KnowWhereGraph

Overview & Step-by-Step Examples

Outline:

1. Introduction

The KnowWhereGraph (KWG) is an openly accessible, W3C-standards-based geographically enriched knowledge graph, with associated tools for representing, exploring and visualizing a growing list of human and environmental data. With sophisticated search capabilities and seamless access to numerous public datasets, KWG rapidly raises the situational awareness of data scientists and decision makers by providing detailed area briefings for any place on Earth within seconds.

KWG incorporates custom ontologies and uses a semantically-enabled hierarchical grid for spatial representations. As of April 2022, its size exceeds 12 billion information triples with over 100 classes representing observations of natural hazards (e.g., hurricanes, wildfires, smoke plumes), spatial characteristics related to climate (e.g., temperature, precipitation, air quality), soil properties, crop and land-cover types, demographics, human health, and others (Table 2, *The KnowWhereGraph Schema*). In addition, new information is continually being incorporated as needed, to support extended and new Use Cases.

Two main application areas serve as initial testbeds for KWG: humanitarian disaster relief, and food-related supply chains, including agricultural land use. With guidance from humanitarian organization Direct Relief, we are developing tools to give their responders rapid access to area briefings following a disaster. These include information on previous disaster events, physical properties of the affected regions such as climate variables, medical and transportation infrastructure, and socio-economic and health characteristics of the human populations affected. In addition, KWG can recommend medical and other experts who are likely relevant to a

particular disaster response effort by leveraging schema connections between disasters, types of expertise, and places.

In partnership with the Food Industry Association (FMI), we are using KWG to connect data related to climate, soil properties, crop and land-cover types, disasters and other topics that are traditionally siloed. One key goal is to enhance strategic planning during disasters by providing online analysis, forecasting, and alerts to ensure key stakeholders throughout the supply chain are ready with backup strategies to keep products moving. KWG can also help farmers and growers to identify mitigation strategies and build resilience in the face of such events.

The project is funded by the National Science Foundation as part of its Convergence Accelerator program's Open Knowledge Network initiative, Award nos. 2033521 and 1936677. The team includes members from academia (University of California, Santa Barbara; Kansas State University; Michigan State University; Arizona State University; University of Southern California), the nonprofit sector (Direct Relief); industry (Esri; Oliver Wyman; Hydronos Labs), and the US federal government (US Geological Survey; US Department of Agriculture).

For further information, see

- Know, Know Where, KnowWhereGraph: A Densely Connected, Cross-Domain Knowledge Graph and Geo-Enrichment Service Stack for Applications in Environmental Intelligence (*AI Magazine*, March 2022)
- KnowWhereGraph Drives Cross-Domain Knowledge (*ArcUser*, Spring 2021)

