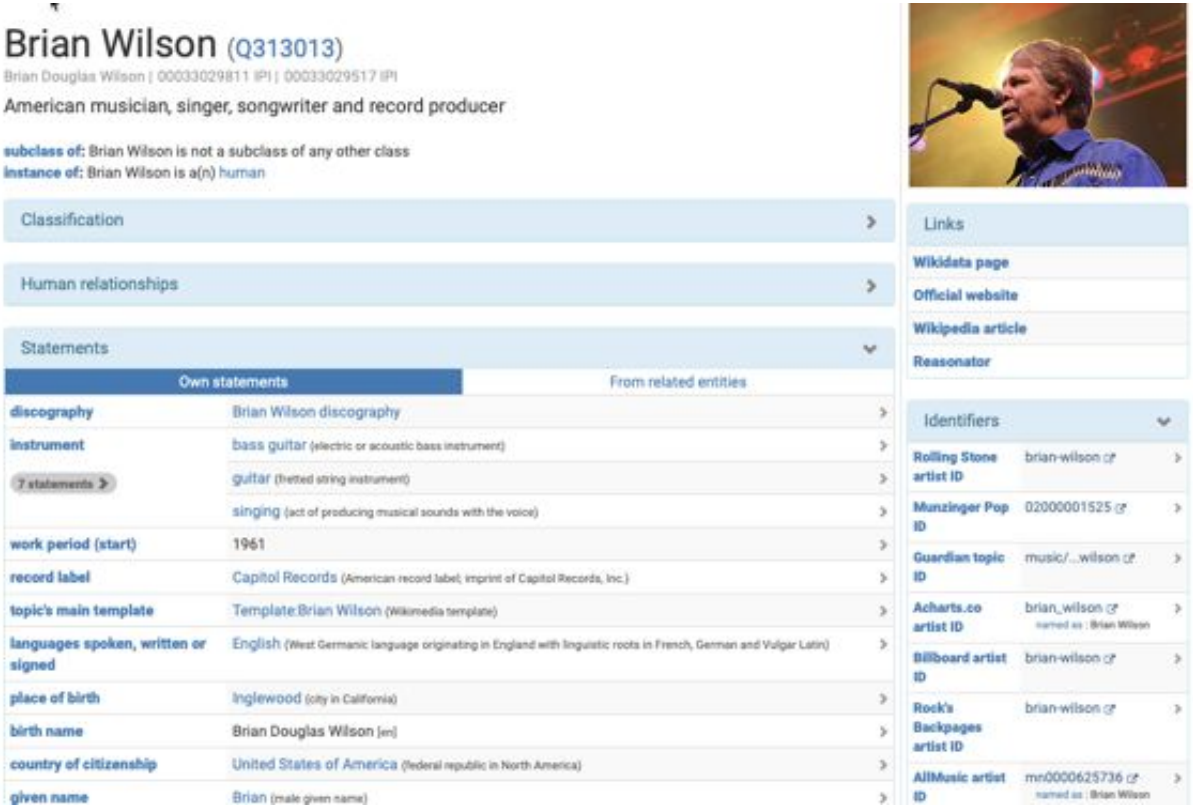
Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson named as : Brian Wilson
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 named as : Brian Wilson

- Class
- Properties
- Values

occupation: Singer

Lots of Cues



Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements | From related entities

discography	Brian Wilson discography
instrument	bass guitar (electric or acoustic bass instrument)
7 statements	guitar (fretted string instrument)
	singing (act of producing musical sounds with the voice)
work period (start)	1961
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)
topic's main template	Template:Brian Wilson (Wikimedia template)
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)
place of birth	Inglewood (city in California)
birth name	Brian Douglas Wilson [en]
country of citizenship	United States of America (federal republic in North America)
given name	Brian (male given name)

