

Semantic Web Project

Maxime
Philippe





Introduction and Objectives

- Investigating Utility of Querying Over Linked Data
- Utilizing Wikidata
- Exploring Answers In-Depth

Wikidata

- A Free and Open Knowledge Base
- Query Service: SPARQL
- wdt: for Quick, General Info VS p: for Detailed, In-Depth Data

| | |
|-----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| award received |  Nobel Prize in Literature  edit |
| point in time | 1997 |
| prize money | 7,500,000 Swedish krona |
| award rationale | <p>who emulates the jesters of the Middle Ages in scourging authority and upholding the dignity of the downtrodden (English)</p> <p>som i medeltida gycklares efterföljd gisslar makten och upprättar de utsattas värdighet (Swedish)</p> |
| | ► 4 references |

Question 1

Find that landmark article on data integration written by an Indian researcher in the 1990s.

```
SELECT DISTINCT ?item ?itemLabel ?date ?awardLabel ?work ?workTitle WHERE {  
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]". }  
  {  
    SELECT DISTINCT ?item ?date ?award ?work ?workTitle WHERE {  
      ?item wdt:P31 wd:Q5. # nature of element is Human  
      ?item wdt:P27 wd:Q668. # country of citizenship is India  
      ?item (wdt:P106/(wdt:P279*)) wd:Q1650915. # Occupation is (subclass of)  
researcher  
      ?item (wdt:P101/(wdt:P279*)) wd:Q2374463. # field of activity is (subclass  
of) Data Science  
      ?item p:P166 ?award_statement. # awards received as statement  
      ?award_statement ps:P166 ?award. # award recieved as ?award  
      ?award_statement pq:P1686 ?work. # for work ?work  
      ?work wdt:P1476 ?workTitle. # with title ?workTitle  
      ?award_statement pq:P585 ?date. # at date ?date  
      FILTER(?date < "+2000-01-01T00:00:00Z"^^xsd:dateTime). # ?date before 2000  
      FILTER(?date > "+1989-01-01T00:00:00Z"^^xsd:dateTime). # ?date after 1989  
    }  
  }  
}
```

Results

Table

1 résultat en 733

| item | itemLabel | date | awardLabel | work | workTitle |
|---------------------------|------------------|----------------|----------------|------------------------------|----------------------------------------------------------------------------------|
| wd:Q93034 | Ravindran Kannan | 1 janvier 1991 | prix Fulkerson | wd:Q55871716 | A random polynomial-time algorithm for approximating the volume of convex bodies |

Justification

Statements

instance of



human

edit

► 1 reference

+ add value

image



RavindranKannan.tiff
198 × 297; 113 KB

▼ 0 references

edit

+ add reference

+ add value

country of citizenship



India

occupation



mathematician

field of work



informatics

award received



Knuth Prize

point in time

2011

► 1 reference



Fulkerson Prize

point in time

1991

together with

Martin Dyer

Alan M. Frieze

for work

A random polynomial-time algorithm for approximating the volume of convex bodies

Question 2


Who is the French archer who has won an Olympic event and received the Honor Legion ?

```
SELECT DISTINCT ?item ?itemLabel WHERE {  
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]". }  
  {  
    SELECT DISTINCT ?item WHERE {  
      ?item wdt:P31 wd:Q5. # nature of element is Human  
      ?item wdt:P27 wd:Q142. # country of citizenship is France  
      ?item wdt:P106 wd:Q13382355. # occupation is archery (job not hobby)  
      ?item p:P1344 ?event_statement. # participate in events as statement  
      ?event_statement (ps:P1344/(wdt:P31*)) wd:Q18536594. # event is (instance  
of) Olympic event  
      ?event_statement pq:P166 wd:Q15243387. # award received is Olympic gold  
medal  
      ?item (wdt:P166/(wdt:P279*)) wd:Q10855271. # award received is (subclass of)  
Knight of the Legion of Honour  
    }  
  }  
}
```

Results


| Table ? | |
|-----------------------------|-----------------|
| item | itemLabel |
| wd:Q1311057 | Sébastien Flute |

Justification

instance of  human

award received  Knight of the Legion of Honour

country of citizenship  France

occupation  archer



participant in

 2000 Summer Olympics

▼ 0 references

 1996 Summer Olympics

▼ 0 references

























 archery at the 1992 Summer Olympics – men's individual
award received Olympic gold medal