2. The Graph

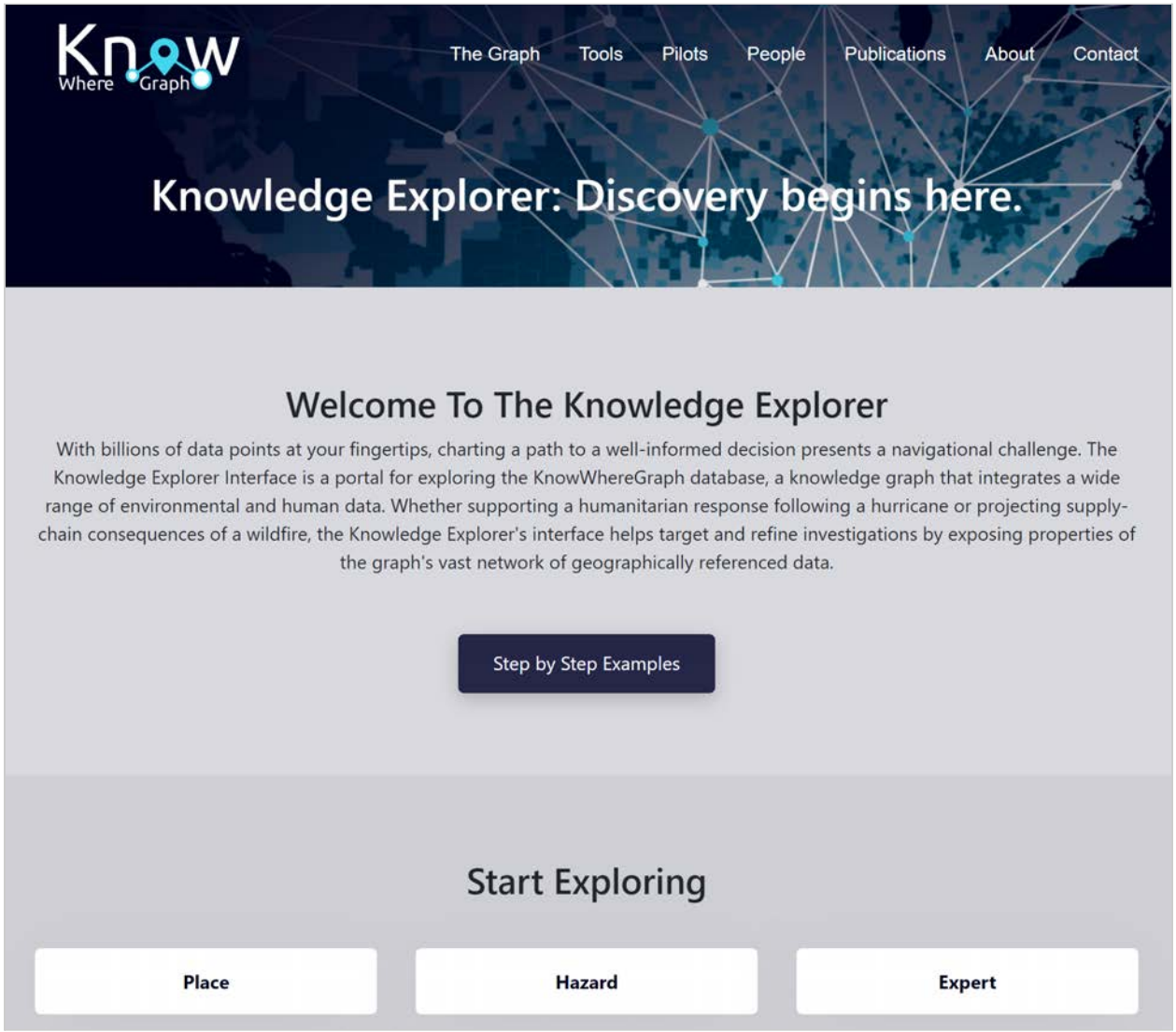
The semantic backbone of the KWG graph is the KnowWhereGraph Schema, which describes concepts of individual domain datasets as extensions of concepts from open standard ontologies, including SOSA, GeoSPARQL, and OWL-Time. The main goals of this schema are to connect previously isolated datasets and establish cross-dataset interoperability. For example, it connects data characterizing a hurricane's trajectory and wind speed with data quantifying human and property damage. We accomplish this by modeling hazard events and environmental observations, as well as their connections via place and time. KWG is a triplestore graph serving up a SPARQL endpoint running on GraphDB Enterprise software, that is kindly provided by Ontotext.

For general purposes, the graph can currently be accessed in two ways:

- Browsing via the Knowledge Explorer semantic browsing tool (Section 3)
- Direct query using SPARQL (Section 4)

