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)

CC-BY Shrestha Ghosh. Slides freely available on tutorial website
What is count information?

Relation between an entity and a set of entities
What is count information?

Relation between an entity and a set of entities

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

**Enumerating Predicates**

- employer

**Objects**

- Noam Chomsky
- Esther Duflo
What is count information?

Relation between an entity and a set of entities

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

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

- Count: 1074
- Objects:
  - Noam Chomsky
  - Esther Duflo
- Predicates:
  - Counting: employer
  - Enumerating: employees

Enumerating Predicates

Counting Predicates

MIT
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

Only entities

(\(?x, \text{employer}, \text{MIT}\))

returns a handful of names from KB

Counts and entities benefit from each other
Count information for recall assessment

Only entities

(\( ?x, \text{employer}, \text{MIT} \))

returns a handful of names from KB

Only counts

(\( \text{MIT}, \text{employees}, ?y \))

gives no insight about the entities

Counts and entities benefit from each other
Count information for recall assessment

Only entities

(\(x, \text{employer}, \text{MIT}\))
returns a handful of names from KB

Only counts

(\(\text{MIT, employees, } y\))
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?)

Tim Berners-Lee

number of children

2

Ada Lovelace

child

Anne Blunt

Ralph King-Milbanke

Byron King-Noel
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?)

Enumeration is often of known entities
Count information for recall assessment

Count information can highlight KB inconsistencies

Definitely incomplete!
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

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

Subject: Noam Chomsky
Predicate: number_of_children

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

<table>
<thead>
<tr>
<th>Cardinals</th>
<th>Ordinals</th>
</tr>
</thead>
<tbody>
<tr>
<td>two sons,</td>
<td>second son,</td>
</tr>
<tr>
<td>three books</td>
<td>third book</td>
</tr>
</tbody>
</table>
Task 1: Identify the count tokens and the compositional cues.

**COUNT** tokens are **linguistically diverse**

<table>
<thead>
<tr>
<th>Cardinals</th>
<th>Ordinals</th>
<th>Number-related terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>two sons,</td>
<td>second son,</td>
<td>twins, trilogy</td>
</tr>
<tr>
<td>three books</td>
<td>third book</td>
<td></td>
</tr>
</tbody>
</table>

Count information from text

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

**COUNT tokens are linguistically diverse**

<table>
<thead>
<tr>
<th>Cardinals</th>
<th>Ordinals</th>
<th>Number-related terms</th>
<th>Indefinite Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>two sons,</td>
<td>second son,</td>
<td>twins, trilogy</td>
<td>a son,</td>
</tr>
<tr>
<td>three books</td>
<td>third book</td>
<td></td>
<td>the book</td>
</tr>
</tbody>
</table>
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
Jolie brought her **six** children: **twins**, one daughter and **three** adopted children to the gala.

**Subject**: Angelina Jolie  
**Predicate**: number_of_children

**Task 2**: Consolidate count tokens

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

1. **Sum up compositional cues**
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**

<table>
<thead>
<tr>
<th>Cardinal</th>
<th>&gt;&gt; number-related terms</th>
<th>&gt;&gt; Ordinals</th>
<th>&gt;&gt; indefinite article</th>
</tr>
</thead>
<tbody>
<tr>
<td>two children</td>
<td>&gt;&gt; twins</td>
<td>&gt;&gt; second child</td>
<td>&gt;&gt; a child</td>
</tr>
</tbody>
</table>
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

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

+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
   
   \[\text{They didn’t have any children } \rightarrow \text{They have no children}\]
   
   \[\text{He has never been married } \rightarrow \text{He has been married 0 times}\]
   
   \[\text{The marriage was without children } \rightarrow \text{The marriage was with no children.}\]

