



**Aiteam**  
NATIONAL & KAPODISTRIAN  
UNIVERSITY OF ATHENS

# Recent advances in question answering for geospatial knowledge graphs

Manolis Koubarakis

Dept. of Informatics and Telecommunications  
National and Kapodistrian University of Athens

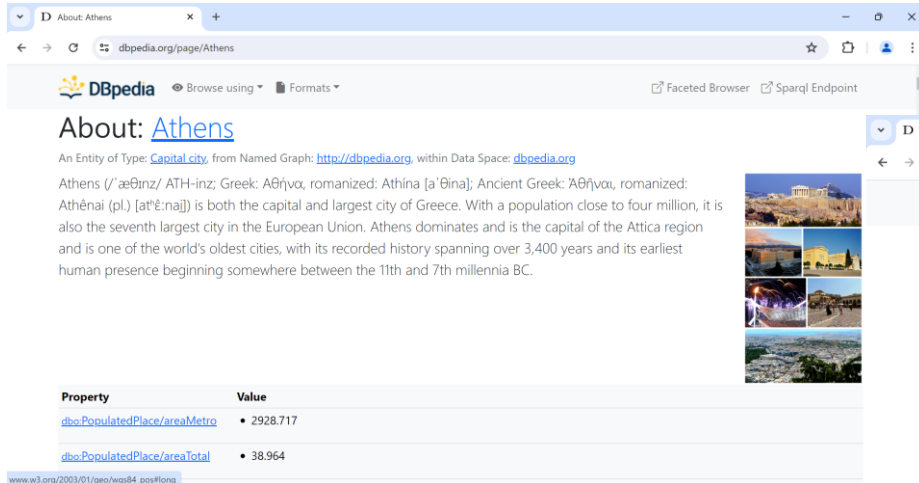
GeoLD workshop @ ESWC May 26, 2024

# Talk Outline

- Geospatial knowledge graphs
- Older question answering engines
- Recent question answering engines
  - The engine of Hamzei et al. (2022)
  - GeoQA2
  - EarthQA
  - GeoQAMap
  - GeoKBQA
- Benchmark datasets for evaluating question answering engines
  - GeoQuestions201
  - GeoQuestions1089
  - The dataset of Yang et al. (2024)
- Open problems

# Geospatial Knowledge Graphs

- **DBpedia** (lat/long coordinates, some cardinal directions, some topological relations implicitly)



DBpedia

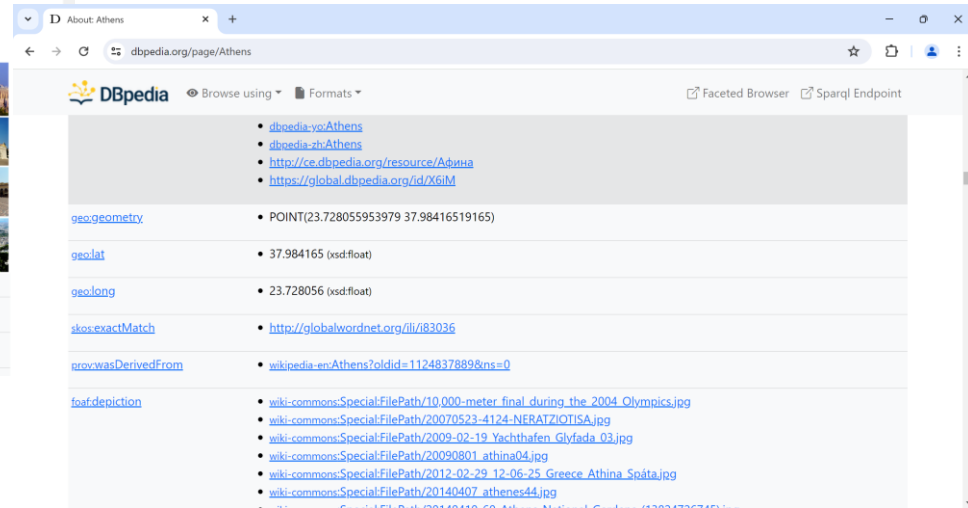
## About: Athens

An Entity of Type: [Capital city](#), from Named Graph: <http://dbpedia.org>, within Data Space: [dbpedia.org](#)

Athens (/ˈæθɪnz/ ATH-inz; Greek: Αθήνα, romanized: Athina [aˈθina]; Ancient Greek: Ἀθῆναι, romanized: Athēnai (pl.) [atʰɛːnai]) is both the capital and largest city of Greece. With a population close to four million, it is also the seventh largest city in the European Union. Athens dominates and is the capital of the Attica region and is one of the world's oldest cities, with its recorded history spanning over 3,400 years and its earliest human presence beginning somewhere between the 11th and 7th millennia BC.

Property	Value
<a href="#">dbpedia:PopulatedPlace/areaMetro</a>	• 2928.717
<a href="#">dbpedia:PopulatedPlace/areaTotal</a>	• 38.964

8,621,058 geentities



DBpedia

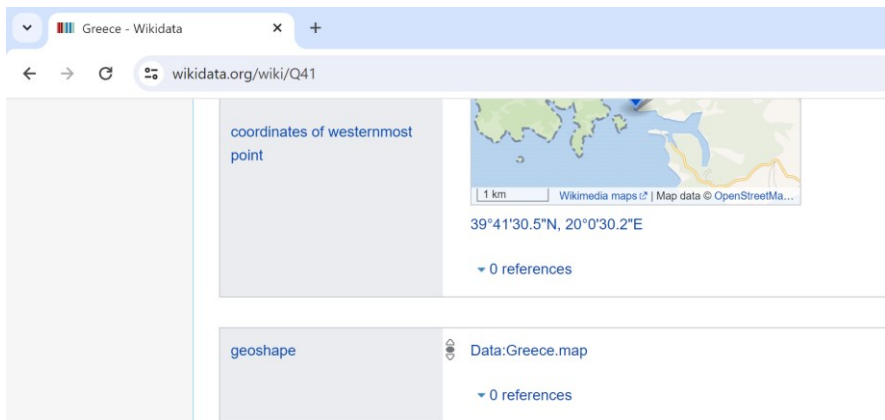
- [dbpedia-yo:Athens](#)
- [dbpedia-zh:Athens](#)
- <http://ce.dbpedia.org/resource/Афина>
- <https://global.dbpedia.org/id/X6iM>

<a href="#">geo:geometry</a>	• POINT(23.728055953979 37.98416519165)
<a href="#">geo:lat</a>	• 37.984165 (xsd:float)
<a href="#">geo:long</a>	• 23.728056 (xsd:float)
<a href="#">skos:exactMatch</a>	• <a href="http://globalwordnet.org/ll/83036">http://globalwordnet.org/ll/83036</a>
<a href="#">rdfs:wasDerivedFrom</a>	• <a href="#">wikipedia-en:Athens?oldid=1124837889&amp;ns=0</a>
<a href="#">foaf:depiction</a>	• <a href="#">wiki-commons:Special:FilePath/10_000-meter_final_during_the_2004_Olympics.jpg</a>
	• <a href="#">wiki-commons:Special:FilePath/20070523-4124-NERATZ/OTISA.jpg</a>
	• <a href="#">wiki-commons:Special:FilePath/2009-02-19_Yachthafen_Glyfada_03.jpg</a>
	• <a href="#">wiki-commons:Special:FilePath/20090801_athina04.jpg</a>
	• <a href="#">wiki-commons:Special:FilePath/2012-02-29_12-06-25_Greece_Athina_Spata.jpg</a>
	• <a href="#">wiki-commons:Special:FilePath/20140407_athens44.jpg</a>
	• <a href="#">wiki-commons:Special:FilePath/20140410_09_Athina_Museum_Gallery_1309173710.jpg</a>

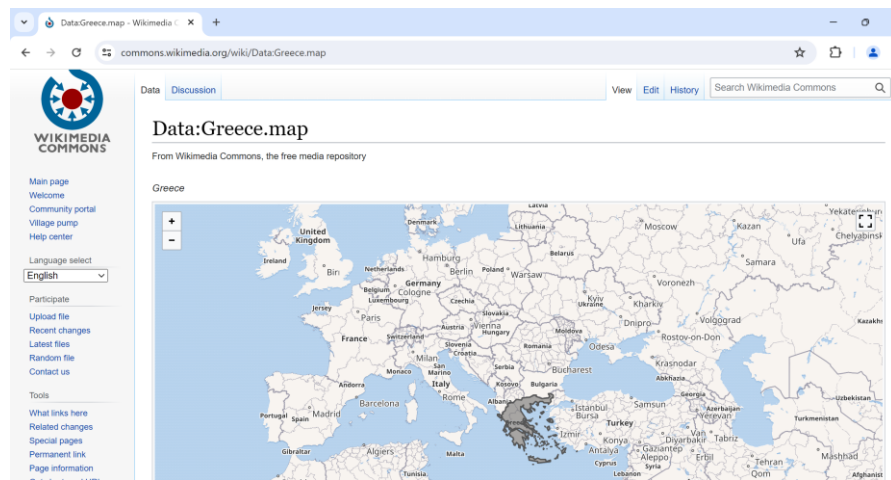
# Geospatial Knowledge Graphs

8,621,058 geentities

- **Wikidata** (same as DBpedia but also geoshapes). Geoshapes encode complex geometries e.g., multipolygons.



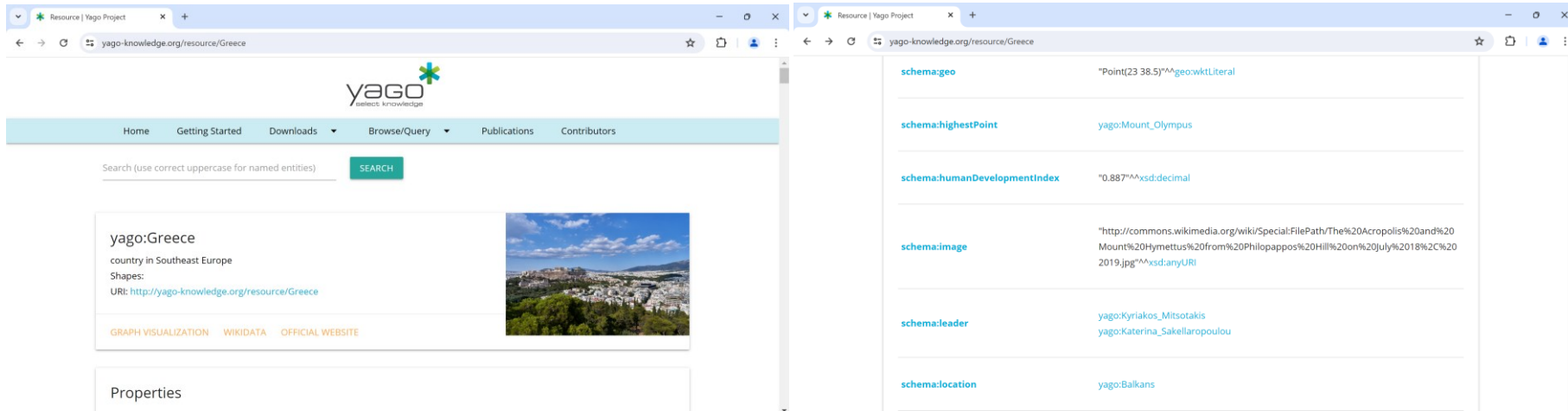
The screenshot shows the Wikidata page for Q41 (Greece). The browser address bar displays "wikidata.org/wiki/Q41". The page content includes a section for "coordinates of westernmost point" with a map of Greece and the coordinates  $39^{\circ}41'30.5''\text{N}, 20^{\circ}0'30.2''\text{E}$ . Below this is a section for "geoshape" with the label "Data:Greece.map" and a link to "0 references".