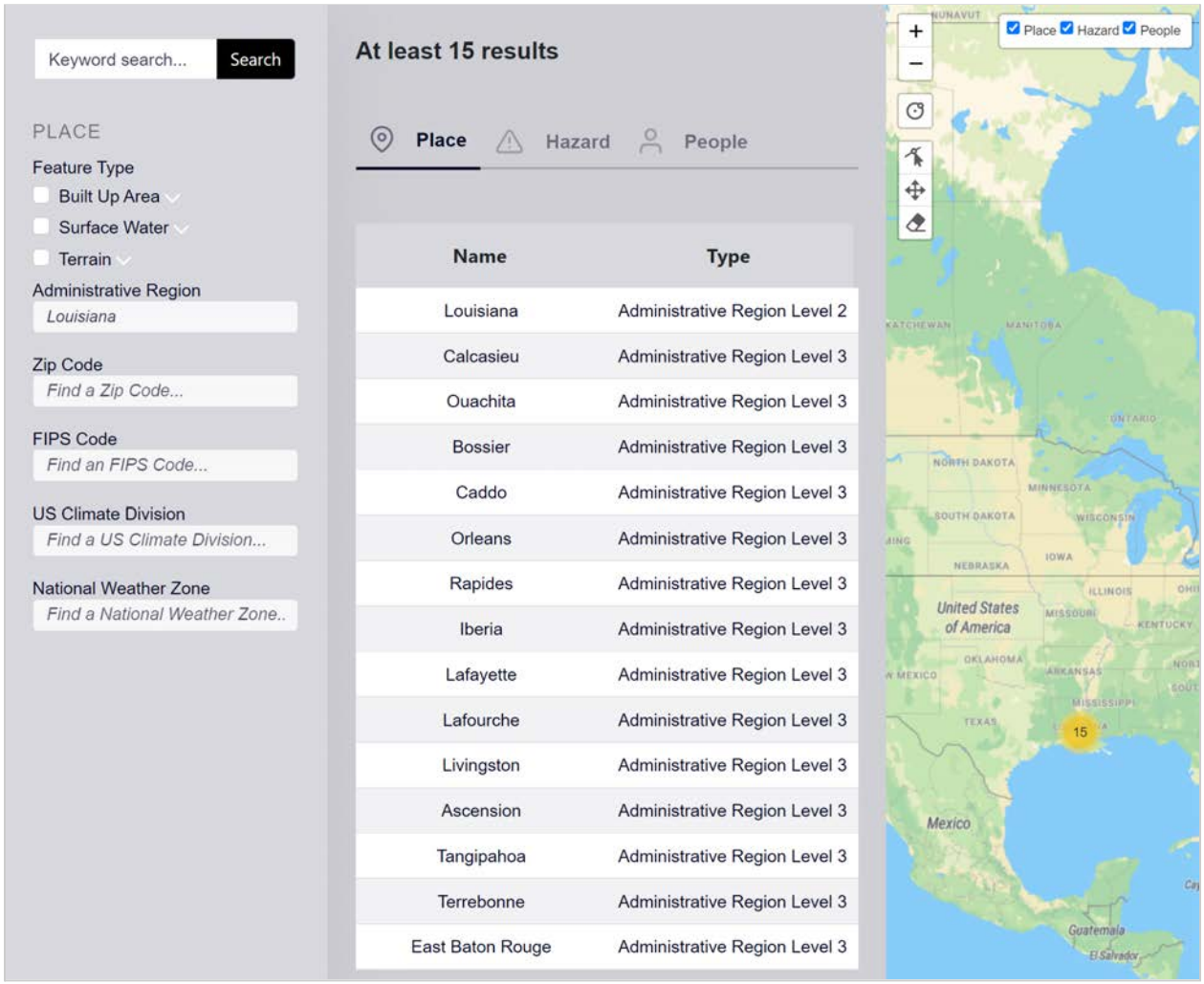
3. The Knowledge Explorer

Access the main landing page of the Knowledge Explorer at https://stko-kwg.geog.ucsb.edu.



