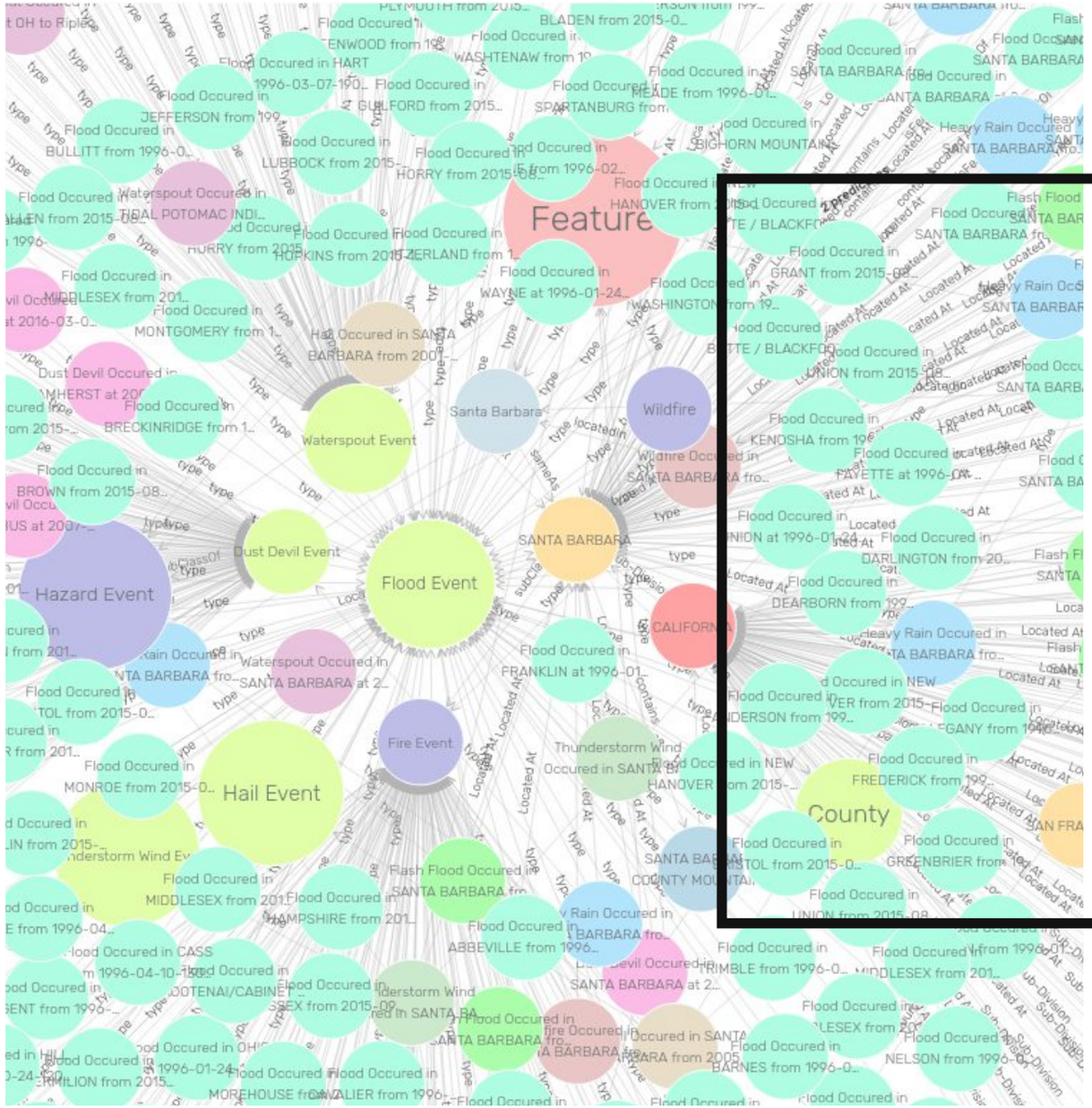
Krzysztof Janowicz, UCSB and University of Vienna (PI)
Pascal Hitzler, K-State (Co-PI)
Wenwen Li, ASU (Co-PI)
Mark Schildhauer, NCEAS (Co-PI)
Dean Rehberger, MSU (Co-PI)

KnowWhereGraph and Beyond: Can Knowledge Graphs provide Situational Awareness for Epidemic Intelligence?