The screenshot shows the Wikimedia Commons page for "Data:Greece.map". The browser address bar displays "commons.wikimedia.org/wiki/Data:Greece.map". The page features the Wikimedia Commons logo and a navigation menu. The main content area shows a map of Greece with various cities and regions labeled, including Athens, Thessaloniki, and Izmir. The map is titled "Data:Greece.map" and is described as "From Wikimedia Commons, the free media repository".

# Geospatial Knowledge Graphs

- **YAGO2** (geoentities from Wikipedia and GeoNames, lat/long coordinates only, some topological relations implicitly)



The screenshot displays the YAGO2 web interface for the entity 'yago:Greece'. The left pane shows the entity's name, description ('country in Southeast Europe'), a photograph of a cityscape, and a search bar. The right pane lists various properties and their values:

schema:geo	"Point(23 38.5)"^^geo:wktLiteral
schema:highestPoint	yago:Mount_Olympus
schema:humanDevelopmentIndex	"0.887"^^xsd:decimal
schema:image	"http://commons.wikimedia.org/wiki/Special:FilePath/The%20Acropolis%20and%20Mount%20Hymettus%20from%20Philopappos%20Hill%20on%20July%2018%2C%202019.jpg"^^xsd:anyURI
schema:leader	yago:Kyriakos_Mitsotakis yago:Katerina_Sakellariopoulou
schema:location	yago:Balkans

10 million geoentities

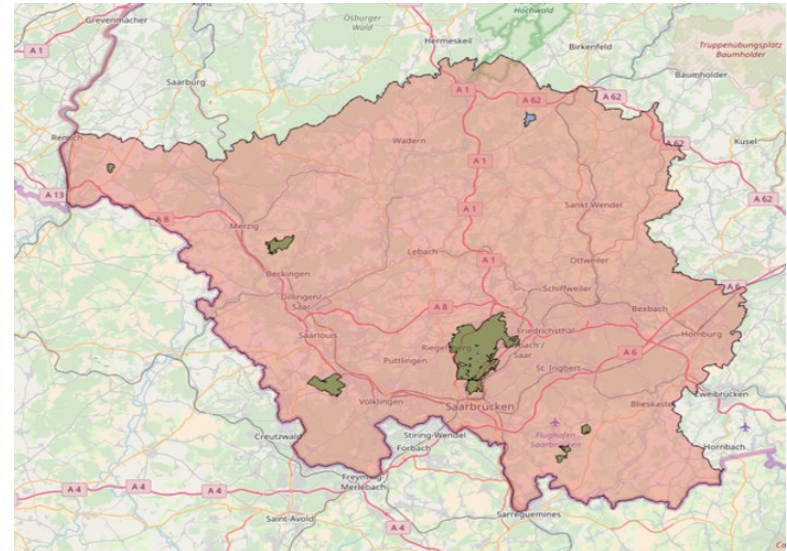
Hoffart et al. *YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia*. *Artif. Intell.* 194: 28-61 (2013)

# Geospatial Knowledge Graphs

- **YAGO2geo** (adds to YAGO2 more complex geometries from administrative geodata of Greece/UK/Ireland/USA, GADM and some OpenStreetMap features)

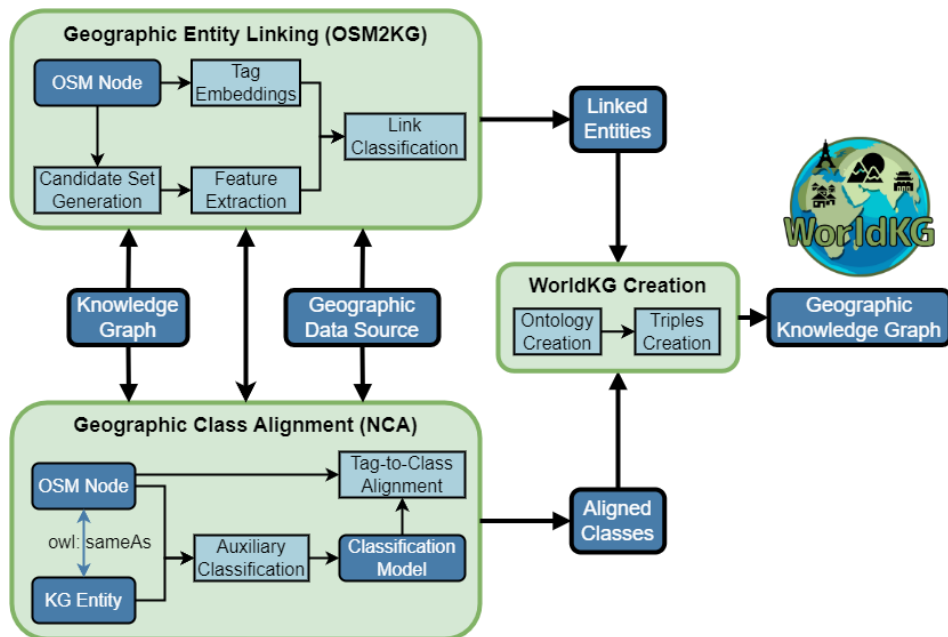
Polygons and multipolygons: 640,000  
Linestrings: 137,000

Karalis et al. *Extending the YAGO2 Knowledge Graph with Precise Geospatial Knowledge*. ISWC 2019.



# Geospatial Knowledge Graphs

- **WorldKG** (makes OpenStreetMap data available as a knowledge graph)



Dsouza, et al. *WorldKG: A World-Scale Geographic Knowledge Graph*. CIKM 2021.

# WorldKG Statistics

Quantity	Count
Total triples	<u>828,550,751</u>
Total entities	<u>113,444,975</u>
Top-level classes	<u>33</u>
Sub-classes	<u>1,143</u>
Unique properties	<u>1,820</u>
Links to Wikidata classes	<u>40</u>
Links to DBpedia classes	<u>21</u>

Only points in the  
SPARQL endpoint  
currently.

In total, WorldKG covers 113,444,975 geographic entities, clearly more than Wikidata (8,621,058) and DBpedia (~~8,621,058~~).

1,224,403



# Geospatial Knowledge Graphs

- **KnowWhereGraph** (integrates thematic and place-centric datasets into a geospatial knowledge graph).
- Pilot scenarios in disaster relief, agricultural land use and food-related supply chains.

**Multipolygons: 4,000**  
**Polygons: 2.1 million**  
**Linestrings: 730,000**  
**Points: 3.9 million**

Janowicz et al. *Know, Know Where, Knowwheregraph: A Densely Connected, Cross-Domain Knowledge Graph and Geo-Enrichment Service Stack for Applications in Environmental Intelligence*. *AI Mag.* 43(1): 30-39 (2022)

Thematic Datasets					Place-Centric Datasets		
Dataset Name/ Theme	Source Agency	Key Attributes	Spatial Coverage	Temporal Coverage	Place-Centric Dataset	Defining Authority	Spatial Coverage
Soil Properties	USDA	soil type, farmland class	Targeted regions in US	Current	S2 Cells	Google	Lvl 9 (Global), Lvl 13 (US),
Wildfires	USGS, USDA, USFS, NIFC	wildfire type, burn severity, num. acres burned, contained date	US	1984–current	Global Administrative Regions	University of Berkeley, Museum of Vertebrate Zoology and the International Rice Research Institute	Global
Earthquakes	USGS	magnitude, length, width, geometry	Global (mag. over 4.5)	2011-01-01 to 2022-01-18			
Climate Hazards	NOAA	injuries, deaths, property damages	US	1950–2022	US Federal Judicial District	DoJ, ESRI	US
Expert - Covid-19 Mobility	Direct Relief (DR)	name, affiliation, expertise	Global	2021			
Expert - General	KWG, UC System, DR, Semantic Scholar	name, affiliation, expertise with spatiotemporal scopes	Global	unlimited	National Weather Zones	NOAA	US
Cropland Types	USDA	crop types (raster data)	US	2008-2021	FIPS Codes	NRCS	US
Air Qual. Obs.	U.S. EPA	AQI value, CO concentration	US	1980–2022	Designated Market Area	Nielen	US
Smoke Plumes	NOAA	daily smoke plumes extent	US	2010-2022	ZIP	ZCTA	US
Climate Observations	NOAA	temperature, precipitation, PDSI, PHSI	US	1950 - 2022	Climate Division	NOAA	US
Disaster Declaration	FEMA	designated area, program, amount approved, program designated date	US	1953 - 2022	Census Metropolitan Area	US Census	US
Smoke Plume Extents	NOAA	Smoke extent	US	2017 - 2022	Drought Zone	NDMC, USDA, NOAA	US
BlueSky Forecasts	Bluesky	PM10, PM5	US	2022-03-07	Geographic Name Information System	USGS	US
Transportation (highway network)	DOT	road type, road length, road sign	US	2014			
Public Health Observations	CDC, US Census, University of Wisconsin Population Health Institute	below poverty level, diabetes, obesity, mental health provider rate, annual mammogram, vaccinated, injury death	US	2017, 2021			
Public Health Infrastructure	HIFLD	pharmacies, hospitals, dialysis centers, public health departments	US	2017 - 2022			
Social Vulnerability	CDC/ATSDR	social vulnerability index	US	2018			
Hurricane Tracks	NOAA	max wind speed, min pressure	US	1851-2020			

# Older Geospatial Question Answering Engines

- Hybrid geo-spatial query methods on the semantic web with a spatially-enhanced index of DBpedia
  - Younis et al., 2012
- Template-based question answering over linked geospatial data (the engine **GeoQA**)
  - Punjani et al., 2018 and 2020
- Neural factoid geospatial question answering
  - Li et al., 2021