Question 4


What are the 20 books that have the most pages, which is lower than 20,000, and that are written by at least one author who has a child ?


```
SELECT DISTINCT ?item ?itemLabel ?numberOfPages WHERE {  
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]". }  
  {  
    SELECT DISTINCT ?item ?numberOfPages WHERE {  
      ?item wdt:P31 wd:Q571. # nature of element is book  
      ?item wdt:P50 ?author. # author is ?author  
      ?author wdt:P40 _:anyValueP40. # has child exist  
      ?item wdt:P1104 ?numberOfPages. # number of pages is ?numberOfPages  
      FILTER(?numberOfPages < 20000) # ?numberOfPages is less than 20 000  
    }  
  }  
}  
ORDER BY DESC (?numberOfPages)  
LIMIT 20
```

Results

| Table  | | | 20 résultats en 2893 ms |  Code |  Télécharger |  Lien |
|-----------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|---------------|-------------------------|------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|
| item | itemLabel | numberOfPages | | | | |
|  wd:Q1943729 | Systema Mycologicum | 1866 | | | | |
|  wd:Q111602524 | Q111602524 | 1801 | | | | |
|  wd:Q51407079 | Manuel de conchyliologie et de paleontologie conchyliologique : ou, Histoire naturelle des mollusques vivants et fossiles | 1486 | | | | |
|  wd:Q106584307 | Q106584307 | 1246 | | | | |
|  wd:Q113367291 | Q113367291 | 951 | | | | |
|  wd:Q99227983 | Q99227983 | 786 | | | | |
|  wd:Q99912674 | Principes du droit des gens | 779 | | | | |
|  wd:Q51393152 | A monograph of the British fossil corals | 736 | | | | |
|  wd:Q100255508 | Q100255508 | 735 | | | | |
|  wd:Q51392595 | Grundzüge der Anatomie der wirbellosen Thiere | 644 | | | | |
|  wd:Q112578977 | Q112578977 | 644 | | | | |
|  wd:Q51469440 | Epicrisis Systematis mycologici | 610 | | | | |
|  wd:Q114950569 | Q114950569 | 610 | | | | |
|  wd:Q4378679 | Histoire de l'Égypte ancienne | 593 | | | | |
|  wd:Q111594361 | Q111594361 | 592 | | | | |
|  wd:Q112766602 | Q112766602 | 592 | | | | |
|  wd:Q86907717 | Q86907717 | 587 | | | | |
|  wd:Q106583829 | Q106583829 | 584 | | | | |
|  wd:Q106407347 | Q106407347 | 577 | | | | |
|  wd:Q110249810 | Dieu, la science, les preuves : l'aube d'une révolution | 577 | | | | |

Justification

instance of  book

number of pages  1,486

author 

Paul Henri Fischer

▼ 0 references

 Joseph Wilson Lowry


▼ 0 references

 Samuel P. Woodward

child 

Bernard Henry Woodward

▼ 0 references

 Horace Bolingbroke Woodward

▼ 0 references

 Bernard Barham Woodward

Finding a difficult question even for ChatGPT

P

Which film has the shortest title in the history of the Cannes Film Festival's Palme d'Or winners and didn't get the award the same year as he got published?



As of my last update in January 2022, the film with the shortest title among Cannes Film Festival's Palme d'Or winners is "MAS*H," directed by Robert Altman. The film won the award in 1970. However, the movie was actually released in the same year that it won the award, so it doesn't meet your criterion of not winning the award in the same year it was published.



It's worth noting that my data may be outdated, so you may want to check the most current list of Palme d'Or winners for any updates.

Question

Which film has the shortest title in the history of the Cannes Film Festival's Palme d'Or winners and didn't get the award the same year as he got published?

```
SELECT DISTINCT ?item ?itemLabel ?title ?titlelength ?award_year ?earliest_publication_year WHERE {  
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]". }  
  {  
    SELECT ?item ?title ?titlelength ?award_year (MIN(?publication_year) as ?earliest_publication_year) WHERE {  
      ?item wdt:P31 wd:Q11424. # nature of element is Movie  
      ?item wdt:P1476 ?title. # with title ?title  
      ?item wdt:P577 ?publication_date. # published the ?publication_date  
      ?item p:P166 ?award_statement. # award statement  
      ?award_statement ps:P166 wd:Q179808. # award is Palme d'Or  
      ?award_statement pq:P585 ?award_date. # given the ?award_date  
      BIND(STRLEN(?title) as ?titlelength). # length of the title  
      BIND(YEAR(?publication_date) as ?publication_year). # year of the publication date  
      BIND(YEAR(?award_date) as ?award_year). # year of the Palme d'Or award date  
    } GROUP BY ?item ?title ?titlelength ?award_year  
    HAVING ( ?earliest_publication_year != ?award_year) # didn't get the award the same year as he got published  
  }  
}  
ORDER BY ASC (?titlelength)  
LIMIT 5
```