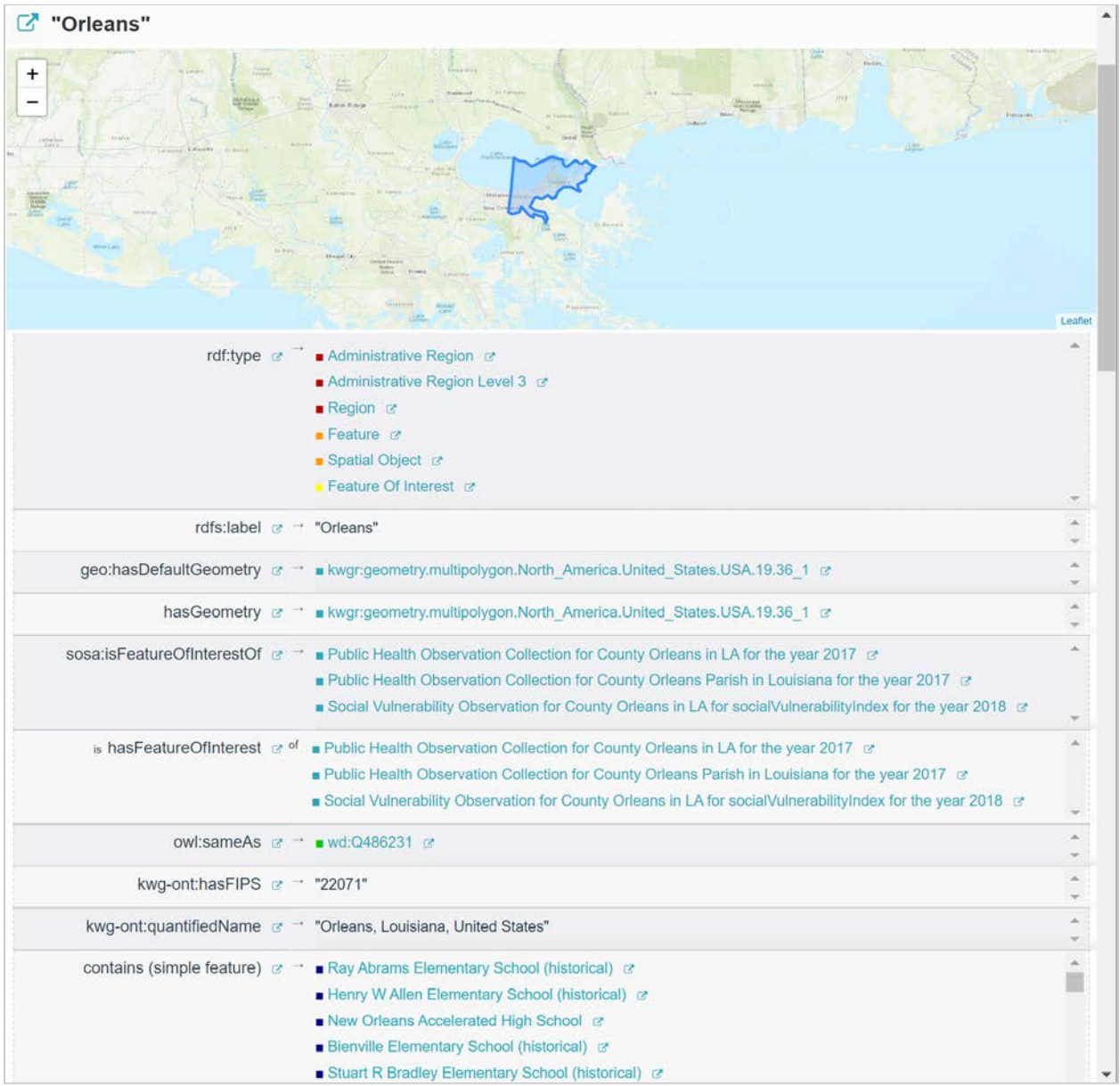
3.1. Explore by Place

The *Place* tab helps to answer the question, “What is here?” Clicking the *Place* button brings you to the Knowledge Explorer’s faceted search interface. Place identifiers such as the name of a county, state or other administrative region, a ZIP code or a US Climate Division can be specified as places of interest. A search for *Louisiana* in the *Administrative Region* search box yields a list of parishes (Administrative Region Level 3) and a state (Administrative Region Level 2), as pictured below.



The interactive map at right shows place markers corresponding to the locations of the results listed in the center panel. Click on a marker to identify it.