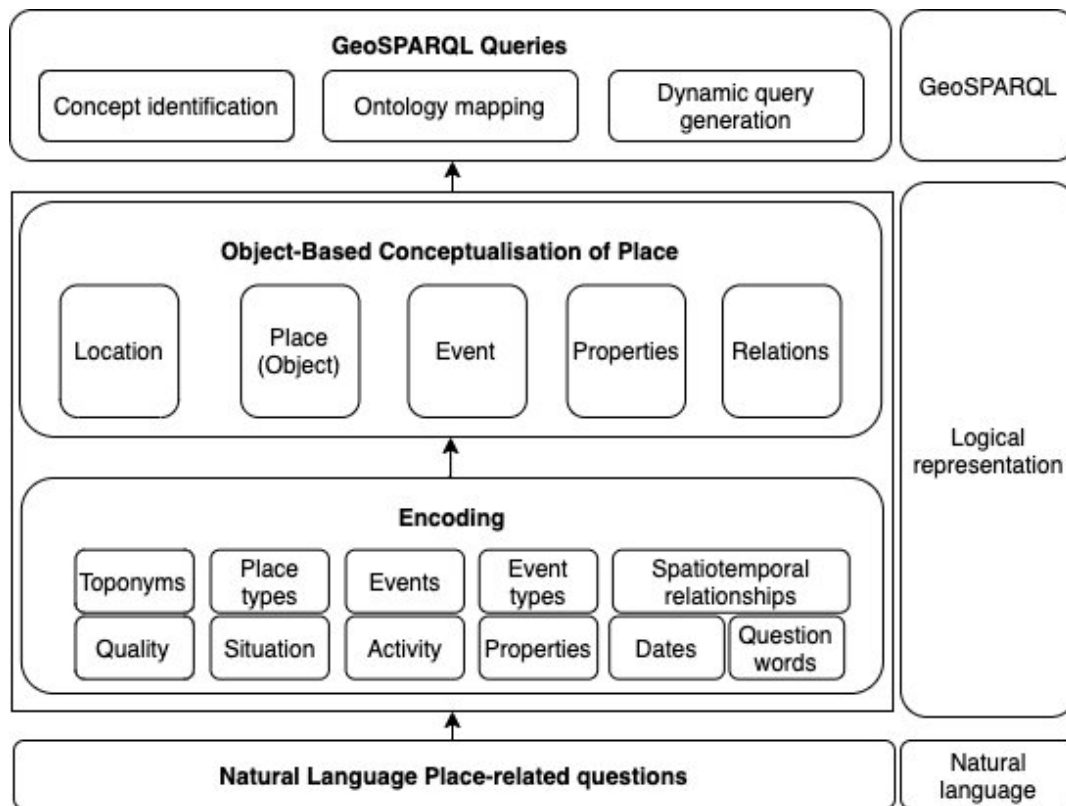


# More Recent Geospatial Question Answering Engines

- The engine of Hamzei et al. (2022)
- GeoQA2
- EarthQA
- GeoQAMap
- GeoKBQA

# The engine of Hamzei et al. (2022)

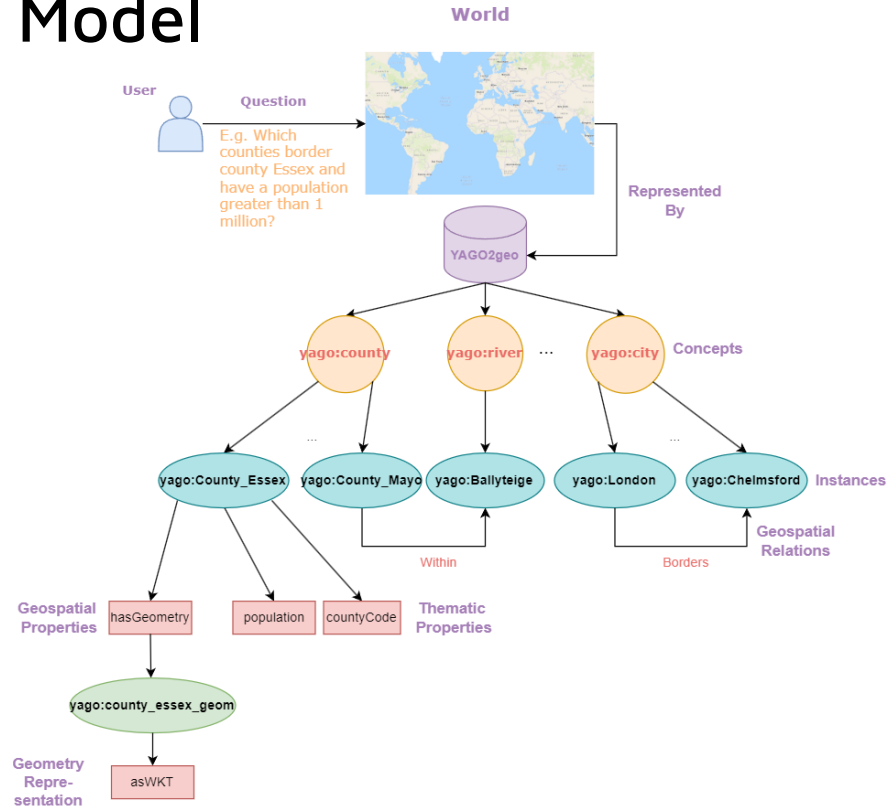
- It has a comprehensive conceptual model
- It targets an extension of the YAGO2geo knowledge graph
- It is template-based



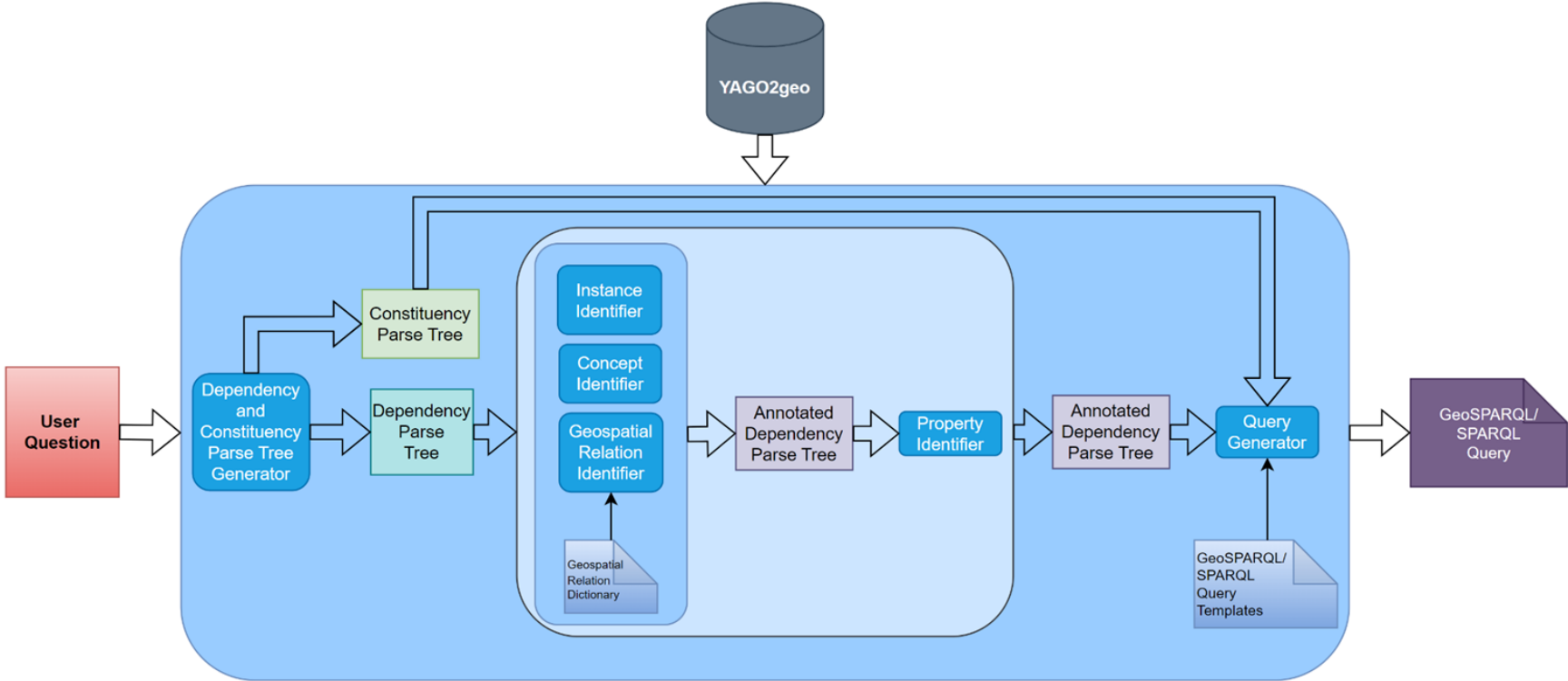
Hamzei, et al. *Translating Place-Related Questions to GeoSPARQL Queries*. The Web Conference 2022.

# GeoQA2 – Conceptual Model

- It is an extension and reengineering of GeoQA.

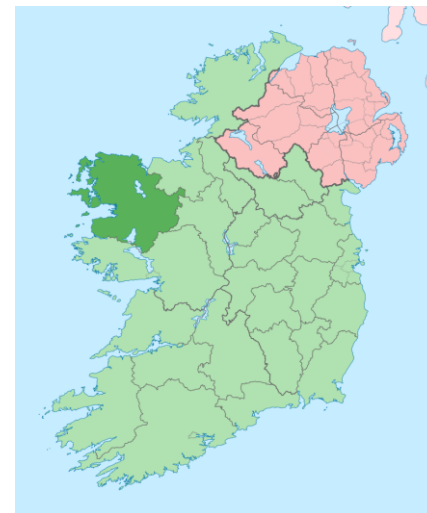


# GeoQA2 - Software Architecture



# GeoQA2 - Example

- Which **lakes** are **in counties** that **border** with **County Mayo**?
  - Instances: `yago:County_Mayo`
  - Concepts: `y2geoo:OSI_County`, `y2geoo:OSM_lake`
  - Relations: `geof:sfContains`, `geof:sfTouches`
  - Properties: -



