



Knowing Where via Knowledge Graph-based GeoEnrichment

Building Tools on Top of the KnowWhereGraph

Zilong Liu
University of Vienna

Knowledge Explorer

The Knowledge Explorer interface is a portal for exploring the KnowWhereGraph database. It provides a set of filters related to attributes-or facets-of the various datasets in the graph that allow a user to quickly refine their search. Clickable links in the search results encourage the user to explore data tangentially related to their initial query that may lead to new insights. Whether...

[Learn More](#)

Geo-Enrichment

Inspired by Esri's Geo-Enrichment services and the Linked Data cloud, KnowWhereGraph starts with the concept of Geo-Enrichment, augments it with Semantic Web and Knowledge Graph (KG) technology, and proposes the idea of Knowledge Graph based Geo-Enrichment which aims at semantically enriching a given geospatial dataset with information from a cross-domain, up-to-date...

[Learn More](#)

GeoGraphVis

The GeoGraphVis tool, developed to support humanitarian response efforts following a disaster, visualizes data from KnowWhereGraph on a map. It allows a user to explore a hurricane's trajectory and impacts (e.g., death, property damage, etc.), health indicators related to the affected populations (e.g., prevalence of diabetes and obesity), as well as information on experts whose areas of expertise...

[Learn More](#)

Cropland Impact Assessment Tool

The Crop Impact Assessment Tool enables one to explore the geographic footprint of past wildfires and their smoke plumes in relation to agricultural distribution, processing and retail facilities. Smoke and ash plumes can contaminate crops across areas far larger than those directly affected by wildfires, suppressing production of safe and consumable leafy greens. This tool enables...

[Learn More](#)

Exploring 12.5+ billion statements with Faceted Search

A Need for Information Retrieval

- The **development of tools** that support knowledge discovery **lags** the pace of knowledge graph development, in general...
 - and with respect to **spatial and temporal data**, in particular.
- We need to effectively utilize, navigate, and understand the statements...
 - contained within **the KnowWhereGraph**.

Knowledge Explorer (Liu et al.)

- Enables more search functionalities under complex search conditions.
- Helps address a wide range of tasks ranging from cross-gazetteer place search to disaster management to expert search.

A faceted search interface for...

**interactively
browsing and
navigating the
KnowWhereGraph**

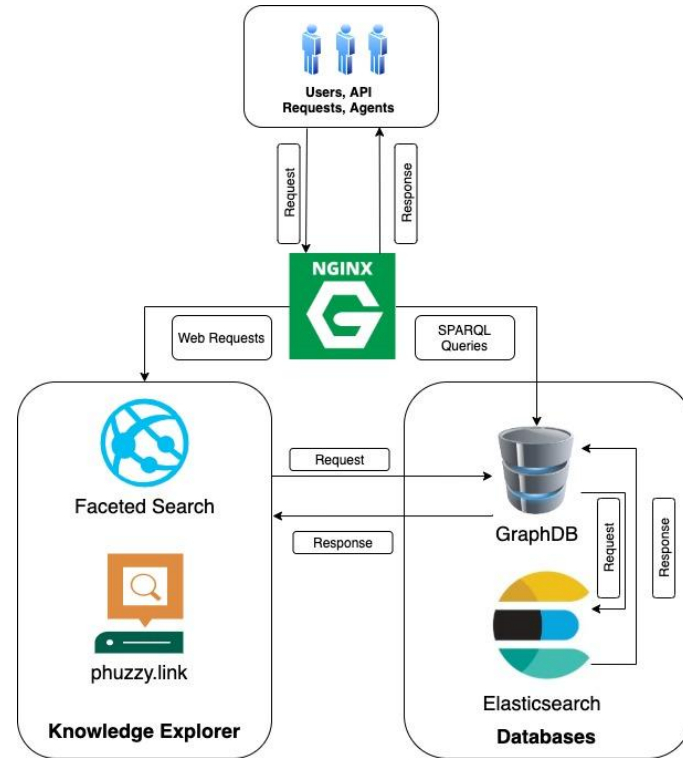
A dark blue banner with a network graph background. The text "Knowledge Explorer: Discovery begins here." is centered in white.

Knowledge Explorer: Discovery begins here.

