Completeness, Recall and Negation in Open-World Knowledge Bases

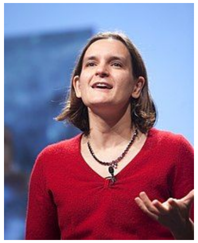
Simon Razniewski, Hiba Arnaout, Shrestha Ghosh, Fabian Suchanek

1. Introduction and Foundations (Simon)
2. Predictive recall assessment (Fabian)
3. Counts from text and KB (Shrestha)
4. Negation (Hiba)
5. Relative completeness & Wrap-up (Simon)

What is count information?

Relation between an entity and a set of entities

Noam Chomsky



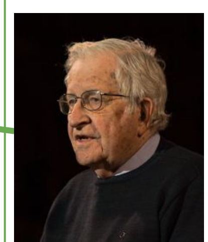
Esther Duflo

What is count information?

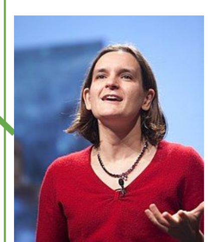
Relation between an entity and a set of entities

Enumerating Predicates
employer

Objects



Noam Chomsky



Esther Duflo

Expressed as **entities** or objects in the set

What is count information?

Relation between an entity and a set of entities

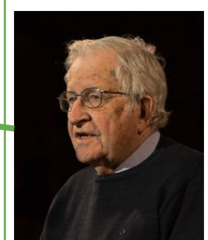
1074

Count

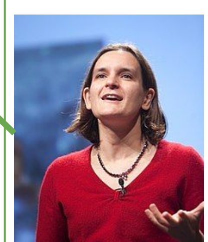
Expressed as **count** or cardinality of the set

Counting Predicates
employees

Objects



Noam Chomsky



Esther Duflo

Expressed as **entities** or objects in the set

Enumerating Predicates
employer

1. Count information for recall assessment
2. How can we extract count information from text?
3. Variants of count information in KB
4. How much count information is accounted for?
5. Counts for KB curation

Count information for recall assessment

Counts and entities benefit from each other

Only entities

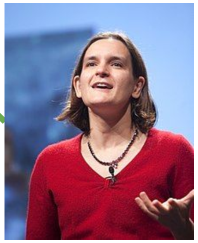
(?x, employer, MIT)

returns a handful of names from KB

Enumerating Predicates
employer

Objects

Noam Chomsky



Esther Duflo

Count information for recall assessment

Counts and entities benefit from each other

Only entities

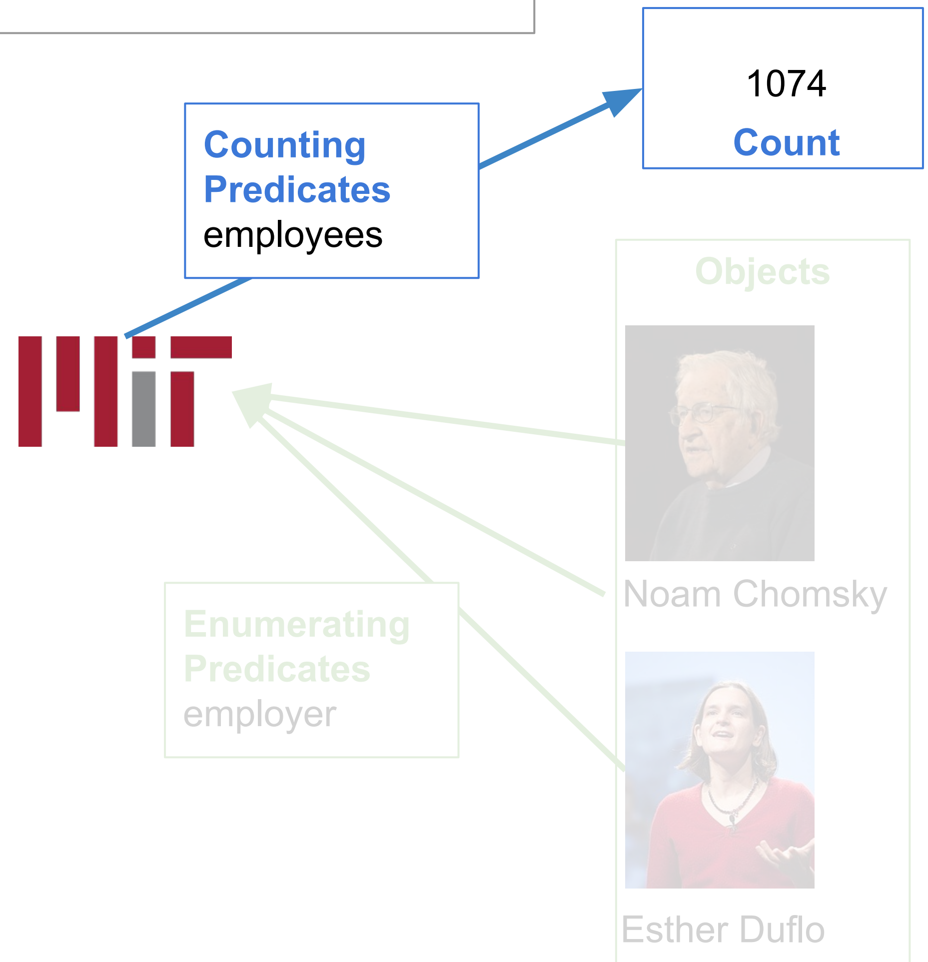
(?x, employer, MIT)

returns a handful of names from KB

Only counts

(MIT, employees, ?y)

gives no insight about the entities



Count information for recall assessment

Only entities

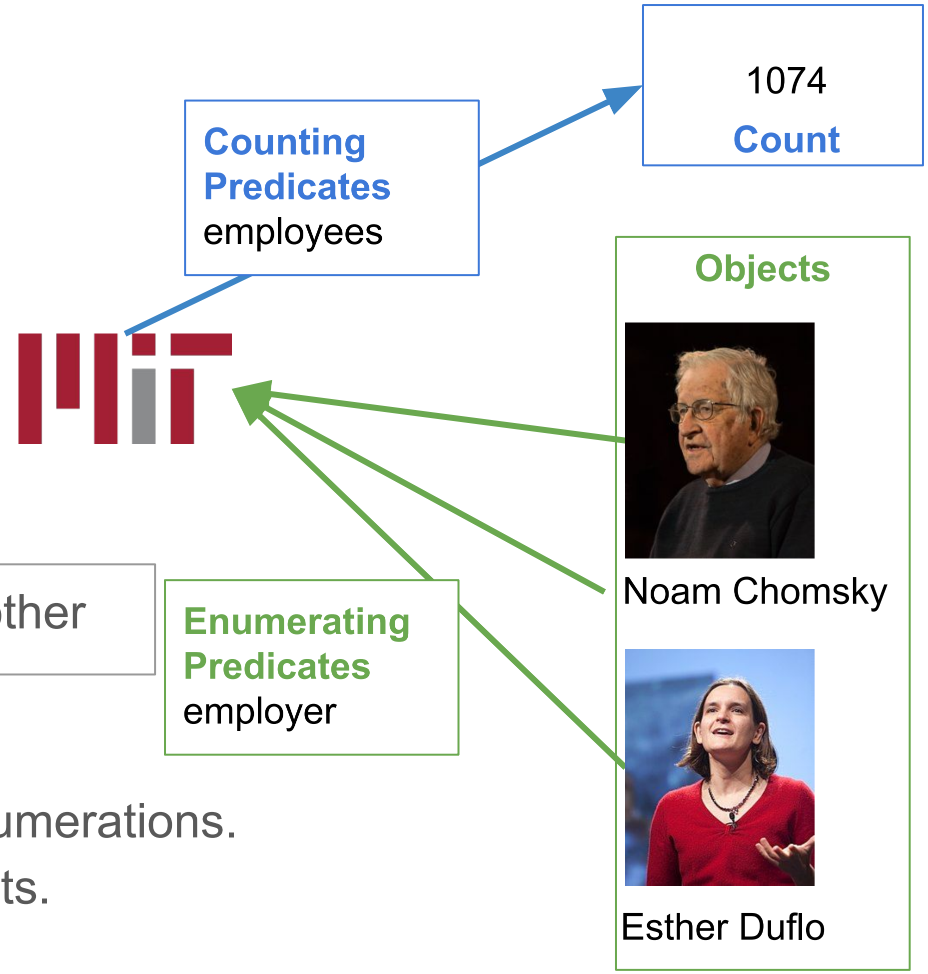
(?x, employer, MIT)

