Completeness, Recall, and Negation in Open-World Knowledge Bases




Simon Razniewski, Hiba Arnaout, Shrestha Ghosh, Fabian Suchanek

max planck institut informatik

TELECOM ParisTech

On the Limits of Machine Knowledge: Completeness, Recall and Negation in Web-scale Knowledge Bases

Simon Razniewski, Hiba Arnaout, Shrestha Ghosh, Fabian Suchanek

1. Introduction & Foundations (Simon) – 9:00-9:30 CEST
2. Predictive recall assessment (Fabian) – 9:30-10:10
3. Counts from text and KB (Shrestha) – 10:10-10:45
4. Negation (Hiba) – 10:45-11:25
5. Relative completeness & Wrap-up (Simon) – 11:25-12:00

max planck institut informatik

TELECOM ParisTech

Machine knowledge in action

Traditional search

Google

physics nobel prize winners

All | News | Images | Videos | Maps | More

https://en.wikipedia.org › wiki › List_of_Nobel_laureat...

List of Nobel laureates in Physics - Wikipedia

John Bardeen is the only **laureate** to win the prize twice—in 1956 and 1972. Marie Skłodowska-Curie also won two **Nobel Prizes**, for **physics** in 1903 and ...

Andrea M. Ghez · Donna Strickland · Jim Peebles · Shuji Nakamura

https://en.wikipedia.org › wiki › Nobel_Prize_in_Physics

Nobel Prize in Physics - Wikipedia

Three **Nobel Laureates** in **Physics**. Front row L-R: Albert A. Michelson (1907 **laureate**), Albert Einstein (1921 **laureate**) and Robert A. Millikan (1923 **laureate**).

First awarded: 1901 **Most awards:** John Bardeen (2)

Most recently awarded to: Roger Penrose, ... **Awarded for:** Outstanding contributions for...

https://www.britannica.com › ... › International Relations

Winners of the Nobel Prize for Physics | Britannica

year	name	country*
1901	Wilhelm Conrad Röntgen	Germany
1902	Hendrik Antoon Lorentz	Netherlands
1902	Pieter Zeeman	Netherlands

View 213 more rows

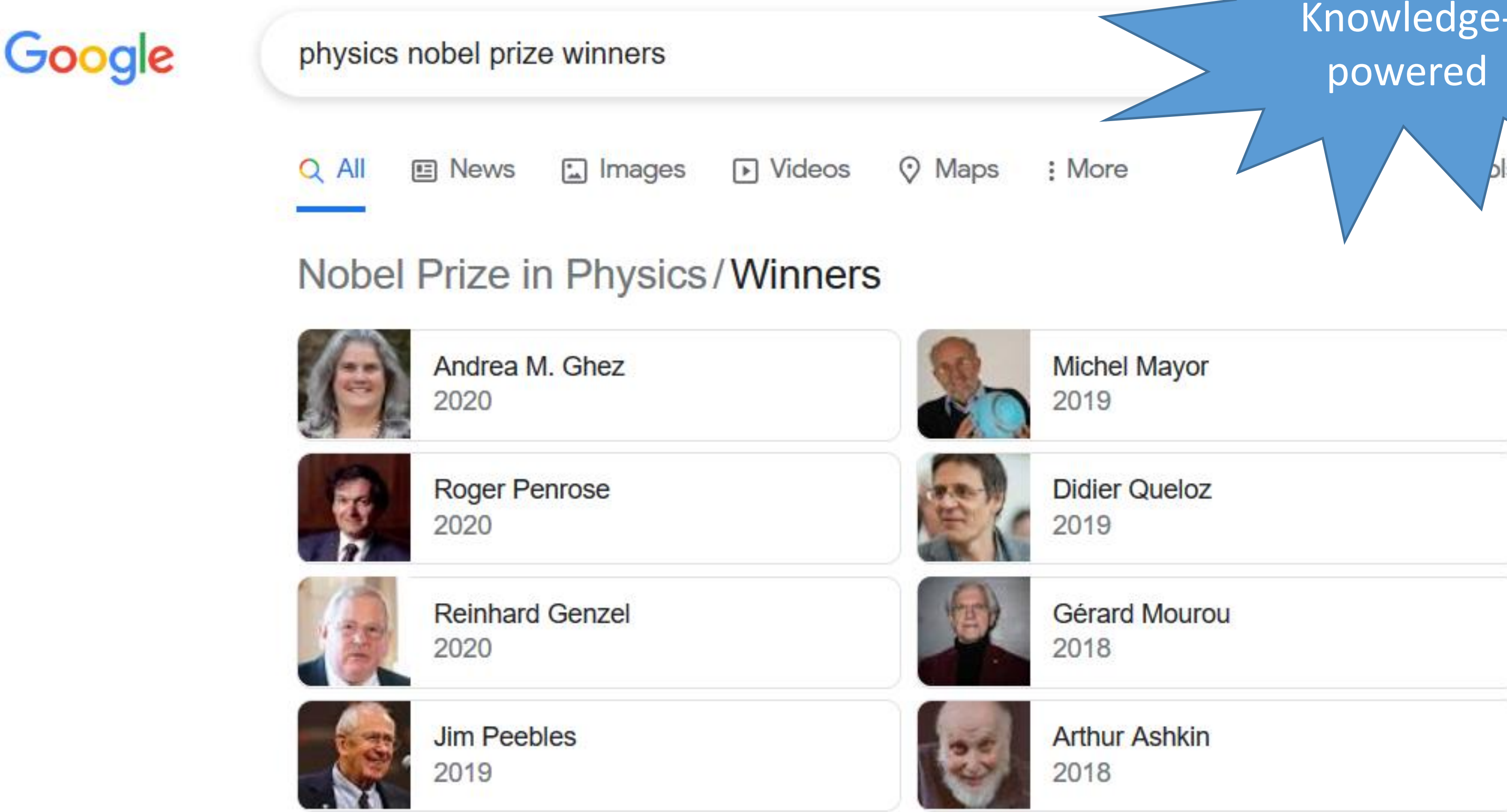
https://www.research-in-germany.org › nobel-laureates