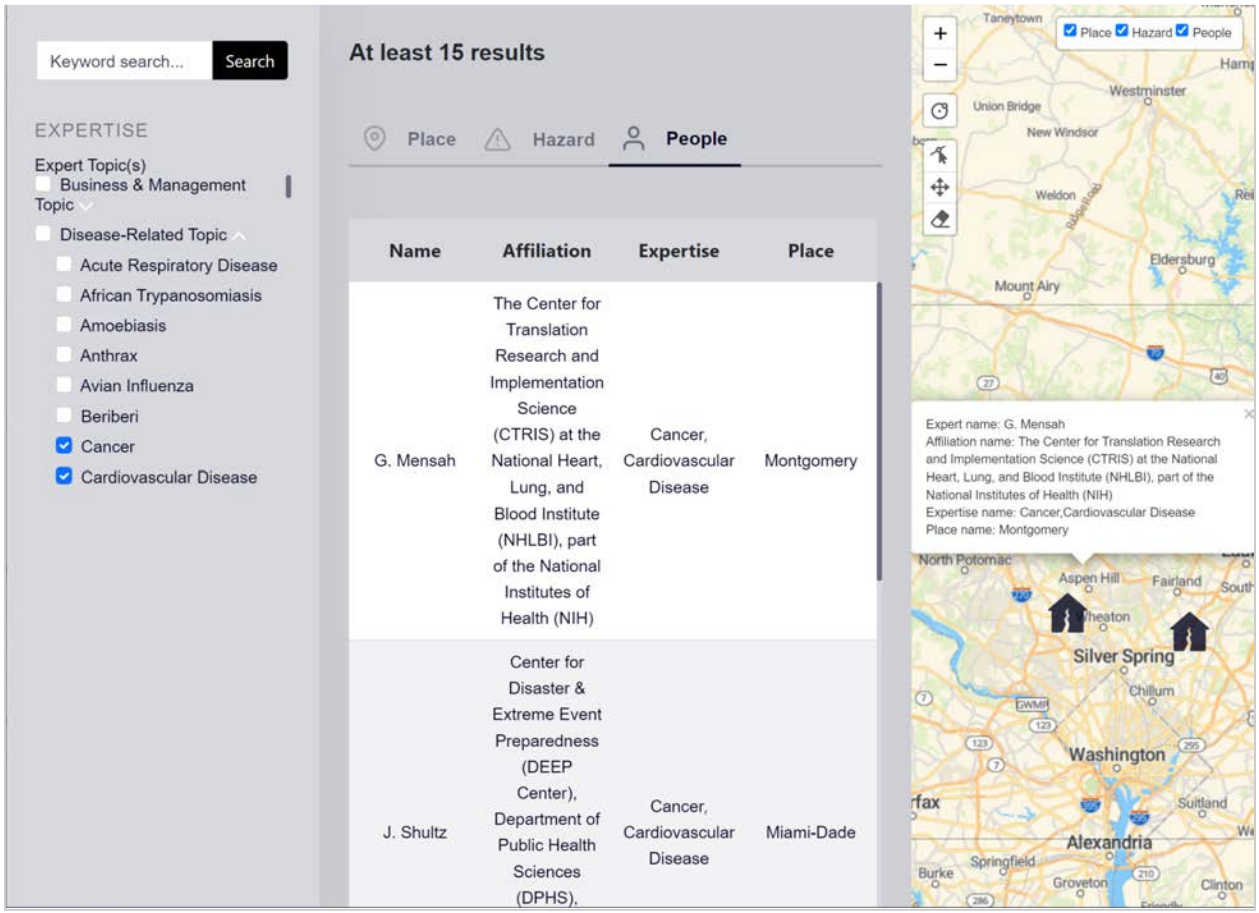
Selecting one of the results in the center pane (e.g., *Orleans*) brings you to the Knowledge Explorer's description page, below, which lists a variety of information (entities and predicates) associated with the selected place result. Partway down the page is a list of police departments, hospitals, schools and other features of interest that Orleans *contains*.



Clicking on a feature, e.g., *Ray Abrams Elementary School (historical)*, brings you to a similar page listing a variety of information about that entity. You can continue to browse the structure and contents of the graph by clicking on links in this "follow your nose"-style search.