<https://stko-kwg.geog.ucsb.edu>

Architecture

- Built with **Angular JS** **which we have recently changed to Angular 2*
- Deployed with **Node.js**
- Backed by **GraphDB** enhanced by **ElasticSearch**
- Map display and control with **Leaflet**
- URI dereferencing with extended **Phuzzy.link** framework (Regalia et al.)
- Source code, setup instructions, and further documentation on GitHub:
<https://github.com/KnowWhereGraph/kwg-faceted-search>



Query Generation

- Faceted search queries are generated **modularly** by a function that calls sub-functions to construct their subsets, ...
 - and these queries can also be **expanded to more complex ones.**
- Not all geospatial relations that directly link places from different gazetteers have been computed and stored in the triplestore... → so we implemented an **S2 cell-based information retrieval method.**

In the KnowWhereGraph, S2 cells connect places through pre-computed RCC8 relations (Shimizu et al).

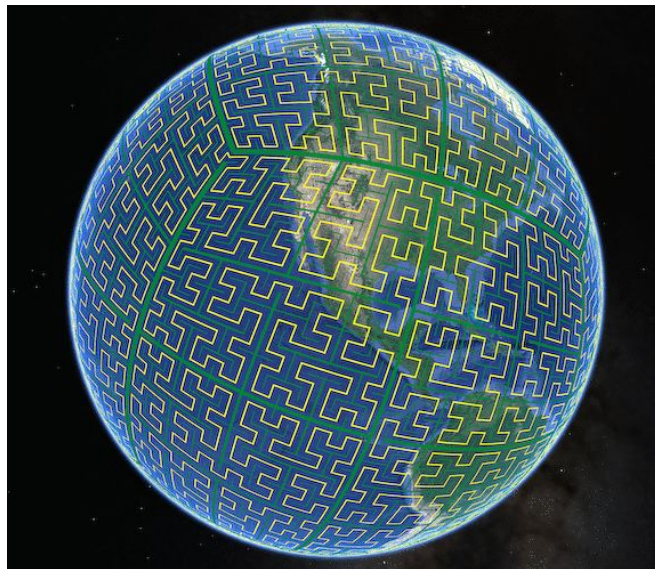
- The S2 cells where geographic features of interest are located were identified first.
- Then, we examined these cells to determine whether they are topologically linked to the administrative region of interest.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX usgs: <http://gnis-ld.org/ld/usgs/ontology/>
PREFIX kwg-ont: <http://stko-kwg.geog.ucsb.edu/ld/ontology/>
PREFIX kwgr: <http://stko-kwg.geog.ucsb.edu/ld/resource/>
select distinct ?entity ?label ?quantifiedName ?type ?typeLabel
{
  ?entity rdf:type ?type;
          rdfs:label ?label.
  ?type rdfs:label ?typeLabel.
  values ?type {usgs:Airport usgs:Building usgs:School}
  ?entity kwg-ont:sfWithin ?s2cell.
  ?s2cell rdf:type kwg-ont:KWGCellLevel13;
          kwg-ont:spatialRelation ?placesConnectedToS2.
  ?placesConnectedToS2 kwg-ont:sfWithin ?superPlacesConnectedToS2.
  filter (?placesConnectedToS2 in (kwgr:Earth.North_America.United_States.USA
  ↪ .48_1) || ?superPlacesConnectedToS2 in (kwgr:Earth.North_America.
  ↪ United_States.USA.48_1))
}
```

Listing 1: An example SPARQL query retrieving airports, buildings, and schools in Washington, USA

S2 Cell-based Information Retrieval

- Our querying strategy is to consider **hierarchical grids as intermediate nodes to obtain possible answers.**
- We leverage discrete global grids as a unified spatial reference system to **link disconnected geographic entities** in the graph.



Elasticsearch Index Management