German Nobel laureates - Research in Germany

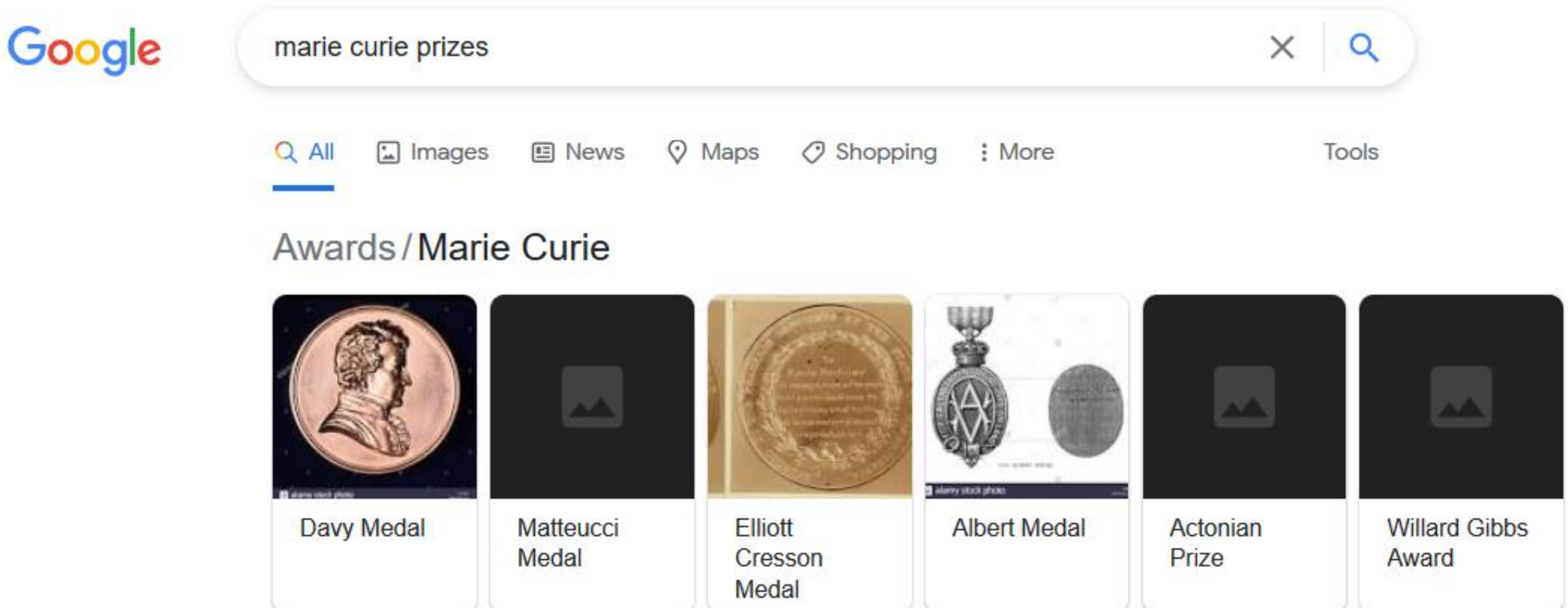
J. Georg Bednorz: 1987 - Physics ... An unusual approach made Georg Bednorz a pioneer in the field of superconductivity – and **Physics Nobel Prize laureate** in ...

Machine knowledge in action

Knowledge-powered



Machine knowledge in action



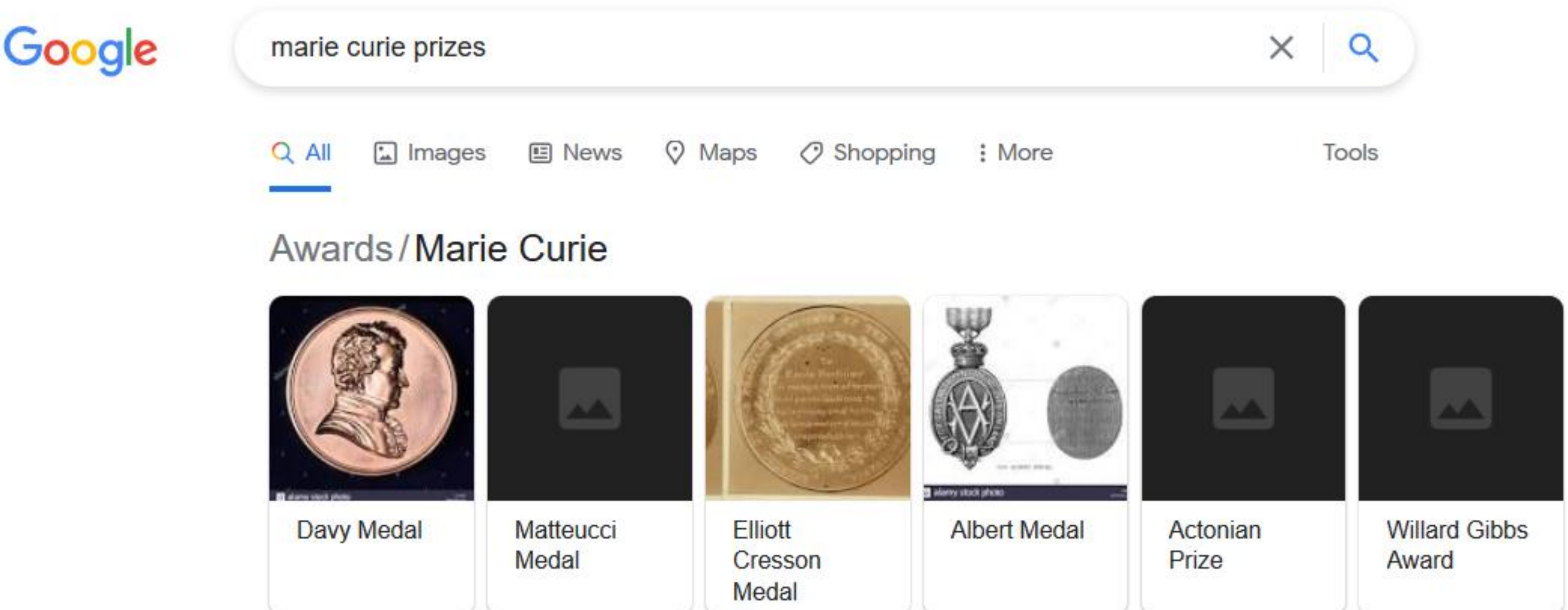
Machine knowledge is awesome

- Reusable, scrutable asset for knowledge-centric tasks
  - Semantic search & QA
  - Entity-centric text analytics
  - Distant supervision for ML
  - Data cleaning

- Impactful projects at major public and commercial players
  - Wikidata, Google KG, Microsoft Satori, …

- Strongly rooted in semantic web community
  - Linked data, vocabularies, ontologies, indexing and querying, …

But:
Machine Knowledge is incomplete



Nobel Prize (2x)

Machine knowledge is incomplete (2)

Wikidata KB:

Semantic Web Journal has only published 84 articles ever

- https://scholia.toolforge.org/venue/Q15817015

Most cited papers on data integration have <38 citations

- https://scholia.toolforge.org/topic/Q386824

But: Machine knowledge is one-sided

- In KB:
  - *Nicola Tesla received title of IEEE fellow*
  - *Vietnam is a member of ASEAN*
  - *iPhone has 12MP camera*

- Not in KB:
  - *Nicola Tesla did not receive the Nobel Prize*
  - *Switzerland is not a member of the EU*
  - *iPhone 12 has no headphone jack*

Why is this problematic? (1) Querying

- Decision making more and more data-driven

- Analytical queries paint wrong picture of reality
  - *E.g., SW journal deemed too small*

- Instance queries return wrong results
  - *E.g., wrongly assuming certain authors never published in SW Journal*

Why is this problematic? (2) Data curation

- Effort priorization fundamental challenge in human-in-the-loop curation
  - *Should we spend effort on obtaining data for SWJ or TKDE?*

- Risk of effort duplication if not keeping track of completed areas
  - *Spending effort on collecting data … already present*

Why is this problematic? (3)
Summarization and decision making

Booking.com

Bathroom
- Toilet paper
- Towels
- Private bathroom
- Toilet
- Free toiletries
- Hairdryer
- Shower

Bedroom
- Linen
- Wardrobe or closet
- Alarm clock

Room Amenities
- Sock
- C

P

Pets a
applic

A
- C
- G

A

Me
- Flat-screen TV
- Satellite channels
- Radio
- Telephone
- TV
- Pay-per-view channels