returns a handful of names from KB

Only counts

(MIT, employees, ?y)

gives no insight about the entities



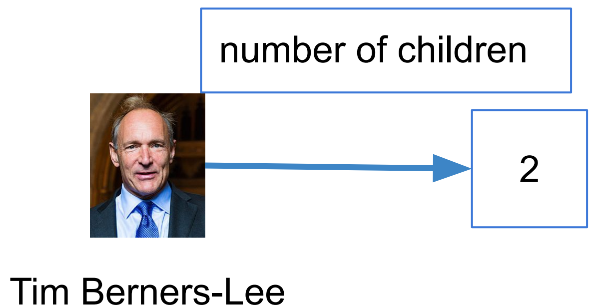
Counts and entities benefit from each other

Count and Entities

- Counts enhance incomplete entity enumerations.
- Representative entities enhance counts.

Count information for recall assessment

KB mixes counts with standard facts



How many children does Tim Berners-Lee have?

2 (KB fact)

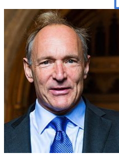
How many children did Ada Lovelace have?

3 (Maybe?)

Count information for recall assessment

KB mixes counts with standard facts

number of children



Tim Berners-Lee → 2

How many children does Tim Berners-Lee have?

2 (KB fact)

child

Ada Lovelace →

- Anne Blunt
- Ralph King-Milbanke
- Byron King-Noel

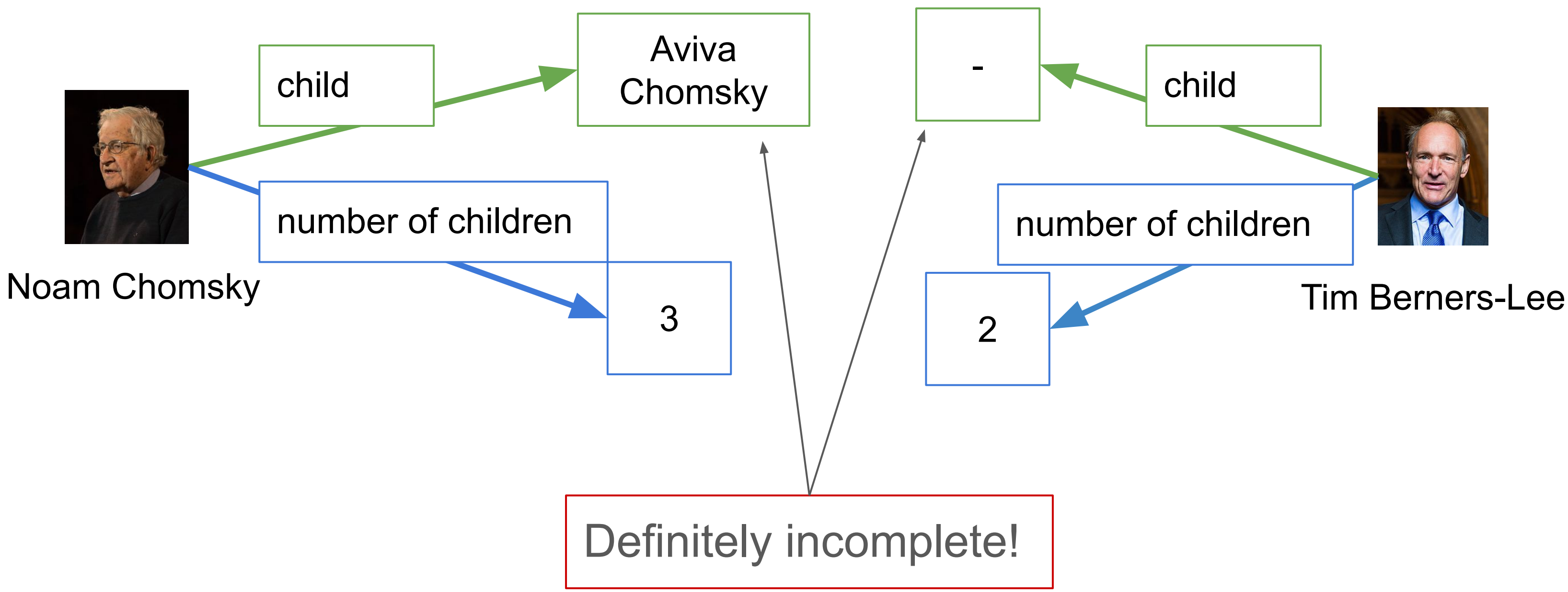
How many children did Ada Lovelace have?

3 (Maybe?)

Enumeration is often of known entities

Count information for recall assessment

Count information can highlight KB inconsistencies



1. Count information for recall assessment
2. How can we extract count information from text?
3. Variants of count information in KB
4. How much count information is accounted for?
5. Counts for KB curation

Count information from text

Problem: Counting Quantifier Extraction

Input:

- a text about a subject S
- a predicate P

Task: Determine the number of objects in which S stands in relation with P

Subject: Noam Chomsky
Predicate: number_of_children

Chomsky was married to Carol. They had three children together
3

Count information from text

Task 1: Identify the count tokens and the compositional cues.

Sequence Labelling of tokens in a sentence on subject S and predicate P with:

- COUNT - for counts
- COMP - for compositional cues
- O - all other tokens

Subject: Noam Chomsky
Predicate: number_of_children

Chomsky was married to Carol. They had three children together
O O O O O O O COUNT O O

Count information from text

Task 1: Identify the count tokens and the compositional cues.

Sequence Labelling of tokens in a sentence on subject S and predicate P with:

- COUNT - for counts
- COMP - for compositional cues
- O - all other tokens

Subject: Angelina Jolie
Predicate: number_of_children

Jolie	has	three	sons	and	three	daughters.
O	O	COUNT	O	COMP	COUNT	O

Count information from text

Task 1: Identify the count tokens and the compositional cues.