- Although GraphDB does not support native text retrieval, we integrated it with Elasticsearch to enable **full-text search and result ranking**.
- **An Elasticsearch GraphDB connector instance**
 - ◆ We enabled all literal fields to be analyzed, and synchronized all values of those properties with multiple values.
 - ◆ The index is set to be automatically updated to be synchronized with the graph as it evolves.

Dereferencing

- We extended Phuzzy.link's support for information display about densely linked entities and customized its display for specific triple patterns.
- This system also supports a **“follow-your-nose” type of browsing** for both machines and humans, with content negotiation in place for easy exploration of related entities.
- We **linked faceted search results to the dereferencing interface** so that users can find additional information about a particular entity which, in turn, opens more possibilities for subsequent searches.

User Interface Interaction

Step by Step Examples

Start Exploring

Place Hazard Expert

Keyword searching

Keyword search... Search

At least 33 results Result Count

Place Hazard People

Facets Selection

EXPERTISE

- Disease-Related Topic
- Acute Respiratory Disease
- African Trypanosomiasis
- Amoebiasis
- Anthrax
- Avian Influenza
- Beriberi
- Cancer
- Cardiovascular Disease
- Chagas Disease
- Chikungunya Fever
- Engineering-Related Topic
- Environmental- & Agricultural-Related Topic
- General Medical & Public Healthcare-Related Topic
- Government & Legal-Related Topic
- Hazard Or Disaster-Related Topic
- Hurricane
- Information Technology-Related Topic

Table Display

Name	Affiliation	Expertise	Place
J. Shutz	Center for Disaster & Extreme Event Preparedness (DEEP Center), Department of Public Health Sciences (DPHS), University of Miami Miller School of Medicine	Acute Respiratory Disease, Cancer, Cardiovascular Disease, Cholera, Covid 19, Diabetes, Disaster Medical Assistance, Disaster Medicine, Disaster Mental Health, Disease	Miami-Dade, Florida, United States
L. Ivers	Massachusetts General Hospital Center for Global Health, Harvard Medical School	Acute Respiratory Disease, Cancer, Cardiovascular Disease, Cholera, Covid 19, Diabetes, Diarrhoeal Diseases, Disease, Epidemiology, Human Immunodeficiency Virus	Norfolk, Massachusetts, United States

Record information

Expert name: E. Ford
