## Querying the Lattice of Measurements

Here we describe the queries in natural and in SPARQL Language that can be answered with our approach. The queries are categorized into two different groups:

* **Q1-Q6**: These queries shows how we can exploit the measurements in order to discover which dataset is relevant to our dataset.
* **Q7-Q12**: These queries are a global entity lookup service. Indeed, one can find which datasets contain an entity and the equivalent URIs of this entity.

The queries can be executed in <http://62.217.127.118:8890/fct/demo_queries.vsp> .

Users can also run their own queries in <http://62.217.127.118:8890/sparql> .

## Queries expressed in Natural language

The table below shows the Queries in ***Natural Language.***

#### Demo Queries

|  |  |
| --- | --- |
| Q1. | Give me the top-K (here K=10) most connected datasets to my dataset (here FishBase) |
| Q2. | Give me all the connected sources with FishBase, and how many URIs these datasets share with datasets from the geographical domain |
| Q3. | Give me all pairs of sources that were not connected, but now they are connected due to closure and the number of their common real world objects |
| Q4. | Give me the increase of the commonalities of all pairs of sources  due to closure in descending order |
| Q5. | Give me the K (here K=4) datasets that maximize the pluralism factor of the entities in my dataset (here FishBase) |
| Q6. | Give me the connected triads from datasets coming from three different domains |
| Q7. | Give me all the datasets that contain information about <http://dbpedia.org/resource/Aristotle> |
| Q8. | Give me all the equivalent URIs of <http://dbpedia.org/resource/Aristotle> |
| Q9. | Give me all the datasets from the publication domain containing information about yago:Socrates |
| Q10. | Give me all the URIs that are equivalent with the URIs of my dataset (here FishBase) |
| Q11. | Give me the datasets that contain information for both Aristotle and Socrates |
| Q12. | Give me all the common RWO between Wikidata, DBpedia and Yago |

#### Supplementary Queries

|  |  |
| --- | --- |
| S1. | Give me all the domains |
| S2. | Give me all the datasets and their domain |

## Queries expressed in SPARQL Language

The table below shows the demo Queries in ***SPARQL Language. One can modify these queries in order to find information about the datasets or the entities of his interest.***

|  |
| --- |
| ***Q1.* *Give me the top-K (here K=10) most connected datasets to my dataset (here FishBase)*** |
| select ?Dataset ?commonRWO where  {  ?m1 rdf:type voidwh:Measurement .  ?m1 voidwh:usesMetrics voidwh:commonPairsRWO .  ?m1 voidwh:over <http://www.fishbase.org> .  ?m1 voidwh:over ?Dataset .  ?m1 voidwh:producesValue ?commonRWO .  ?m1 voidwh:accordingTo voidwh:EntityMatching .  filter(?Dataset!=<http://www.fishbase.org>)  } order by desc(?v)  limit 10 |

|  |
| --- |
| ***Q2. Give me all the connected sources with FishBase, and how many URIs these datasets share with datasets from the geographical domain*** |
| select ?Dataset1 ?Dataset2 ?commonRWO where  {  ?m1 rdf:type voidwh:Measurement .  ?m1 voidwh:usesMetrics voidwh:commonPairsRWO .  ?m1 voidwh:over <http://www.fishbase.org> .  ?m1 voidwh:over ?Dataset1 .  ?m1 voidwh:producesValue ?v .  ?m1 voidwh:accordingTo voidwh:EntityMatching .  ?m2 rdf:type voidwh:Measurement .  ?m2 voidwh:usesMetrics voidwh:commonPairsRWO .  ?m2 voidwh:over ?Dataset1.  ?m2 voidwh:over ?Dataset2 .  ?m2 voidwh:producesValue ?commonRWO .  ?m2 voidwh:accordingTo voidwh:EntityMatching .  ?Dataset2 dcterms:subject <http://dbpedia.org/resource/Location> .  filter(?Dataset1!=<http://www.fishbase.org>) .filter(?Dataset1!=?Dataset2)  } order by desc(?commonRWO) |