COUNT tokens are **linguistically diverse**

Cardinals

two sons,
three books

Count information from text

Task 1: Identify the count tokens and the compositional cues.

COUNT tokens are **linguistically diverse**

Cardinals	Ordinals
two sons, three books	second son, third book

Count information from text

Task 1: Identify the count tokens and the compositional cues.

COUNT tokens are **linguistically diverse**

Cardinals	Ordinals	Number-related terms
two sons, three books	second son, third book	twins, trilogy

Count information from text

Task 1: Identify the count tokens and the compositional cues.

COUNT tokens are **linguistically diverse**

Cardinals	Ordinals	Number-related terms	Indefinite Articles
two sons, three books	second son, third book	twins, trilogy	a son, the book

Count information from text

Task 1: Identify the count tokens and the compositional cues.

COMP cues for counts occur

- between consecutive count tokens, as
- **comma**-separated, **and**-separated counts

Subject: Angelina Jolie
Predicate: number_of_children

Jolie brought her twins , one daughter and three adopted children to the gala.

COMP COMP

Count information from text

Task 2: Consolidate count tokens

Return **a single answer** per text, given subject-predicate pair

1. Sum up compositional cues
2. Select prediction per type
3. Rank mention types

Count information from text

Task 2: Consolidate count tokens

Return a single answer per text, given subject-predicate pair

1. **Sum up compositional cues**

6

Jolie brought her six children: twins , one daughter and three adopted children to the gala.

Subject: Angelina Jolie
Predicate: number_of_children

Count information from text

Task 2: Consolidate count tokens

Return a single answer per text, given subject-predicate pair

1. Sum up compositional cues
2. **Select prediction per type**

6 (cardinal) **6** (cardinal)

Jolie brought her six children: twins , one daughter and three adopted children to the gala.

Subject: Angelina Jolie
Predicate: number_of_children

6 (cardinal)

Count information from text

Task 2: Consolidate count tokens

Return a single answer per text, given subject-predicate pair

1. Sum up compositional cues
2. Select prediction per type
3. **Rank mention types**

cardinal >>	number-related terms >>	ordinals >>	indefinite article
two children >>	twins >>	second child >>	a child

Count information from text

Task 2: Consolidate count tokens

Return a single answer per text, given subject-predicate pair

1. Sum up compositional cues
2. Select prediction per type
3. **Rank mention types**

cardinal >> number-related terms >> ordinals >> indefinite article

two children >> twins >> second child >> a child

Jolie brought her six children: twins , one daughter and three adopted children to the gala.

Subject: Angelina Jolie
Predicate: number_of_children

6 (cardinal)

Count information from text

Training data generation: Incompleteness-aware distant supervision

Input: KB, count predicate P

Output:

- all subjects S and the count
- all sentences about S containing cardinal mentions similar to the KB count

Count information from text

Training data generation: Incompleteness-aware distant supervision

Input: KB, count predicate P

Output:

- all subjects S and the count
- all sentences about S containing cardinal mentions similar to the KB count

all counts tokens

Count information from text

Training data generation: Incompleteness-aware distant supervision

Input: KB, count predicate P

Output:

- all subjects S and the count
- all sentences about S containing cardinal mentions similar to the KB count

all counts tokens

+ve: equal to or representative of KB count
-ve : otherwise and all non-numerals
Ignore: candidate counts > KB counts

Count information from text

Ground Truth

Use KB information as Ground Truth

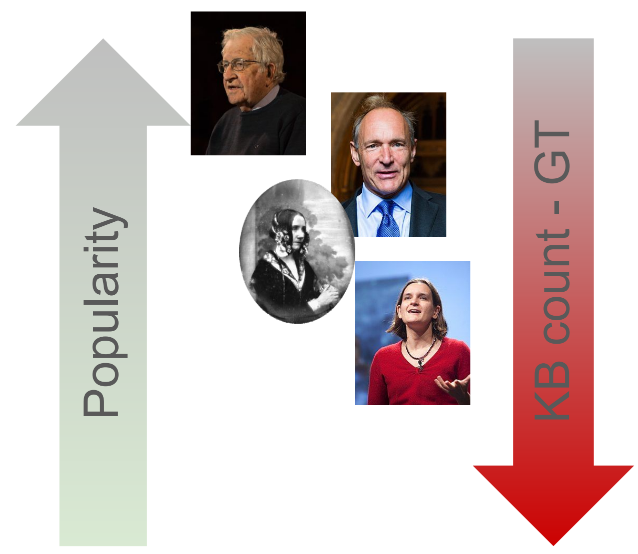
Challenges

KB incompleteness negatively impacts training quality

Solution

Consider only popular KB entities

Set upper bound for predicate count value = 99th percentile of KB predicate value distribution



Count information from text

Challenge

Counting cardinality when it is Zero

Count information from text

Challenge

Counting cardinality when it is Zero

Solution

Focus on

i) Negation determiners: ‘no’ and ‘any’

ii) Non-existence-proving adverbs: ‘without’ and ‘never

Count information from text

Challenge

Counting cardinality when it is Zero

Solution

Focus on

i) Negation determiners: ‘no’ and ‘any’

ii) Non-existence-proving adverbs: ‘without’ and ‘never

No training - Labelling only when applying models

1. Text preprocessing

They didn’t have any children	→ *They have no children*
He has never been married	→ *He has been married 0 times*
The marriage was without children	→ *The marriage was with no children.*

2. Textual occurrences of ‘no’ and ‘0’ → CARDINAL (0)

Count information from text

Relation	Baseline [22]			CINEX-CRF			CINEX-CRF (per type)					
							Cardinals		Numt.+Art.		Ordinals	
	P	Cov	MAE	P	Cov	MAE	P	Contr	P	Contr	P	Contr
containsWork	42.0	**29.0**	3.7	**49.2**	**29.0**	**2.6**	55.0	33.9	62.5	40.7	20.0	25.4
hasMember	11.8	6.0	3.8	**64.3**	**18.0**	**1.2**	62.5	28.6	65.0	71.4	0	0
containsAdmin	51.8	14.5	7.3	**78.6**	**22.0**	**1.7**	85.7	87.5	33.3	10.7	0	1.8
hasChild	37.0	**22.0**	**2.2**	**50.0**	19.5	2.3	67.3	70.5	6.3	20.5	14.3	9.0
hasSpouse	26.8	11.0	1.3	**58.1**	**12.5**	**0.5**	75.0	18.6	43.8	37.2	63.2	44.2
hasZeroChild				92.3	18.8	-						
hasZeroSpouse				71.9	13.7	-						

Performance of CINEX in consolidation of counting quantifier mentions on Wikidata.

Paramita Mirza, Simon Razniewski, Fariz Darari, Gerhard Weikum
[Enriching Knowledge Bases with Quantifiers]
International Semantic Web Conference (ISWC) 2018.