Affiliation name: Division of Adult and Community Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention
Expertise name: Cancer, Cardiovascular Disease, Diabetes, Disease, Epidemiology, Glucose, Health Disparities, Health Emergency/Infectious Disease, Influenza
Place name: Jackson, Michigan, United States

Map Visualization

Overview & Step-by-Step Examples

Outline:

1. Introduction
2. The Graph
3. The Knowledge Explorer
 - 3.1. Explore by Place
 - 3.2. Explore by Disease
 - 3.3. Explore by Hazard
4. Example SPARQL Queries
5. Providing Feedback

1. Introduction

The KnowWhereGraph (KWG) is an openly accessible, W3C-standards-based geographically enriched knowledge graph, with associated tools for representing, exploring and analyzing a growing list of human and environmental data. With sophisticated search capabilities and seamless access to numerous public datasets, KWG rightly raises the situational awareness of data scientists and decision makers by providing selected area linkages 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 data exceeds 2 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, [in KnowWhereGraph: Science](#)). 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 geo-relevant 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 knowledge following a disaster. These include information on previous disaster events, physical

kwgr:Earth.North_America.United_States.USA.10.43_1

kwg-ont:hasFIPS or "12086"

kwg-ont:qualifiedName or "Miami-Dade, Florida, United States"

contains (simple feature) or

- Arch Creek Natural Bridge (historical)
- Homestead Air Force Base (historical)
- Pinecrest
- Aventura
- Norland
- Ludlum
- Country Walk
- East Penton
- Fisher Island
- Fountainbleau

contains (simple feature) or

- Smoke Plume Snapshot 0 on 2017-05-14
- Smoke Plume Snapshot 0 on 2019-04-15
- Smoke Plume Snapshot 0 on 2019-04-26
- Smoke Plume Snapshot 0 on 2020-04-09
- Smoke Plume Snapshot 0 on 2020-04-21
- Smoke Plume Snapshot 0 on 2020-04-25
- Smoke Plume Snapshot 0 on 2020-09-16
- Smoke Plume Snapshot 0 on 2021-03-07
- Smoke Plume Snapshot 0 on 2021-03-18
- Smoke Plume Snapshot 0 on 2021-03-18
- Smoke Plume Snapshot 0 on 2021-03-25

Dereferencing

stko-kwg.phsu.ucsf.edu/#result_search?tab=people&pg=20&expert-topic=acuteRespiratoryDisease,topic=africanTrypanosomiasis,topic=amoebiasis,1...

National Center for Disease, Avian

Knowledge Explorer Demo



**Knowledge Explorer
Tool**

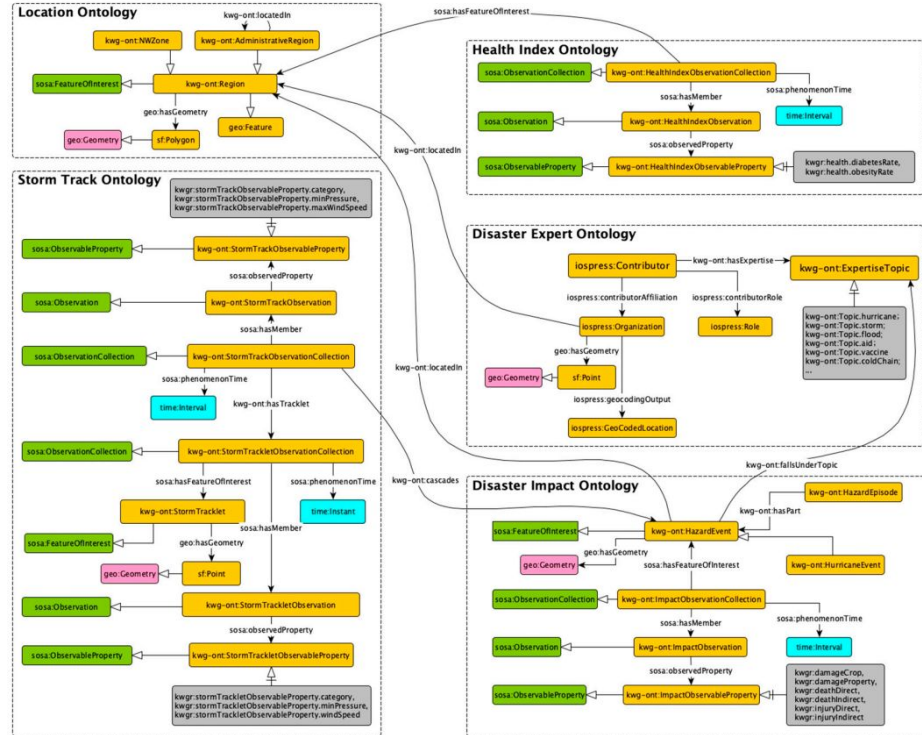