Links

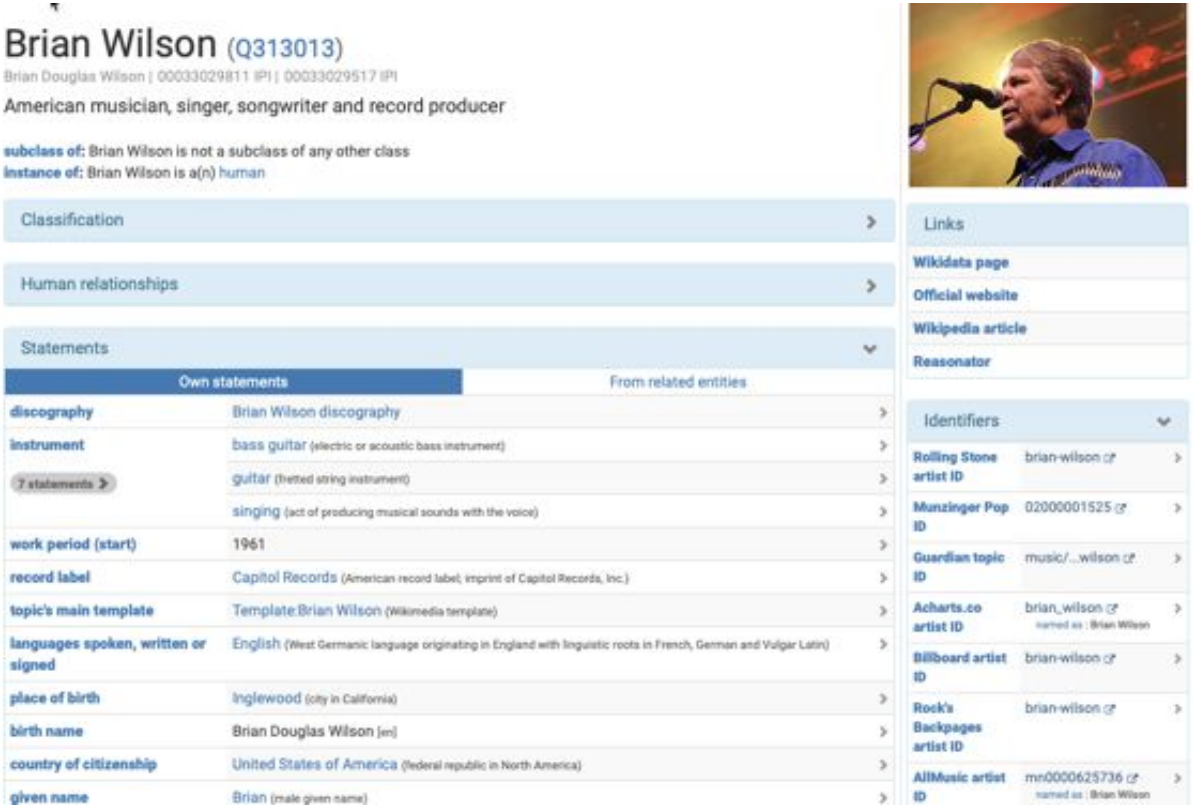
- **Wikidata page**
- **Official website**
- **Wikipedia article**
- **Reasonator**

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson named as : Brian Wilson
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 named as : Brian Wilson

- Class
- Properties
- Values

Lots of Cues



Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements | From related entities

discography	Brian Wilson discography
instrument	bass guitar (electric or acoustic bass instrument)
7 statements	guitar (fretted string instrument)
	singing (act of producing musical sounds with the voice)
work period (start)	1961
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)
topic's main template	Template:Brian Wilson (Wikimedia template)
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)
place of birth	Inglewood (city in California)
birth name	Brian Douglas Wilson [en]
country of citizenship	United States of America (federal republic in North America)
given name	Brian (male given name)

Links

Wikidata page
Official website
Wikipedia article
Reasonator

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson named as : Brian Wilson
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 named as : Brian Wilson

- Class
- Properties
- Values

Record Label: ...

Lots of Cues

Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements		From related entities
discography	Brian Wilson discography	
instrument	bass guitar (electric or acoustic bass instrument)	
7 statements	guitar (fretted string instrument)	
	singing (act of producing musical sounds with the voice)	
work period (start)	1961	
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)	
topic's main template	Template:Brian Wilson (Wikimedia template)	
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)	
place of birth	Inglewood (city in California)	
birth name	Brian Douglas Wilson [en]	
country of citizenship	United States of America (federal republic in North America)	
given name	Brian (male given name)	