Food & Drink
- On-site coffee house
- Chocolate or cookies Additional charge
- Fruits Additional charge

Safety & security
- Fire extinguishers
- CCTV outside property
- CCTV in common areas
- Smoke alarms
- 24-hour security
- Safety deposit box

General
- Paid WiFi
- Mini-market on site
- Vending machine (drinks)
- Designated smoking area
- Air conditioning
- free room
- Facilities for disabled guests
- Ironing facilities
- Non-smoking rooms
- Iron
- Air conditioning

Accessibility
- Visual aids: Tactile signs
- Visual aids: Braille
- Lower bathroom sink
- Higher level toilet
- Toilet with grab rails
- Wheelchair accessible

Wellness
- Fitness
- Full body massage Additional charge
- Hand massage Additional charge
- Head massage Additional charge
- Couples massage Additional charge
- Foot massage Additional charge
- Neck massage Additional charge
- Back massage Additional charge
- Spa/wellness packages
- Steam room
- Spa Facilities
- Light therapy
- Facial treatments
- Beauty Services
- Sun loungers or beach chairs
- Pool/beach towels
- Hot tub/jacuzzi
- Massage Additional charge
- Spa and wellness centre Additional charge
- Fitness centre
- Sauna Additional charge

Languages spoken
- English

No free WiFi!

Camera
- Pro 12MP camera system: Ultra Wide, Wide, and Telephoto cameras
- Ultra Wide: ƒ/2.4 aperture and 120° field of view
- Wide: ƒ/1.6 aperture
- Telephoto: ƒ/2.2 aperture
- 2.5x optical zoom in, 2x optical zoom out; 5x optical zoom range
- Digital zoom up to 12x
- Night mode portraits enabled by LiDAR Scanner
- Portrait mode with advanced bokeh and Depth Control
- Portrait Lighting with six effects (Natural, Studio, Contour, Stage, Stage Mono, High-Key Mono)
- Dual optical image stabilization (Wide and Telephoto)
- Sensor-shift optical image stabilization
- Five-element lens (Ultra Wide); six-element lens (Telephoto); seven-element lens (Wide)
- Brighter True Tone flash with Slow Sync
- Panorama (up to 63MP)
- Sapphire crystal lens cover
- 100% Focus Pixels (Wide)
- Night mode (Ultr Wide)
- Deep Fusion (hoto)
- 720p HD video recording at 30 fps
- Sensor-shift optical image stabilization for video (Wide)
- Optical image stabilization for video (Wide)
- 2.5x optical zoom in, 2x optical zoom out; 5x optical zoom range
- Digital zoom up to 7x
- Audio zoom
- Brighter True Tone flash
- QuickTake video
- Slo-mo video support for 1080p at 120 fps or 240 fps
- Time-lapse video with stabilization
- Night mode Time-lapse
- Extended dynamic range for video up to 60 fps
- Cinematic video stabilization (4K, 1080p, and 720p)
- Continuous autofocus video

No headphone jack

Topic of this tutorial

How to know how much a KB knows?

How to = techniques
How much knows = completeness/recall/coverage bookkeeping/estimation
KB = General world knowledge repository

What this tutorial offers

- Logical foundations
  - Setting and formalisms for describing KB completeness (part 1)
- Predictive assessment
  - How (in-)completeness can be statistically predicted (Part 2)
- Count information
  - How count information enables (in-)completeness assessment (Part 3)
- Negation
  - How salient negations can be derived from incomplete KBs (Part 4)
- Relative recall
  - How to define and measure recall in without gold standard (Part 5)

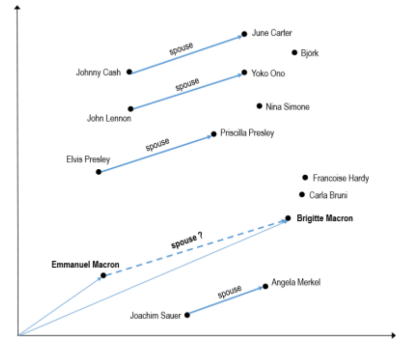
Goals:

1. Systematize the topic and its facets
2. Lay out assumptions, strengths and limitations of approaches
3. Provide a practical toolsuite

Relevant research domains

- Semantic Web
- Logics
- Statistics
- Machine Learning
- Natural language processing

What this tutorial is NOT about



TransE - a KBC model

- Knowledge base completion (KBC)
  - "How to make KBs more complete"
- Related: Understanding of completeness is needed to know when/when not to employ KBC
  - KBC naively is open-ended

  → Understanding of completeness needed to "stop"

Beatles members:	
John Lennon	36%
Paul McCartney	23%
George Harrison	18%
Bob Dylan	5%
Ringo Starr	3%
Elvis Presley	2%
Yoko Ono	2%

- But:
  - Heuristic, error-prone KBC not always desired
  - Completeness awareness != actionable completion
- Literature on knowledge graph completion, link prediction, missing value imputation, etc.
  - E.g., Rossi, Andrea, et al. Knowledge graph embedding for link prediction: A comparative analysis *TKDD 2021*

https://www.mpi-inf.mpg.de/kr-2021-tutorial

On the Limits of Machine Knowledge: Completeness, Recall and Negation in Web-scale Knowledge Bases

Simon Razniewski, Hiba Arnaout, Shrestha Ghosh, Fabian Suchanek

1. Introduction & Foundations (Simon) – 9:00-9:30 CEST
2. Predictive recall assessment (Fabian) – 9:30-10:10
3. Counts from text and KB (Shrestha) – 10:10-10:45
4. Negation (Hiba) – 10:45-11:25
5. Relative completeness & Wrap-up (Simon) – 11:25-12:00

max planck institut informatik

TELECOM ParisTech

Knowledge base - definition

Given set **E** (entities), **L** (literals), **P** (predicates)

- Predicates are positive or negated properties
  - *bornIn, notWonAward, …*

- An assertion is a triple (s, p, o) $\in$ **E** $\times$ **P** $\times$ (**E**$\cup$**L**)
- A practically available KB $\mathbf{K}^a$ is a set of assertions
- The ``ideal'' (complete) KB is called $\mathbf{K}^i$
- Available KBs are incomplete: $\mathbf{K}^a \subseteq \mathbf{K}^i$

Knowledge bases (KBs aka. KGs)

subject-predicate-object triples about entities, attributes of and relations between entities

+ composite objects

predicate (subject, object)

type (Marie Curie, physicist)
subtypeOf (physicist, scientist)

taxonomic knowledge

placeOfBirth (Marie Curie, Warsaw)
residence (Marie Curie, Paris)
¬placeOfBirth (Marie Curie, France)

factual knowledge

discovery (Polonium, 12345)
discoveryDate (12345, 1898)
discoveryPlace (12345, Paris)
discoveryPerson (12345, Marie Curie)

spatio-temporal & contextual knowledge

atomicNumber (Polonium, 84)
halfLife (Polonium, 2.9 y)

expert knowledge

History of knowledge bases

Cyc

WordNet

guitarist
$\subset$ {player,musician}
$\subset$ artist

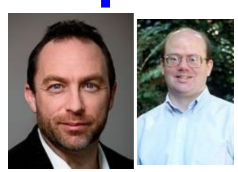
{player,footballer}
$\subset$ athlete

$\forall$ x: human(x) $\Rightarrow$
  ($\exists$ y: mother(x,y) $\wedge$
  $\exists$ z: father(x,z))