1. Count information for recall assessment
2. How can we extract count information from text?
3. Variants of count information in KB
4. How much count information is accounted for?
5. Counts for KB curation

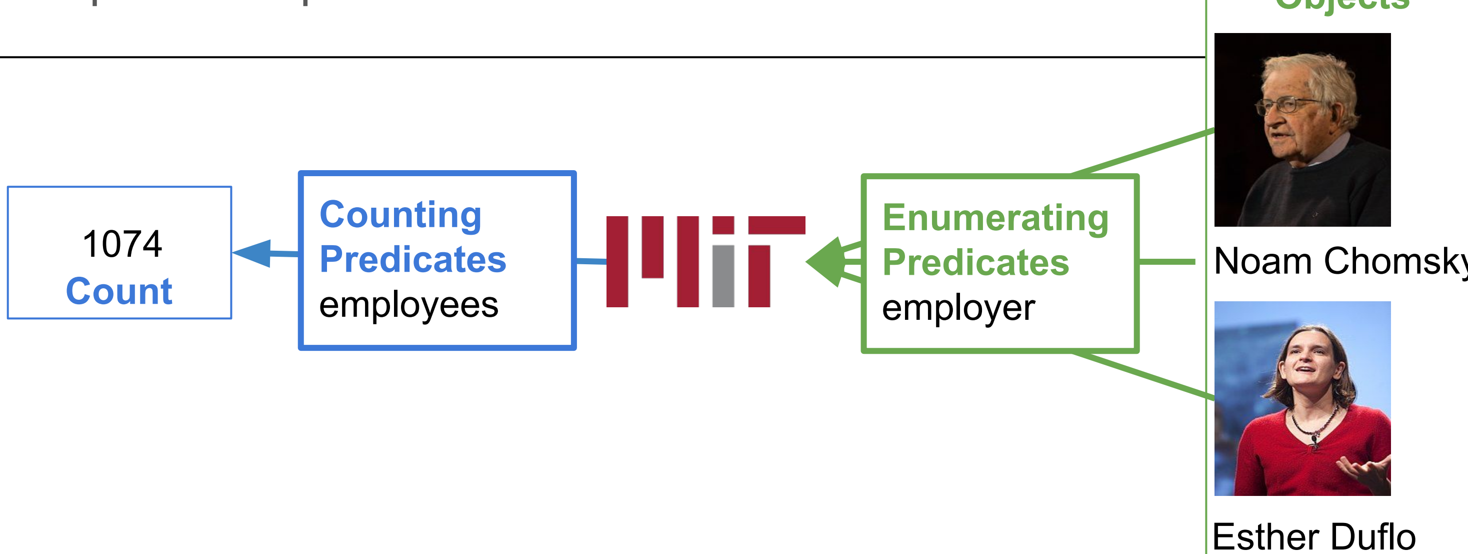
Count information in KB

Problem: Identification of semantically related count predicates

Input:

- a set of KB triples *(s,p,o)*
- and its inverse predicate triples *(s,p^{-1},o)*

Task: Determine counting and enumerating predicates and semantically related predicate pairs.



Count information in KB

Task 1: Identification of the count predicates - counting and enumerating

Count information in KB

Task 1: Identification of the count predicates - counting and enumerating

List of frequent KB predicates

academic_staff, staff, faculty	number_of_children		wins, doubles_titles, singles_titles
		...	
work_institution^{-1}, workplace^{-1}, work_institutions^{-1}	child		gold^{-1}

Count information in KB

Task 1: Identification of the two variants of count predicates

Counting Predicates

academic_staff, staff, faculty	number_of_children	...	wins, doubles_titles, singles_titles

Enumerating Predicates

work_institution^{-1}, workplace^{-1}, work_institutions^{-1}	child	...	gold^{-1}

Challenge: The separation is not clear.

Not all counting predicates store (single) integers

Not all enumerating predicates store entities

Count information in KB

Task 1: Identification of the two variants of count predicates

Counting Predicates

academic_staff, staff, faculty | number_of_children | ... | wins, doubles_titles, singles_titles

Enumerating Predicates

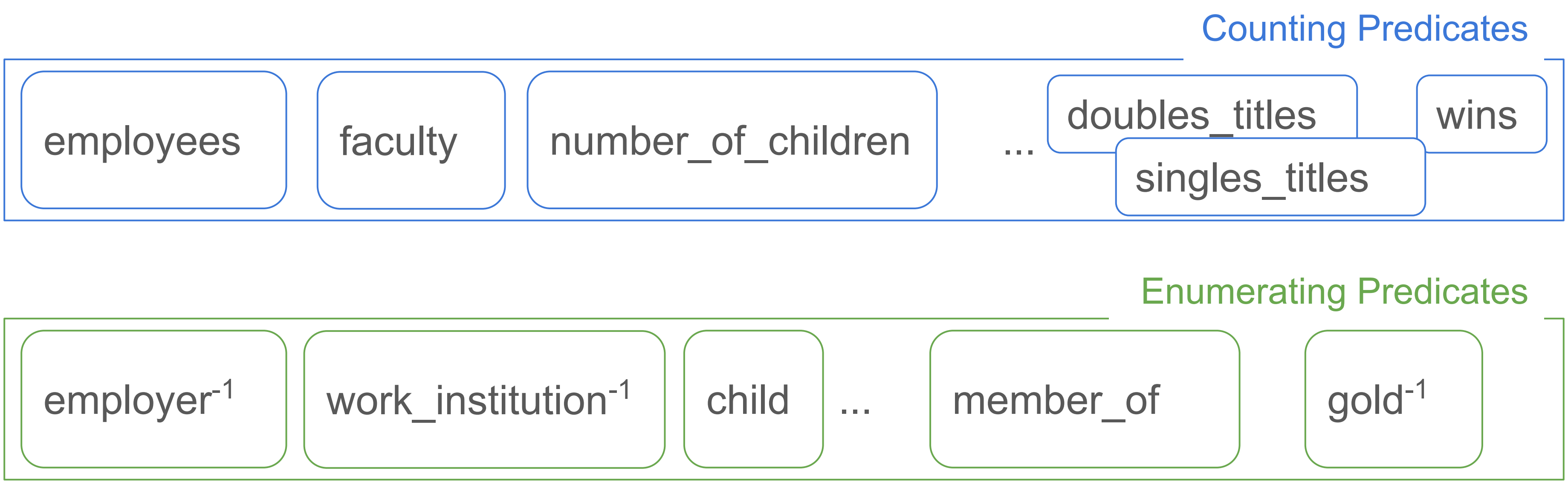
work_institution^{-1}, workplace^{-1}, work_institutions^{-1} | child | ... | gold^{-1}

Supervised Classification using:

- **Textual Features** - count predicates are more often used in plural form
- **Type Information** - classes of subject and objects
- **KB statistics** - #objects per subject, datatype distribution of the objects

