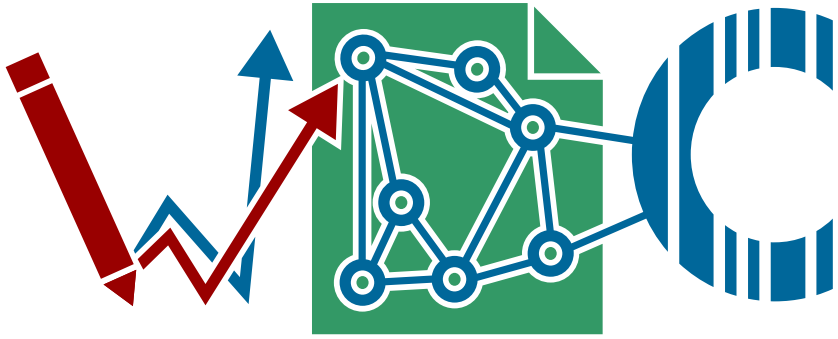


WIKI DATA CON

Wikidata & languages 2019

Barriers to Using Wikidata as a Knowledge Base

Peter F. Patel-Schneider

WikidataCon 2019, 25–26 October, Berlin

Barriers to Using Wikidata as a Knowledge Base

Abstract:
At first glance Wikidata appears to be an excellent source of general background knowledge. It is large, contains generally high-quality information, contains a large ontology, and can be freely used. However, it is difficult to use Wikidata as a knowledge base of general-purpose information. The worst barriers to using Wikidata involve its ontology and qualifiers. More care in building and describing the ontology and more uniformity in modeling would go a long way to make Wikidata more useful as a source of knowledge.

Who am I?

WIKI DATA CON
Wikidata & languages 2019

- A long-time researcher in Knowledge Representation
  - Working on Description Logics: the *logic* underlying OWL
  - Designed much of the W3C OWL Web Ontology Language
- Worker in Silicon Valley
  - Building large general-purpose knowledge store
    - Knowledge Graph, if you like, but better
  - Put together from open-source knowledge bases
    - General-purpose—like DBpedia, Freebase
    - Special-purpose—like Open Street Map, Open Movie Database
  - Using knowledge store in other systems

My Current Use of Wikidata

WIKI DATA CON
Wikidata & languages 2019

- Main source for general-purpose background knowledge
- Main source for organization of knowledge
- Supporting a conversational agent
  - Handle definite descriptions
    - What's that movie directed by the wife of the director of Avatar?
    - Who's the woman director that was married to the director of Titanic?
  - Direct user utterances to appropriate back-end
    - Which back-end can find train trips from San Francisco to Palo Alto

I regularly run into problems when trying to use Wikidata this way.
I have suggestions on how to improve Wikidata for these purposes.

Suggestions (Summary)

WIKI DATA CON
Wikidata & languages 2019

- Community: Better documentation and modelling. More uniformity. Expend more effort at the beginning and less later.
- Formal: Design a formal logic for Wikidata, including time and qualifiers.
- Tools: Build tools that show implications, contradictions, empty classes, guides for users.
- Bots: Require rerun and repair capability for bots.

These are, of course, only my suggestions.
Others experiencing similar problems may have different suggestions.

Suggestions (Formal)

WIKI DATA CON
Wikidata & languages 2019

- Create a formal logic for Wikidata, and keep it up to date
  - Include disjointness, relative ranges
  - Include class definitions
  - Include unusual existing Wikidata constructs, including constraints
  - Include qualifiers, particularly temporal qualifiers
- Having a logic means that implications of claims can be determined
  - Is an instance of a subclass an instance of the superclass?
  - What is true at a particular point in time.

Suggestions (Tools)

WIKI DATA CON
Wikidata & languages 2019

- Tools that show implications of claims, according to current logic
- Tools that show claims relevant at a particular time (or other context)
- Tools that show contradictions and necessarily empty classes
  - Similar to what constraint checkers do, but (probably) more principled

Suggestions (Community)

WIKI DATA CON
Wikidata & languages 2019

- Add information to fill out domains, but carefully
- Value curation as much or more than adding information
- Don't use alternative sources when better source is available
- Partner with efforts that have complete information in a domain
- Describe classes better
- Add disjointness statements between classes
- Take more care in adding ontological claims
- Model similar entities and situations the same way
- Reign in bots—require more care, rerun capability

Examining the Problems of Wikidata (as a Source of General-Purpose Knowledge)

WIKI DATA CON
Wikidata & languages 2019

- All my analysis is currently anecdotal, but as a user of Wikidata as a Knowledge Base I regularly hit problems.
  - I connect information together and make (simple) inferences that expose problems.
  - I combine information from different areas that model differently.
  - I use lots of information—much to much to curate carefully.
- Wikidata is GREAT! It's useful in *many* ways
  - Wikidata is improving—errors are being fixed.
- Wikidata is not be great enough
  - Using Wikidata as a source of knowledge requires a lot of effort
    - Too much effort for small players?
- We need to make Wikidata better
  - Fix the problems in Wikidata as a source of general-purpose knowledge
  - Put in place processes to reduce future problems

NB: A comprehensive study to quantify the problems is warranted.