$\forall$ x,u,w: (mother(x,u) $\wedge$
  mother(x,w)
  $\Rightarrow$ u=w)

Manual compilation

Wikipedia



6 Mio. English articles
40 Mio. contributors

Automation and human-in-the-loop

DBpedia

yago select knowledge

freebase

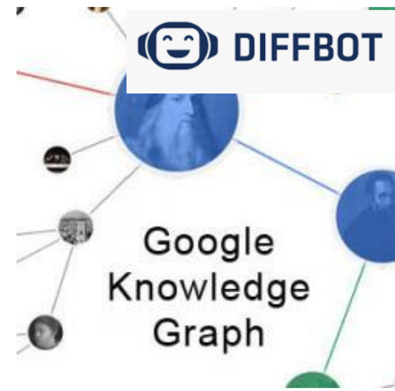
WolframAlpha computational knowledge engine

amazon

bing

WIKIDATA

DIFFBOT



BOSCH

Bloomberg

Alibaba.com

1985 1990 2000 2005 2010 2020

KB scale and use cases

Wikidata (open)
- 95 M items
- 1.1 B statements

WIKIDATA

Google KG
- 5 B items
- 500 B statements

Google Knowledge Graph

Major use cases:
- semantic search & QA
- language understanding
- distant supervision for ML
- data cleaning

KB incompleteness is inherent

Why?

Reality

Einstein received the Nobel Prize in 1921, the Copley medal, the Prix Jules Jansen, the Medal named after Max Planck, and several others

Doc

1. Sources incomplete

Honorary doctorate, UMadrid
Gold medal, Royal Astronomic Society
Benjamin Franklin Medal,
…
¬NobelPrizeFor(Einstein, RelativityTheory)
¬NobelPrizeFor(Einstein, ElectricToaster)
…

2. Negations quasi-infinite

Knowledge base construction

4. Extraction resource-bounded

Award(Einstein, NobelPrize)
~~Award(Einstein, Copley medal)~~
Award(Einstein, Prix Jules Jansen)
Friend(Einstein, Max Planck)

3. Extractors imperfect

Weikum et al.
[Machine Knowledge: Creation and Curation of Comprehensive Knowledge Bases]
FnT 2021

Resulting challenges

1. Available KBs are incomplete

$$\mathbf{K^a} << \mathbf{K^i}$$

2. Available KBs hardly store negatives

$$\mathbf{K^{a^-}} \approx \emptyset$$

Formal semantics for incomplete KBs: Closed vs. open-world assumption

won	
name	**award**
Brad Pitt	Oscar
Marie Curie	Nobel Prize
Berners-Lee	Turing Award

	Closed-world assumption	**Open-world assumption**
won(BradPitt, Oscar)?	→ *Yes*	→ *Yes*
won(Pitt, Nobel Prize)?	→ *No*	→ ***Maybe***

- Databases traditionally employ closed-world assumption
- KBs (semantic web) necessarily operate under open-world assumption

Open-world assumption

me of Thrones directed by Shakespeare?

World-aware AI?
Practically useful paradigm?

- Q: *Trump broth*

  KB: *Maybe*

The logicians way out – completeness metadata

- Need power to express
  both maybe and no
  (*Some paradigm which allows both open- and closed-world interpretation of data to co-exist)*

- Approach: Completeness assertions [Motro 1989]

won	
name	**award**
Brad Pitt	Oscar
Marie Curie	Nobel Prize
Berners-Lee	Turing Award

Completeness assertion:
`wonAward is complete for Nobel Prizes`

won(Pitt, Oscar)? → *Yes*

won(Pitt, Nobel)? → *No* (CWA)

won(Pitt, Turing)? → *Maybe* (OWA)

The power of completeness metadata

Know what the KB knows:

→ Locally, $K^a = K^i$

Absent assertions are really false:

→ Locally, $s \neg\in K^a$ implies $s \neg\in K^i$

Completeness metadata: Formal view

Complete (won(name, award); award = ‘Nobel’)

Implies constraint on possible state of $\mathbf{K}^a$ and $\mathbf{K}^i$

$$won^i(name, \text{‘Nobel’}) \rightarrow won^a(name, \text{‘Nobel’})$$

(tuple-generating dependency)

Darari et al.
Completeness Statements about RDF Data Sources and Their Use for Query Answering
ISWC 2013

Cardinality assertions: Formal view

- *"Nobel prize was awarded 603 times"*

→ $|won^i(name, \text{'Nobel'})| = 603$

→ Allows counting objects in $\mathbf{K}^a$
  - Equivalent count → Completeness assertion
  - Otherwise, fractional coverage/recall information
    - *"93% of awards covered"*

- Grounded in number restrictions/role restrictions in Description Logics

B. Hollunder and F. Baader
[Qualifying Number Restrictions in Concept Languages]
KR 1991

Formal reasoning with completeness metadata

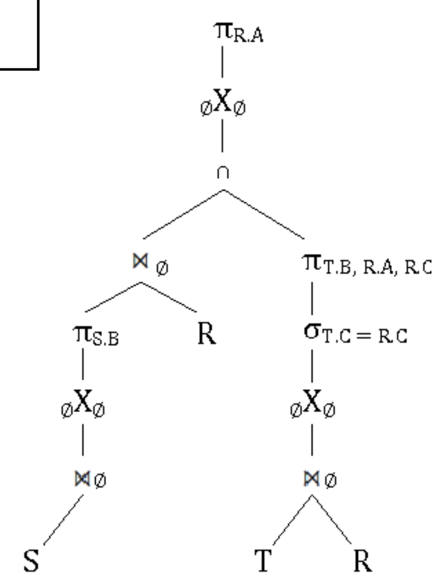
Problem: **Query completeness reasoning**

Input:

- Set of completeness assertions for base relations
- Query Q

Task:

- Compute completeness assertions that hold for result of Q



Formal reasoning with completeness metadata

Work	Description Language	Results
Motro, TODS 1989	Views	Algorithm
Fan & Geerts, PODS 2009	Various query languages (CQ-Datalog)	Decidability/ Complexity
Razniewski & Nutt 2011	Join queries	Complexity
Lang et al., SIGMOD 2014	Selections	Algorithm
Razniewski et al., SIGMOD 2016	Selections	Algorithm, computational completeness

Where can completeness metadata come from?