EIOS Nov. 2022

KnowWhereGraph A6677

We provide area briefings within seconds for any region on Earth


Feature
Flood Event
Hazard Event
Hail Event
Fire Event
Dust Devil Event
Waterspout Event
Wildfire
Santa Barbara
SANTA BARBARA
County

WHAT

WHY

Data Acquisition Bottleneck

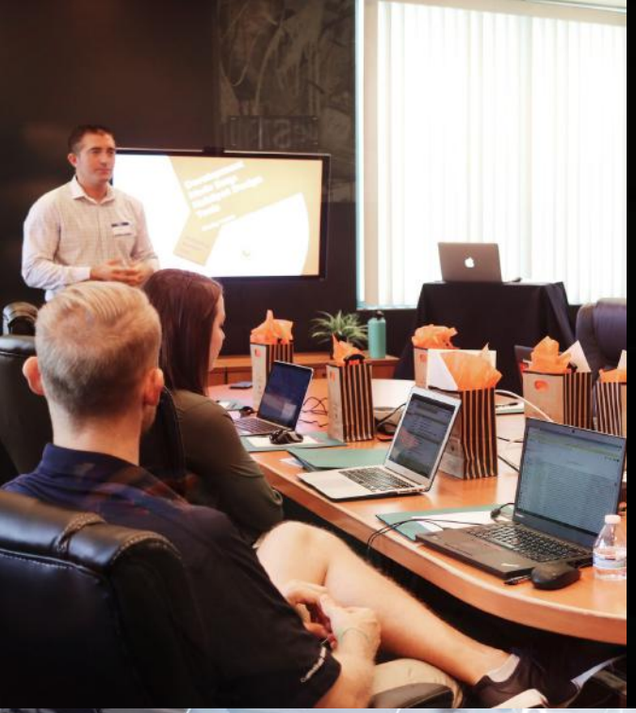
The typical data science project spends a **majority** of all resources on data discovery, entry, cleaning, and integration— instead of actual analysis

80%

Processing Data

20%

Deriving Insights



Our Value Proposition

KnowWhereGraph aims at **providing area briefings within seconds** for any region on Earth to answer questions such as

- "***What is here?***"
- "***What happened here before?***"
- "W*ho knows more?*"
- "*How does it compare to other regions or previous events?*"

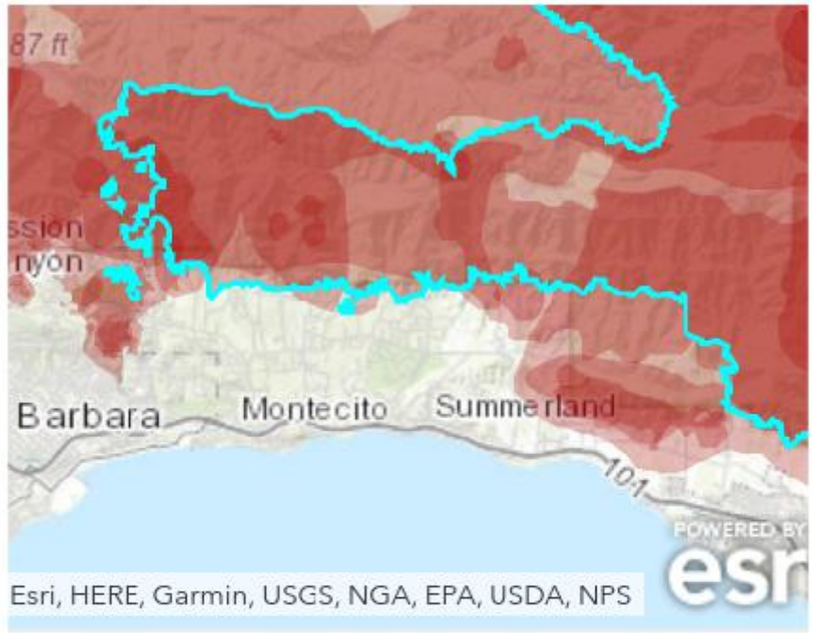
By doing so, we assist decision-makers and data scientists in rapidly enriching their data with millions of connected, up-to-date facts at the human-environment interface to gain the situational awareness required for good decision-making.

Our ultimate goal is to flip the 80/20 bottleneck on its head.