A KnowWhereGraph- and Geovisualization-Empowered Cyberinfrastructure

Supporting Disaster Response

- When a disaster occurs, it is important to acquire information about the potentially affected area, its infrastructure, and its people in order to...
 - develop **situational awareness** and plan a response to address the humanitarian needs of the affected population.
- This requires rapid assembly of multi-source geospatial data that need to be organized and...
 - **visualized** in a way to support disaster-relief efforts.

An Integrated Schema for Disaster Management

- Location Ontology
- Storm Track Ontology
- Health Index Ontology
- Disaster Expert Ontology
- Disaster Impact Ontology



Descriptions and Sources of Datasets

Ontology	Dataset	Publisher	Description	URI
Storm Track	Atlantic hurricane database	National Oceanic and Atmospheric Administration (NOAA)	A dataset with data collected every six hours, containing information on location, maximum winds, central pressure, and size of known tropical cyclones and subtropical cyclones.	https://www.nhc.noaa.gov/data/#hurdat (accessed on 16 January 2023)
Disaster Impact	Storm event database	National Oceanic and Atmospheric Administration (NOAA)	A dataset covering storm and other significant weather event reports with statistics on personal injuries and damage estimates.	https://www.ncdc.noaa.gov/stormevents/ (accessed on 16 January 2023)
Health Index	Public health data	University of Wisconsin Population Health Institute	A database containing public health factors, such as a food environment index based on counties.	https://www.countyhealthrankings.org/explore-health-rankings (accessed on 16 January 2023)
		Centers for Disease Control and Prevention	A database containing public health factors, such as diabetes rates and obesity rates	https://nccd.cdc.gov/DHDSPAtlas/ (accessed on 16 January 2023)
Disaster Expert	Expert data	Semantic Scholar	A dataset on experts' areas of expertise, work affiliations, and locations.	https://www.semanticscholar.org/product/api (accessed on 16 January 2023)
Location	Global administrative areas	Database of Global Administrative Areas (GADM)	Boundaries of the world's administrative regions.	https://gadm.org/data.html (accessed on 16 January 2023)
	Weather forecast zones	National Weather Service (NWS)	Boundaries of areas used by NWS for forecasts and warnings as well as map backgrounds.	https://www.weather.gov/gis/PublicZones (accessed on 16 January 2023)