Links

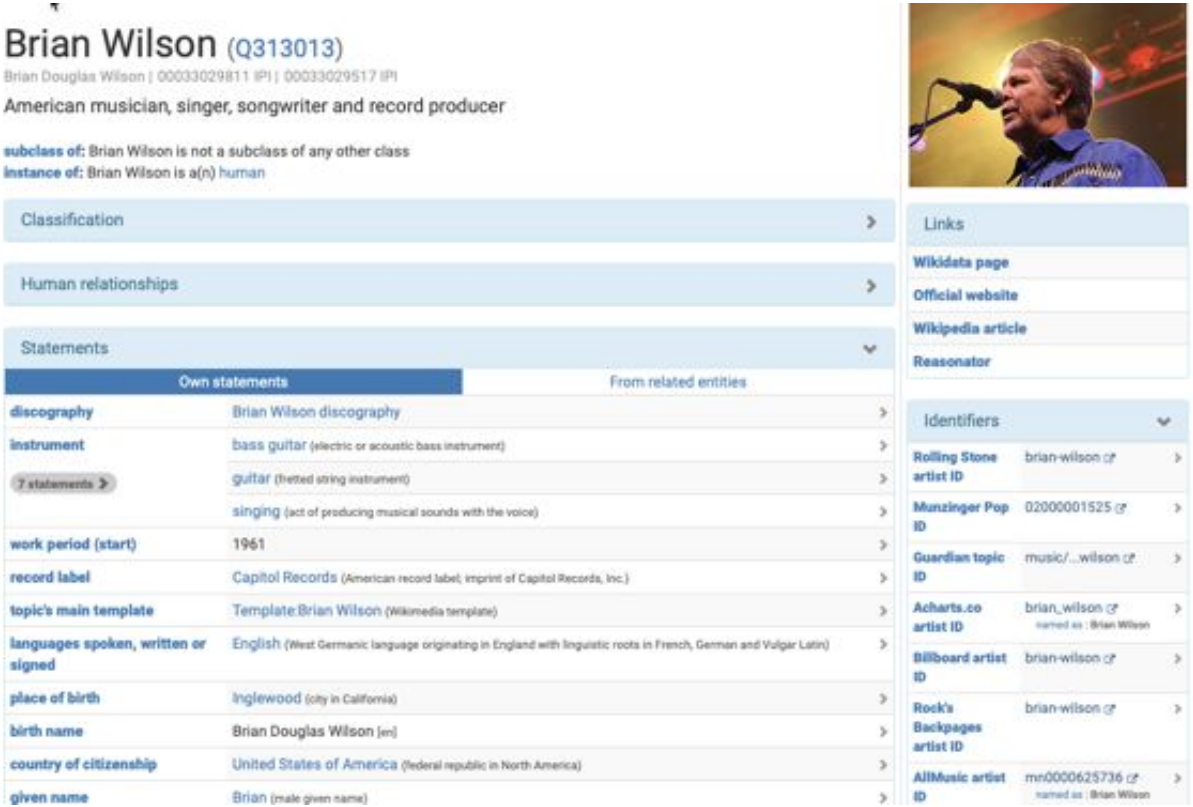
- **Wikidata page**
- **Official website**
- **Wikipedia article**
- **Reasonator**

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson (named as : Brian Wilson)
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 (named as : Brian Wilson)

- Class
- Properties
- Values

Lots of ~~Cues~~ Features



Brian Wilson (Q313013)

Brian Douglas Wilson | 00033029811 IPI | 00033029517 IPI

American musician, singer, songwriter and record producer

subclass of: Brian Wilson is not a subclass of any other class
instance of: Brian Wilson is a(n) human

Classification

Human relationships

Statements

Own statements | From related entities

discography	Brian Wilson discography
instrument	bass guitar (electric or acoustic bass instrument)
7 statements	guitar (fretted string instrument)
	singing (act of producing musical sounds with the voice)
work period (start)	1961
record label	Capitol Records (American record label; imprint of Capitol Records, Inc.)
topic's main template	Template:Brian Wilson (Wikimedia template)
languages spoken, written or signed	English (West Germanic language originating in England with linguistic roots in French, German and Vulgar Latin)
place of birth	Inglewood (city in California)
birth name	Brian Douglas Wilson [en]
country of citizenship	United States of America (federal republic in North America)
given name	Brian (male given name)

Links

- **Wikidata page**
- **Official website**
- **Wikipedia article**
- **Reasonator**

Identifiers

Rolling Stone artist ID	brian-wilson
Munzinger Pop ID	02000001525
Guardian topic ID	music/...wilson
Acharts.co artist ID	brian_wilson named as : Brian Wilson
Billboard artist ID	brian-wilson
Rock's Backpages artist ID	brian-wilson
AllMusic artist ID	mn0000625736 named as : Brian Wilson

- Class
- Properties
- Values

What to do with all those Features?

What to do with all those Features?

If labelled data -> Machine Learning

What to do with all those Features?