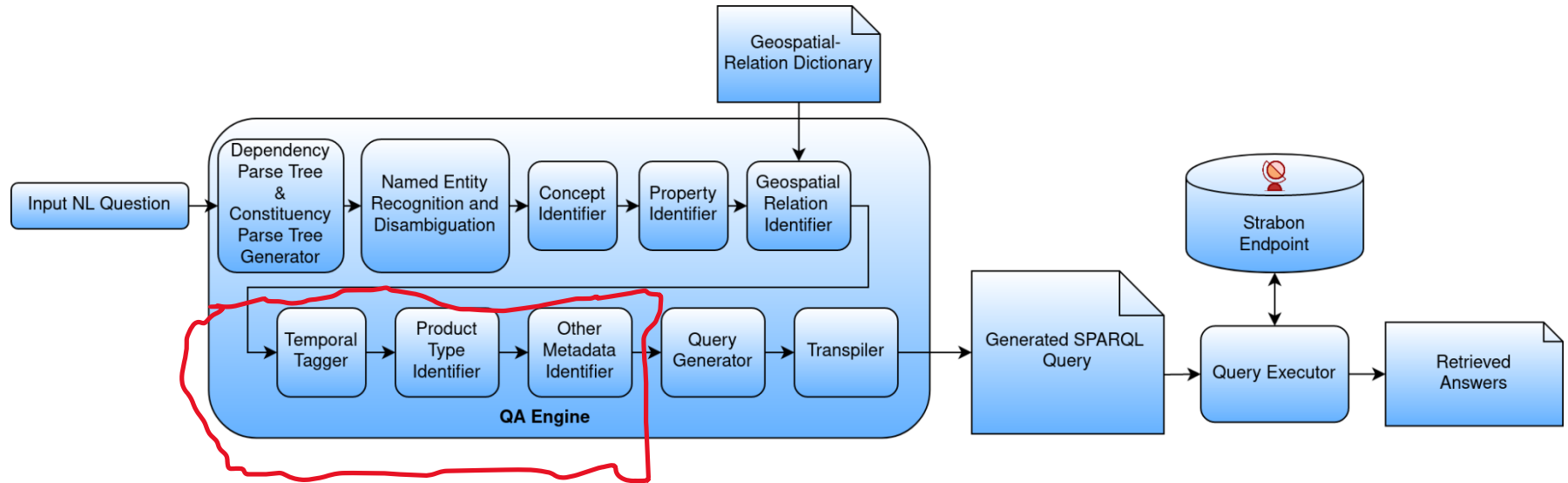
# Generated GeoSPARQL Query

```
SELECT DISTINCT ?c17 WHERE {  
  ?c17 a <http://kr.di.uoa.gr/yago2geo/ontology/OSI\_County>.  
  ?c17 geo:hasGeometry ?cGeometry17.  
  ?cGeometry17 geo:asWKT ?cWKT17.  
  ?c6 a <http://kr.di.uoa.gr/yago2geo/ontology/OSM\_lake>.  
  ?c6 geo:hasGeometry ?cGeometry6.  
  ?cGeometry6 geo:asWKT ?cWKT6.  
  <http://yago-knowledge.org/resource/County\_Mayo> geo:hasGeometry ?iGeometry1.  
  ?iGeometry1 geo:asWKT ?iWKT1 .  
  FILTER (geof:sfTouches (?iWKT1, ?cWKT17))  
  FILTER (geof:sfContains (?cWKT17, ?cWKT6))  
}
```



# EarthQA

- It is a question answering engine for **Earth observation data archives**.
- It has been developed as an extension of GeoQA2.



# EarthQA in Operation

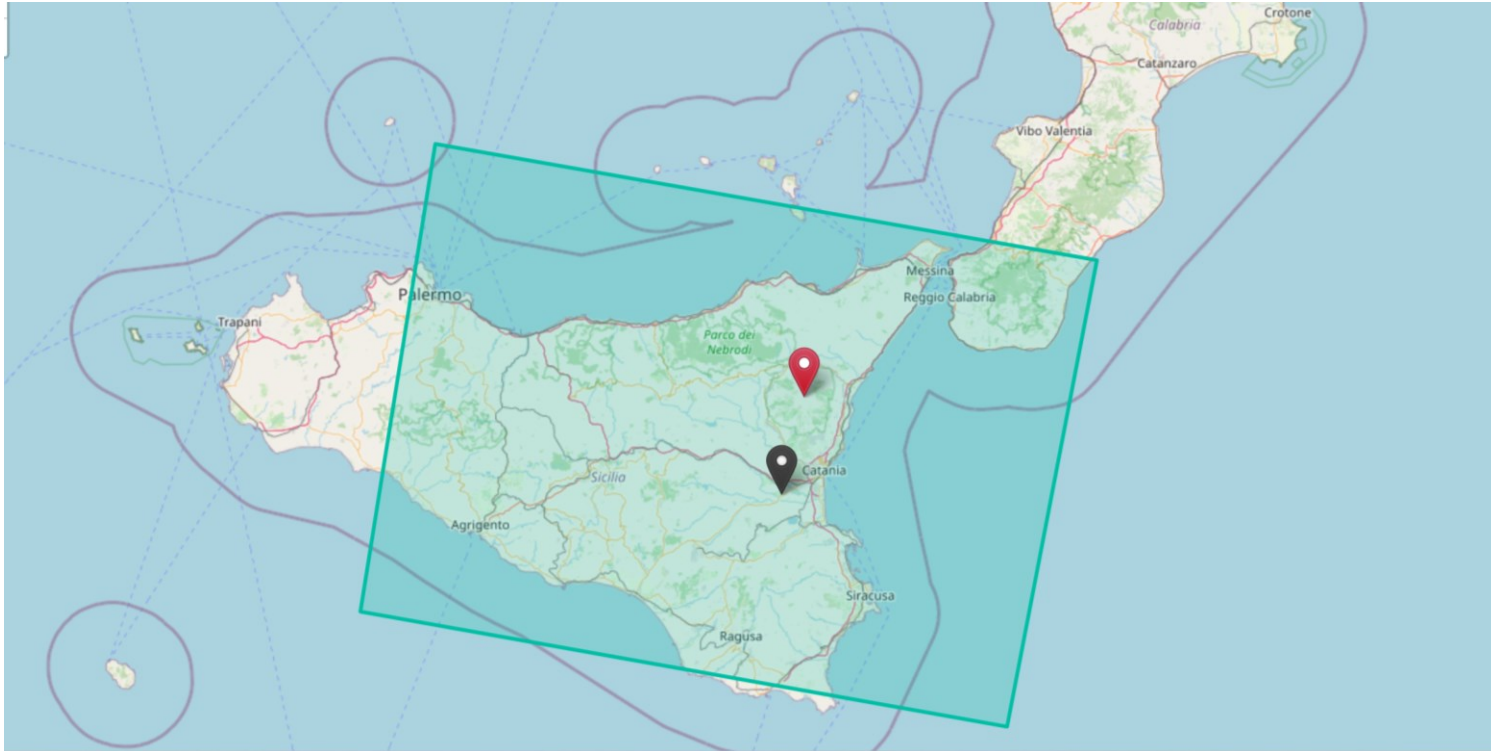
The screenshot displays the EarthQA web interface. At the top left is the EarthQA logo. On the top right are navigation links for "How It Works", "About", a settings gear icon, and a UK flag. A search bar contains the query "Find Sentinel-1 products that show Etna in March 2018" with a search icon to its right. Below the search bar is a link for "Advanced Options".

The "Info" section is expanded, showing the following details:

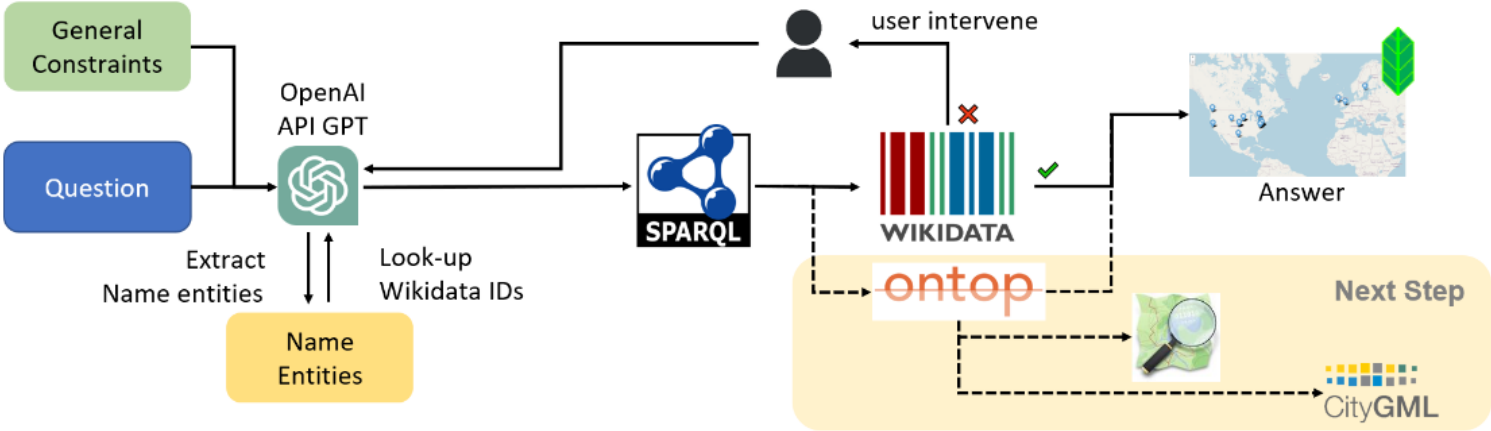
Generated Query.	Source	Query
Query #1 of 1 total queries	EODATA	<pre>select distinct ?x where { ?hex &lt;http://ws.creodias.eu/metadata/attribute#feature&gt; ?x . ?hex ?pred &lt;http://dbpedia.org/resource/Mount_Etna&gt; . ?x &lt;http://ws.creodias.eu/metadata/attribute#title&gt; ?title . ?x &lt;http://ws.creodias.eu/metadata/attribute#geometry&gt; ?geom . ?x &lt;http://ws.creodias.eu/metadata/attribute#mission&gt;</pre>

The "Map" section is visible at the bottom, showing a map of the region around Calabria, Italy, with labels for Catanzaro and Crotona. A zoom control with "+" and "-" buttons is on the left side of the map.

# EarthQA in Operation



# GeoQAMap

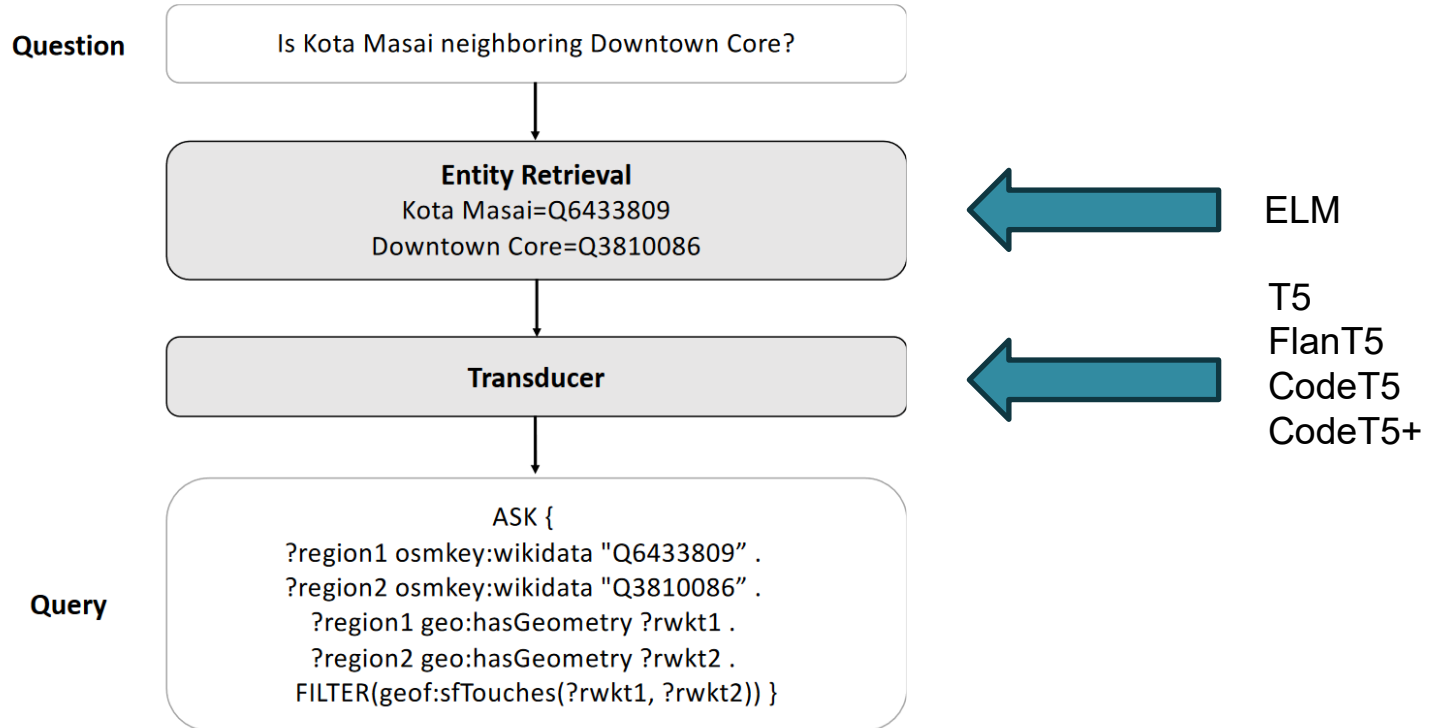


Feng et al. *GeoQAMap - Geographic Question Answering with Maps*. GIScience 2023.