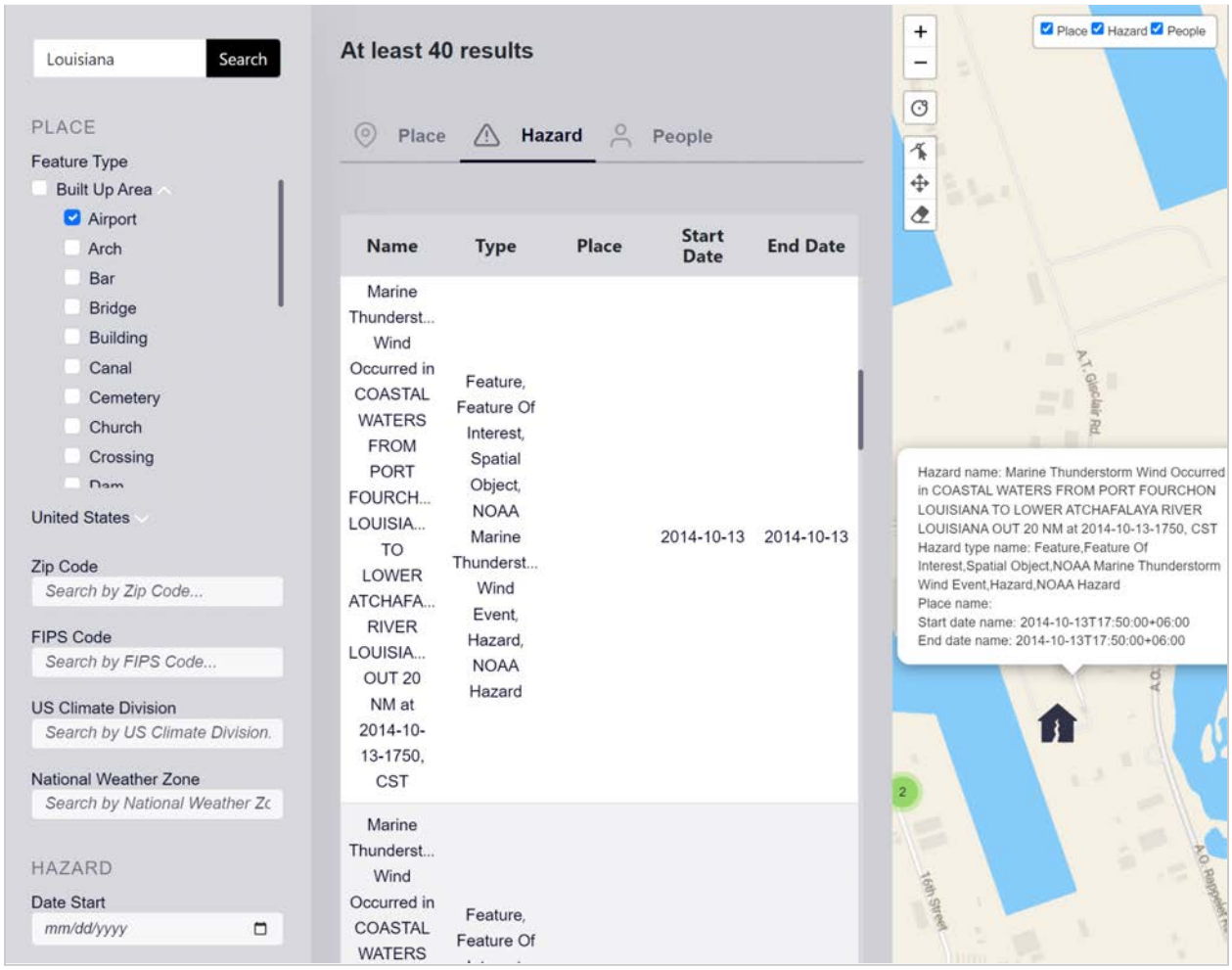
3.2. *Explore by People*

To see the types of information in the graph related to a particular person, select the *People* tab of the Knowledge Explorer's faceted search page. To find experts in a particular topic, filter the results by expanding and selecting expertise topics in the left-hand pane, as *Cancer* and *Cardiovascular Disease* are selected in the example below:



3.3. *Explore by Hazard*

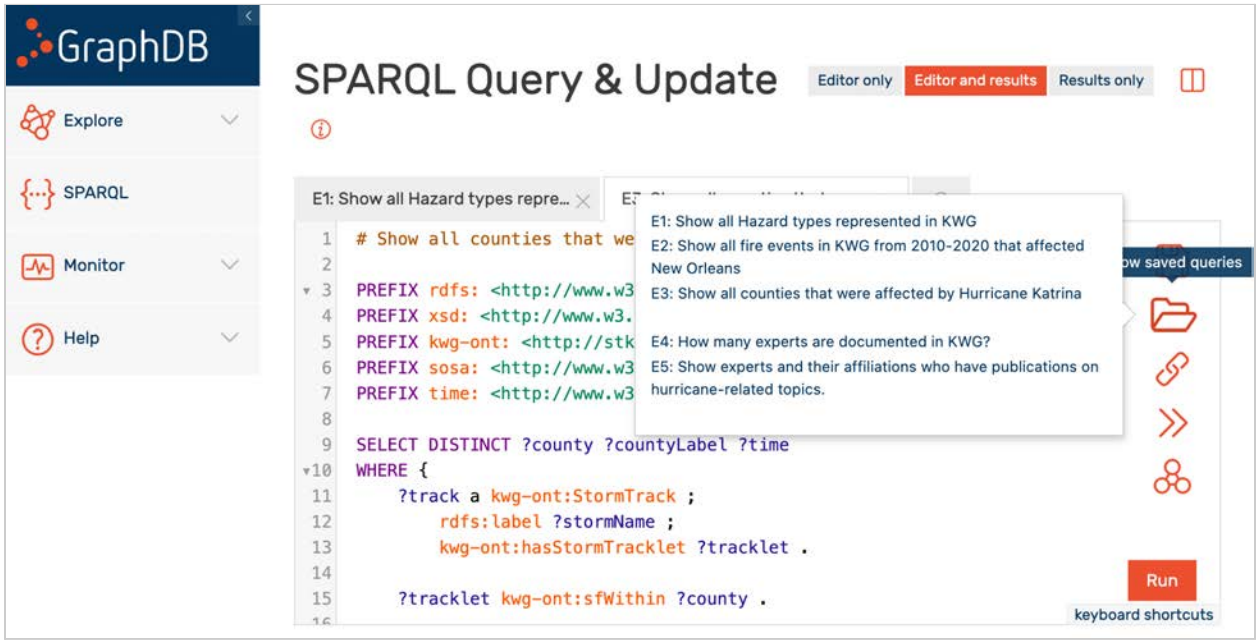
To browse information on documented wildfires, hurricanes and other hazard events, select the *Hazard* tab of the Knowledge Explorer. View particular hazard events (e.g., *Earthquake*, *Fire*) that affected particular places and feature types (e.g., *Airport*), by expanding the drop-down menus and selecting the desired filter criteria.



4. Example SPARQL Queries

To query the graph directly using SPARQL, visit our SPARQL endpoint. Our SPARQL endpoint runs a customized version of Ontotext's GraphDB Workbench software, which, in addition to providing a SPARQL query interface, provides both tabular and visual browsing interfaces. See the Exploring Data section of the GraphDB documentation for more information on how to access and use these interfaces.

In the following subsections, we provide several example competency questions and their corresponding SPARQL queries. These can be accessed under the *Saved SPARQL queries* section of the SPARQL endpoint home page or by clicking the folder icon on the right side of the SPARQL Query & Update window and selecting a query (pictured below). These queries (questions) were developed in consultation with our partners in the fields of agriculture, disaster response, and natural resource management.