Categorizing Observed Problems of Wikidata (as a Source of General-Purpose Knowledge)

WIKI DATA CON
Wikidata & languages 2019

- Missing Information
- Factual Errors **(caused by bots)**
- **Ontology Problems**
- **Qualifier Problems**

Missing Information

WIKI DATA CON
Wikidata & languages 2019

- Wikidata is incomplete. (Surprise!)
- This is (partly) by design. See the notability policy.
- Wikidata is particularly lacking in areas that are not *things*, e.g., actions.

Missing Information

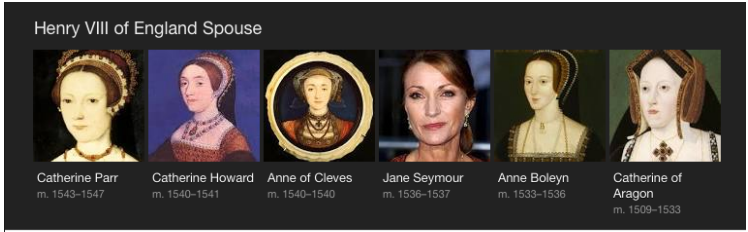
- Wikidata is incomplete. (Surprise!)
- This is (partly) by design. See the notability policy.
- Wikidata is particularly lacking in areas that are not *things*, e.g., actions.

Suggestions:

- Add more information to Wikidata (duh!), but carefully.
- Partner with groups using Wikibase for specialized knowledge.

Factual Errors

- Factual errors are a problem everywhere.



- Wikidata included
  - Some aspects of the Wikidata ecosystem make errors even more problematic

An (Oldish) Wikidata Factual Error

Corpse Bride (Q34417)

Revision as of 18:41, 23 August 2017 by Sanechthon (talk | contribs) (*Created claim: Property:P3077: 158222*)
(diff) ← Older revision | Latest revision (diff) | Newer revision → (diff)

2005 British-American stop-motion-animated fantasy film
Tim Burton's Corpse Bride

In more languages Configure

Language	Label	Description	Also known as
English	Corpse Bride	2005 British-American stop-motion-animated fantasy film	Tim Burton's Corpse Bride
Spanish	Corpse Bride	No description defined	
Traditional Chinese	No label defined	No description defined	
Chinese	怪誕屍新娘	No description defined	

All entered languages

Statements

instance of: stop-motion animated film
1 reference
imported from Wikimedia project: English Wikipedia

director: Tim Burton
10 references

director: Mike Johnson
7 references

director: Mike Johnson
7 references

Mike Johnson (Q6847408)

stop motion animator

edit

In more languages Configure

Language	Label	Description	Also known as
English	Mike Johnson	stop motion animator	
Spanish	Mike Johnson	No description defined	
Traditional Chinese	No label defined	No description defined	
Chinese	No label defined	No description defined	

All entered languages

Statements

instance of: human
1 reference
edit
+ add value

Mike Johnson (Q1137104)

American singer-songwriter, guitarist and bassist

edit

In more languages Configure

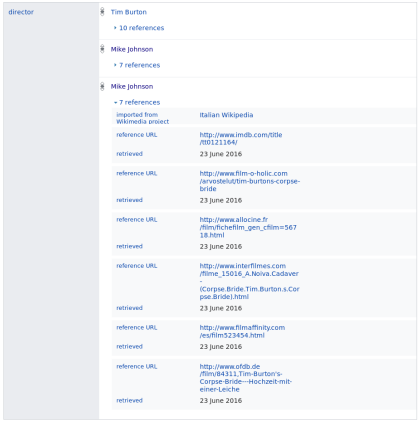
Language	Label	Description	Also known as
English	Mike Johnson	American singer-songwriter, guitarist and bassist	
Spanish	Mike Johnson	No description defined	
Traditional Chinese	No label defined	No description defined	
Chinese	No label defined	No description defined	