# Comments

- GeoQAMap is work in progress.
- The paper (Feng et al. 2023) discusses how to deal with questions of types C and F of the benchmark GeoQuestions1089 (details later).

# GeoKBQA



# Comments

- A benchmark dataset for evaluation geospatial question answering engines has been constructed (details later).
- GeoKBQA uses ELM for entity recognition and disambiguation and different versions of T5 for query generation.
- No comparison with other systems is being done.

# Evaluating Geospatial Question Answering Engines

- I will discuss three benchmark datasets that have been used for the evaluation of the above geospatial question answering engines:
  - GeoQuestions201
  - GeoQuestions1089
  - The dataset of Yang et al. (2024)



# GeoQuestions201

- **GeoQuestions201** contains 201 question-query-answer triples
- Questions crowdsourced from students
- Target knowledge graph created by interlinking:
  - **DBpedia**
  - **OpenStreetMap** (UK and Ireland)
  - **Global Administrative Areas** (UK and Ireland)
- Wide variety of questions
- Has been used as a gold standard to test the effectiveness of GeoQA, Li et al. (2021) and Hamzei et al. (2022).
- Available at [https://geoqa.di.uoa.gr/geospatial\\_gold\\_standard.html](https://geoqa.di.uoa.gr/geospatial_gold_standard.html) .

# The GeoQuestions1089 dataset

- **GeoQuestions1089** contains 1089 question-query-answer triples
- Questions crowdsourced from students
- Expanding on **GeoQuestions201**, in both complexity and scope
- Target knowledge graph:
  - **YAGO2**
  - **YAGO2geo** (UK, Ireland, USA and Greece)
- Available at <https://github.com/AI-team-UoA/GeoQuestions1089> .

# GeoQuestions1089 categories

The questions of dataset **GeoQuestions1089** fall under the following nine categories:

A. Asking for a thematic or a spatial attribute of a **feature**

Example:

- Where is **Loch Goil** located?
- What is the geometry of **Loch Goil** ?

# GeoQuestions1089 categories

B. Asking whether a **feature** is in a **geospatial relation** with another **feature** or **features**

Example:

- Is Liverpool **east of** Ireland?
- Is New York City **in** France?

# GeoQuestions1089 categories

C. Asking for features of a given class that are in a **geospatial relation** with another **feature**

Examples:

- Which **counties border** county **Lincolnshire**?
- Which **hotels are less than 2km from** **George Best Belfast City Airport**?

# GeoQuestions1089 categories

D. Asking for features of a given class that are in a **geospatial relation** with any features of another class

Example:

- Which churches are **near** castles?

# GeoQuestions1089 categories

E. Asking for features of a given class that are in a **geospatial relation** with unspecified features of another class, and either one or all, is/are in another **geospatial relation** with a **feature** specified explicitly

Examples:

- Which churches are **near** a castle in **Scotland**?
- In **Greece**, which beaches are **near** villages?

# GeoQuestions1089 categories

F. As in categories C, D and E above, plus more thematic and/or geospatial **attributes** of the features expected as answers

Example:

- Which mountains in Scotland are higher than 1000 meters?



# GeoQuestions1089 categories

G. Questions with **quantities** and **aggregates**

Examples:

- What is the total area of lakes in Oxfordshire?
- How many lakes are there in Epirus?
- How many people live in villages in Attica?

# GeoQuestions1089 categories

H. Questions with **superlatives** or **comparatives**

Examples:

- Which is the largest island in Greece?
- Is the largest island in France larger than Crete?

# GeoQuestions1089 categories

I. Questions with **quantities/aggregates**, and **superlatives/comparatives**

Examples:

- Which city in the UK has the most hospitals?
- Is the total size of lakes in Greece larger than lake Loch Lomond in Scotland?

# Comparison to GeoQuestions201

- **GeoQuestions201:**
  - mostly simple questions that can be answered with simple queries
- **GeoQuestions1089**
  - numerous complex questions that use more advanced SPARQL constructs to be answered
    - **nested queries**
      - Which is the 10th largest town in Alaska?
    - **NOT EXISTS filters**
      - Which islands don't have any lakes but have forests?
    - **arithmetic calculations**
      - How many times bigger is the Republic of Ireland than Northern Ireland?

# Using the dataset to evaluate geospatial QA engines

- We used the dataset **GeoQuestions1089** to benchmark the QA engines **GeoQA2** and the **system of Hamzei et al. (2022)**.

# Evaluation: Methodology

- Methodology and Metrics:
  - The engines generate a query from a natural language question.
  - If the generation is successful it is run against an RDF store .
  - The result is compared to the gold result included in [GeoQuestions1089](#). For the generated query to be considered correct, the results must match exactly.

# Evaluation: Results

Category	Hamzei et al.			GeoQA2		
	Generated Queries	Correct Answers	Correct Answers*	Generated Queries	Correct Answers	Correct Answers*
Type-A	89.71%	10.85%	12.10%	84%	47.42%	56.45%
Type-B	95.68%	53.23%	55.63%	76.25%	58.99%	77.35%
Type-C	97.75%	30.33%	31.03%	79.21%	44.38%	56.02%
Type-D	100%	12%	12.00%	56%	12%	21.42%
Type-E	99.25%	7.40%	7.46%	80%	31.85%	39.81%
Type-F	79.16%	4.10%	5%	66.66%	16.66%	25%
Type-G	98.27%	11.49%	11.69%	74.13%	32.18%	43.41%
Type-H	97.18%	7.74%	7.97%	71.12%	26.05%	36.63%
Type-I	92%	0%	0.00%	84%	20%	23.80%
Total	<b>95.77%</b>	18.97%	19.81%	76.99%	<b>38.54%</b>	<b>50.06%</b>

\* percentage of correct answers over the number of generated queries

# Comments

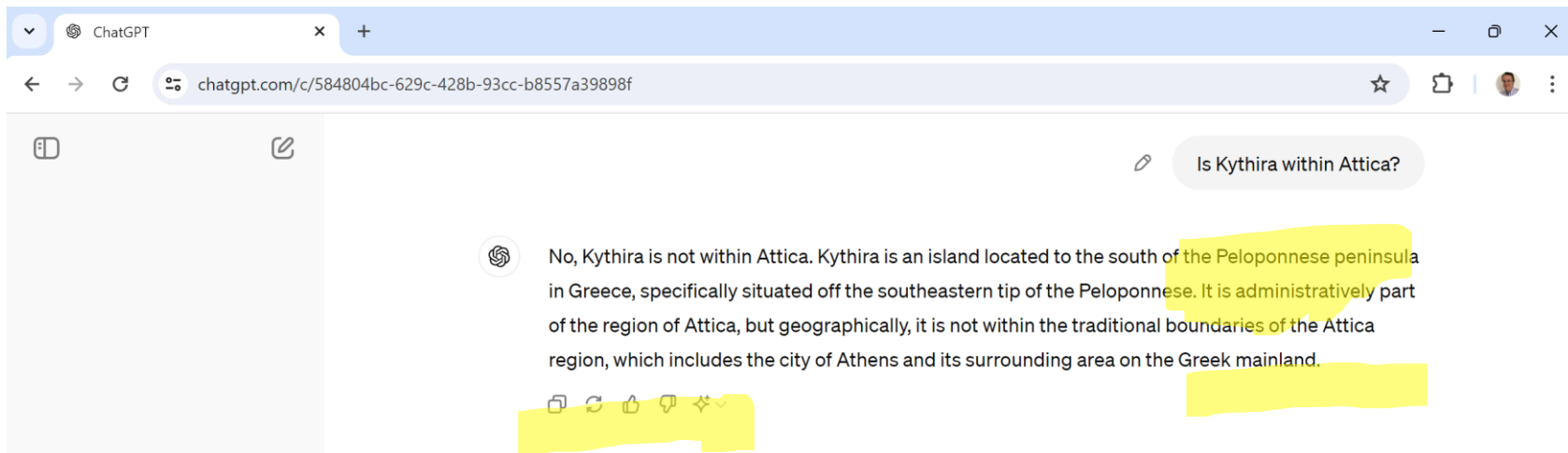
- Consider the user question: “Where is Dublin located?”
- Questions:
  - Shall we include Dublin, Ireland?
  - Shall we include Dublin, Ohio?
  - Shall we include the Kingdom of Dublin? ([https://en.wikipedia.org/wiki/Kingdom\\_of\\_Dublin](https://en.wikipedia.org/wiki/Kingdom_of_Dublin))
- Answer by Hamzei et al. : Yes, and all the other Dublins that exist in the knowledge graph!
- Answer by GeoQA: only Dublin, Ireland should be included.



# ChatGPT for GeoQuestions1089

- It is interesting to see whether chatbots such as **ChatGPT**, **Gemini** or **Claude** can answer correctly the questions of the benchmark GeoQuestions1089.

