

# **Wikidata's SPARQL**

## **introduction for beginners**

CC BY-SA  
Lucie-Aimée Kaffee

Frimelle  
@lucielinde

# What is turtle?

RDF triple language

Subject Predicate Object

as Web URI

Alan Turing (Q7251) instance of (P31) human (Q5)

RDF

<http://www.wikidata.org/entity/Q7251> <http://www.wikidata.org/prop/direct/P31> <http://www.wikidata.org/entity/Q5>

turtle

@prefix wd: <<http://www.wikidata.org/entity/>>

@prefix wdt: <<http://www.wikidata.org/prop/direct/>>

wd:Q7251 wdt:P31 wd:Q5

# What is SPARQL?

RDF query language

to form complex queries with SELECT and retrieve data from an endpoint

Return **all items**, which are **instance of** **human**, limited to 200 items.

```
PREFIX wd: <http://www.wikidata.org/entity/>
```

```
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
```

```
SELECT ?subject WHERE { ?subject wdt:P31 wd:Q5 . } LIMIT 200
```

# Exploring Wikidata's linked data

Show all data for Alan Turing (Q7251) limited to 100

```
PREFIX entity: <http://www.wikidata.org/entity/>
```

```
SELECT ?predicate ?object WHERE {  
  entity:Q7251 ?predicate ?object .  
}  
LIMIT 100
```

[On Wikidata Query Service Beta](#)

## Extract of the result

predicate	object
rdfs:label	Alan Turing
rdfs:label	آلان تورنج
skos:altLabel	Алан Цюрынг
<http://www.wikidata.org/prop/direct/P101>	<http://www.wikidata.org/entity/Q21198>

Turing's field of work (P101) is computer science (Q21198)

# Exploring Wikidata's linked data

Show all fields of work of Alan Turing

```
PREFIX entity: <http://www.wikidata.org/entity/>
PREFIX p: <http://www.wikidata.org/prop/direct/>
```

```
SELECT ?object WHERE {
  entity:Q7251 p:P101 ?object .
} LIMIT 10
```

[On Wikidata Query Service Beta](#)

## Result

object
<http://www.wikidata.org/entity/Q21198>
<http://www.wikidata.org/entity/Q897511>

```
PREFIX entity: <http://www.wikidata.org/entity/>
PREFIX p: <http://www.wikidata.org/prop/direct/>
```

```
SELECT ?label WHERE {
  entity:Q7251 p:P101 ?object .
  ?object rdfs:label ?label FILTER (LANG(?label) = "en") .
} LIMIT 100
```

[On Wikidata Query Service Beta](#)

## Result

label
computer science
cryptanalysis

Turing's field of work (P101) is computer science (Q21198) and cryptanalysis (Q897511)

# Exploring Wikidata's linked data

Data linked to Alan Turing (Q7251) limited to 100

```
PREFIX entity: <http://www.wikidata.org/entity/>
PREFIX p:      <http://www.wikidata.org/prop/direct/>
```

```
SELECT ?subject ?predicate WHERE {
  ?subject ?predicate entity:Q7251 .
} LIMIT 100
```

[On Wikidata Query Service Beta](#)

## Extract of the result

subject	predicate
<http://www.wikidata.org/entity/Q163310>	<http://www.wikidata.org/prop/direct/P185>
<http://www.wikidata.org/entity/Q4489310>	<http://www.wikidata.org/prop/direct/P138>
<http://www.wikidata.org/entity/statement/q92741-7ED21209-4E23-4EB1-886A-772CC6A5350F>	<http://www.wikidata.org/prop/statement/P185>
<http://www.wikidata.org/entity/statement/q4489310-98f175a6-4ff2-6148-505c-6e6c0c7dfb68>*	<http://www.wikidata.org/prop/statement/P138>

# Truthy and full statements

represents statements in two formats

## truthy

contain only the statements of best rank

best rank are all statements of preferred rank, if those exist or all statements of normal rank

qualifiers are ignored

Prefix *wdt*: (<http://www.wikidata.org/prop/direct/>) - e.g. *wdt:P101*

## full statements

represent all data about a statement

represented in a separate node with the unique id of the statement - e.g. *q92741-7ED21209-4E23-4EB1-886A-772CC6A5350F*

Prefix *p*: (<http://www.wikidata.org/prop/>) - e.g. *p:P101*

# Full statements

qualifier and references are objects of the full statement

*wds:Q64-01f1575b-4bf3-982b-34d6-102fbe912ae2 a wikibase:Statement ;*

*wikibase:rank wikibase:NormalRank ;*

*psv:P1082 wdv:0419376cf1da9a8efcbb5ab7c87ded58 ;*

*ps:P1082 10000 ;*

*pqv:P585 wdv:04a483576f9618f0263dd20cb503ad32 ;*

*pq:P585 1680-01-01T00:00:00Z .*

*Q64-01f1575b-4bf3-982b-34d6-102fbe912ae2 (one statement for population of Berlin) is type of statement ;*

*has a NormalRank ; (this statement wouldn't show up in truthy since population has a preferred rank)*