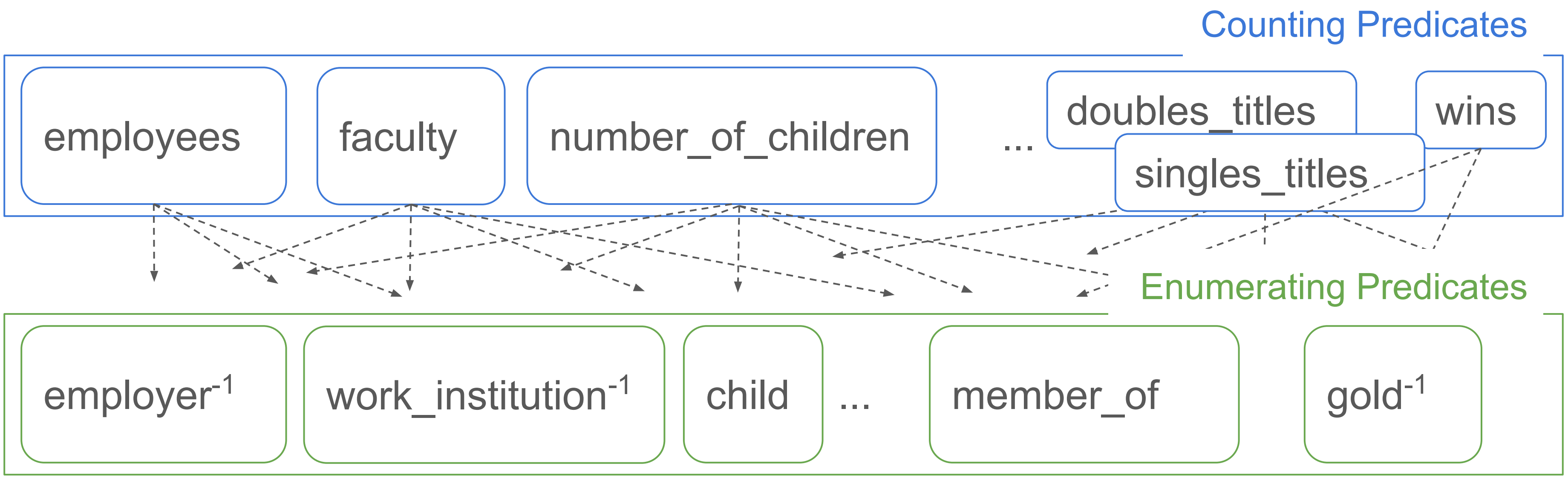
Count information in KB

Task 2: **Align** pairs of counting and enumerating predicates



Count information in KB

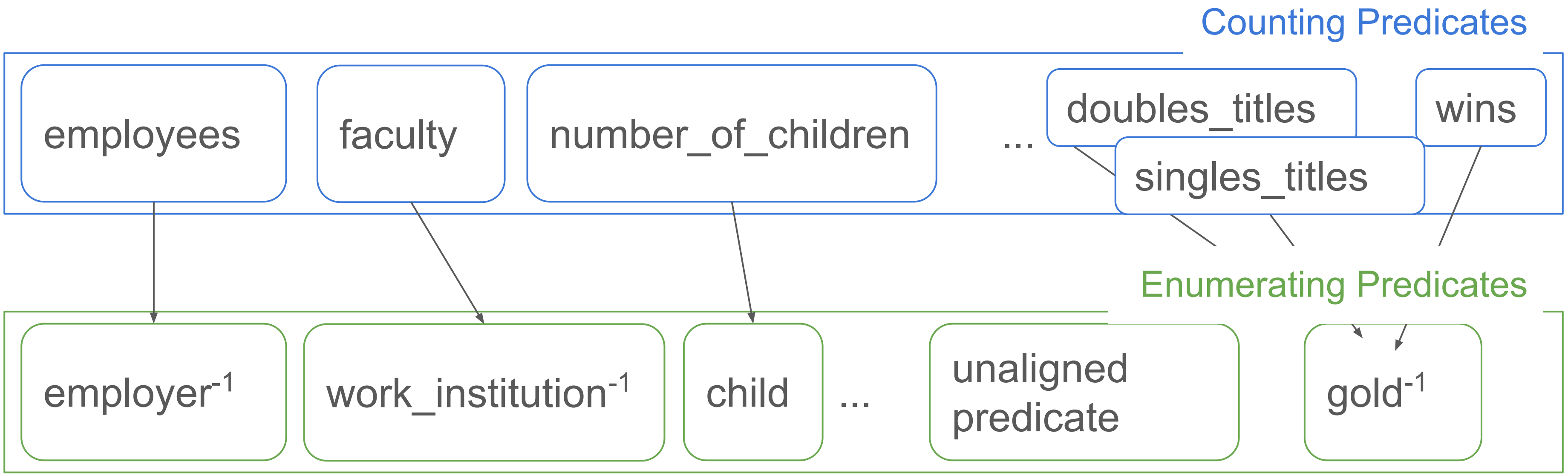
Task 2: **Align** pairs of counting and enumerating predicates



Challenge: KB facts are sparse and unclean.

Institutions can use faculty_size, employees or staff to mean the same thing.

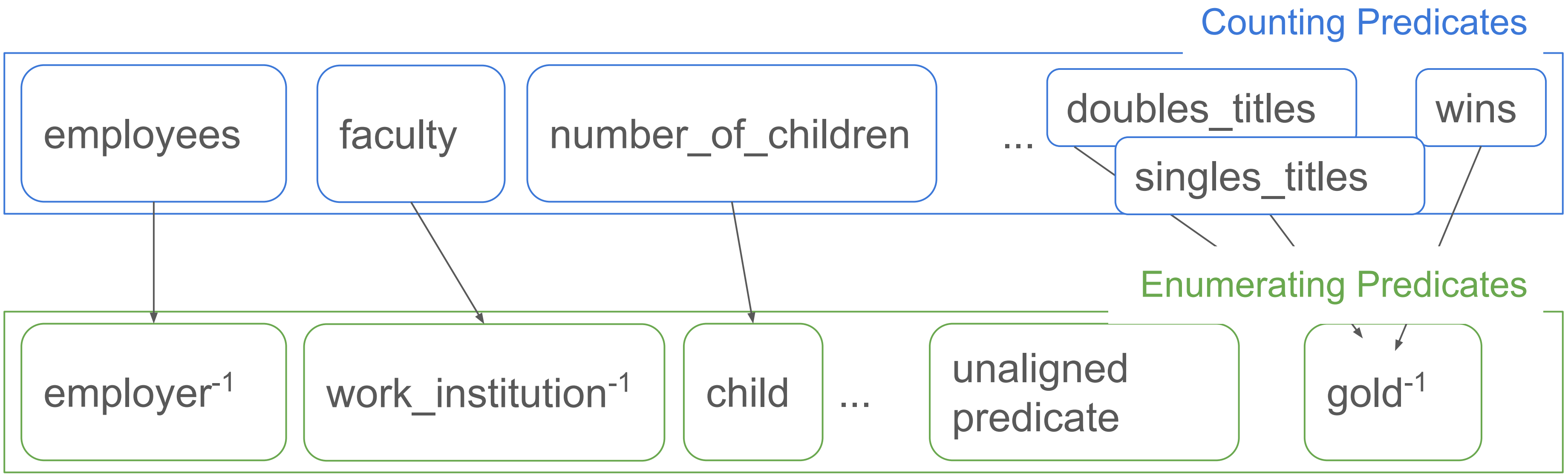
Count information in KB



Heuristics used for the predicate pair ***(e,c)***, where ***e*** stores entities and ***c*** counts.