If labelled data -> Machine Learning

Human?	occ:Singer?	Record Label?	...	Chef?
1	1	1	...	0

Weights	20	30	10	...	0.5

Confidence = 60

What to do with all those Features?

If labelled data -> Machine Learning

What to do with all those Features?

If labelled data -> Machine Learning

If not ->

Image Source: icon-library.net

What to do with all those Features?

If labelled data -> Machine Learning

If not -> Heuristics!

2. tf-idf

TF-IDF

TF-IDF is a measure of originality of a word by comparing the number of times a word appears in a doc with the number of docs the word appears in.

$$\text{TF-IDF} = \text{TF}(t, d) \times \text{IDF}(t)$$

Term frequency: Number of times term t appears in a doc, d

Inverse document frequency

$$\log \frac{1 + n}{1 + df(d, t)} + 1$$

n ← # of documents

df(d, t) ← Document frequency of the term t

Image Source: becominghuman.ai blog

properties / entities	genre	family name	record label	disco-graphy	Dbo: MusicalArtist	TF/IDF	Levenshtein
Q313013 (Brian Wilson, musician)	1	1	1	1	1	0.98	1.0
Q913269 (Brian Wilson, baseball player)	0	1	0	0	0	0.64	1.0
Q1135582 (Super Flurry Animals, band)	1	0	1	1	1	0.23	1.0
Q7642367 (Super Flurry Animals Discography)	0	0	0	0	0	0.0	0.61
Q185343 (Mark Knopfler, musician)	1	1	1	1	1	0.99	1.0
DF = document frequency	52	31	36	15	49		
IDF = log	3.20	1.85	1.65	3.46	2.11		

Labels	Candidates	
John	dbr:Pope_John dbo:Person dbo:Saint dbp:Religion	dbr:John_Lennon dbo:Person dbo:Singer dbp:Albums
Saint Francis	dbr:St_Francis dbo:Person dbo:Saint dbp:Religion	dbr:Pope_Francis dbo:Person dbo:Saint dbp:Religion
Saint Madonna	dbr:Madonna dbo:Person dbo:Singer dbp:Albums	dbr:Saint_Madonna dbo:Person dbo:Saint dbp:Religion
Victor	dbr:Saint_Victor dbo:Person dbo:Saint dbp:Religion	
St Mary	dbr:Mary dbo:Person dbo:Saint dbp:Religion	dbr:Mother_Mary

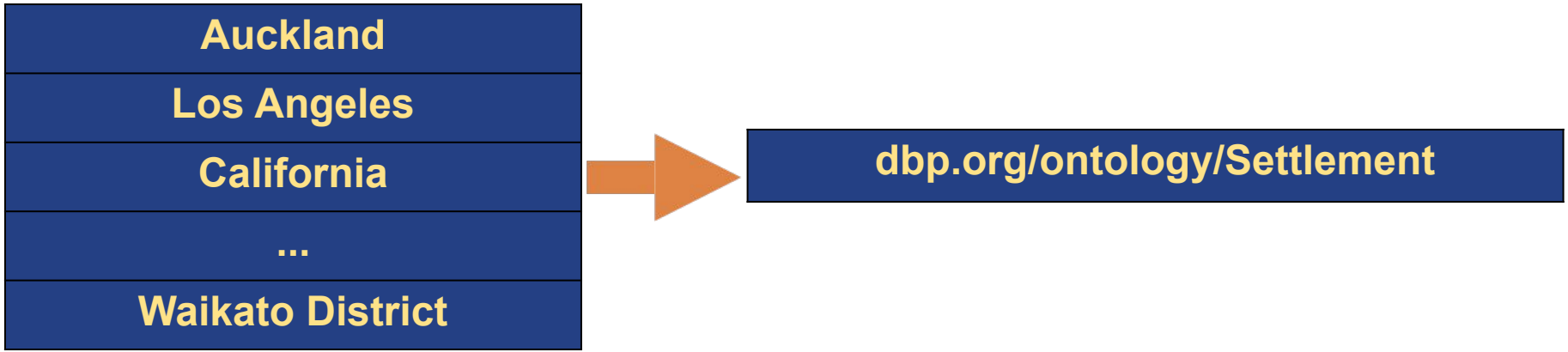
Step 1: Semantic Feature Extraction

Semantic Feature	dbo: Person	dbo: Saint	dbo: Singer	dbp: Religion	dbp: Albums	
(TF) Term Frequency	5	4	1	4	1	
Document Frequency	8	6	2	6	2	
(IDF) Inverse Document Frequency	0.05	0.18	0.65	0.18	0.65	
(TFIDF) TF ⊗ IDF	0.25	0.72	0.65	0.72	0.65	Score $TFIDF \cdot v_i^T$
(v_1) dbr: madonna	1	0	1	0	1	1.55
(v_2) dbr: saint_ madonna	1	1	0	1	0	1.69