Overcoming the 80/20 Bottleneck with (Esri's) **GeoEnrichment**

Enriched California Fire Perimeters 1878 2018 Burn areas (Features: 368, Selected: 1)

YEAR_	STATE	AGENCY	UNIT_ID	FIRE_NAME	ALARM_DATE	CONT_DATE	
2017	CA	USF	VNC	THOMAS	12/3/2017, 4:00 PM	1/11/2018, 4:00 PM	281,893.00



Have a smartphone	2019 Seniors (Age 65+)
8,288.00	2,480.00

Enrich (your own) data by adding (demographic) facts about **your spatial area of interest**

GeoEnrichment Limitations

Pros

- On-demand
- Well-curated
- Apportioned
- GIS-ready

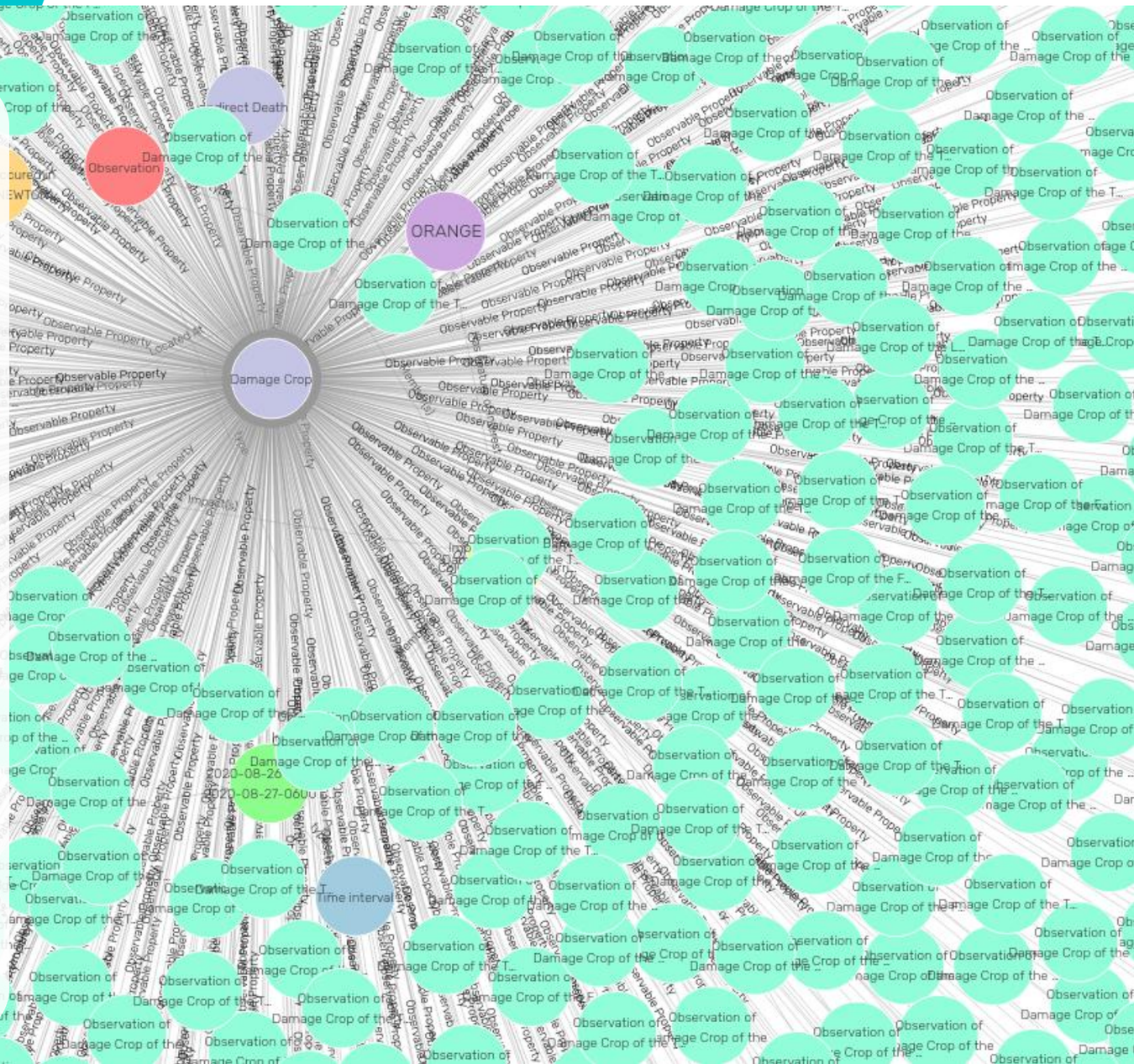
Cons

- Pre-defined categories
- Essentially closed data silos
- Flat, tabular data
- Not always up-to-date
- Does not scale
- Limited support for automated integration