1. Predicate pair co-occurrences - #subjects ***e*** and ***c*** co-occur

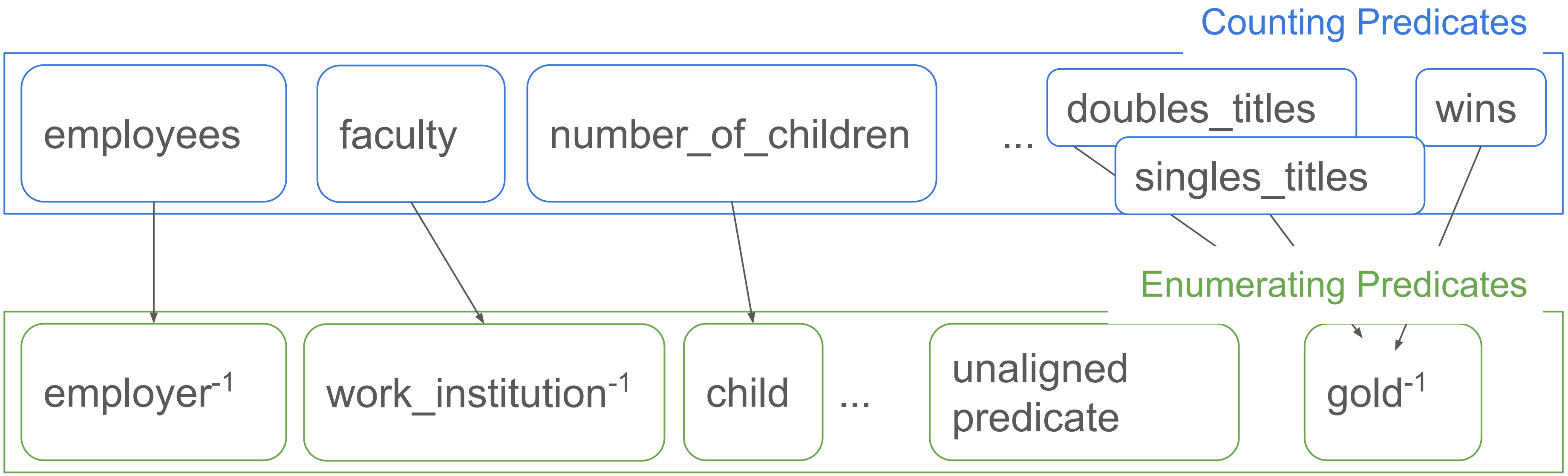
Count information in KB



Heuristics used for the predicate pair ***(e,c)***, where ***e*** stores entities and ***c*** counts.

1. Predicate pair co-occurrences - #subjects ***e*** and ***c*** co-occur
2. Value distribution - number of objects of ***e*** compared to count in ***c***

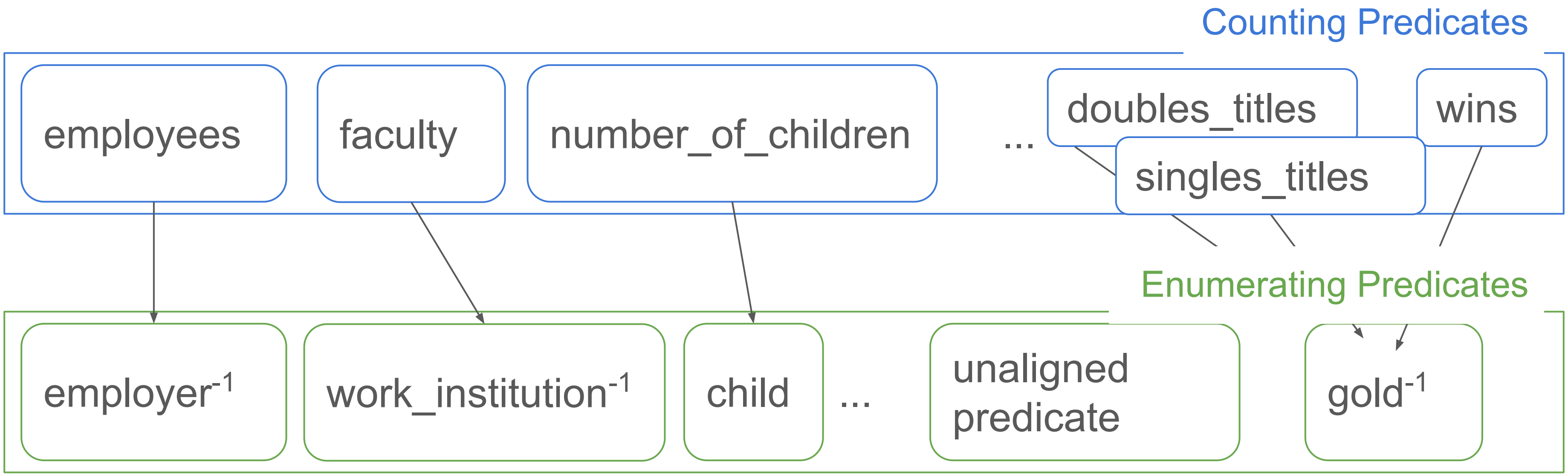
Count information in KB



Heuristics used for the predicate pair ***(e,c)***, where ***e*** stores entities and ***c*** counts.

1. Predicate pair co-occurrences - #subjects ***e*** and ***c*** co-occur
2. Value distribution - number of objects of ***e*** compared to count in ***c***
   a. is it equal for all subjects?
   b. is there any correlation?

Count information in KB



Heuristics used for the predicate pair ***(e,c)***, where ***e*** stores entities and ***c*** counts.

1. Predicate pair co-occurrences - #subjects ***e*** and ***c*** co-occur
2. Value distribution - number of objects of ***e*** compared to count in ***c***
    a. is it equal for all subjects?
    b. is there any correlation?
2. Linguistic similarity - do ***e*** and ***c*** talk share topical similarity?

Count information in KB

Training data generation: Crowd-sourced annotation of randomly selected predicate subsets

Challenges: KB predicates rarely have clean values

Count information in KB

Training data generation: Crowd-sourced annotation of randomly selected predicate subsets

Challenges: KB predicates rarely have clean values

- Cannot rely only on #triples per subject for enumerating predicates
- Integer value for a predicate ↛ Counting predicate (seat number, codes, IDs)
- Need for human in the loop

Count information in KB

Training data generation: Crowd-sourced annotation of randomly selected predicate subsets

Challenges: KB predicates rarely have clean values

- Cannot rely only on #triples per subject for enumerating predicates
- Integer value for a predicate ↛ Counting predicate (seat number, codes, IDs)
- Need for human in the loop

Input: Predicate P, 5 KB triples per predicate

Output:

- Graded relevance score for each P
- +ve: Average score from 3 users is between [0.6, 1.0]
- -ve: Average score between [0, 0.4]

Count information in KB

Training data generation: Crowd-sourced annotation of randomly selected predicate subsets

Challenges: KB predicates rarely have clean values

- Cannot rely only on #triples per subject for enumerating predicates
- Integer value for a predicate ↛ Counting predicate (seat number, codes, IDs)
- Need for human in the loop

Input: Predicate P, 5 KB triples per predicate

Output:

to weed out predicates without clear polarity

- Graded relevance score for each P
- +ve: Average score from 3 users is between [0.6, 1.0]
- -ve: Average score between [0, 0.4]

Count information in KB

Ground truth data: Crowd-sourced annotation of the top enumerating (counting) predicates aligned to randomly selected counting (enumerating) predicates

Input: Counting predicate C and top aligned predicates of the other set (E_1, E_2, ..) returned by all heuristics.

faculty_size

work_institution^{-1}

works_at^{-1}

employer^{-1}

Count information in KB

Ground truth data: Crowd-sourced annotation of the top enumerating (counting) predicates aligned to randomly selected counting (enumerating) predicates

Input: Counting predicate C and top aligned predicates of the other set (E_1, E_2, ..) returned by all heuristics.

faculty_size

work_institution^{-1}

works_at^{-1}

employer^{-1}

Output:

- Graded relevance score for each pair (C, E_1), (C, E_2), …
- Determine top-3 aligned predicates for C

Repeat this for enumerating predicates to get their top-3 aligned counting predicates.