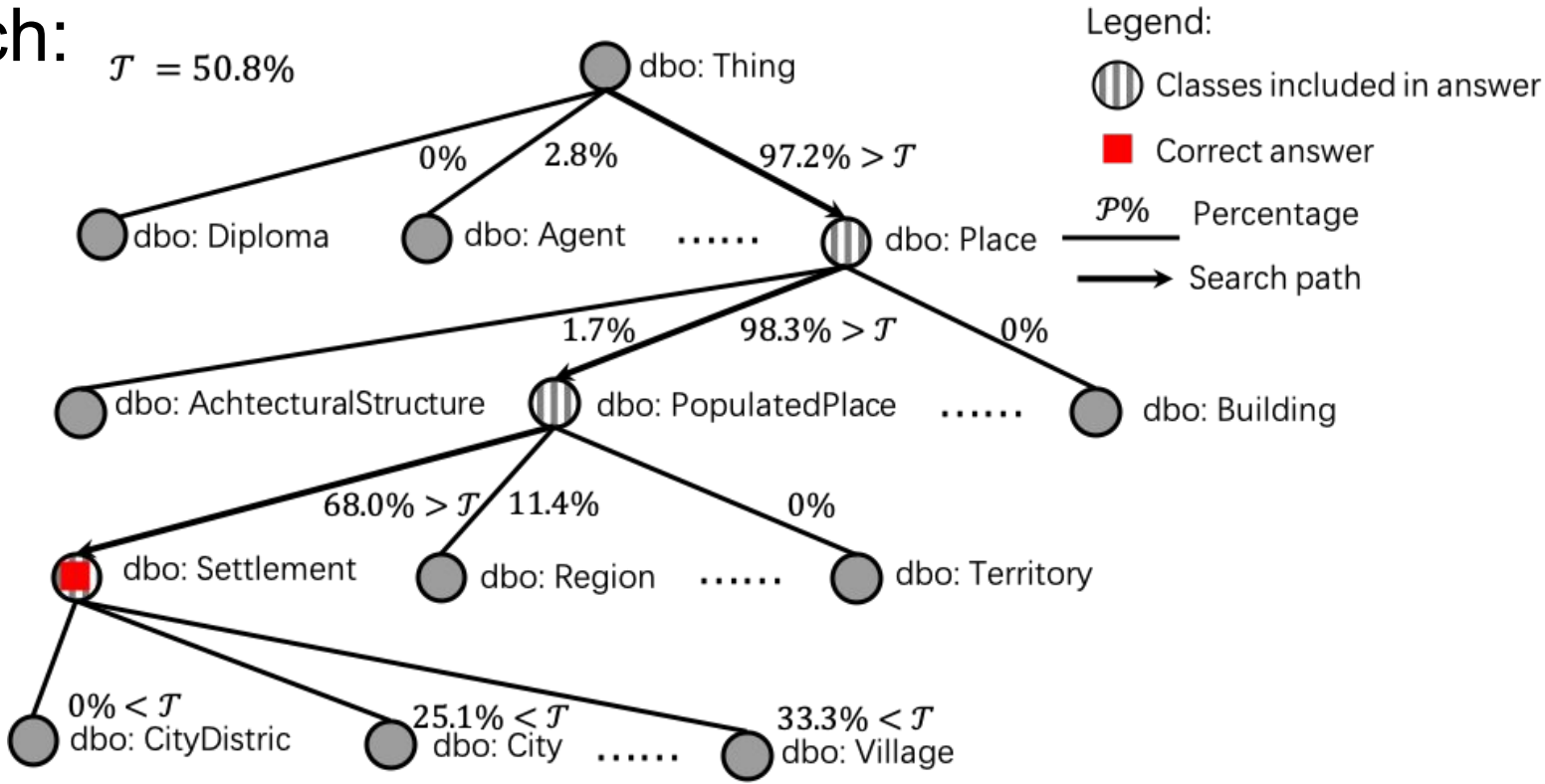
Step 2: Candidate Selection (TFIDF)