- Data creators should pass them along as metadata
- Or editors should add them in curation steps

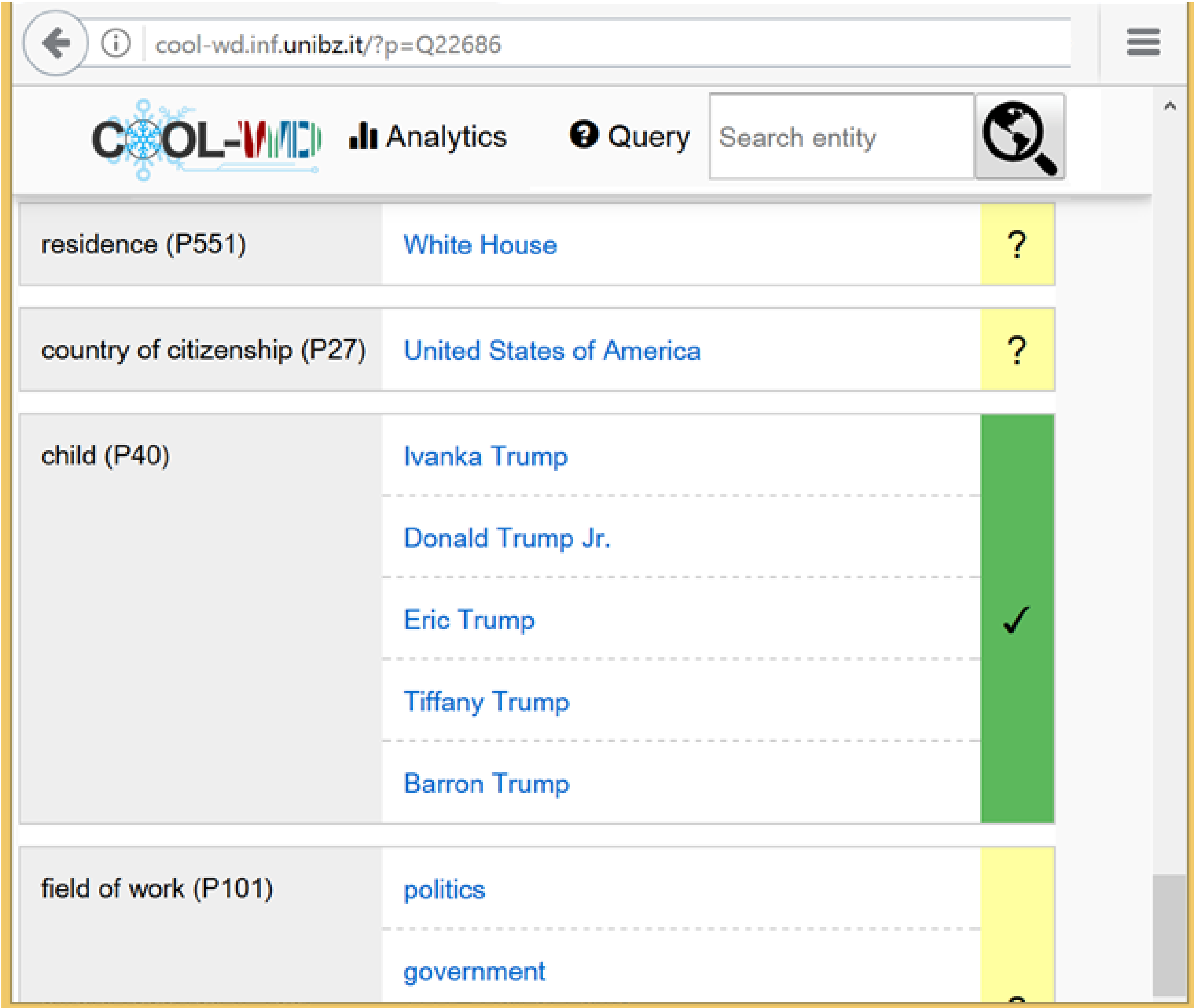
Abingdon	4. Residential triangle, Longmead etc.		Pub is only restaurant? Footways that link stuff, stubbed in places.
Shippon	5. Whole village, minus the barracks		Mostly done here.

This is a complete **list of compositions by Maurice Ravel**,

28	*Tout est lumière*	soprano, mixed choir, and orchestra	1901	• Prix de Rome competition
29	*Myrrha*, cantata	soprano, tenor, baritone, and orchestra	1901	text: Fernand Beissier; • Prix de Rome competition
31	*Semiramis*	cantata	1902	• student competition; • partially lost

- E.g., COOL-WD tool

Darari et al.
COOL-WD: A Completeness Tool for Wikidata
ISWC 2017


cool-wd.inf.unibz.it/?p=Q22686
COOL-WD
Analytics
Query
Search entity
residence (P551)
White House
?
country of citizenship (P27)
United States of America
?
child (P40)
Ivanka Trump
Donald Trump Jr.
Eric Trump
✓
Tiffany Trump
Barron Trump
field of work (P101)
politics
government

But…

- Requires human effort
  - Soliciting metadata more demanding than data
  - Automatically created KBs do not even have editors

Remainder of this tutorial:

How to automatically acquire information about what a KB knows

Takeaway Part 1: Foundations

- KBs are pragmatic collections of knowledge
  - Issue 1: Inherently incomplete
  - Issue 2: Hardly store negative knowledge
- Open-world assumption (OWA) as formal interpretation leads to counterintuitive results
- Metadata about completeness or counts as way out

Next: How to use predictive models to derive completeness metadata

On the Limits of Machine Knowledge: Completeness, Recall and Negation in Web-scale Knowledge Bases

Simon Razniewski, Hiba Arnaout, Shrestha Ghosh, Fabian Suchanek

1. Introduction & Foundations (Simon) – 9:00-9:30 CEST
2. Predictive recall assessment (Fabian) – 9:30-10:10
3. Counts from text and KB (Shrestha) – 10:10-10:45
4. Negation (Hiba) – 10:45-11:25
5. Relative completeness & Wrap-up (Simon) – 11:25-12:00

max planck institut
informatik

TELECOM ParisTech

Relative completeness

- So far: “Gold” yardstick reality
  - *Do we have all the Turing award winners?*
  - *How many of the 923 Nobel prize winners do we have?*

- Now: Pragmatic “silver” yardsticks
  1. How much textual information does the KB cover?
  2. How well are entities covered relative to others?
  3. How well does the KB support queries?

How much textual information does the KB cover?

- Yardstick: Descriptive text

→How much of text is covered in KB?

Augusta Ada King, Countess of Lovelace (*née* **Byron**; 10 December 1815 – 27 November 1852) was a British mathematician, known for her work on Charles Babbage's proposed mechanical general-purpose computer, the Analytical Engine.

Ada Lovelace (Q7259)

instance of	human
country of citizenship	United Kingdom
significant person	Charles Babbage

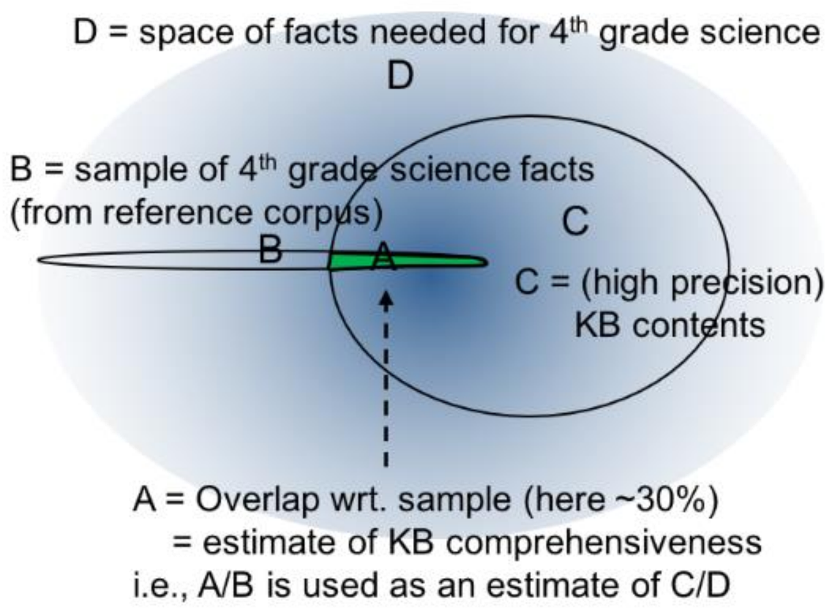
Challenge: How to measure?