All entered languages

Statements

instance of: human
1 reference
edit
+ add value

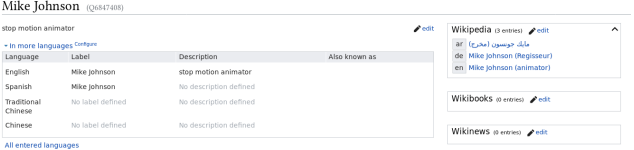
An (Oldish) Wikidata Factual Error



director

- Tim Burton
  - 10 references
- Mike Johnson
  - 7 references
- Mike Johnson
  - 7 references

imported from Wikimedia project	Italian Wikipedia
reference URL	http://www.imdb.com/title/tt0121164/
retrieved	23 June 2016
reference URL	http://www.film-o-holic.com/arvostelut/tim-burtons-corpse-bride
retrieved	23 June 2016
reference URL	http://www.allocine.fr/film/fichefilm_gen_cfilm=56718.html
retrieved	23 June 2016
reference URL	http://www.interfilmes.com/filme_15016_A.Noiva.Cadaver-(Corpse.Bride.Tim.Burton.s.Corpse.Bride).html
retrieved	23 June 2016
reference URL	http://www.filmaffinity.com/es/film523454.html
retrieved	23 June 2016
reference URL	http://www.ofdb.de/film/84311,Tim-Burton's-Corpse-Bride---Hochzeit-mit-einer-Leiche
retrieved	23 June 2016



Mike Johnson (Q6847408)

stop motion animator — edit

In more languages Configure

Language	Label	Description	Also known as
English	Mike Johnson	stop motion animator	
Spanish	Mike Johnson	No description defined	
Traditional Chinese	No label defined	No description defined	
Chinese	No label defined	No description defined	

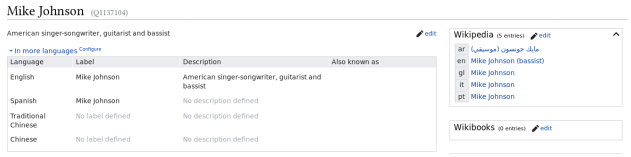
All entered languages

Wikipedia (3 entries) edit

ar	مايك جونسون (مخرج)
de	Mike Johnson (Regisseur)
en	Mike Johnson (animator)

Wikibooks (0 entries) edit

Wikinews (0 entries) edit



Mike Johnson (Q1137104)

American singer-songwriter, guitarist and bassist — edit

In more languages Configure

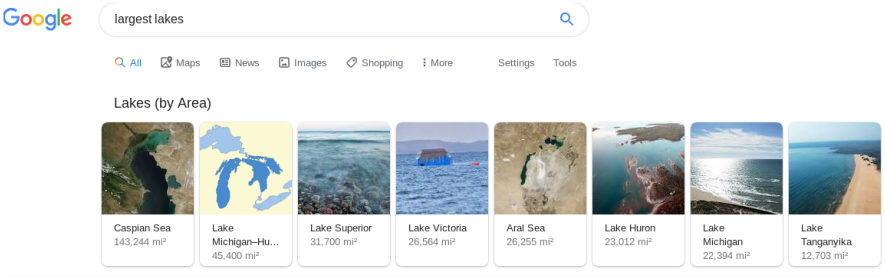
Language	Label	Description	Also known as
English	Mike Johnson	American singer-songwriter, guitarist and bassist	
Spanish	Mike Johnson	No description defined	
Traditional Chinese	No label defined	No description defined	
Chinese	No label defined	No description defined	

Wikipedia (5 entries) edit

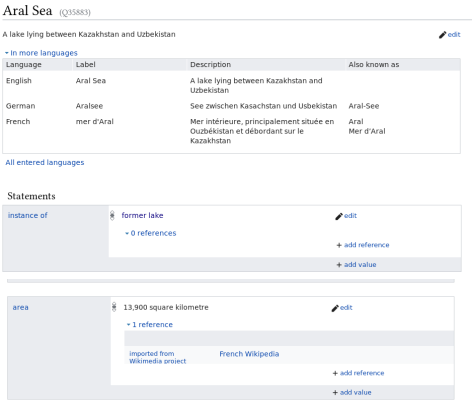
ar	مايك جونسون (موسيقي)
en	Mike Johnson (bassist)
gl	Mike Johnson
it	Mike Johnson
pt	Mike Johnson

Wikibooks (0 entries) edit

Factual Errors—Stale Information



Factual Errors—Stale Information



Aral Sea (Q35883)

A lake lying between Kazakhstan and Uzbekistan edit

in more languages

Language	Label	Description	Also known as
English	Aral Sea	A lake lying between Kazakhstan and Uzbekistan	
German	Aralsee	See zwischen Kasachstan und Usbekistan	Aral-See
French	mer d'Aral	Mer intérieure, principalement située en Ouzbékistan et débordant sur le Kazakhstan	Aral Mer d'Aral