3. CTA and CPA

Objective: CTA



Approach: CTA



CPA

Column 1	Column 2	Candidates
M. Thatcher (Q17421946, **Q7416**, Q512101)	Lincolnshire (**Q23090**)	dbo:birthplace
W. Churchill (Q27436889, Q22003261, Q19864690)	Oxfordshire (**Q23169**, Q23217)	dbo:deathplace
N. Chamberlain ()	Birmingham (Q19444, Q79867, Q223429)	(empty) → dbo:birthplace
W. Gladstone (**Q160852**, Q41777394)	Liverpool (**Q24826**)	dbo:birthplace, dbo:deathplace
B. Disraeli (**Q82006**, Q269211, Q66649130)	Middlesex (**Q19186**)	dbo:birthplace

Results: CEA

Round 1		Round 2		Round 3		Round 4	
f1	precision	f1	precision	f1	precision	f1	precision
0.884	0.908	0.826	0.852	0.857	0.866	0.804	0.814

4. Shortcomings

Shortcomings

Results on Round #3

Δ	#	Participants	F1 Score	Precision
▲	01	MTab	0.970	0.970
▼	02	IDLab	0.962	0.964
●	03	ADOG	0.912	0.913
▲	04	tabularisi	0.857	0.866

Shortcomings

Another pass needed

Shortcomings

Another pass needed

Custom handling of data types

Shortcomings

Another pass needed

Custom handling of data types

Intra-row information

5. Analysis

Analysis: # Rows

Analysis: # Rows

Experiments on # Rows				
Files	**# Predictions**	**Precision**	**Recall**	**F1 score**
# Rows > 100	235,914	0.877	0.446	0.591
100 > # Rows > 10	171,378	0.847	0.313	0.457
10 > # Rows	29,873	0.675	0.044	0.082

Analysis: # Rows

Table 2. CTA performance growth rate on different ranges

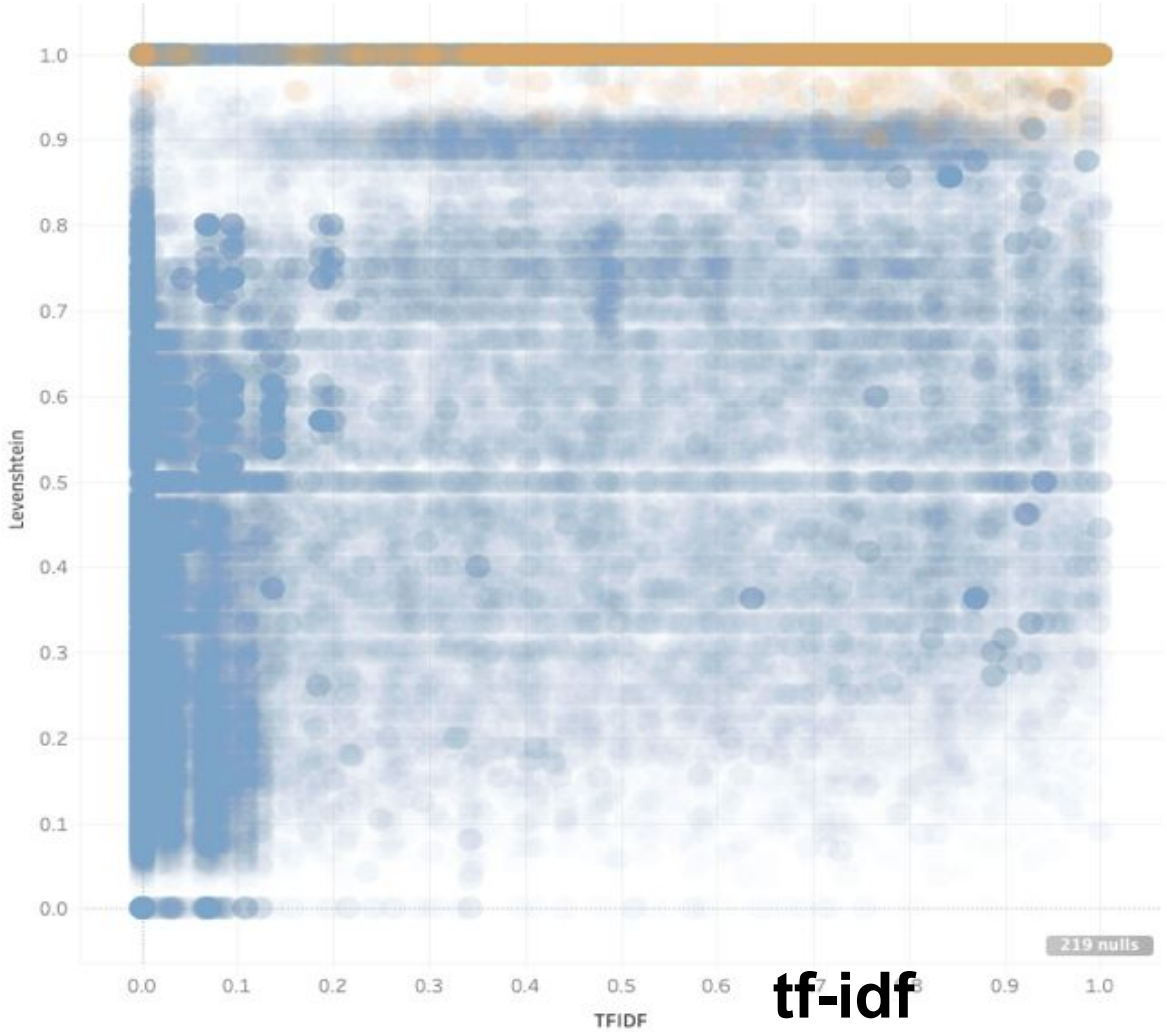
Ranges	Columns	Growth rate
(0,2]	0	NA
(2,3]	322	NA
(3,4]	520	3.2%
(4,5]	447	11.6%
(6,10]	2582	5.7%
(10,20]	3394	43.5%
(20,50]	2540	20.6%
(51,100]	1340	13.7%
(101,+∞)	1722	8.4%