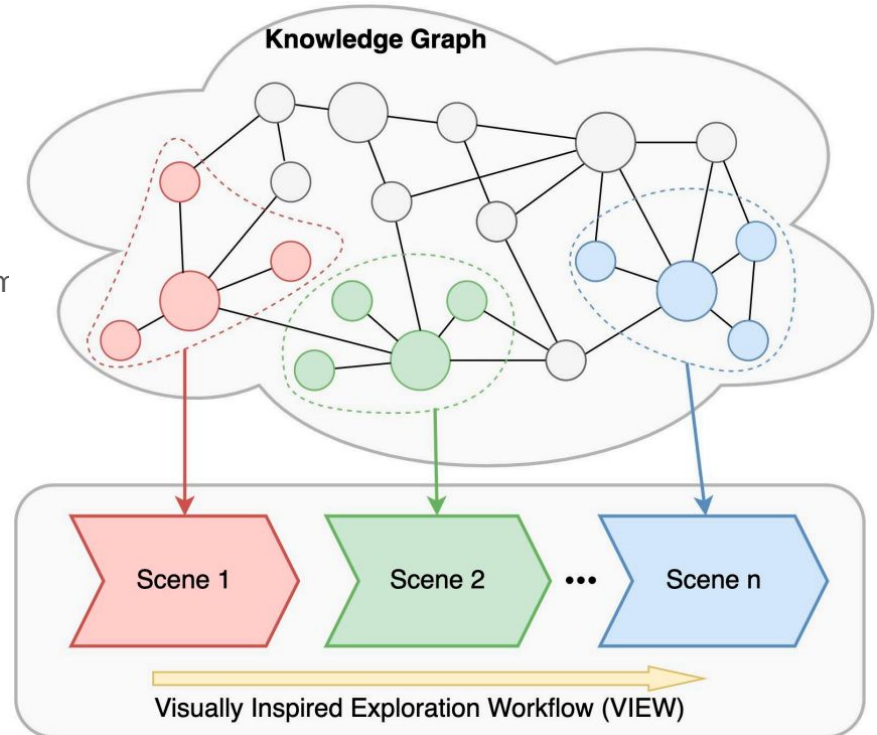
Scene-based Visualization Strategy

Our visualizations system both selects subgraphs and integrates them to meet decision needs.

- The visual scenes are **not** necessarily organized as a single decision path.
- The VIEW can be split into multiple branches to form a **decision tree**.

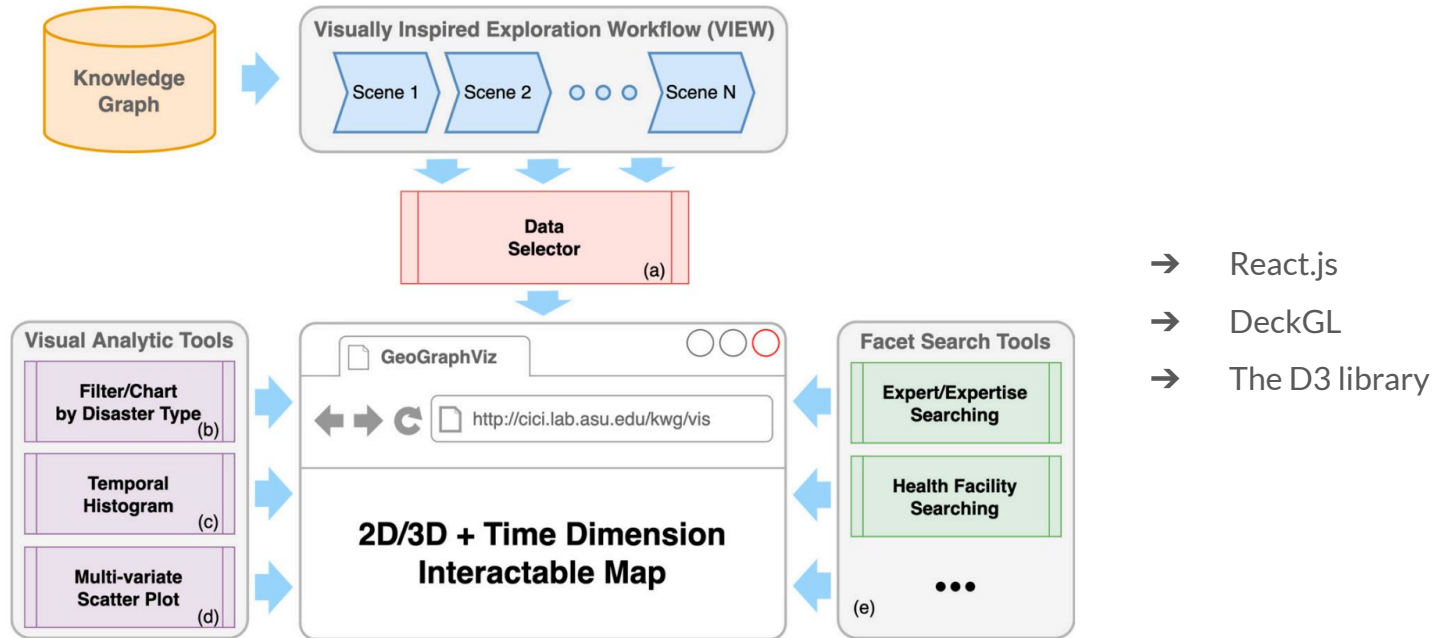
In a disaster relief scenario, three scenes are used to **contextualize disasters and potential risks**.

- A fourth scene is used to **find experts** with specialized knowledge to facilitate an effective response.



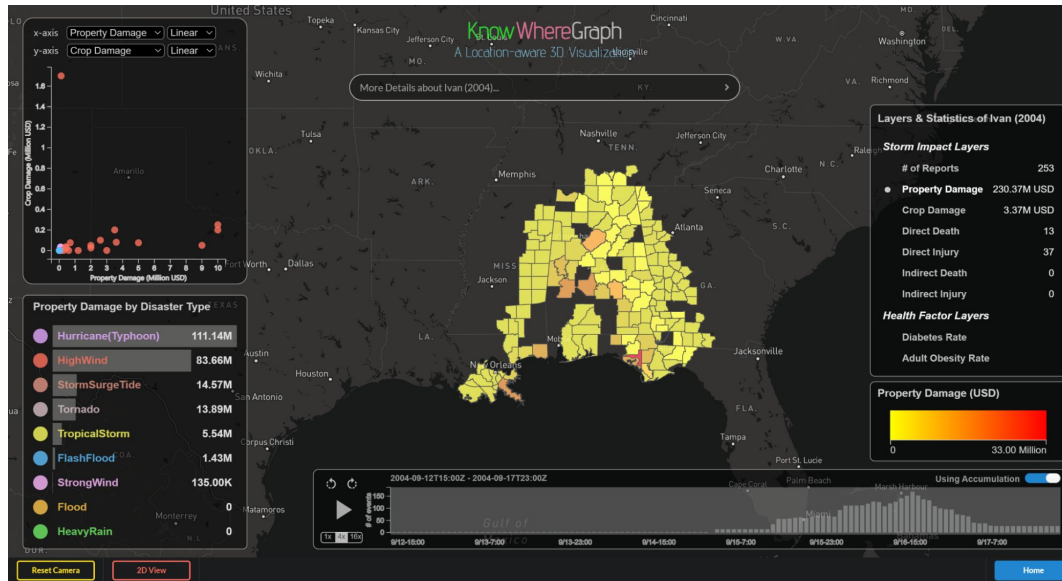
1. Identify subgraphs based on a humanitarian relief use case
2. Map subgraphs to scenes in a visualization workflow

The GeoGraphVis Framework



GeoGraphVis Resources

- Hosted on <http://cici.lab.asu.edu/kwg/vis/>
- Github repository: <https://github.com/KnowWhereGraph/GeoGraphVis>



GeoGraphVis Demo



Cropland Impact Assessment Tool

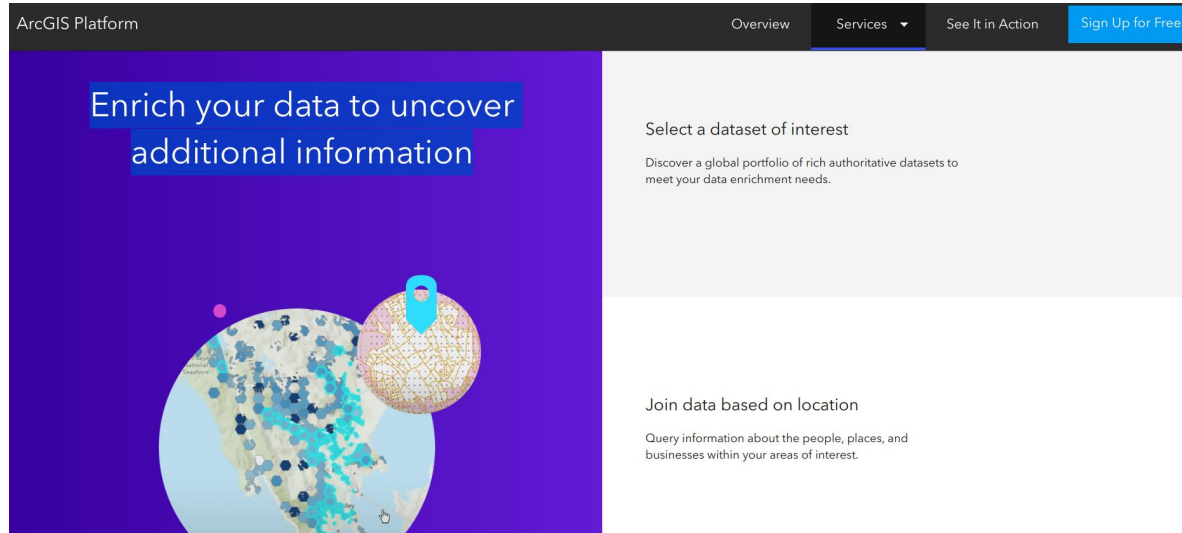
Cropland Impact Assessment Tool Demo



Using GeoEnrichment Extensions for QGIS and ArcGIS

GeoEnrichment in GI Systems Supports Decision-Making.

- Esri's GeoEnrichment service enables analysts to enrich their local data on-demand with a range of up-to-date demographic **variables apportioned to their area of concern and need.**

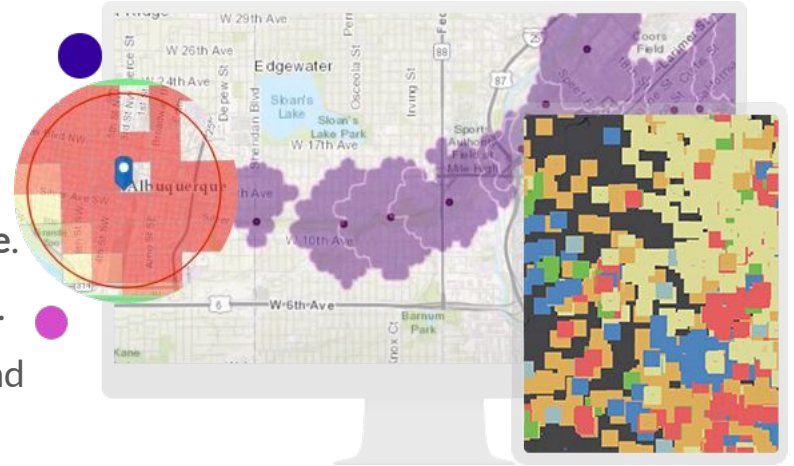


The screenshot shows the ArcGIS Platform website for the GeoEnrichment service. The navigation bar includes 'ArcGIS Platform', 'Overview', 'Services' (with a dropdown arrow), 'See It in Action', and a blue 'Sign Up for Free' button. The main content area is split into two columns. The left column has a dark blue background with the text 'Enrich your data to uncover additional information' and a circular graphic showing a map of the United States with various data points and a magnified view of a specific location. The right column has a light gray background and contains two sections: 'Select a dataset of interest' with a sub-description 'Discover a global portfolio of rich authoritative datasets to meet your data enrichment needs.', and 'Join data based on location' with a sub-description 'Query information about the people, places, and businesses within your areas of interest.'

<https://www.esri.com/en-us/arcgis/products/arcgis-platform/services/geoenrichment>

Advantages of GeoEnrichment Services