All entered languages

Statements

instance of — former lake — edit

0 references

+ add reference

+ add value

area — 13,900 square kilometre — edit

1 reference

imported from Wikimedia project: French Wikipedia

+ add reference

+ add value

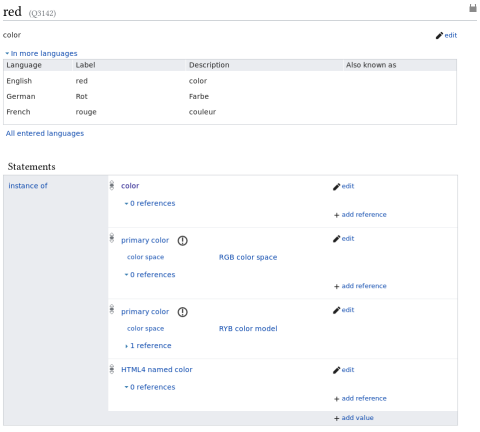
Coordonnées	45° nord, 60° est
Type	Endoréique
Origine	Naturel
Superficie · **Maximale** · **Minimale**	8 303 km² (août 2015) 68 900 km² (1960) 7 297 km² (septembre 2014)[1]
Longueur	428 km (1960)
Largeur	284 km (1960)
Altitude	53,4 m (1960) 42 m (nord) en 2007 29 m (sud) en 2007
Profondeur · **Maximale**	68 m

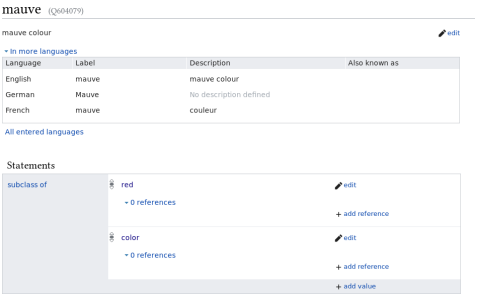
Factual Errors—Suggestions

WIKI DATA CON
Wikidata & languages 2019

- Bots should be careful when adding information, particularly when using names or adding references.
  - Even well-written bots can make a lot of errors if their assumptions are or become invalid.
- Don't use information from alternative sources when a better source has the information.
- Require that bots can be fixed and rerun, removing erroneous information they added.
- Require that bots be rerun, removing information no longer found in source.
- Build a community that values fixing problems

Ontology Problems—Non-uniform Modeling





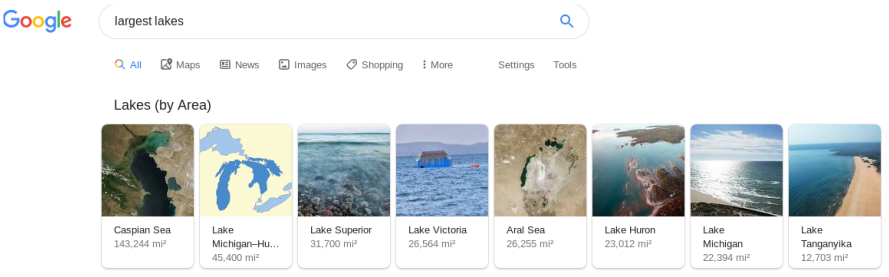
- red instance of color, mauve subclass of color and a subclass of one of its instances
- Doesn't make any sense—what are colors and what are not colors?

Ontology Problems—Non-uniform Modeling

wd:Q35637	Nobel Peace Prize	part of	Nobel Prize
wd:Q44585	Nobel Prize in Chemistry	instance of	Nobel Prize
wd:Q44585	Nobel Prize in Chemistry	part of	Nobel Prize
wd:Q47170	Nobel Prize in Economics	instance of	Nobel Prize
wd:Q47170	Nobel Prize in Economics	part of	Nobel Prize
wd:Q37922	Nobel Prize in Literature	subclass of	Nobel Prize
wd:Q38104	Nobel Prize in Physics	subclass of	Nobel Prize
wd:Q38104	Nobel Prize in Physics	part of	Nobel Prize
wd:Q80061	Nobel Prize in Physiology or Medicine	subclass of	Nobel Prize
wd:Q10751662	Нобелевская премия 2008	instance of	Nobel Prize

Thanks to Aidan Hogan for bringing this up.

Ontology Problems—Questionable Instances of Classes



Where did this “extra” Great Lake come from?