4.1. Show all Hazard types represented in KWG.

```
## Show all Hazard types represented in KWG.

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX kwg-ont: <http://stko-kwg.geog.ucsb.edu/lod/ontology/>

SELECT DISTINCT ?type ?label
WHERE {
    ?type rdfs:subClassOf kwg-ont:Hazard .
    ?type rdfs:label ?label .
}
```

4.2. Show all fire events in KWG from 2010-2020 that affected New Orleans.

```
## Show all fire events in KWG from 2010-2020 that affected New Orleans.

PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX time: <http://www.w3.org/2006/time#>
PREFIX kwg-ont: <http://stko-kwg.geog.ucsb.edu/lod/ontology/>
PREFIX sosa: <http://www.w3.org/ns/sosa/>
PREFIX elastic: <http://www.ontotext.com/connectors/elasticsearch#>
PREFIX elastic-index: <http://www.ontotext.com/connectors/elasticsearch/instance#>

SELECT DISTINCT ?fire ?description ?type ?startTime ?endTime
WHERE {
    ?search a elastic-index:kwg_fs_index ;
        elastic:query "Orleans" ;
        elastic:entities ?fire .

    ?fire a ?eventType ;
        rdfs:label ?description ;
        sosa:isFeatureOfInterestOf ?observationCollection .

    ?eventType rdfs:subClassOf ?superClass ;
               rdfs:label ?type .

    values ?superClass { kwg-ont:Fire }

    ?observationCollection sosa:phenomenonTime ?time .

    ?time time:hasBeginning/time:inXSDDateTime | time:inXSDDate ?startTime ;
          time:hasEnd/time:inXSDDateTime | time:inXSDDate ?endTime .

    FILTER (?startTime > "2010-01-01"^^xsd:date && ?endTime > "2010-01-01"^^xsd:date &&
"2020-01-01"^^xsd:date > ?startTime && "2020-01-01"^^xsd:date > ?endTime)
}
```

4.3. Show all counties that were affected by Hurricane Katrina.

```
## Show all counties that were affected by Hurricane Katrina.

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX kwg-ont: <http://stko-kwg.geog.ucsb.edu/lod/ontology/>
PREFIX sosa: <http://www.w3.org/ns/sosa/>
PREFIX time: <http://www.w3.org/2006/time#>

SELECT DISTINCT ?county ?countyLabel ?time
WHERE {
    ?track a kwg-ont:StormTrack ;
        rdfs:label ?stormName ;
        kwg-ont:hasStormTracklet ?tracklet .

    ?tracklet kwg-ont:sfWithin ?county .

    ?county a kwg-ont:AdministrativeRegion_3 ;
        rdfs:label ?countyLabel .

    ?oc sosa:hasFeatureOfInterest ?tracklet ;
        sosa:phenomenonTime ?time .

    ?time time:inXSDgYear ?year .

    FILTER regex(?stormName, "^KATRINA")
    FILTER (?year = "2005"^^xsd:gYear)
}
```

4.4. How many experts are documented in KWG?

```
## How many experts are documented in KWG?

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX kwg-ont: <http://stko-kwg.geog.ucsb.edu/lod/ontology/>
PREFIX iospress: <http://ld.iospress.nl/rdf/ontology/>

SELECT (COUNT( DISTINCT ?expert) as ?count)
WHERE {
    ?expert a iospress:Contributor.
    ?expert rdfs:label ?label .
    ?expert kwg-ont:hasExpertise ?expertise .
    ?expertise rdfs:label ?expertLabel .
}
```

4.5. Show experts and their affiliations who have publications on hurricane-related topics.

```
## Show experts and their affiliations who have publications on hurricane-related topics.

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX kwg-ont: <http://stko-kwg.geog.ucsb.edu/lod/ontology/>
PREFIX iospress: <http://ld.iospress.nl/rdf/ontology/>

SELECT ?expert ?label ?expertiseTopicLabel ?organization ?orgName
WHERE {

    ?expert a iospress:Contributor ;
        rdfs:label ?label ;
        kwg-ont:hasExpertise ?expertise ;
        iospress:contributorAffiliation ?organization .

    ?expertise rdfs:label ?expertiseTopicLabel .

    ?organization iospress:organizationName ?orgName .

    FILTER regex(?expertiseTopicLabel, "^Hurricane")
}
```

5. Providing Feedback

We hope to collect constructive feedback from KWG users of all levels. After browsing the graph and trying the tools, please consider answering a few questions or simply provide comments in this survey. We would love to hear and learn from you about your experience in scaling and sustaining such graphs and about your ideas for new data and denser links among datasets. KWG and the Open Knowledge Network are community efforts; please get involved. Thanks for your time and interest.