- Data are always up-to-date and does **not** age on the analyst's hard disk.
- In times of misinformation and information overload, the data come from a **trusted resource**.
- The data are **tailored** to the analyst's study area.
- The data can be **directly** processed, analyzed, and displayed by GI systems.



<https://www.esri.com/en-us/arcgis/products/arcgis-platform/services/geo-enrichment>

GI System Functionalities

- Spatial analysis
- Mapping patterns of geographic processes
- Geo-data management
- ...

Capabilities

ArcGIS offers unique capabilities and flexible licensing for applying location-based analytics to your business practices. Gain greater insights using contextual tools to visualize and analyze your data. Collaborate and share via maps, apps, dashboards and reports.



Spatial Analysis & Data Science

Connect the seemingly disconnected with the most comprehensive set of analytical methods and spatial algorithms available. Use location as the connective thread to uncover hidden patterns, improve predictive modeling, and create a competitive edge. Leverage the power of spatial analysis and data science on demand and at scale.

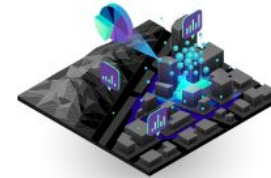
[Learn more →](#)



Field Operations

Location is at the heart of field activities. Focused ArcGIS applications can be used stand-alone or in combination to support field workflows and enable office and field personnel to work in unison, using the same authoritative data.

[Learn more →](#)



Mapping

Maps help you spot spatial patterns in your data so you can make better decisions and take action. Maps also break down barriers and facilitate collaboration. ArcGIS gives you the ability to create, use, and share maps on any device.

[Learn more →](#)

<https://www.esri.com/en-us/arcgis/about-arcgis/overview>



3D GIS

3D GIS brings real-world context to your maps and data. Instantly transform your data into smart 3D models and visualizations that help you analyze and solve problems and share ideas and concepts with your team and customer.

[Learn more →](#)



Imagery & Remote Sensing

ArcGIS gives you everything you need to manage and extract answers from imagery and remotely sensed data. It includes imagery tools and workflows for visualization and analysis, and access to the world's largest imagery collection.

[Learn more →](#)



Data Collection & Management