Ontology Problems—Great Lakes

lake (Q23397)

body of relatively still water, localized in a basin

edit

▾ In more languages

Language	Label	Description	Also known as
English	lake	body of relatively still water, localized in a basin	
German	See	von Land umgebenes Stillgewässer mit oder ohne Zu- und Abfluss	Landsee Binnensee
French	lac	étendue d'eau située à l'intérieur de terres	

All entered languages

Statements

subclass of

still waters — edit

▾ 0 references

+ add reference

landform — edit

▾ 0 references

+ add reference

+ add value

image

edit

Bariloche- Argentina2.jpg
800 × 600; 165 KB

▾ 0 references

+ add reference

+ add value

Lake Michigan–Huron (Q1377732)

one of the Great Lakes of North America
Lake Michigan-Huron | Michigan-Huron | Michigan-Huron

edit

▾ In more languages

Language	Label	Description	Also known as
English	Lake Michigan-Huron	one of the Great Lakes of North America	Lake Michigan-Huron Michigan-Huron Michigan-Huron
German	No label defined	See in Kanada	
French	lac Michigan-Huron	lac américain	

All entered languages

Statements

instance of

lake — edit

▸ 1 reference

+ add value

part of

Great Lakes — edit

▾ 0 references

+ add reference

+ add value

image

edit

Great Lakes Lake Michigan-Huron.png
581 × 406; 7 KB

▸ 1 reference

Ontology Problems—Great Lakes

Lake Michigan (Q1169)

one of the Great Lakes of North America

Michigan

edit

▸ In more languages

Language	Label	Description	Also known as
English	Lake Michigan	one of the Great Lakes of North America	Michigan
German	Michigansee	Großer See in Nordamerika	
French	lac Michigan	lac d'Amérique du Nord	

All entered languages

Statements

instance of	lake ▸ 2 references	edit
	+ add value	
part of	Lake Michigan–Huron ▾ 0 references + add reference	edit
	Great Lakes ▾ 0 references + add reference	edit
	+ add value	
image	 Great Lakes Lake Michigan.png 581 × 406; 10 KB	edit

Lake Huron (Q1383)

one of the Great Lakes of North America

Huron

edit

▸ In more languages

Language	Label	Description	Also known as
English	Lake Huron	one of the Great Lakes of North America	Huron
German	Huronsee	einer der 5 Großen Seen in Nordamerika	
French	lac Huron	lac d'Amérique du Nord	

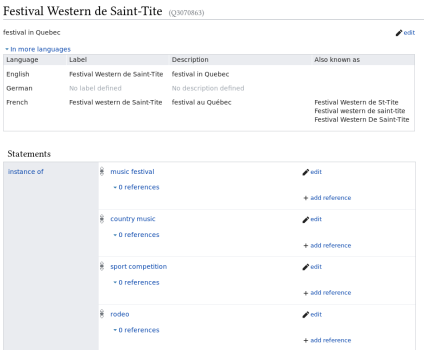
All entered languages

Statements

instance of	lake ▸ 1 reference	edit
	+ add value	
part of	Lake Michigan–Huron ▾ 0 references + add reference	edit
	Great Lakes ▾ 0 references + add reference	edit
	+ add value	
image	 Lake Huron NASA 2011.jpg 1,617 × 1,573; 885 KB ▾ 0 references + add reference	edit

Ontology Problems—Vague Classes

- action (Q4026292)—something an agent can do or perform



Festival Western de Saint-Tite (Q3070863)

festival in Quebec

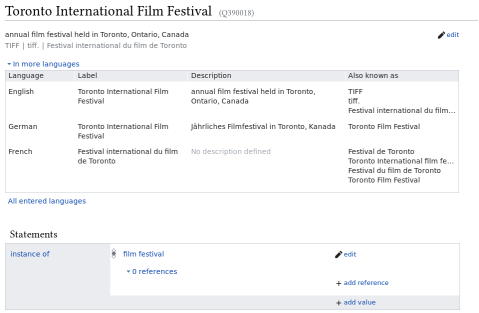
edit

In more languages

Language	Label	Description	Also known as
English	Festival Western de Saint-Tite	festival in Quebec	
German	No label defined	No description defined	
French	Festival western de Saint-Tite	festival au Québec	Festival Western de St-Tite Festival western de saint-tite Festival Western De Saint-Tite

Statements

instance of	music festival 0 references	edit add reference
	country music 0 references	edit add reference
	sport competition 0 references	edit add reference
	rodeo 0 references	edit add reference



Toronto International Film Festival (Q390018)

annual film festival held in Toronto, Ontario, Canada

TIFF | tiff. | Festival international du film de Toronto

edit

In more languages

Language	Label	Description	Also known as
English	Toronto International Film Festival	annual film festival held in Toronto, Ontario, Canada	TIFF tiff. Festival international du film...
German	Toronto International Film Festival	jährliches Filmfestival in Toronto, Kanada	Toronto Film Festival
French	Festival international du film de Toronto	No description defined	Festival de Toronto Toronto International film fe... Festival du film de Toronto Toronto Film Festival

All entered languages

Statements

instance of	film festival 0 references	edit add reference
		add value

- Festival Western de Saint-Tite is an instance of action; TIFF isn't
  - The first is a rodeo and thus a sport and thus an activity and thus an action
  - This result is at questionable
  - The difference between the two festivals is definitely a problem