# Examples

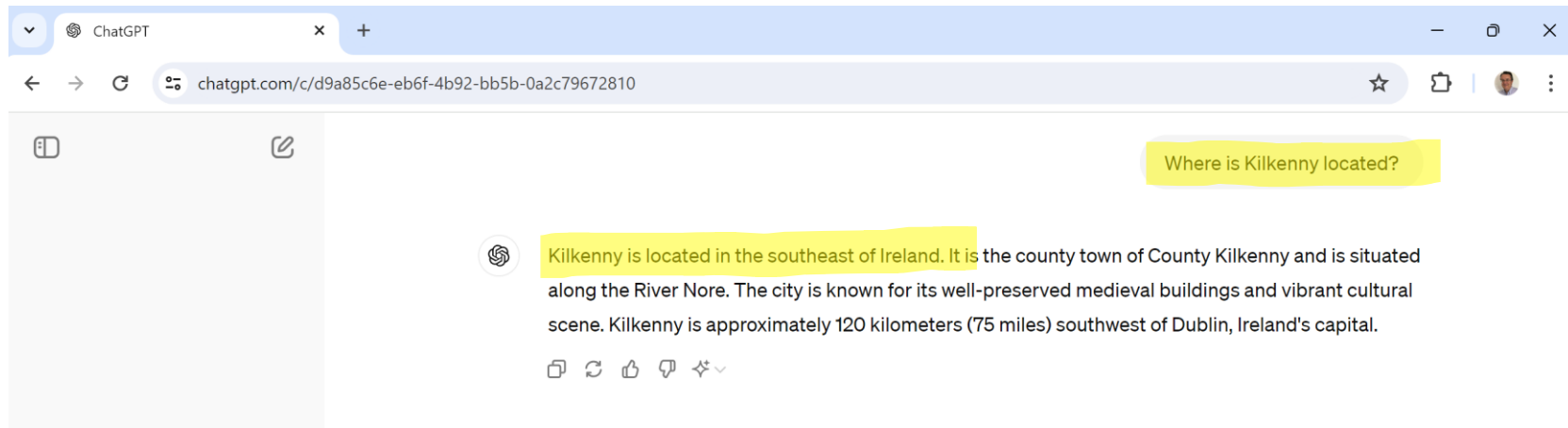


The screenshot shows a browser window with a single tab titled "ChatGPT". The address bar contains the URL "chatgpt.com/c/584804bc-629c-428b-93cc-b8557a39898f". The interface includes a sidebar on the left with a list icon and a share icon. The main chat area shows a user question in a grey bubble: "Is Kythira within Attica?". The AI response, preceded by the OpenAI logo, states: "No, Kythira is not within Attica. Kythira is an island located to the south of the Peloponnese peninsula in Greece, specifically situated off the southeastern tip of the Peloponnese. It is administratively part of the region of Attica, but geographically, it is not within the traditional boundaries of the Attica region, which includes the city of Athens and its surrounding area on the Greek mainland." Below the response are icons for copy, refresh, thumbs up, thumbs down, and star. Several areas of the screenshot are highlighted in yellow, including the AI response text and the interaction icons.

Category B question.

Correct answer.

# Examples



The screenshot shows a browser window with the ChatGPT interface. The address bar displays the URL `chatgpt.com/c/d9a85c6e-eb6f-4b92-bb5b-0a2c79672810`. The user's question, "Where is Kilkenny located?", is highlighted in yellow. The AI's response is also partially highlighted in yellow, showing the beginning of the answer: "Kilkenny is located in the southeast of Ireland. It is the county town of County Kilkenny and is situated along the River Nore. The city is known for its well-preserved medieval buildings and vibrant cultural scene. Kilkenny is approximately 120 kilometers (75 miles) southwest of Dublin, Ireland's capital." Below the response are icons for copy, refresh, thumbs up, thumbs down, and a dropdown menu.

Category A question.

Partially correct answer. It would be nice to have the exact geometry of Kilkenny as well (this is what the benchmark has as answer).

# Examples

ChatGPT ▾

How many lakes overlap with Greek municipalities?



In Greece, there are several lakes that overlap with municipalities. Here are some of the notable examples:

1. **Lake Trichonida** - This is the largest natural lake in Greece and it overlaps with the municipalities of Agrinio, Thermo, and Makryneia in the Aetolia-Acarnania regional unit.
2. **Lake Vegoritida** - This lake is situated in the northern part of Greece and overlaps with the municipalities of Amyntaio, Edessa, and Arnissa in the Florina and Pella regional units.
3. **Lake Prespa** - Located in the northwest, Lake Prespa overlaps with the municipalities of Prespes in the Florina regional unit and has shared borders with North Macedonia and Albania.
4. **Lake Volvi** - Found in the northern part of Greece, this lake overlaps with the municipalities of Volvi and Lagkadas in the Thessaloniki regional unit.
5. **Lake Pamvotida (Ioannina Lake)** - This lake is located in the Epirus region and overlaps with the municipality of Ioannina.
6. **Lake Orestiada** - Situated in the northwest, this lake overlaps with the municipality of Kastoria.

These examples demonstrate that several Greek lakes overlap with one or more municipalities, impacting local administration and often requiring cooperative management across municipal boundaries.



Category G  
question.

Incorrect answer.

# Example – Yesterday around Noon

The screenshot shows a web browser window with the URL `chatgpt.com/c/6c5b8c37-0f7b-43ff-be6c-fe800d3229eb`. The page title is "ChatGPT". On the left sidebar, under "Today", there is a chat entry titled "Corsica smaller than Crete". The main chat area shows a user question: "Is the largest island in France larger than Crete?". The AI response is: "No, the largest island in France, Corsica, is not larger than Crete. Corsica has an area of approximately 8,680 square kilometers, whereas Crete has an area of around 8,336 square kilometers. Therefore, Corsica is slightly smaller than Crete." Below the response are icons for voice, copy, refresh, thumbs up, and thumbs down. A red text overlay at the bottom right of the chat area reads: "Category H question. Incorrect."

# Example (cont'd)

The screenshot shows a web browser window with a single tab titled "ChatGPT". The address bar contains the URL "chatgpt.com/c/add989a7-67cf-4810-b9bd-dce6438bc53f". The browser's sidebar on the left lists recent conversations under "Today" and "Previous 7 Days". The main chat area shows a user message "are you sure?" and a response from ChatGPT. The response includes an apology and corrected information about the areas of Crete and Corsica. At the bottom, there is a feedback prompt: "Is this conversation helpful so far?" with thumbs up and thumbs down icons, and a close button.

ChatGPT

are you sure?

I apologize for the error. Let's correct that information:

Crete is indeed larger than Corsica. The correct areas are:

- **Crete:** Approximately 8,336 square kilometers.
- **Corsica:** Approximately 8,722 square kilometers.

Therefore, Corsica, the largest island in France, is actually smaller than Crete.

Is this conversation helpful so far?

# Let's Try Another Example!

The screenshot shows a web browser window with the URL `chatgpt.com/c/9b6892d9-88ff-425e-ba6e-ea732cde4efd`. The page title is "ChatGPT". On the left sidebar, there is a "Today" section with a highlighted item "Crete Larger Than Cyprus" and other items like "Corsica Vs Crete" and "Corsica smaller than Crete". The main chat area shows a user message: "Is the largest island of Greece larger than Cyprus?". The AI response is: "Yes, the largest island of Greece, Crete, is larger than Cyprus. Crete has an area of approximately 8,336 square kilometers (3,219 square miles), making it the largest Greek island and the fifth largest island in the Mediterranean Sea. In comparison, Cyprus has an area of about 9,251 square kilometers (3,572 square miles), which makes it larger than Crete. So, while Crete is the largest island in Greece, Cyprus is larger than Crete." Below the response are icons for copy, share, and regenerate. At the bottom, there is a text input field with the placeholder "Message ChatGPT" and a send button. A footer note states "ChatGPT can make mistakes. Check important info." and a help icon is visible in the bottom right corner.

ChatGPT

Today

Crete Larger Than Cyprus

Corsica Vs Crete

Corsica smaller than Crete

Previous 7 Days

Kilkenny: Located in Ireland

User Request: Summarize conversati

Kythira not in Attica

Upgrade plan  
Get GPT-4, DALL-E, and more

Is the largest island of Greece larger than Cyprus?

Yes, the largest island of Greece, Crete, is larger than Cyprus.

Crete has an area of approximately 8,336 square kilometers (3,219 square miles), making it the largest Greek island and the fifth largest island in the Mediterranean Sea.

In comparison, Cyprus has an area of about 9,251 square kilometers (3,572 square miles), which makes it larger than Crete.

So, while Crete is the largest island in Greece, Cyprus is larger than Crete.

Message ChatGPT

ChatGPT can make mistakes. Check important info.

# Let's Check Wikipedia

ChatGPT Corsica - Wikipedia en.wikipedia.org/wiki/Corsica

*main article: history of Corsica*

## Contents hide

- (Top)
- > **History**
- > **Geography**
  - Climat
- > **Ecology**
- Demographics
- Immigration
- Languages
- Cuisine
- Art
- Sport
- Administration
- Economy
- > **Transport**
- Politics
- See also
- Notes
- Bibliography

## Prehistory and antiquity [ edit ]

*Main article: Prehistory of Corsica*

Corsica has been occupied since the **Mesolithic** era. The permanent human presence in Corsica is documented in the **Neolithic** period from the 6th millennium BC.<sup>[4]</sup>


After a brief occupation by the **Carthaginians**, *colonization by the ancient Greeks*, and an only slightly longer occupation by the **Etruscans**, it was incorporated by the **Roman Republic** at the end of the **First Punic War** and, with **Sardinia**, in 238 BC became a **province of the Roman Republic**.<sup>[5]</sup> The Romans, who built a colony in **Aléria**, considered Corsica as one of the most backward regions of the Roman world. The island produced sheep, honey, **resin** and wax, and exported many slaves, not well considered because of their fierce and rebellious character.<sup>[5]</sup> Moreover, it was known for its cheap wines, exported to Rome, and was used as a place of relegation, one of the most famous exiles being the Roman philosopher **Seneca**.<sup>[6]</sup>

Corsica was integrated into **Roman Italy** by **Emperor Diocletian** (r. 284–305). Administratively, the island was divided into *pagi*, which in the Middle Ages became the *pievi*, the basic administrative units of the island until 1768.<sup>[5]</sup>

## Middle Ages [ edit ]

*Main article: Medieval Corsica*

In the fifth century, the western half of the Roman Empire collapsed, and the island was invaded by the **Vandals** and the **Ostrogoths**.<sup>[5]</sup> Briefly recovered by the **Byzantine Empire**, it soon became part of the **Kingdom of the Lombards**. This made it a dependency of the **March of Tuscany**, which used it as an outpost against the **Saracens**.<sup>[7]</sup> **Pepin the Short**, king of the **Frank**s and **Charlemaone's** father, expelled the Lombards and nominally granted Corsica to



Location of Corsica within France

<b>Country</b>	<span><span><span></span></span><span> </span></span> France
<b>Prefecture</b>	Ajaccio
<b>Departments</b>	2 <span>[hide]</span> <div>Corse-du-Sud (2A) Haute-Corse (2B)</div>
<b>Government</b> <ul style="list-style-type: none"><li>• <b>President of Executive Council</b></li><li>• <b>President of the Corsican Assembly</b></li><li>• <b>Prefect</b></li></ul>	Gilles Simeoni (Femu a Corsica) Marie-Antoinette Maupertuis Amaury de Saint-Quentin
<b>Area</b> <ul style="list-style-type: none"><li>• <b>Total</b></li></ul>	8,722 km <sup>2</sup> (3,368 sq mi)
<b>Population</b> (2024) <sup>[1]</sup> <ul style="list-style-type: none"><li>• <b>Total</b></li><li>• <b>Density</b></li></ul>	355,528 41/sq mi (16/sq mi)



# Let's Check Wikipedia

island in the Mediterranean Sea, after [Greece](#), [Sardinia](#), [Cyprus](#), and [Corsica](#). Crete rests about 160 km (99 mi) south of the Greek mainland, and about 100 km (62 mi) southwest of [Anatolia](#). Crete has an area of 8,450 km<sup>2</sup> (3,260 sq mi) and a coastline of 1,046 km (650 mi). It bounds the southern border of the [Aegean Sea](#), with the [Sea of Crete](#) (or North Cretan Sea) to the north and the [Libyan Sea](#) (or South Cretan Sea) to the south. Crete covers 260 km from west to east but is narrow from north to south, spanning three longitudes but only half a latitude.