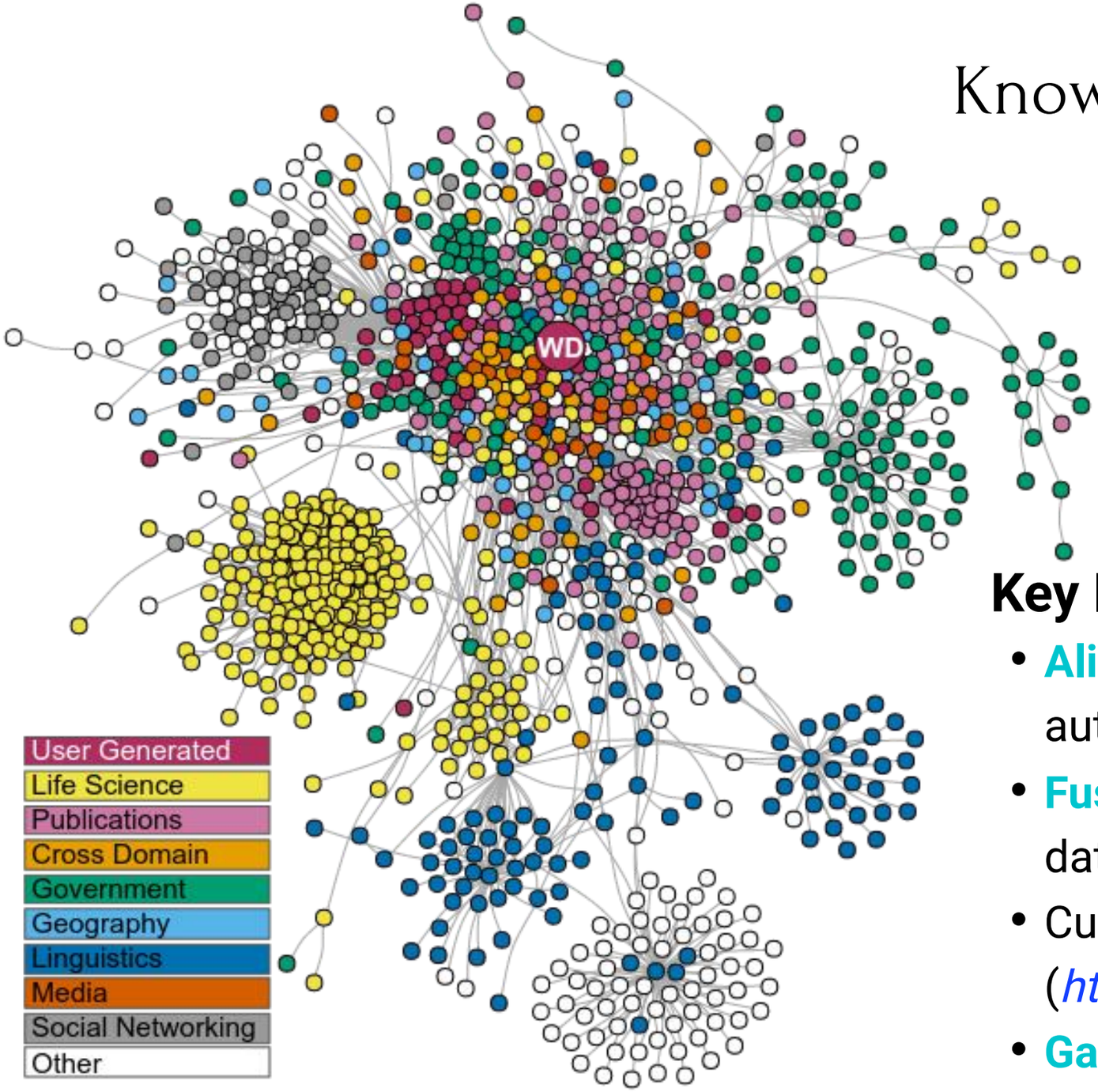
Knowledge Graphs Connect Data Across Themes

Knowledge Graphs

- **Emphasize relationships** over (in addition to) attributes
- **Connect data** (not datasets) **across themes** and silos
- Break apart the data-metadata distinction by **making data smart**, self-descriptive
- **Machine-readable**, rich semantics enables **conflation** and inference
- While **places** are key nodes in KG, we **know almost nothing** about them