*the population (P1082) is connected to a full value ;*

*the population is 10000 ;*

*the qualifier for point (P585) in time is connected to a full value ;*

*the qualifier point in time (P585) implies this statement is valid for 1680 .*



# Full statements

qualifier and references are objects of the full statement

*wds:Q64-01f1575b-4bf3-982b-34d6-102fbe912ae2 a wikibase:Statement ;*

*wikibase:rank wikibase:NormalRank ;*

*psv:P1082 wdv:0419376cf1da9a8efcbb5ab7c87ded58 ;*

*ps:P1082 10000 ;*

*pqv:P585 wdv:04a483576f9618f0263dd20cb503ad32 ;*

*pq:P585 1680-01-01T00:00:00Z .*

*Q64-01f1575b-4bf3-982b-34d6-102fbe912ae2 (one statement for population of Berlin) is type of statement ;*

*has a NormalRank ; (this statement wouldn't show up in truthy since population has a preferred rank)*

*the population (P1082) is connected to a full value ;*

*the population is 10000 ;*

*the qualifier for point (P585) in time is connected to a full value ;*

*the qualifier point in time (P585) implies this statement is valid for 1680 .*

# Full statements

qualifier and references are objects of the full statement

*wds:Q64-01f1575b-4bf3-982b-34d6-102fbe912ae2* a *wikibase:Statement* ;

*wikibase:rank* *wikibase:NormalRank* ;

*psv:P1082 wdv:0419376cf1da9a8efcbb5ab7c87ded58* ;

*ps:P1082 10000* ;

*pqv:P585 wdv:04a483576f9618f0263dd20cb503ad32* ;

*pq:P585 1680-01-01T00:00:00Z* .

*Q64-01f1575b-4bf3-982b-34d6-102fbe912ae2* (one statement for population of Berlin) is type of statement ;

has a *NormalRank* ; (this statement wouldn't show up in truthy since population has a preferred rank)

the population (*P1082*) is connected to a full value ;

the population is 10000 ;

the qualifier for point (*P585*) in time is connected to a full value ;

the qualifier point in time (*P585*) implies this statement is valid for 1680 .

# Full statements

qualifier and references are objects of the full statement

*wds:Q64-01f1575b-4bf3-982b-34d6-102fbe912ae2* a *wikibase:Statement* ;

*wikibase:rank* *wikibase:NormalRank* ;

*psv:P1082 wdv:0419376cf1da9a8efcbb5ab7c87ded58* ;

*ps:P1082 10000* ;

*pqv:P585 wdv:04a483576f9618f0263dd20cb503ad32* ;

*pq:P585 1680-01-01T00:00:00Z* .

*Q64-01f1575b-4bf3-982b-34d6-102fbe912ae2* (one statement for population of Berlin) is type of statement ;

has a *NormalRank* ; (this statement wouldn't show up in truthy since population has a preferred rank)

the population (P1082) is connected to a full value ;

the population is 10000 ;

the qualifier for point (P585) in time is connected to a full value ;

the qualifier point in time (P585) implies this statement is valid for 1680 .

# Example Query

Who's birthday is today?

```
PREFIX entity: <http://www.wikidata.org/entity/>
PREFIX t:      <http://www.wikidata.org/prop/direct/>
PREFIX rdfs:   <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd:    <http://www.w3.org/2001/XMLSchema#>
|
SELECT ?entity (year(?date) as ?year) WHERE {
  ?entityS t:P569 ?date .
  ?entityS rdfs:label ?entity .

  FILTER (datatype(?date) = xsd:dateTime)
  FILTER (month(?date) = month(now()))
  FILTER (day(?date) = day(now()))
} LIMIT 100
```

P569: date of birth

[On Wikidata Query Service Beta](#)

# Ressources

## [Wikidata Query Service Beta](http://wdqs-beta.wmflabs.org/)

<http://wdqs-beta.wmflabs.org/>

## [Mediawiki: RDF Dump Format](https://www.mediawiki.org/wiki/Wikibase/Indexing/RDF_Dump_Format)

[https://www.mediawiki.org/wiki/Wikibase/Indexing/RDF\\_Dump\\_Format](https://www.mediawiki.org/wiki/Wikibase/Indexing/RDF_Dump_Format)

## [Prefixes used for Wikidata](https://www.mediawiki.org/wiki/Wikibase/Indexing/RDF_Dump_Format#Prefixes_used)

[https://www.mediawiki.org/wiki/Wikibase/Indexing/RDF\\_Dump\\_Format#Prefixes\\_used](https://www.mediawiki.org/wiki/Wikibase/Indexing/RDF_Dump_Format#Prefixes_used)

## [Mediawiki: Query Examples](https://www.mediawiki.org/wiki/Wikibase/Indexing/SPARQL_Query_Examples)

[https://www.mediawiki.org/wiki/Wikibase/Indexing/SPARQL\\_Query\\_Examples](https://www.mediawiki.org/wiki/Wikibase/Indexing/SPARQL_Query_Examples)

## [github: wikidata query rdf docs](https://github.com/wikimedia/wikidata-query-rdf/tree/master/docs)

<https://github.com/wikimedia/wikidata-query-rdf/tree/master/docs>

## [Wikidata turtle dumps](https://dumps.wikimedia.org/wikidatawiki/entities/)

<https://dumps.wikimedia.org/wikidatawiki/entities/>

## [W3C RDF Primer](http://www.w3.org/2007/02/turtle/primer/)

<http://www.w3.org/2007/02/turtle/primer/>