Crete and a number of islands and islets that surround it constitute the Region of Crete (Greek: Περιφέρεια Κρήτης), which is the southernmost of the 13 [top-level administrative units of Greece](#), and the fifth most populous of Greece's regions. Its capital and largest city is [Heraklion](#), on the north shore of the island. As of 2021, the region had a population of 624,408.<sup>[4]</sup> The [Dodecanese](#) are located to the northeast of Crete, while the [Cyclades](#) are situated to the north, separated by the [Sea of Crete](#). The [Peloponnese](#) is to the region's northwest.

Crete was the center of [Europe](#)'s first advanced civilization, the [Minoans](#), from 2700 to 1420 BC. The Minoan civilization was overrun by the [Mycenaean civilization](#) from mainland Greece. Crete was later ruled by Rome, then successively by the [Byzantine Empire](#), [Andalusian Arabs](#), the [Venetian Republic](#), and the [Ottoman Empire](#). In 1898 Crete, whose people had for some time wanted to join the Greek state, achieved independence from the Ottomans, formally becoming the [Cretan State](#). Crete became part of Greece in December 1913.

The island is mostly mountainous, and its character is defined by a high mountain range crossing from west to east. It includes Crete's highest point, [Mount Ida](#), and the range of the [White Mountains](#) (Lefka Ori) with 30 summits above 2,000 metres (6,600 ft) in altitude and the [Samaria Gorge](#), a [World Biosphere Reserve](#). Crete forms a significant part of the economy and cultural heritage of Greece, while retaining its own local cultural traits (such as its own [poetry](#) and [music](#)). The [Nikos Kazantzakis](#) airport at Heraklion and the [Daskalogiannis](#) airport at [Chania](#) serve international travelers. The [Minoan palace](#) at [Knossos](#) is also located in Heraklion.<sup>[5]</sup>

**Geography**

<b>Location</b>	<a href="#">Eastern Mediterranean</a>
<b>Coordinates</b>	<a href="#">35°12.6′N 24°54.6′E</a>
<b>Area</b>	8,450 km <sup>2</sup> (3,260 sq mi)
<b>Area rank</b>	88
<b>Highest elevation</b>	2,456 m (8058 ft)
<b>Highest point</b>	<a href="#">Mount Ida</a> (Psiloritis)

# Let's Check Wikipedia

led to the capture of the present-day territory of [Northern Cyprus](#) and the [displacement](#) of over 150,000 Greek Cypriots<sup>[24][25]</sup> and 50,000 Turkish Cypriots.<sup>[26]</sup> A separate Turkish Cypriot state in the north was [established by unilateral declaration](#) in 1983; the move was widely condemned by the [international community](#), with Turkey alone [recognising](#) the new state. These events and the resulting political situation are matters of a [continuing dispute](#).

Cyprus is a major tourist destination in the Mediterranean.<sup>[27][28][29]</sup> The country has an [advanced high-income economy](#). The Republic of Cyprus has been a member of the [Commonwealth](#) since 1961 and was a [founding member](#) of the [Non-Aligned Movement](#) until it [joined the European Union](#) on 1 May 2004.<sup>[30]</sup> On 1 January 2008, the Republic of Cyprus joined the [eurozone](#).<sup>[31]</sup>

## Etymology

The earliest attested reference to *Cyprus* is the 15th century BC [Mycenaean Greek](#) 𐀓𐀆𐀗𐀆𐀓, *ku-pi-ri-jo*,<sup>[32]</sup> meaning "Cypriot" (Greek: Κύπριος), written in [Linear B](#) syllabic script.<sup>[33]</sup> The classical Greek form of the name is Κύπρος (*Kýpros*).

The etymology of the name is unknown. Suggestions include:

- the Greek word for the [Mediterranean cypress tree](#) (*Cupressus sempervirens*), κυπάρισσος (*kypárisσos*)
- the Greek name of the [henna tree](#) (*Lawsonia alba*), κύπρος (*kýpros*)
- an [Eteocypriot](#) word for copper. It has been suggested, for example, that it has roots in the [Sumerian](#) word for copper (*zubar*) or for [bronze](#) (*kubar*), from the large deposits of copper ore found on the island.<sup>[34]</sup>

Through overseas trade, the island has given its name to the [Classical Latin](#) word for copper

<b>Vernaculars</b>	Cypriot Greek · Cypriot Turkish
<b>Ethnic groups</b>	Greek Cypriots Turkish Cypriots Armenian Cypriots Maronite Cypriots
<b>Religion</b> (2020; including Northern Cyprus)	72.3% Christianity 25.0% Islam 1.9% no religion 0.8% other <sup>[2]</sup>
<b>Demonym(s)</b>	Cypriot
<b>Government</b>	Unitary presidential republic
<span> </span> • <b>President</b>	Nikos Christodoulides
<span> </span> • <b>Vice-President</b>	Vacant <sup>[b]</sup>
<span> </span> • <b>President of the House of Representatives</b>	Annita Demetriou
<b>Legislature</b>	House of Representatives
<b>Independence</b> from the United Kingdom	
<span> </span> • <b>London-Zürich Agreements</b>	19 February 1959
<span> </span> • <b>Independence proclaimed</b>	16 August 1960
<span> </span> • <b>Independence Day</b>	1 October 1960
<span> </span> • <b>Joined the EU</b>	1 May 2004
<b>Area</b>	
<span> </span> • <b>Total</b> <sup>[c]</sup>	9,251 <span> </span> km <sup>2</sup> (3,572 <span> </span> sq <span> </span> mi) (162nd)
<span> </span> • <b>Water</b> (%)	0.11 <sup>[3]</sup>

<https://en.wikipedia.org/wiki/Cyprus#>

# Example (cont'd) - 18:28 yesterday

ChatGPT

chatgpt.com/c/c4aebc3e-9a68-4d3c-8710-1d927154e780

Today

Largest French island: Corsica. ...

Crete Larger Than Cyprus

Corsica Vs Crete

Corsica smaller than Crete

Previous 7 Days

Kilkenny: Located in Ireland

User Request: Summarize conversat

Kythira not in Attica

Is the largest island of France larger than Crete?

Yes, the largest island of France, Corsica, is larger than Crete. Here are the details:

- **Corsica:** It has an area of approximately 8,680 square kilometers.
- **Crete:** It has an area of approximately 8,336 square kilometers.

Therefore, Corsica is indeed larger than Crete.

Is this conversation helpful so far?

# Example (cont'd) – 18:30 yesterday

The screenshot shows a web browser window with a single tab titled "ChatGPT". The address bar displays the URL "chatgpt.com/c/355b4b39-8b9d-4820-96cc-111c5706063f". The interface includes a sidebar on the left with a list of chat sessions under "Today" and "Previous 7 Days". The main chat area shows a user's question: "is the largest island of Greece larger than Cyprus:". The AI's response is: "Yes, the largest island of Greece, Crete, is larger than Cyprus. Here are the specifics: • **Crete:** The largest island of Greece, with an area of approximately 8,336 square kilometers. • **Cyprus:** An independent country and an island in the eastern Mediterranean Sea, with an area of about 9,251 square kilometers. Therefore, Cyprus is actually larger than Crete." Below the response are icons for voice, copy, refresh, thumbs up, and thumbs down. At the bottom, there is a feedback prompt: "Is this conversation helpful so far?" with thumbs up, thumbs down, and close buttons.

ChatGPT

is the largest island of Greece larger than Cyprus:

Yes, the largest island of Greece, Crete, is larger than Cyprus.

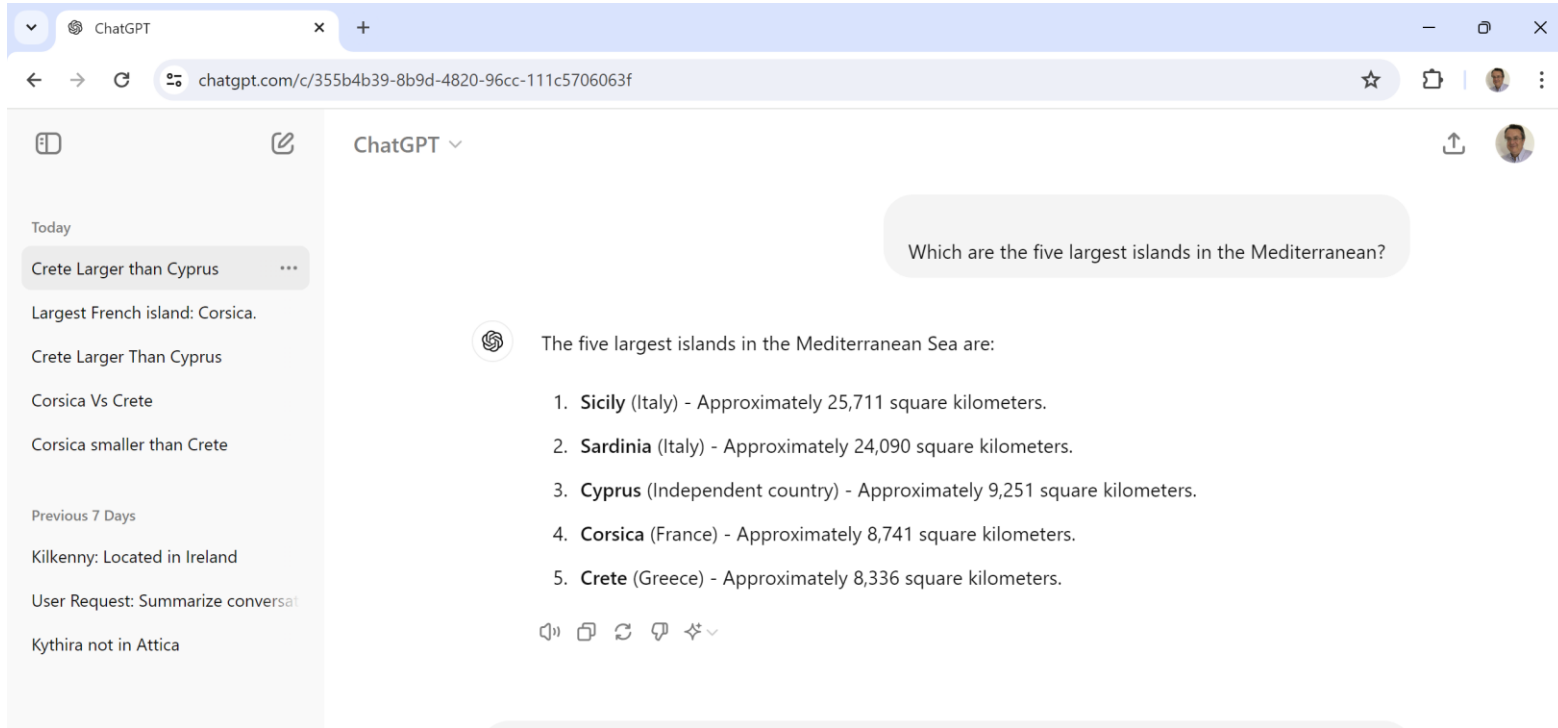
Here are the specifics:

- **Crete:** The largest island of Greece, with an area of approximately 8,336 square kilometers.
- **Cyprus:** An independent country and an island in the eastern Mediterranean Sea, with an area of about 9,251 square kilometers.