Simple science texts

Reference corpus:

~1.2M sentences comprising elementary science textbooks, dictionary definitions of all fourth grade vocabulary words, simple Wikipedia pages for all fourth grade vocabulary words

→ 4000 extracted sample triples (B)



Dalvi, B., Tandon, N., & Clark, P. (2017). [Domain-Targeted, High Precision Knowledge Extraction](). TACL.

Simple science texts (2)

KB	Precision	Coverage of Tuple-Expressible Science Knowledge (Recall on science KB)
WebChild	89%	3.4%
NELL	85%	0.1%
ConceptNet	40%	8.4%
ReVerb-15M	55%	11.5%
Our KB	81%	23.2%

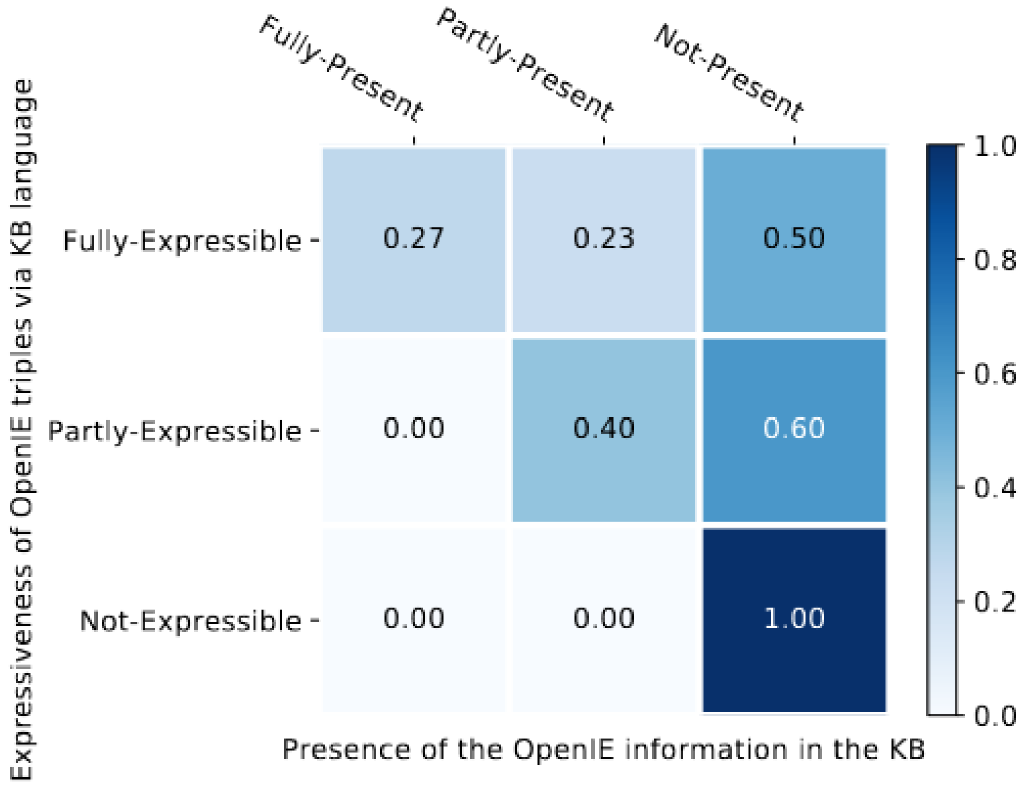
Wikipedia pages

Triples from Wikipedia pages (OPIEC corpus)

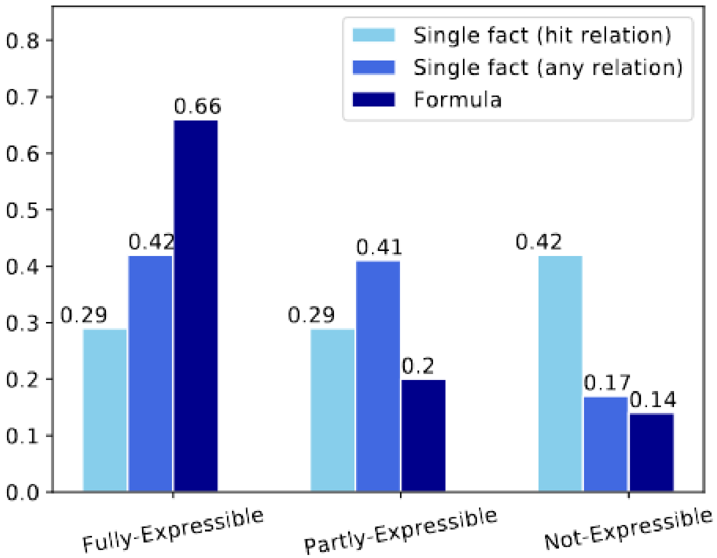
- subcorpus of 5.8M assertions where both arguments are disambiguated

K. Gashteovski et al.
[On Aligning OpenIE Extractions with Knowledge Bases: A Case Study]
Eval4NLP 2020

Wikipedia pages (2)



Schema recall



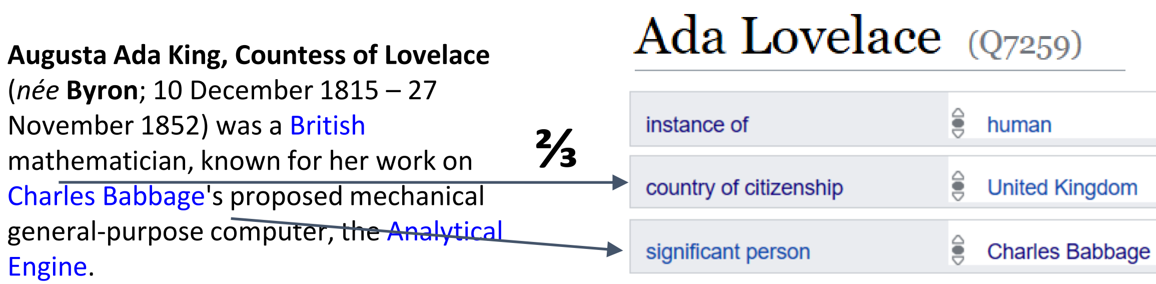
Can an OPIEC triple be expressed in DBpedia?