2. Textual occurrences of ‘no’ and ‘0’ → CARDINAL (0)
Count information from text

<table>
<thead>
<tr>
<th>Relation</th>
<th>Baseline [22]</th>
<th>CINEX-CRF</th>
<th>CINEX-CRF (per type)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>Cov</td>
<td>MAE</td>
</tr>
<tr>
<td>containsWork</td>
<td>42.0</td>
<td>29.0</td>
<td>3.7</td>
</tr>
<tr>
<td>hasMember</td>
<td>11.8</td>
<td>6.0</td>
<td>3.8</td>
</tr>
<tr>
<td>containsAdmin</td>
<td>51.8</td>
<td>14.5</td>
<td>7.3</td>
</tr>
<tr>
<td>hasChild</td>
<td>37.0</td>
<td>22.0</td>
<td>2.2</td>
</tr>
<tr>
<td>hasSpouse</td>
<td>26.8</td>
<td>11.0</td>
<td>1.3</td>
</tr>
<tr>
<td>hasZeroChild</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hasZeroSpouse</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

Paramita Mirza, Simon Razniewski, Fariz Darari, Gerhard Weikum
Enriching Knowledge Bases with Quantifiers
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.

1074 Count

**Counting Predicates**
- employees

**Enumerating Predicates**
- employer

**Objects**
- Noam Chomsky
- Esther Duflo
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*

<table>
<thead>
<tr>
<th>academic_staff, staff, faculty</th>
<th>number_of_children</th>
<th>wins, doubles_titles, singles_titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>work_institution(^{-1}), workplace(^{-1}), work_institutions(^{-1})</td>
<td>child</td>
<td>gold(^{-1})</td>
</tr>
</tbody>
</table>
## Task 1: Identification of the two variants of count predicates

<table>
<thead>
<tr>
<th>Academic Staff</th>
<th>Staff</th>
<th>Faculty</th>
</tr>
</thead>
<tbody>
<tr>
<td>academic_staff</td>
<td>staff</td>
<td>faculty</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Children</th>
<th>...</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Work Institution</th>
<th>Workplace</th>
<th>Work Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>work_institution</td>
<td>workplace</td>
<td>work_institutions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Child</th>
<th>...</th>
<th>Gold</th>
</tr>
</thead>
<tbody>
<tr>
<td>child</td>
<td>...</td>
<td>gold</td>
</tr>
</tbody>
</table>

Counting Predicates:
- wins
- doubles_titles
- singles_titles

Enumerating Predicates:
- gold$^{-1}$
Count information in KB

Task 1: Identification of the two variants of count predicates

<table>
<thead>
<tr>
<th>academic_staff, staff, faculty</th>
<th>number_of_children</th>
<th>...</th>
<th>wins, doubles_titles, singles_titles</th>
</tr>
</thead>
</table>

| work_institution⁻¹, workplace⁻¹, work_institutions⁻¹ | child | ... | gold⁻¹ |

Supervised Classification using:

- **Textual Features** - count predicates are more often used in singular 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
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 \(c\)
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 $\rightarrow$ 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 $\rightarrow$ 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 $\rightarrow$ 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

**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 (E₁, E₂, ..) 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 \( (E_1, E_2, \ldots) \) returned by all heuristics.

**Output:**
- Graded relevance score for each pair \((C, E_1), (C, E_2), \ldots\)
- Determine top-3 aligned predicates for C

Repeat this for enumerating predicates to get their top-3 aligned counting predicates.
Count information in KB

<table>
<thead>
<tr>
<th>Model</th>
<th>Recall</th>
<th>Precision</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>40.6</td>
<td>40.6</td>
<td>40.6</td>
</tr>
<tr>
<td>Logistic</td>
<td>55.6</td>
<td>51.7</td>
<td>53.5</td>
</tr>
<tr>
<td>Prior</td>
<td>55.6</td>
<td>51.0</td>
<td>53.5</td>
</tr>
<tr>
<td>Lasso</td>
<td>51.1</td>
<td>59.6</td>
<td>55.0</td>
</tr>
<tr>
<td>Neural</td>
<td>53.0</td>
<td>49.6</td>
<td>51.2</td>
</tr>
</tbody>
</table>