Analysis: Custom Handling

Ablations with/without Abbreviations					
Candidates	**Cells**	**# Predictions**	**Precision**	**Recall**	**F1 score**
API, ES	All	419,921	0.766	0.694	0.728
API, ES	abbreviations only	63,207	*0.201*	0.027	0.048
API, ES	All except abbreviations	356,714	**0.866**	0.666	0.753
ES-abbr	abbreviations only	80,451	**0.788**	0.137	0.233
API, ES, ES-abbr	All	437,165	**0.852**	0.802	**0.826**

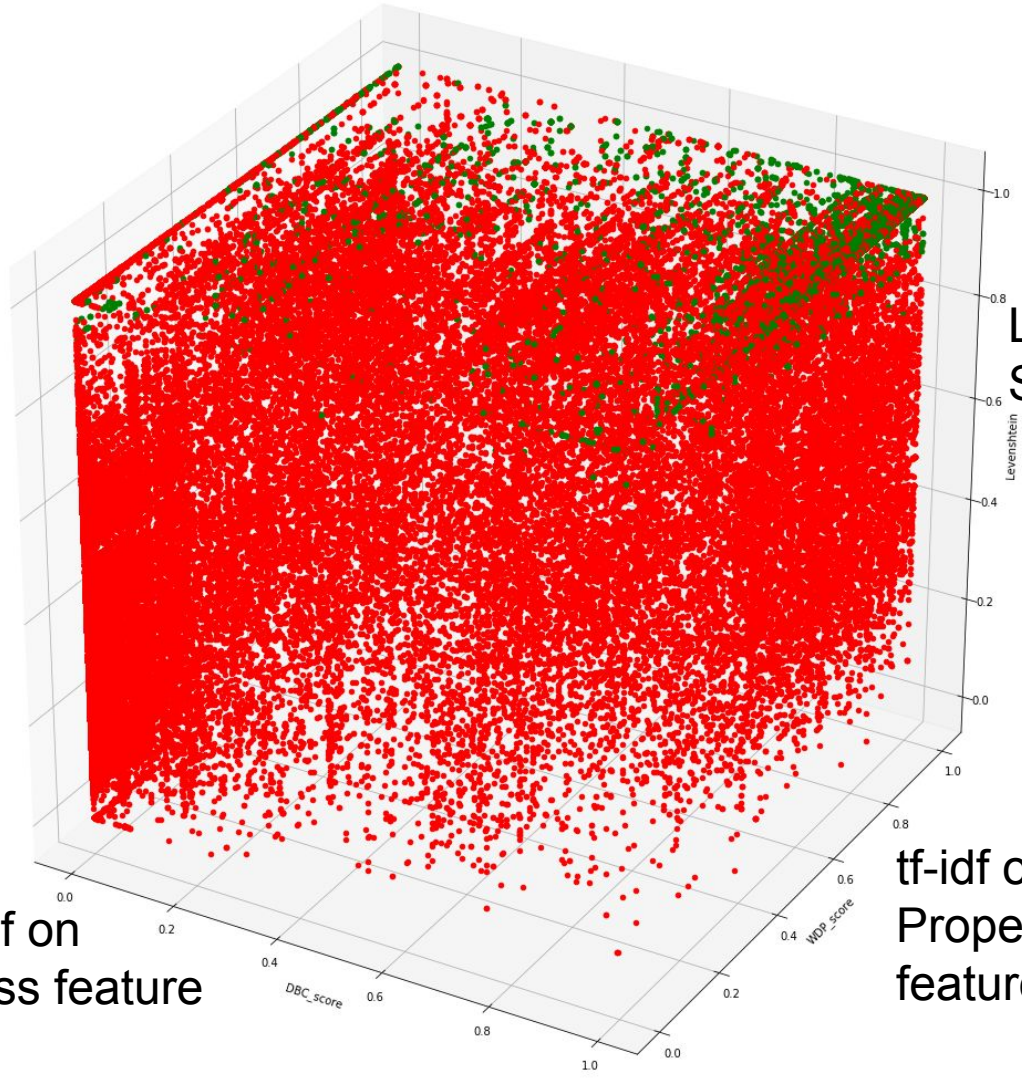
Analysis: Embeddings

Levenshtein Similarity



tf-idf

Analysis: Embeddings



Levenshtein Similarity

tf-idf on Class feature

tf-idf on Property feature

USC Viterbi
School of Engineering
Information Sciences Institute

Takeaways

Takeaways

- Lots of Semantic Cues (not just classes)

Takeaways

- Lots of Semantic Cues (not just classes)
- When no data -> TF-IDF

Takeaways

- Lots of Semantic Cues (not just classes)
- When no data -> TF-IDF
- Revising always good

Takeaways

- Lots of Semantic Cues (not just classes)
- When no data -> TF-IDF
- Revising always good
- Over-revising is an overkill (PSL)

Takeaways

- Lots of Semantic Cues (not just classes)
- When no data -> TF-IDF
- Revising always good
- Over-revising is an overkill (PSL)
- String Similarity $\perp$ Semantic Similarity

Fin.

Thank You
kia mihi

Avijit Thawani

PhD student
with Pedro Szekely
and Jay Pujara

thawani@isi.edu

Appendix

PSL

Graphical Model

= Several passes!

Probabilistic Soft Logic

PSL is a

- Probabilistic Programming Language for easily defining