#	OIE triple	KB formula
t_1	Temporal annotation	
	(Coral Fang; "was released by"; Sire Records) *Time: (in, 2003)*	(Coral Fang; dbo:recordLabel; Sire Records) $\wedge$ (Coral Fang; dbo:releaseDate; 2003)
t_2	Complex formula	
	(Garrett Davis; "was Rep. from"; Kentucky)	(G. D.; dbo:profession; State representative) $\wedge$ [(G. D.; dbo:region; K.) $\vee$ (G. D.; dbo:state; K.)]
t_3	Existential quantification	
	(Franz Liszt; "transcribed piece for"; Piano solo)	$\exists x$: (F. L.; dbo:write; x) $\wedge$ (x; dbo:genre; P. solo)

Temporal development

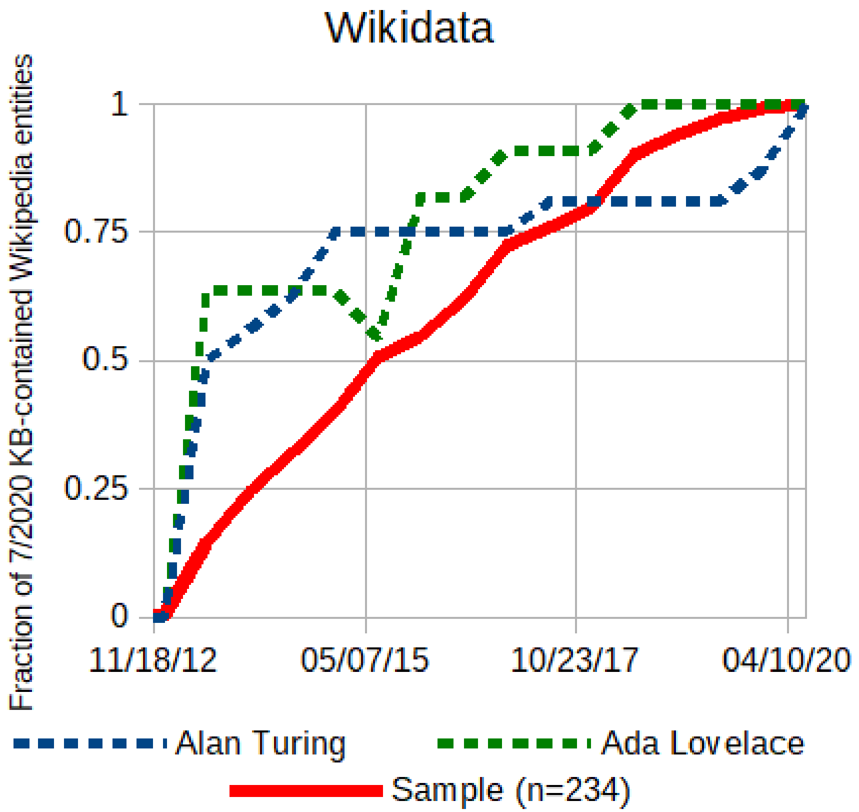
How does KB coverage change over time?

Criterion*:*
How many of the hyperlinked entities in a Wikipedia article occur also as objects in the entity's Wikidata article



Razniewski and Das
[Structured knowledge: Have we made progress?](#)
[An empirical study of KB coverage over 19 years](#)
CIKM 2020

Temporal development (2)



→ KBs get better
→ Absolute coverage still low (5-10%)

Relative completeness

- So far: Yardstick was reality
  - *Do we have all the Turing award winners?*
  - *How many of the 923 Nobel prize winners do we have?*

- Now: Pragmatic “silver” yardsticks
  1. How much textual information does the KB cover?
  2. How well are entities covered relative to others?
  3. How well does the KB support queries?

Entity comparison - idea

 Coverage(Wikidata for Putin)?

 There are more than 3000 properties one can assign to Putin…

 How well is data about him compared to others?

 Compared to whom?

Entity comparison – idea (2)

Quantify based on comparison with other similar entities

Ingredients:

- Similarity metric — *Who is similar to Trump?*
- Data quantification — *How much data is good/bad?*

- Deployed in Wikidata as Relative Completeness (Recoin)

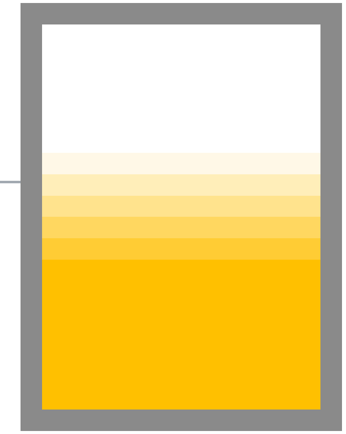
Balaraman et al. *Recoin: Relative Completeness in Wikidata*, **Wiki Workshop 2018**

Edsger W. Dijkstra (Q8556)

Dutch computer scientist

sex or gender	male ▸ 5 references
country of citizenship	Kingdom of the Netherlands ▸ 2 references
name in native language	Edsger Wybe Dijkstra (Dutch) ▸ 1 reference
given name	Edsger

Edsger W. Dijkstra (Q8556)



Dutch computer scientist

▾ Recoin: Most relevant properties which are absent

Property ID	Label	Relative	Add Claim
P937	work location	13.4%	+
P1343	described by source	11.6%	+
P512	academic degree	10.04%	+
P39	position held	9.06%	+
P102	member of political party	6.24%	+
P856	official website	5.46%	
P140	religion	4.13%	+
P22	father	3.48%	+
P551	residence	3.25%	+
P40	child	3.22%	+

Advanced property ranking methods

- Razniewski, Simon, et al. "Doctoral advisor or medical condition: Towards entity-specific rankings of knowledge base properties." *ADMA*, 2017.
- Gleim, Lars C., et al. "SchemaTree: Maximum-Likelihood Property Recommendation for Wikidata." *ESWC*, 2020.
- Luggen, Michael, et al. "Wiki2Prop: A Multimodal Approach for Predicting Wikidata Properties from Wikipedia." *The Web Conference*. 2021.

Relative completeness

- So far: Yardstick was reality
  - *Do we have all the Turing award winners?*
  - *How many of the 923 Nobel prize winners do we have?*

- Now: Pragmatic “silver” yardsticks
  1. How much textual information does the KB cover?
  2. How well are entities covered relative to others?
  3. How well does the KB support queries?

How well does the KB support queries?

- Is the KB ~~complete?~~

  useful

Usage = querying

Demand-weighted completeness prediction

- Alexa query logs
- Queries: Entity+property→value lookups

- Peering of entities via classes
- Predict property-query frequency of novel entities by interpolating from peers

```
baracKObama:
    hasHeight:      0.16
    hasBirthdate:   0.12
    hasBirthplace:  0.08
    hasSpouse:      0.07
    hasChild:       0.05
```

- **Task:** Given an entity E in a KB, and query usage data of the KB, predict the distribution of relations that E must have in order for 95% of queries about E to be answered successfully

Hopkinson et al.
[Demand-Weighted Completeness Prediction for a Knowledge Base](#)
NAACL 2018

Demand-weighted completeness prediction (2)

- Neural models can successfully predict query loads
- Query loads can be used
  to assess demand-weighted completeness