With ArcGIS, you can easily collect, crowdsourcing, store, access, and share your data efficiently and securely. You can integrate data stored in your business systems and geo-enable any data from any source.

[Learn more →](#)

Still... There are Key Limitations (Janowicz et al., 2022).

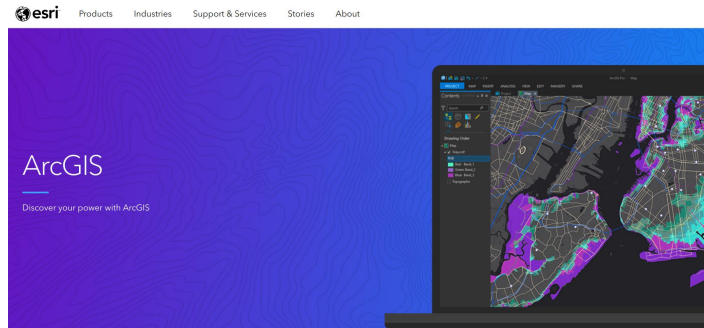
- They only serve data for a **small set of predefined categories**, such as demographic data.
- They are **closed** data silos that encode just **one domain/cultural perspective**.
- Because they are centrally maintained, **scalability and timely updates become bottlenecks** when those services try to incorporate more (diverse) data.
- They do **not have an integration layer** that enables follow-up queries over the enriched data.

Toward KG-based Geo-enrichment (Services)

- A new approach is needed that combines the strength of geo-enrichment services, that is, **seamless access to contextual information for an analyst's areas of concern**, with a technology that provides **open, densely integrated, cross-domain data across a wide range of perspectives**. (Janowicz 2021)
 - **The KnowWhereGraph** provides a solution.

What We Can Contribute Back to GI Systems.

- Extensions to both commercial and open-source GIS software packages, e.g., ArcGIS and QGIS.



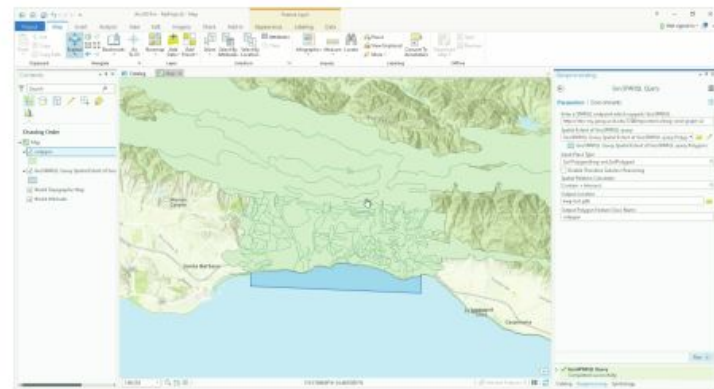
<https://www.esri.com/en-us/arcgis/about-arcgis/overview>



<https://www.qgis.org/en/site/>

Starting from An Early Work

- Region-based spatial data retrieval
- Property enrichment for geographic entities
- Direct relation exploration among geographic entities
- N-degree relation identification



(A) Retrieving soil polygons



(B) Retrieve wildfires that affected soil polygons

FIGURE 3 Our Knowledge Graph-based geo-enrichment toolbox collections for ArcGIS Pro. (A) The GeoSPARQL Query toolbox, (B) The Property Enrichment toolbox

Demo within ArcGIS (&QGIS)



**Geo-Enrichment Extensions
for QGIS & ArcGIS**

Resources on Installation and Usage

- Github repository for the ArcGIS Extension:
<https://github.com/KnowWhereGraph/geoenrichment-qgis>
- Github repository for the ArcGIS Extension:
<https://github.com/KnowWhereGraph/geoenrichment-arcgis>
- [Geo-Enrichment Extension Documentation](#)





Summary and Ongoing Work

- Why spatial is special in knowledge graphs?
- The KWG's ontology, data, and pilots
- Spatial Data Integration via S2 cells
- Querying the KWG & Geovisualization
- Building tools on top of the KWG