Scores for predicting i) Enumerating ii) Counting predicates

<table>
<thead>
<tr>
<th>Metric</th>
<th>Counting @1</th>
<th>Counting @3</th>
<th>Enumerating @1</th>
<th>Enumerating @3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute</td>
<td>0.71</td>
<td>0.56</td>
<td>0.62</td>
<td>0.63</td>
</tr>
<tr>
<td>Jaccard</td>
<td>0.76</td>
<td>0.61</td>
<td>0.69</td>
<td>0.67</td>
</tr>
<tr>
<td>Conditional $C$</td>
<td>0.71</td>
<td>0.56</td>
<td>0.68</td>
<td>0.67</td>
</tr>
<tr>
<td>Conditional $E$</td>
<td>0.76</td>
<td>0.68</td>
<td>0.62</td>
<td>0.63</td>
</tr>
<tr>
<td>P’wiseMI</td>
<td>0.73</td>
<td>0.58</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td>P’fectMR</td>
<td>0.70</td>
<td>0.57</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.77</td>
<td>0.69</td>
<td>0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>P’tileVM</td>
<td>0.72</td>
<td>0.57</td>
<td>0.65</td>
<td>0.65</td>
</tr>
<tr>
<td>CosineSim</td>
<td>0.79</td>
<td>0.61</td>
<td>0.74</td>
<td>0.73</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td><strong>0.84</strong></td>
<td><strong>0.67</strong></td>
<td><strong>0.75</strong></td>
<td><strong>0.75</strong></td>
</tr>
</tbody>
</table>

NDCG scores for predicate alignment

Shrestha Ghosh, Simon Razniewski, Gerhard Weikum
Uncovering Hidden Semantics of Set Information in Knowledge Bases
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%
- 2,205 negative assertions
- 2.5M new count facts increasing KB knowledge by 28.3%

from just 4 Wikidata properties across 10 classes
How much count information is accounted for?

Counts from text

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

2,205 negative assertions

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

from just 4 Wikidata properties across 10 classes

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%**

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

from just 4 Wikidata properties across 10 classes

Paramita Mirza, Simon Razniewski, Fariz Darari, Gerhard Weikum

_Enriching Knowledge Bases with Quantifiers_

### How much count information is accounted for?

<table>
<thead>
<tr>
<th>KB</th>
<th>Enumerating</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBpedia-raw</td>
<td>4,090</td>
</tr>
<tr>
<td>DBpedia mapped</td>
<td>308</td>
</tr>
<tr>
<td>Wikidata-truthy</td>
<td>203</td>
</tr>
<tr>
<td>Freebase</td>
<td>7,614</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>12,215</strong></td>
</tr>
</tbody>
</table>

Number of predicted **enumerating** KB predicates
How much count information is accounted for?

<table>
<thead>
<tr>
<th>KB</th>
<th>Enumerating</th>
<th>Counting</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBpedia-raw</td>
<td>4,090</td>
<td>5,853</td>
</tr>
<tr>
<td>DBpedia mapped</td>
<td>308</td>
<td>898</td>
</tr>
<tr>
<td>Wikidata-truthy</td>
<td>203</td>
<td>1,067</td>
</tr>
<tr>
<td>Freebase</td>
<td>7,614</td>
<td>1,687</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>12,215</strong></td>
<td><strong>9,505</strong></td>
</tr>
</tbody>
</table>

Number of predicted counting KB predicates

Shrestha Ghosh, Simon Razniewski, Gerhard Weikum
*Uncovering Hidden Semantics of Set Information in Knowledge Bases*
# How much count information is accounted for?

Number of predicted count predicates and KB alignments

<table>
<thead>
<tr>
<th>KB</th>
<th>Enumerating</th>
<th>Counting</th>
<th>Alignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBpedia-raw</td>
<td>4,090</td>
<td>5,853</td>
<td>3,703</td>
</tr>
<tr>
<td>DBpedia mapped</td>
<td>308</td>
<td>898</td>
<td>270</td>
</tr>
<tr>
<td>Wikidata-truthy</td>
<td>203</td>
<td>1,067</td>
<td>31</td>
</tr>
<tr>
<td>Freebase</td>
<td>7,614</td>
<td>1,687</td>
<td>274</td>
</tr>
<tr>
<td>Total</td>
<td>12,215</td>
<td>9,505</td>
<td>4,278</td>
</tr>
</tbody>
</table>

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](https://example.com)
How much count information is accounted for?

Open questions and challenges

- #alignments << #counting and #enumerating predicates
  - unaligned count predicate \(\rightarrow\) 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
- 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