Results

| Table | | | | | | 5 résultats en 5 ms | </> Code | ⬇ Télécharger | 🔗 Lien |
|----------------------------|----------------------------|---------------|-------------|------------|---------------------------|---------------------|----------|---------------|--------|
| item | itemLabel | title | titlelength | award_year | earliest_publication_year | | | | |
| Q658204 | Blow-Up | Blowup | 6 | 1967 | 1966 | | | | |
| Q105393862 | Titane | Titane | 6 | 2019 | 2021 | | | | |
| Q1346745 | If.... | If.... | 6 | 1969 | 1968 | | | | |
| Q1353441 | Que le spectacle commence | All That Jazz | 13 | 1980 | 1979 | | | | |
| Q714524 | Quand passent les cigognes | Летят журавли | 13 | 1958 | 1957 | | | | |

award in 2019
published in 2021



Justification

instance of film

title Blowup (English)

publication date

18 December 1966
place of publication United States of America
▶ 1 reference

11 May 1967
place of publication Germany

award received

Palme d'Or
point in time 1967

ChatGPT

- A Large Language Model by OpenAI
- Role in Project: Knowledge Base identification & Query structure
- Limitations: Doesn't give correct values

M

give me knowledge base website where i will be able to find the answer of the following questions using SPARQL : Who is the French archer who has won an Olympic event and received the Honor Legion ?



You can use the Wikidata knowledge base to find information about the French archer who has won an Olympic event and received the Honor Legion using SPARQL. Wikidata is a free and open knowledge base that contains a vast amount of structured data on various topics, including information about individuals, their achievements, and awards.

Here's how you can query this information in SPARQL using the Wikidata SPARQL endpoint:

1. Go to the Wikidata SPARQL query service: <https://query.wikidata.org/>.
2. Use the following SPARQL query to find the French archer who meets your criteria:

sparql

Copy code

```
SELECT ?archer ?archerLabel
WHERE {
  ?archer wdt:P106 wd:Q13331195; # French archer
          wdt:P166 wd:Q207316;   # Olympic event winner
          wdt:P166 wd:Q7737527;  # Legion of Honour recipient
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]" }
}
```

This query looks for individuals who have the occupation of "French archer" (Q13331195), have won an Olympic event (Q207316), and received the Legion of Honour (Q7737527). It also includes a service for labeling to get the names of the individuals.

Conclusion

- ChatGPT : great for starting the project
- Wikidata : perfect for all our questions
& Request assistant + Request examples
BUT badly optimized...

“Blazegraph has a built-in query optimizer which often works well. However, sometimes it is not so successful; in such cases queries may need to be optimised manually.”

source : Wikidata:SPARQL query service/query optimization/hy

| Contents [hide] | |
|-----------------|---------------------------------------------------------------------------------------------------|
| 1 | Optimization strategies |
| 1.1 | Fixed values and ranges |
| 1.2 | Property paths |
| 1.3 | Inverse property paths |
| 1.4 | Order tweaking |
| 1.5 | Delay call to label service |
| 1.6 | Named subqueries |
| 1.6.1 | Further uses of named subqueries |
| 1.7 | Searching labels |
| 1.8 | Use COUNT(*) when possible, and fast range counts |
| 1.9 | Distinct term scan, and group by and count optimization |
| 1.10 | Count sizes of huge subsets intersections with WikibaseCircusSearch |
| 2 | Services |
| 2.1 | GAS Service |
| 3 | Automatic optimisation -- background overview |
| 4 | A query that has difficulties |
| 4.1 | Other optimisation modes |
| 5 | Other queries the optimiser has difficulty with |
| 5.1 | Performance falls off a cliff, when the number in a group of interest exceeds a certain threshold |
| 5.2 | Adding a names and name-labels wrapper to a query for life events makes it time out |
| 5.3 | Adding a property-labels wrapper to a query for properties makes it time out |
| 5.4 | Queries with FILTER NOT EXISTS |
| 5.4.1 | Rules of Thumb |
| 6 | Dealing with cached queries |
| 7 | More... ? |