Ontology Problems—Vague Similar Classes

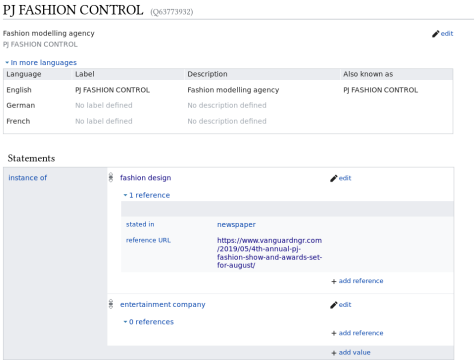
Major subclasses of
geographic entity (Q27096213) place or entity that could be a place

- geographical object (Q618123) point; object that can be uniquely identified with geographical data, may be man-made or naturally-created
- geographic object (Q1503302)
- geographic region (Q82794) 2D or 3D defined space, …
- location (Q17334923) location of physical or non-physical object in space
  - geographic location (Q2221906) point or an area on the Earth's surface or elsewhere
- position (Q23008351) physical extent or point in space …

Eiffel Tower (Q243) is instance of

- tourist attraction (Q570116) and thus geographical object (Q618123)
- tourist attraction (Q570116) and thus geographic object (Q1503302)
- landmark (Q2319498) and thus geographic region (Q82794)
- tower (Q12518) and thus location (Q17334923),
  but not geographic location (Q2221906)

Ontology Problems—Class Errors


PJ FASHION CONTROL (Q63773932)

Fashion modelling agency
PJ FASHION CONTROL

edit

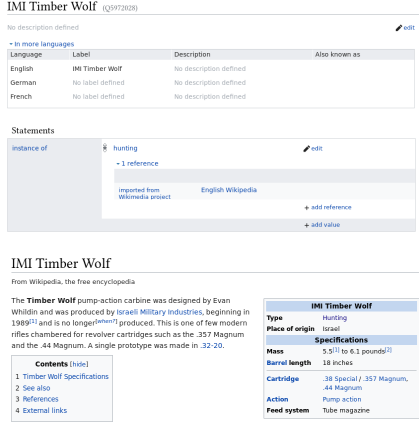
▾ In more languages

Language	Label	Description	Also known as
English	PJ FASHION CONTROL	Fashion modelling agency	PJ FASHION CONTROL
German	No label defined	No description defined	
French	No label defined	No description defined	

Statements

instance of
fashion design edit
▾ 1 reference
stated in newspaper
reference URL https://www.vanguardngr.com/2019/05/4th-annual-gj-fashion-show-and-awards-set-for-august/
+ add reference
entertainment company edit
▾ 0 references
+ add reference
+ add value

(PJ Fashion Control is a modeling agency.)


IMI Timber Wolf (Q5972028)

No description defined

edit

▾ In more languages

Language	Label	Description	Also known as
English	IMI Timber Wolf	No description defined	
German	No label defined	No description defined	
French	No label defined	No description defined	

Statements

instance of
hunting edit
▾ 1 reference
imported from Wikimedia project English Wikipedia
+ add reference
+ add value

IMI Timber Wolf

From Wikipedia, the free encyclopedia

The **Timber Wolf** pump-action carbine was designed by Evan Whildin and was produced by Israeli Military Industries, beginning in 1989[1] and is no longer[when?] produced. This is one of few modern rifles chambered for revolver cartridges such as the .357 Magnum and the .44 Magnum. A single prototype was made in .32-20.

Contents [hide]
1 Timber Wolf Specifications
2 See also
3 References
4 External links

IMI Timber Wolf	
Type	Hunting
Place of origin	Israel
Specifications	
Mass	5.5[2] to 6.1 pounds[2]
Barrel length	18 inches
Cartridge	.38 Special / .357 Magnum, .44 Magnum
Action	Pump action
Feed system	Tube magazine

- Class errors are a particular kind of error, and more problematic than most errors because the instance relationship is so important.

Ontology Problems—Questionable Subclass Chains

wheelchair tennis player (Q18814798) sportsperson who plays wheelchair tennis
⇒paralympic athlete (Q25439032) sportsperson competing at Paralympic Games
⇒profession (Q28640) vocation founded upon specialized educational training
⇒occupation (Q12737077) label applied to a person based on an activity they [do]
⇒job (Q192581) an activity done by a person to earn money
⇒work (Q16532276) an activity done by a person to earn a living
⇒human activities (Q61788060) activity initiated by a human
⇒activity (Q1914636) event; actions that result in changes of state

mason (Q327321) A craftsman who builds walls using stones and/or bricks
⇒profession (Q28640) vocation founded upon specialized educational training
⇒job (Q192581) an activity done by a person to earn money
⇒work (Q16532276) an activity done by a person to earn a living
⇒human activities (Q61788060) activity initiated by a human
⇒activity (Q1914636) event; actions that result in changes of state

The individual subclass relationships might look acceptable but their combination is not.