Count information in KB

Model	Recall	Precision	F1
Random	40.6	40.6	40.6
Logistic	**55.6**	51.7	53.5
Prior	**55.6**	51.0	53.5
Lasso	51.1	**59.6**	**55.0**
Neural	53.0	49.6	51.2

Model	Recall	Precision	F1
Random	12.8	12.8	12.8
Logistic	51.2	19.0	27.7
Prior	48.7	20.2	28.5
Lasso	**71.7**	**23.3**	**35.1**
Neural	35.8	20.8	26.3

Scores for predicting i) Enumerating ii) Counting predicates

Metric	Counting @1	Counting @3	Enumerating @1	Enumerating @3
Absolute	0.71	0.56	0.62	0.63
Jaccard	0.76	0.61	0.69	0.67
$Conditional_C$	0.71	0.56	0.68	0.67
$Conditional_E$	0.76	0.68	0.62	0.63
P'wiseMI	0.73	0.58	0.71	0.70
P'fectMR	0.70	0.57	0.73	0.72
Correlation	0.77	0.69	0.62	0.61
P'tileVM	0.72	0.57	0.65	0.65
CosineSim	0.79	0.61	0.74	0.73
Combined	**0.84**	**0.67**	**0.75**	**0.75**

NDCG scores for predicate alignment

Shrestha Ghosh, Simon Razniewski, Gerhard Weikum
Uncovering Hidden Semantics of Set Information in Knowledge Bases
Journal of Web Semantics (JWS) 2020.

1. Count information for recall assessment
2. How can we extract count information from text?
3. Variants of count information in KB
4. How much count information is accounted for?
5. Counts for KB curation

How much count information is accounted for?

Counts from text

173k new count facts increasing KB knowledge by **77%**

from just 4 Wikidata properties across 10 classes

2,205 negative assertions

2.5M new count facts increasing KB knowledge by **28.3%**

How much count information is accounted for?

Counts from text

173k new count facts increasing KB knowledge by **77%**

from just 4 Wikidata properties across 10 classes

2,205 negative assertions

2.5M new count facts increasing KB knowledge by **28.3%**

for the predicates: *hasSpouse* and *hasChild*

How much count information is accounted for?

Counts from text

173k new count facts increasing KB knowledge by **77%**

from just 4 Wikidata properties across 10 classes

2,205 negative assertions

2.5M new count facts increasing KB knowledge by **28.3%**

for the predicates: *hasSpouse* and *hasChild*

for 110 Wikidata properties-class pairs

Paramita Mirza, Simon Razniewski, Fariz Darari, Gerhard Weikum
[Enriching Knowledge Bases with Quantifiers]
International Semantic Web Conference (ISWC) 2018.

How much count information is accounted for?

KB	Enumerating
DBpedia-raw	4,090
DBpedia mapped	308
Wikidata-truthy	203
Freebase	7,614
Total	12,215

Number of predicted enumerating KB predicates

From more than 36k frequent predicates across KBs including inverses.

How much count information is accounted for?

KB	Enumerating	Counting
DBpedia-raw	4,090	5,853
DBpedia mapped	308	898
Wikidata-truthy	203	1,067
Freebase	7,614	1,687
Total	12,215	9,505

Number of predicted counting KB predicates

From more than 26k frequent predicates across KBs.

How much count information is accounted for?

Number of predicted count predicates and KB alignments

KB	Enumerating	Counting	Alignments
DBpedia-raw	4,090	5,853	3,703
DBpedia mapped	308	898	270
Wikidata-truthy	203	1,067	31
Freebase	7,614	1,687	274
Total	12,215	9,505	4,278

Quite a low number of alignments: indicative of KB sparsity

Shrestha Ghosh, Simon Razniewski, Gerhard Weikum
[Uncovering Hidden Semantics of Set Information in Knowledge Bases]
Journal of Web Semantics (JWS) 2020.

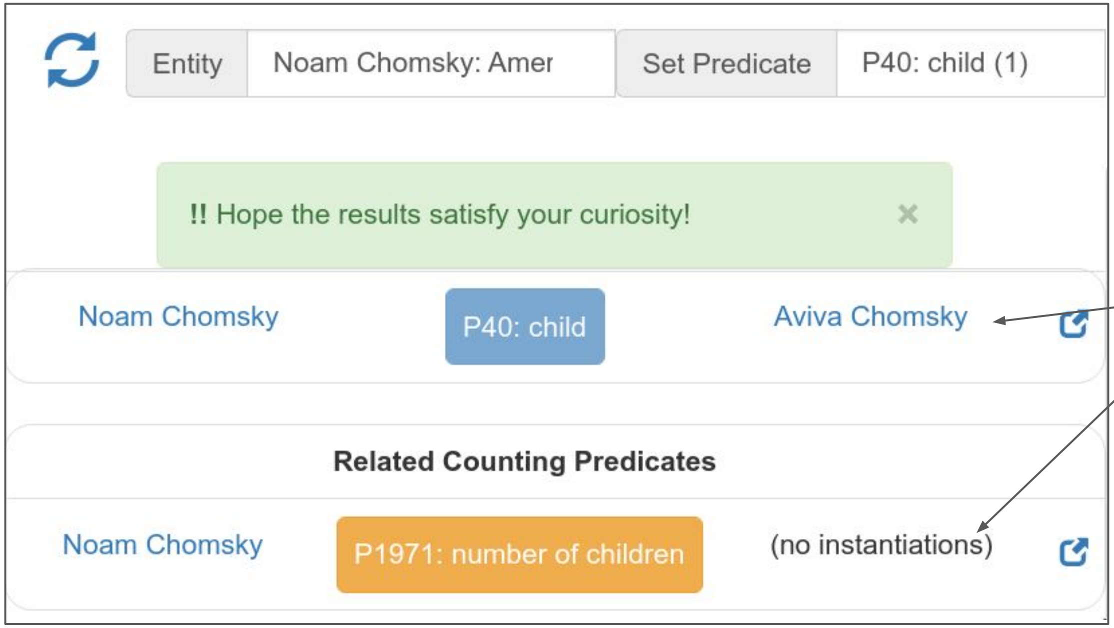
How much count information is accounted for?