→ Unseen sample:
58% complete w.r.t. 95% query goal

Temporal development

300 random questions filtered from large search query logs

- AOL query log
- Bing query log
- Google query suggestion

Human annotator

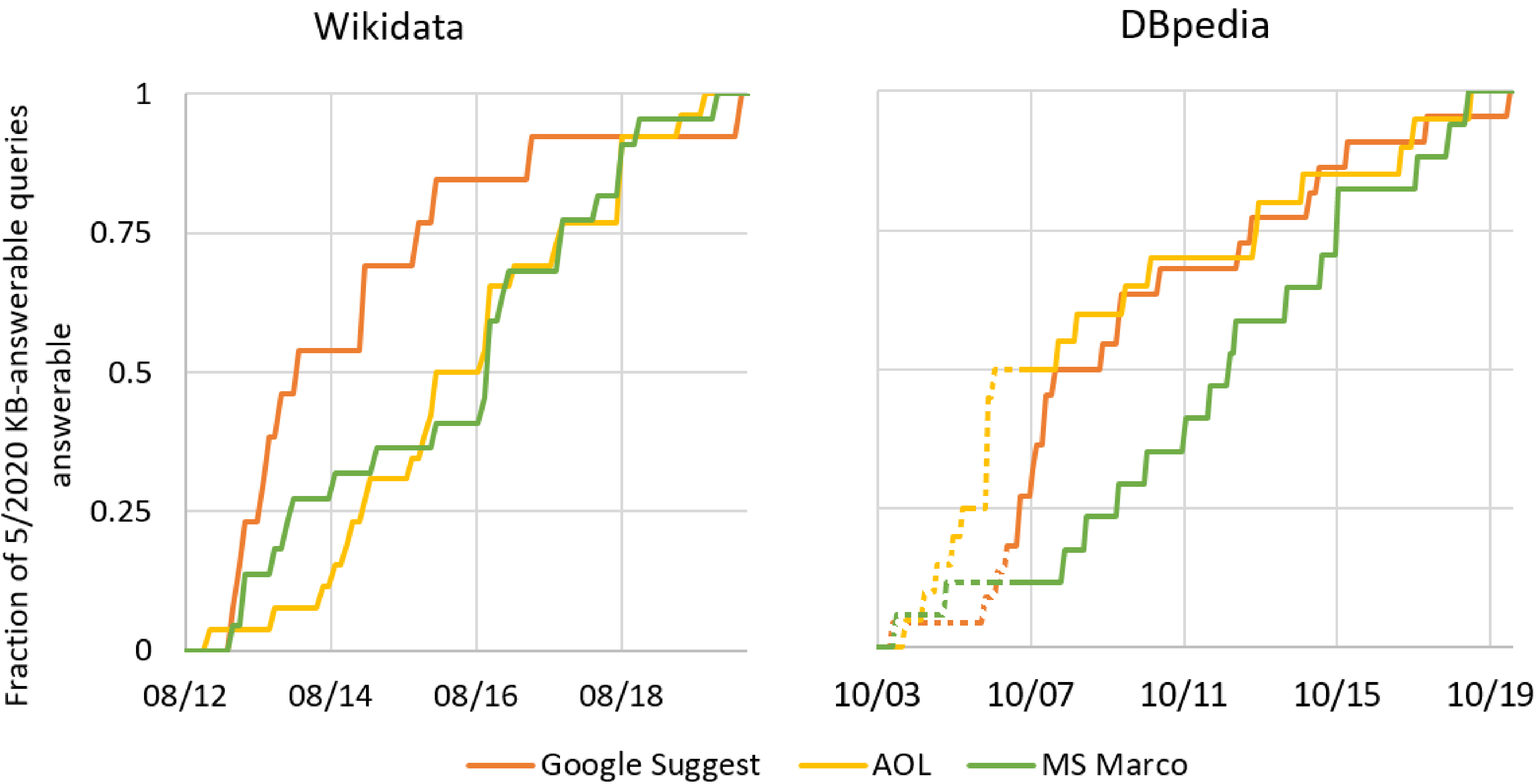
- Task: Can a KB answer a query, and if so, since when?

Query	First answerable
how old is dustin pedroia	May 18, 2017
where is italian job filmed	October 15, 2015
what type of government is ontario	April 22, 2020
what time zone is ohio in	August 31, 2013

Examples Google Suggest/Wikidata

Razniewski and Das
Structured knowledge: Have we made progress?
An empirical study of KB coverage over 19 years
CIKM 2020

Temporal development (2)



Takeaway Part 5: Relative recall

- Real-world recall not always measurable and/or relevant
- Alternative yardsticks:
  - Text
  - Related entities
  - Usage data (query logs)
- Logical next step: Cost/benefit priorization

On the Limits of Machine Knowledge: Completeness, Recall and Negation in Web-scale Knowledge Bases

Simon Razniewski, Hiba Arnaout, Shrestha Ghosh, Fabian Suchanek

1. Introduction & Foundations (Simon) – 9:00-9:30 CEST
2. Predictive recall assessment (Fabian) – 9:30-10:10
3. Counts from text and KB (Shrestha) – 10:10-10:45
4. Negation (Hiba) – 10:45-11:25
5. Relative completeness & Wrap-up (Simon) – 11:25-12:00

Wrap-up: Take-aways

1. KBs are incomplete and limited on the negative side

2. Predictive techniques work
   from a surprising set of paradigms

3. Count information a prime way
   to gain insights into completeness/coverage

4. Salient negations can be heuristically materialized

Wrap-up: Recipes

- Ab-initio KB construction
    1. Intertwine data and metadata collection
    2. Human insertion: Provide tools
    3. Automated extraction: Learn from extraction context

- KB curation
    1. Exploit KB-internal or textual cardinality assertions
    2. Inspect statistical properties on density or distribution
    3. Compute overlaps on pseudo-random samples

Open research questions

1. How are entity, property and fact completeness related?

2. How to distinguish salient negations from data modelling issues?

3. How to estimate coverage of knowledge in pre-trained language models?

4. How to identify most valuable areas for recall improvement?

Wrap-up: Wrap-up

- KBs major drivers of knowledge-intensive applications
- Severe limitations concerning completeness and coverage-awareness
- This tutorial: Overview of problem, techniques and tools to obtain awareness of completeness

Takeaway Part 1: Foundations

- KBs are pragmatic collections of knowledge
  - Issue 1: Inherently incomplete
  - Issue 2: Hardly store negative knowledge
- Open-world assumption (OWA) as formal interpretation leads to counterintuitive results
- Metadata about completeness or counts as way out

https://www.mpi-inf.mpg.de/kr-2021-tutorial

Takeaway: Predictive recall assessment

Using statistical techniques, we can predict more or less

- the recall of facts
  - are we missing objects for a subject?
  - do all subjects have an attribute in the real world?
  - does a text enumerate all objects for a subject?
- the recall of entities
  - is the distribution of entities representative?
  - how many entities are in the real world?

Takeaway: Counts from text and KB

1. Count information comes in two variants
   - Counting predicates - store integer counts
   - Enumerating predicates - store entities
2. Count information in text
   - occurs as cardinals, ordinals, non-numeric noun phrases
   - occurs with compositional cues
3. Count information in KBs
   - is expressed in two variants
   - occurs semantically related count predicates
4. Count information
   - can enrich KB
   - highlight inconsistencies

Takeaway: negation

- **Current KBs lack negative knowledge**
- **Rising interest in the explicit addition of negation to OW KB.**
- **Negations highly relevant in many applications including:**
  - **Commercial decision making (e.g., hotel booking)**
  - **General-domain question answering systems (e.g., is Switzerland a member of the EU?)**
- **Methodologies include:**
  - **Statistical inference**
  - **Text extraction**
  - **Pretrained LMs.**