Ontology Problems—Missing Instances

How many women are in Wikidata? (Noticed by Valeria de Paiva.)

WIKI DATA CON
Wikidata & languages 2019

human (Q5)

human being | humankind | people | homosapiens | person

common name of Homo sapiens, unique extant species of the genus Homo

subclass of: every human is also a(n) person, omnivore, natural person
instance of: human is a(n) group of organisms known by one particular common name

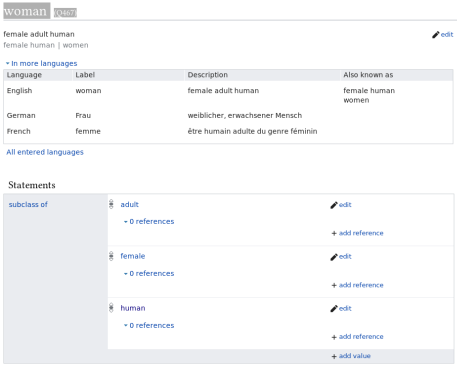
Instances

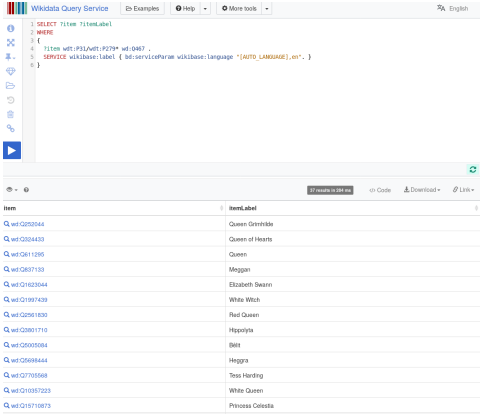
Direct instances	4 Yanka Kupala , František Plánička , Augusto Pinochet , Julian Assange , Blanche of Castile , Adolf Hitler , Piet Kraak , Louis IX of France , Patricio Aylwin , Nicolas Sarkozy , Ricardo Lagos , Eduardo Frei Ruiz-Tagle , Michelle Bachelet , Galileo Galilei , Sebastian Piñera , El Greco , Elvis Presley , Diego Adèle , Claude Monet , Paul Morand , ... further results
All instances	5447239
Typical Properties	sex or gender , occupation , date of birth , country of citizenship , given name , place of birth , family name , date of death , VIAF ID , educated at , ISNI , place of death , Library of Congress authority ID , languages spoken, written or signed , position held , employer , GND ID , ORCID iD , member of sports team , CBDB ID , name in native language , award received , SUDOC authorities ID , sport , NTA ID ,

Classification

Direct superclasses	person 5501810 , omnivore 5447256 , natural person 5447248
Direct subclasses	With instances 22 · With subclasses 44 · All 172 person of color 171 , missing person 108 , anonymous 102 , centenarian 79 , bog body 61 , woman 37 , oldest human 37 , technician 17 , spouse 14 , candidate 10 , women in STEM fields 9 , child 7 , folk saint 5 , migrant 5 , citizen 4 , special purpose artist 4 , unidentified decedent 2 , scholar 2 , acquaintance 2 , Military dependent 2 , cancer survivor 1 , neighbor 1
All subclasses	1117

Ontology Problems—Missing Instances





Ontology Problems—Suggestions

- Document the precise meaning of a class (like Cyc tries to do!)
  - Particularly what are instances and what are not.
  - Particularly what are subclasses and what are not.
  - Provide a special place in Wikidata for these definitional comments.
  - It's extra work at the beginning but a lot of later work can be saved.
- If a class only has a few instances (or subclasses, or parts), make sure that they all exist and are modelled the same way
- Require bots to check their sources for anomalous information
- Allow disjointness claims between classes, and check them when adding instances and subclasses
  - To help prevent class instance errors
  - To help prevent subclass errors
- Have formal definitions of many classes (e.g., woman) and infer their instances
  - Alternatively, prohibit instances of classes like these
  - Provide localized ranges for properties—children of humans are human

Qualifier Problems—Temporal Qualifiers

- Is a claim asserted now or does an item exist now?
- 33 (55) properties belonging to Wikidata property with datatype 'time' (Wikidata property related to time and duration)
- Which to use when?
  - start time (P580), end time (P582), point in time (P585), date of birth (P569), inception (P571), publication date (P577), date of official opening (P1619)
    - Some of these are in subproperty relationships
    - But Wikidata doesn't provide inferencing to show implications of subproperty links
- When to use qualifier and when to use temporal property
  - work period (end) (P2032) and work period (start) (P2031) vs start time (P580) and end time (P582) on occupation (P106)
- External tools can be written against the current set of temporal qualifiers to determine what is true now
  - But what happens when a new temporal qualifier comes along?