|  |
| --- |
| ***Q3. Give me all pairs of sources that were not connected, but now they are connected due to closure and the number of their common real world objects*** |
| select ?Dataset1 ?Dataset2 ?commonRWO  FROM <http://www.ics.forth.gr/isl/LODCloudMetrics> where {  ?m1 a voidwh:Measurement .  ?m1 voidwh:usesMetrics voidwh:commonPairsRWO .  ?m1 voidwh:over ?Dataset1 .  ?m1 voidwh:over ?Dataset2 .  ?m1 voidwh:producesValue ?commonRWO .  ?m1 voidwh:accordingTo voidwh:EntityMatching .  FILTER NOT EXISTS{  ?m2 a voidwh:Measurement .  ?m2 voidwh:usesMetrics voidwh:commonPairsRWO .  ?m2 voidwh:over ?Dataset1.  ?m2 voidwh:over ?Dataset2.  ?m2 voidwh:accordingTo voidwh:EntityMatchingBeforeClosure .  }  .filter(?Dataset1>?Dataset2) } order by desc(?commonRWO) |

|  |
| --- |
| ***Q4. Give me the increase of the commonalities of all pairs of sources due to closure in descending order*** |
| select ?Dataset1 ?Dataset2 (?value-?valueOLD) as ?increase  from <http://www.ics.forth.gr/isl/LODCloudMetrics>  where {  ?m1 rdf:type voidwh:Measurement .  ?m1 voidwh:usesMetrics voidwh:commonPairsRWO .  ?m1 voidwh:accordingTo voidwh:EntityMatching.  ?m1 voidwh:over ?Dataset1 .  ?m1 voidwh:over ?Dataset2 .  ?m1 voidwh:producesValue ?value .  ?m2 rdf:type voidwh:Measurement .  ?m2 voidwh:usesMetrics voidwh:commonPairsRWO.  ?m2 voidwh:accordingTo voidwh:EntityMatchingBeforeClosure.  ?m2 voidwh:over ?Dataset1 .  ?m2 voidwh:over ?Dataset2 .  ?m2 voidwh:producesValue ?valueOLD  .filter(?Dataset1>?Dataset2)  .FILTER(?value>?valueOLD)  } order by desc(?increase) |

|  |
| --- |
| ***Q5.******Give me the K (here K=4) datasets that maximize the pluralism factor of the entities in my dataset (here FishBase)*** |
| select <http://www.fishbase.org> as ?Dataset1 ?Dataset2 ?Dataset3 ?Dataset4 ?commonRWO where{  ?m1 a voidwh:Measurement .  ?m1 voidwh:usesMetrics voidwh:commonQuadsRWO .  ?m1 voidwh:over <http://www.fishbase.org> .  ?m1 voidwh:over ?Dataset2 .  ?m1 voidwh:over ?Dataset3 .  ?m1 voidwh:over ?Dataset4 .  ?m1 voidwh:producesValue ?commonRWO .  ?m1 voidwh:accordingTo voidwh:EntityMatching .  FILTER(?Dataset2!=<http://www.fishbase.org>) .  FILTER(?Dataset3!=<http://www.fishbase.org>) .  FILTER(?Dataset4!=<http://www.fishbase.org>) .  FILTER(?Dataset4<?Dataset3)  .FILTER (?Dataset3<?Dataset2)  } order by desc(?commonRWO) |

|  |
| --- |
| ***Q6. Give me the connected quads from datasets coming from four different domains*** |
| select ?Dataset1 ?Dataset1Domain ?Dataset2 ?Dataset2Domain  ?Dataset3 ?Dataset3Domain ?commonRWO  where{  ?m1 rdf:type voidwh:Measurement .  ?m1 voidwh:usesMetrics voidwh:commonTriadsRWO .  ?m1 voidwh:over ?Dataset1 .  ?Dataset1 dcterms:subject ?Dataset1Domain .  ?m1 voidwh:over ?Dataset2 .  ?Dataset2 dcterms:subject ?Dataset2Domain .  ?m1 voidwh:over ?Dataset3 .  ?Dataset3 dcterms:subject ?Dataset3Domain .  ?m1 voidwh:producesValue ?commonRWO  .filter(str(?Dataset1)<str(?Dataset2)) .  filter(str(?Dataset2)<str(?Dataset3)) .  filter(str(?Dataset3)>str(?Dataset1))  .filter(?Dataset1Domain!=?Dataset2Domain)  .filter(?Dataset1Domain!=?Dataset3Domain)  .filter(?Dataset2Domain!=?Dataset3Domain)  } order by desc(?commonRWO) limit 50 |