- Hinge Loss Markov Random Fields
- using a syntax like First Order Logic.

PSL in one slide

PSL in one slide

Define closed predicates:

- instance(madonna, Singer)
- candidate(R^3C^1, madonna)

instance(st_madonna, Saint) …

candidate(R^3C^1, st_madonna) …

PSL in one slide

Define closed predicates:

- instance(madonna, Singer)
- candidate(R^3C^1, madonna)

instance(st_madonna, Saint) …

candidate(R^3C^1, st_madonna) …

Define open predicates:

- type(C^1, Singer)?
- entity(R^3C^1, madonna)?

type(C^1, Saint)?

entity(R^3C^1, st_madonna)?

PSL in one slide

Define closed predicates:

- instance(madonna, Singer) instance(st_madonna, Saint) …
- candidate(R^3C^1, madonna) candidate(R^3C^1, st_madonna) …

Define open predicates:

- type(C^1, Singer)? type(C^1, Saint)?
- entity(R^3C^1, madonna)? entity(R^3C^1, st_madonna)?

Restrict with PSL rules:

- 10: candidate(R^xC^y, Q^z) -> entity(R^xC^y, Q^z)
- 20: candidate(R^xC^y, Q^z) & type(C^y, T^w) & instance(Q^z, T^w) -> entity(R^xC^y, Q^z)
- entity(R^xC^y, Q^1) & Q^1!=Q^2 -> ! entity(R^xC^y, Q^2) .

PSL output

class(C^1, Singer): 0.12

class(C^1, Saint): 0.89

entity(R^3C^1, madonna): 0.23

entity(R^3C^1, st_madonna): 0.68

1st result baseline

F1:	0.865
Precision:	0.871
Recall:	0.858

(7 datasets annotated by us)

PSL results

F1:	0.903
Precision:	0.910
Recall:	0.896

(7 datasets annotated by us)

PSL without ranked priors

F1:	0.777
Precision:	0.783
Recall:	0.771

(7 datasets annotated by us)