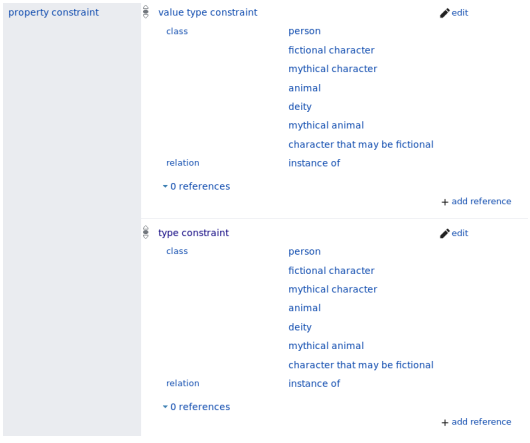
Qualifier Problems—Suggestions

WIKI DATA CON
Wikidata & languages 2019

- Add qualifiers only rarely
- Formalize qualifiers
  - To specify what a qualifier means
- Write tools that show effect of qualifiers
  - For temporal qualifiers—show what is true at a particular time

Constraints—Moving Towards a Solution

- Constraints do provide some facilities for improving Wikidata, but
  - They don't provide inference capabilities
  - They are only soft constraints, with exceptions allows
  - They aren't powerful enough to show class problems
  - They don't provide local ranges (e.g., the child of a human is a human)
  - They often appear to be descriptive, rather than prescriptive
- ShEx or SHACL would provide stronger constraints to partially implement a logic for Wikidata



Recap: Suggestions (Formal)

WIKI DATA CON
Wikidata & languages 2019

- Create a formal logic for Wikidata, and keep it up to date
  - Include disjointness, relative ranges
  - Include class definitions
  - Include unusual existing Wikidata constructs, including constraints
  - Include qualifiers, particularly temporal qualifiers
- Having a logic means that implications of claims can be determined
  - Is an instance of a subclass an instance of the superclass?
  - What is true at a particular point in time.

Recap: Suggestions (Tools)

WIKI DATA CON
Wikidata & languages 2019

- Tools that show implications of claims, according to current logic
- Tools that show claims relevant at a particular time (or other context)
- Tools that show contradictions and necessarily empty classes
  - Similar to what constraint checkers do, but (probably) more principled

Recap: Suggestions (Community)

WIKI DATA CON
Wikidata & languages 2019

- Add information to fill out domains, but carefully
- Value curation as much or more than adding information
- Don't use alternative sources when better source is available
- Partner with efforts that have complete information in a domain
- Describe classes better
- Add disjointness statements between classes
- Take more care in adding ontological claims
- Model similar entities and situations the same way
- Reign in bots—require more care, rerun capability

Make Wikidata Greater, Again!

WIKI DATA CON
Wikidata & languages 2019

With some (significant) additions Wikidata can be an even greater source of knowledge,
but the community needs to decide that (some of) the problems are worth fixing.

Make Wikidata Greater, Again!

WIKI DATA CON
Wikidata & languages 2019

With some (significant) additions Wikidata can be an even greater source of knowledge,
but the community needs to decide that (some of) the problems are worth fixing.

The major web players can spend the effort get around the problems,
but minor players may not be able to by themselves.

Make Wikidata Greater, Again!

WIKI DATA CON
Wikidata & languages 2019

With some (significant) additions Wikidata can be an even greater source of knowledge,
but the community needs to decide that (some of) the problems are worth fixing.

The major web players can spend the effort get around the problems,
but minor players may not be able to by themselves.

Thank you.

References

Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, eds. *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press, 2nd edition, August 2007, pages 458–486.

Cyc. Cycorp, `https://www.cyc.com/`

Aidan Hogan. “Nobel Prizes and consensus in Wikidata.” Wikimedia Foundation, `https://lists.wikimedia.org/pipermail/wikidata/2019-September/013544.html`

Markus Krötzsch, Maximilian Marx, Ana Ozaki, and Veronika Thost. “Attributed Description Logics: Reasoning on Knowledge Graphs” in Jérôme Lang, eds., *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI’18)*, pp. 5309–5313, July 2018. International Joint Conferences on Artificial Intelligence.

Boris Motik, Bernardo Cuenca Grau, Ian Horrocks, Zhe Wu, Achille Fokoue, and Carsten Lutz, eds. *OWL 2 Web Ontology Language Profiles (Second Edition)*. W3C Recommendation, 11 December 2012. W3C.