Open questions and challenges

- #alignments << #counting and #enumerating predicates
  - unaligned count predicate → scope for new predicates
- Clustering similar predicates (faculty ↔ staff size)
  - staff size exists for an entity instead of faculty, then use it
- Cardinality extractors from text individually trained for each predicate
- Extracting negative cardinals is difficult - zero, none, don’t have.
- Enumeration for static (children, spouses) vs dynamic classes (population, books)

1. Count information for recall assessment
2. How can we extract count information from text?
3. Variants of count information in KB
4. How much count information is accounted for?
5. Counts for KB curation

Counts for KB curation

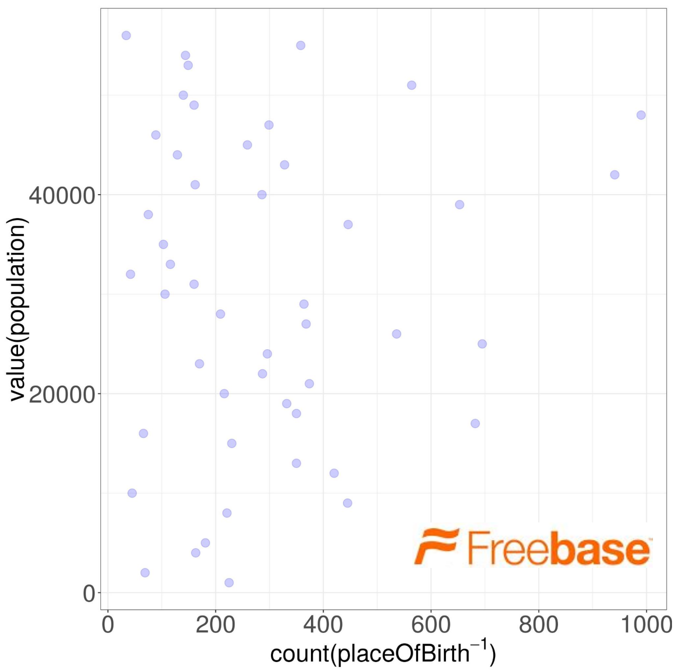


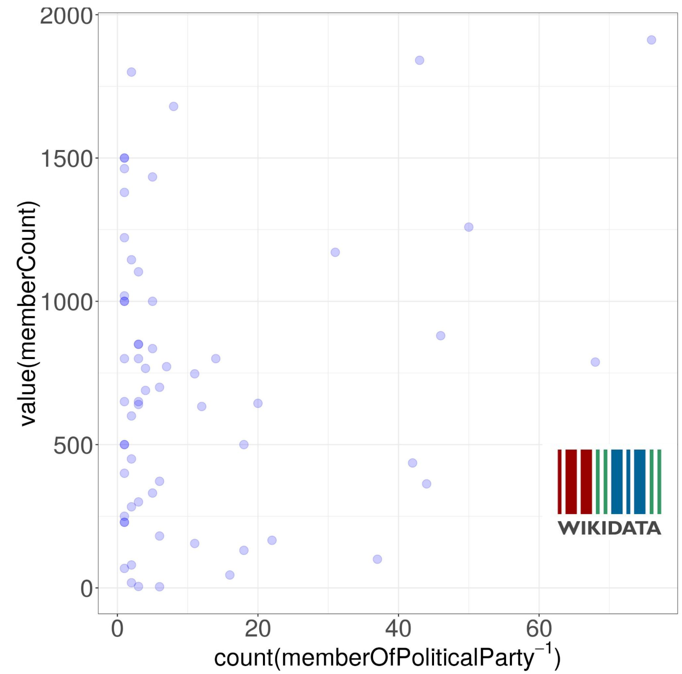
KB inconsistencies are highlighted

https://counqer.mpi-inf.mpg.de/spo

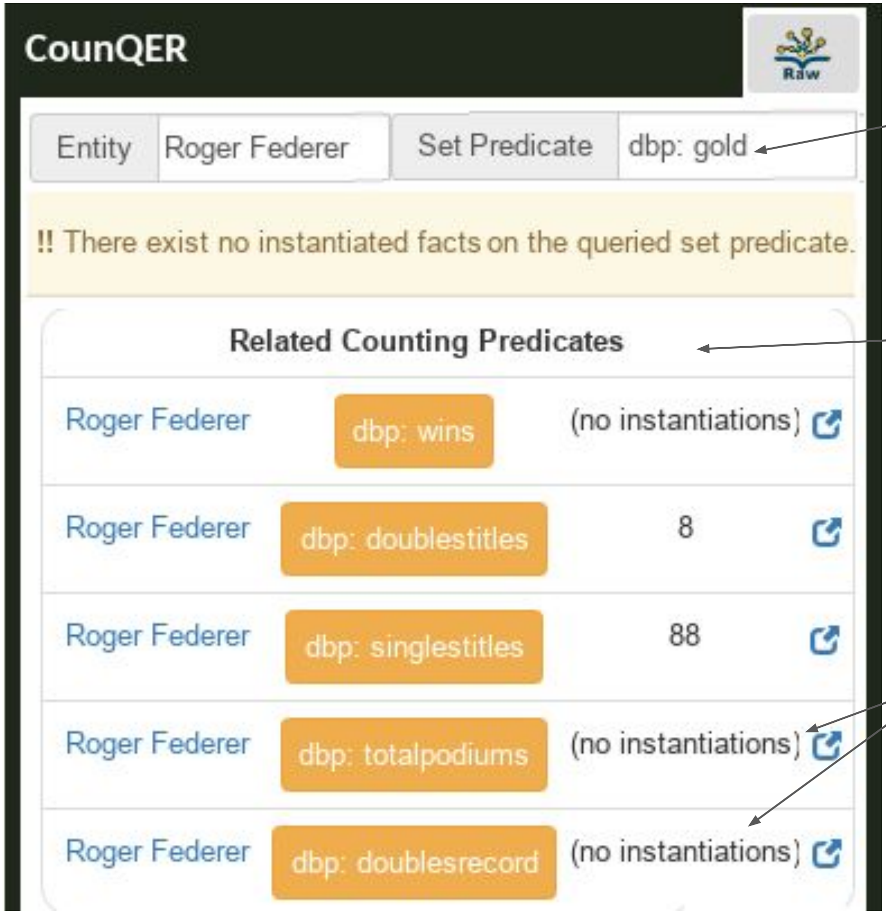
Counts for KB curation

Value distribution of aligned predicates show incompleteness





Counts for KB curation



Enhanced KB question answering

No answer to the original query on enumerating predicate

Related count answers obtained from aligned count predicates

Count predicates which could potentially give more information

Takeaway: Counts from text and KB

1. Count information for recall assessment
    - Counts and entities benefit from each other
    - KB mixes counts with standard facts
    - Counts can improve KB recall
2. Count information in text
    - is linguistically diverse (cardinals, ordinals, ..)
    - used to get the #objects for a given subject and predicate
3. Count information in KBs
    - can be identified by supervised classification
    - occurs as semantically related counting and enumerating predicates
4. KB curation using counts
    - highlights inconsistencies
    - gives value distribution of aligned predicates
    - can enhance KB question answering