Therefore, Cyprus is actually larger than Crete.

Is this conversation helpful so far?

# In Case you are Wondering 😊



The screenshot shows a web browser window with a single tab titled "ChatGPT". The address bar displays the URL "chatgpt.com/c/355b4b39-8b9d-4820-96cc-111c5706063f". The chat interface includes a sidebar on the left with a search icon and a "ChatGPT" header. The sidebar lists several chat topics under "Today" and "Previous 7 Days". The main chat area shows a user question in a grey bubble: "Which are the five largest islands in the Mediterranean?". Below it, the AI response is displayed, starting with the OpenAI logo and the text "The five largest islands in the Mediterranean Sea are:". This is followed by a numbered list of five islands with their approximate areas in square kilometers. At the bottom of the response, there are icons for voice, copy, refresh, share, and a star.

ChatGPT

Today

- Crete Larger than Cyprus
- Largest French island: Corsica.
- Crete Larger Than Cyprus
- Corsica Vs Crete
- Corsica smaller than Crete

Previous 7 Days

- Kilkenny: Located in Ireland
- User Request: Summarize conversat
- Kythira not in Attica

Which are the five largest islands in the Mediterranean?

The five largest islands in the Mediterranean Sea are:

1. **Sicily** (Italy) - Approximately 25,711 square kilometers.
2. **Sardinia** (Italy) - Approximately 24,090 square kilometers.
3. **Cyprus** (Independent country) - Approximately 9,251 square kilometers.
4. **Corsica** (France) - Approximately 8,741 square kilometers.
5. **Crete** (Greece) - Approximately 8,336 square kilometers.

# Results

<b>Category</b>	<b>Correctly Answered</b>	<b>Partially Correctly Answered</b>
Category A	15.4%	59.4%
Category B	61.4%	0%
Category C	31.3%	51.1%
Category D	23.8%	52.4%
Category E	18.8%	63%
Category F	4%	8%
Category G	6.3%	3.4%
Category H	38.7%	0%
Category I	4%	0%
Total	22.6%	26.4%

# Let's Try Gemini!

The screenshot shows a web browser window with the Gemini application. The address bar shows the URL `gemini.google.com/app/2e072d7041e2422e`. The page header includes a hamburger menu, the text "Gemini", and a "Try Gemini Advanced" button with a user profile picture. A dark blue notification bar at the top reads "See the latest updates to the [Gemini Apps Privacy Hub](#)".

The main content area displays a user question: "Is the largest island of France larger than Crete?". The AI response is as follows:

◆ The largest island of France is Corsica, with an area of roughly 8,680 square kilometers [Wikipedia list of islands of France].

Crete, on the other hand, boasts an area of around 8,330 square kilometers [Wikipedia Crete].

Therefore, Corsica is indeed larger than Crete by approximately 350 square kilometers.

Below the response are icons for "Show drafts" and a speaker icon. At the bottom, there is a text input field with the placeholder "Enter a prompt here" and icons for image upload and voice input. A footer note states: "Gemini may display inaccurate info, including about people, so double-check its responses. [Your privacy and Gemini Apps](#)".

# Let's Try Gemini!

The screenshot shows the Gemini web interface in a browser. The browser's address bar displays the URL `gemini.google.com/app/2e072d7041e2422e`. The Gemini header includes a menu icon, the name "Gemini", a "Try Gemini Advanced" button, and a user profile icon. A dark blue notification banner at the top reads "See the latest updates to the [Gemini Apps Privacy Hub](#)".

The main content area shows a user prompt: "Is the largest island of Greece larger than Cyprus?". Below the prompt, the Gemini response is displayed, starting with a blue star icon. The response text is: "Yes, the largest island of Greece, Crete, is larger than Cyprus." This is followed by a bulleted list of facts: "Crete has an area of approximately 8,312 square kilometers [Wikipedia List of European islands by area]." and "Cyprus has an area of roughly 9,251 square kilometers [Wikipedia List of European islands by area].". A follow-up paragraph states: "Even though Cyprus has a higher number, Crete is geographically bigger. There might be some confusion because the political situation in Cyprus is complex, with a portion being claimed by Turkey." To the right of the response, there are "Show drafts" and a speaker icon.

At the bottom, there is a text input field with the placeholder "Enter a prompt here" and icons for image upload and voice input.



# The Dataset of Yang et al. (2024)

- It was created as part of the GeoKBQA engine work.
- It starts from GeoQuestions1089 and:
  - **Reorganizes the GeoQuestions1089 classes** based on their corresponding GeoSPARQL queries **to create templates** that will be used for question generation.
  - **Substitutes entities, classes, and spatial functions** to augment the diversity of schema items and spatial functions, and to expand the scale of the dataset.
  - **Paraphrases** the resulting geospatial questions using ChatGPT (GPT-3.5 Turbo) to enhance natural language diversity.
  - **Selects question–query pairs** based on successful queries to the OpenStreetMap database, choosing pairs **that yield results**.
- 4468 question-query pairs are generated in this way.

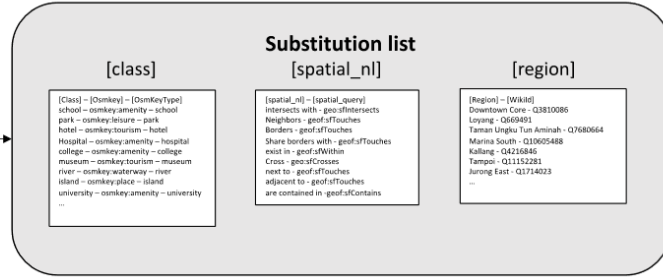
# Example

## [Step 1]

### Which [class] [spatial\_nl] [region]?

```
PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
PREFIX osmkey: <https://www.openstreetmap.org/wiki/Key:>
PREFIX osm: <https://www.openstreetmap.org/>
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX strdf: <http://strdf.di.uoa.gr/ontology#>
```

```
SELECT DISTINCT ?class WHERE {
  ?region1 osmkey:wikidata "[wikid]" .
  ?region1 geo:hasGeometry ?rwtk1 .
  ?class [osmkey] "[osmkeyType]" .
  ?class geo:hasGeometry ?cwk1 .
  FILTER([spatial_query](?rwtk1, ?cwk1))
}
```



## [Step 2]

### Which road cross Jurong Island?

```
PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
PREFIX osmkey: <https://www.openstreetmap.org/wiki/Key:>
PREFIX osm: <https://www.openstreetmap.org/>
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX strdf: <http://strdf.di.uoa.gr/ontology#>
```

```
SELECT DISTINCT ?class WHERE {
  ?region1 osmkey:wikidata "'Q14439130"' .
  ?region1 geo:hasGeometry ?rwtk1 .
  ?class osmkey:route "road" .
  ?class geo:hasGeometry ?cwk1 .
  FILTER(geo:stCrosses(?rwtk1, ?cwk1))
}
```

## [Step 3]

### Which road cross Jurong Island? Can you tell me which road crosses Jurong Island? What road crosses Jurong Island?

```
PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
PREFIX osmkey: <https://www.openstreetmap.org/wiki/Key:>
PREFIX osm: <https://www.openstreetmap.org/>
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX strdf: <http://strdf.di.uoa.gr/ontology#>
```

```
SELECT DISTINCT ?class WHERE {
  ?region1 osmkey:wikidata "'Q14439130"' .
  ?region1 geo:hasGeometry ?rwtk1 .
  ?class osmkey:route "road" .
  ?class geo:hasGeometry ?cwk1 .
  FILTER(geo:stCrosses(?rwtk1, ?cwk1))
}
```

ChatGPT

# Results

- The dataset is used to evaluate GeoKBQA targeting OpenStreetMap data of Singapore. The results are the following:

	<b>T5 Base Model (220M)</b>		<b>T5 Large Model (770M)</b>	
	<b>ELQ</b>	<b>Golden</b>	<b>ELQ</b>	<b>Golden</b>
T5	51.7857	71.6518	75.8929	79.0179
FLAN-T5	62.2768	73.4375	78.5714	80.5804
CodeT5	84.375	83.4821	91.5179	92.6339
CodeT5+	93.3036	92.4107	93.0804	94.1964

# Open Problems

- Answering **spatiotemporal questions** over knowledge graphs
- Example: What are today the names and borders of the countries that made up Yugoslavia in 1991? Show the borders on a map.
- Question: are there spatiotemporal knowledge graphs (i.e., graphs that show the evolution of spatial regions over time)?
  - **GeoChangeViz**. See [https://ceur-ws.org/Vol-3632/ISWC2023\\_paper\\_423.pdf](https://ceur-ws.org/Vol-3632/ISWC2023_paper_423.pdf)

# Open Problems

- Utilizing existing large language models in geospatial question answering engines (e.g., detect query intent, improve GeoSPARQL query generation etc.).

# Open Problems

- Investigating whether current LLMs can do spatial reasoning
  - Cohn et al. (2023) <https://arxiv.org/abs/2304.11164>
  - Anthony Cohn (2023) <https://arxiv.org/abs/2309.15577>

# Open Problems

- Developing foundation models for geodata and use them for toponym and location recognition and disambiguation, entity resolution, spatial reasoning, question answering etc.
  - Mai et al. (2023) <https://arxiv.org/abs/2304.06798>
  - Balsebre et al. (2024) <https://arxiv.org/abs/2403.09059>

# Advertisement – European Summer School in AI 2024

- Takes place in Athens
- July 15 to 26, 2024 (2 weeks)
- 30 week-long courses
- 11 two hour long advanced tutorials
- EurAI travel grants covering registration are available
- Visit <https://essai2024.di.uoa.gr/> for more!

ESSAI & ACAI 2024  
ATHENS - GREECE

European Summer School on Artificial Intelligence

Advanced Course on Artificial Intelligence

15 - 26  
JULY 2024  
ATHENS, GREECE

<https://essai2024.di.uoa.gr/>

Organized Jointly as 4th Tailor Summer School on Artificial Intelligence

ORGANIZERS:  
EurAI TAILOR EEC ATEAM

VENUE:  
National Centre for Scientific Research "Demokritos"

www.essai.gr Email: registration\_essay2024@di.uoa.gr



# Acknowledgement

- This work has been supported by Horizon Europe project STELAR (<https://stelar-project.eu/>) Grant No. 101070122.



# Thank you!

Visit the web page of my group: <https://ai.di.uoa.gr/>  
Follow us on twitter @AITeamUoA, @mkoubarakis