|  |
| --- |
| ***Q7. Give me all the datasets that contain information about <***[***http://dbpedia.org/resource/Aristotle***](http://dbpedia.org/resource/Aristotle)***>*** |
| select distinct ?dataset where {  <http://dbpedia.org/resource/Aristotle> dcterms:identifier ?SID  . ?equivalentURI dcterms:identifier ?SID  .?equivalentURI dcterms:provenance ?dataset  } |

|  |
| --- |
| ***Q8. Give me all the equivalent URIs of <***[***http://dbpedia.org/resource/Aristotle***](http://dbpedia.org/resource/Aristotle)***>*** |
| select distinct ?equivalentURI where {  <http://dbpedia.org/resource/Aristotle> dcterms:identifier ?SID  .?equivalentURI dcterms:identifier ?SID  } |

|  |
| --- |
| ***Q9. Give me all the datasets from the publication domain containing information about yago:Socrates*** |
| select distinct ?dataset where{  <http://yago-knowledge.org/resource/Socrates> dcterms:identifier ?SID .  ?equivalentURI dcterms:identifier ?SID .  ?equivalentURI dcterms:provenance ?dataset .  ?dataset dcterms:subject <http://dbpedia.org/resource/Publication>  } |

|  |
| --- |
| ***Q10. Give me all the URIs that are equivalent with the URIs of my dataset (here FishBase)*** |
| select ?myDatasetURI ?Dataset2 ?Dataset2URI where{  ?myDatasetURI dcterms:identifier ?SID .  ?Dataset2URI dcterms:identifier ?SID .  ?myDatasetURI dcterms:provenance <http://www.fishbase.org> .  ?Dataset2URI dcterms:provenance ?Dataset2 .  filter(?Dataset2!=<http://www.fishbase.org>)  } LIMIT 50 |

|  |
| --- |
| ***Q11. Give me the datasets that contain information for both*** [***Aristotle***](Aristotle) ***and*** [***Socrates***](http://dbpedia.org/resource/Socrates) |
| select distinct ?Dataset where {  <http://dbpedia.org/resource/Aristotle> dcterms:identifier ?SID .  ?uri1 dcterms:identifier ?SID .  ?uri1 dcterms:provenance ?Dataset .  <http://dbpedia.org/resource/Socrates> dcterms:identifier ?SID2 .  ?uri2 dcterms:identifier ?SID2 .  ?uri2 dcterms:provenance ?Dataset .  } |

|  |
| --- |
| ***Q12. Give me all the common RWO between Wikidata, DBpedia and Yago*** |
| select ?Wikipedia ?DBpedia ?Yago where{  ?Wikipedia dcterms:identifier ?SID .  ?DBpedia dcterms:identifier ?SID .  ?Yago dcterms:identifier ?SID.  ?Wikipedia dcterms:provenance <http://www.wikidata.org/> .  ?DBpedia dcterms:provenance <http://dbpedia.org/> .  ?Yago dcterms:provenance <http://yago-knowledge.org> .  } limit 100 |

#### Supplementary Queries

|  |
| --- |
| ***S1. Give me all the domains*** |
| select distinct ?domain where{  ?dataset rdf:type void:Dataset .  ?dataset dcterms:subject ?domain  } |

|  |
| --- |
| ***S2. Give me all the datasets*** |
| select distinct ?dataset where{  ?dataset rdf:type void:Dataset  } |