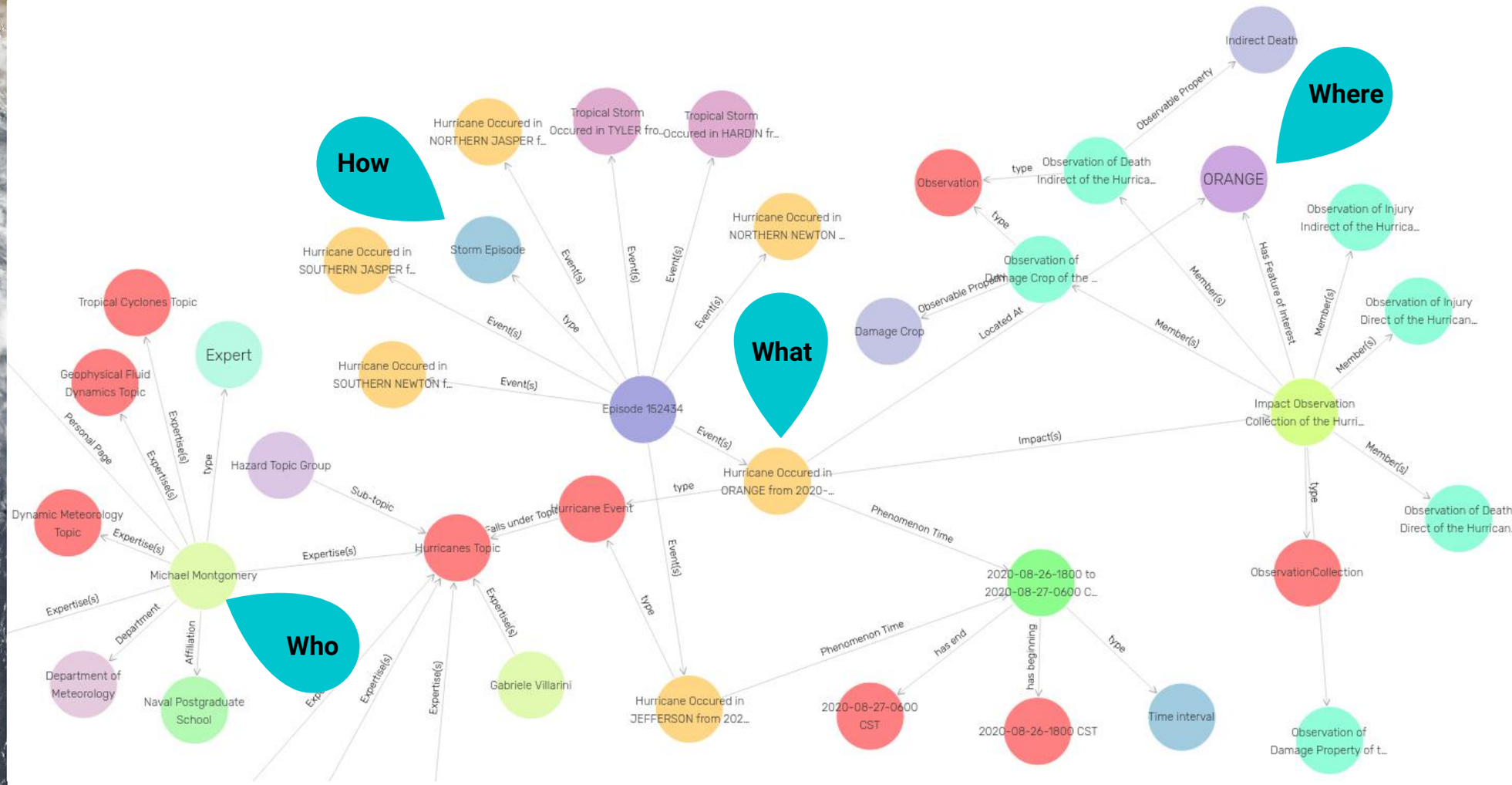
Knowledge Graph Ecosystem



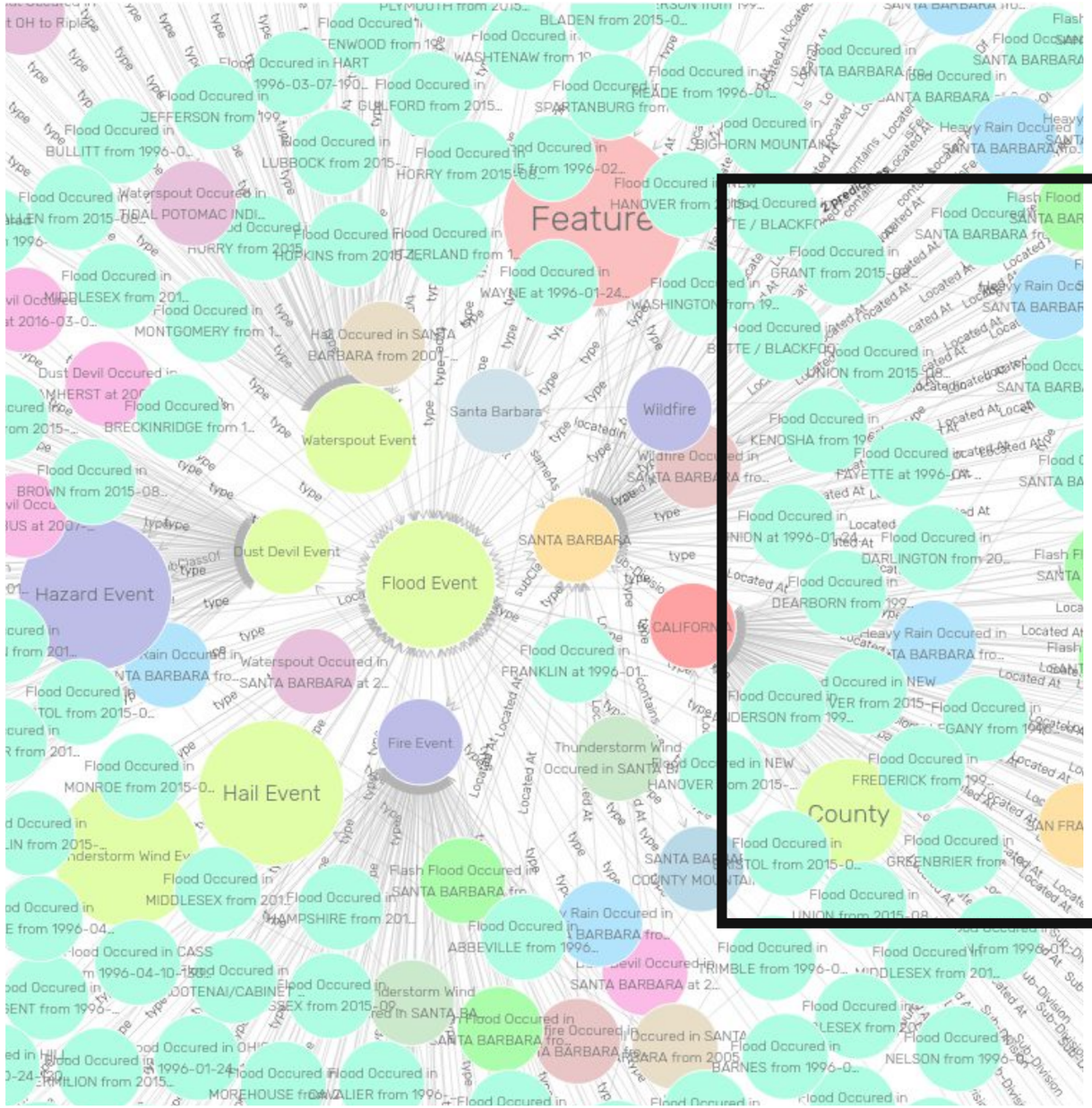
Key Ideas

- **Align vocabularies** across graphs (semi-automatically)
- **Fuse data** across graphs to access billions of data items at your **fingertips**
- Currently we establish links to **SPOKE** (*https://spoke.ucsf.edu/*) and **UFOKN**
- **Gaining a more holistic understanding!**

KnowWhereGraph = Knowledge Graph + GeoEnrichment



We are the first to **combine geoenrichment with knowledge graphs** to power **environmental intelligence** applications



WHO BENEFITS

Pilots and Prototypes

Farm to Table Supply Chain & Sustainability

Enhance the sustainability, efficiency, and safety of consumer food supply in collaboration with the *Food Industry Association*.

Land Valuation and Risk of Default

Driver-based land potential assessment for model based valuation and risk assessment for agricultural credit applications & loan portfolio monitoring.

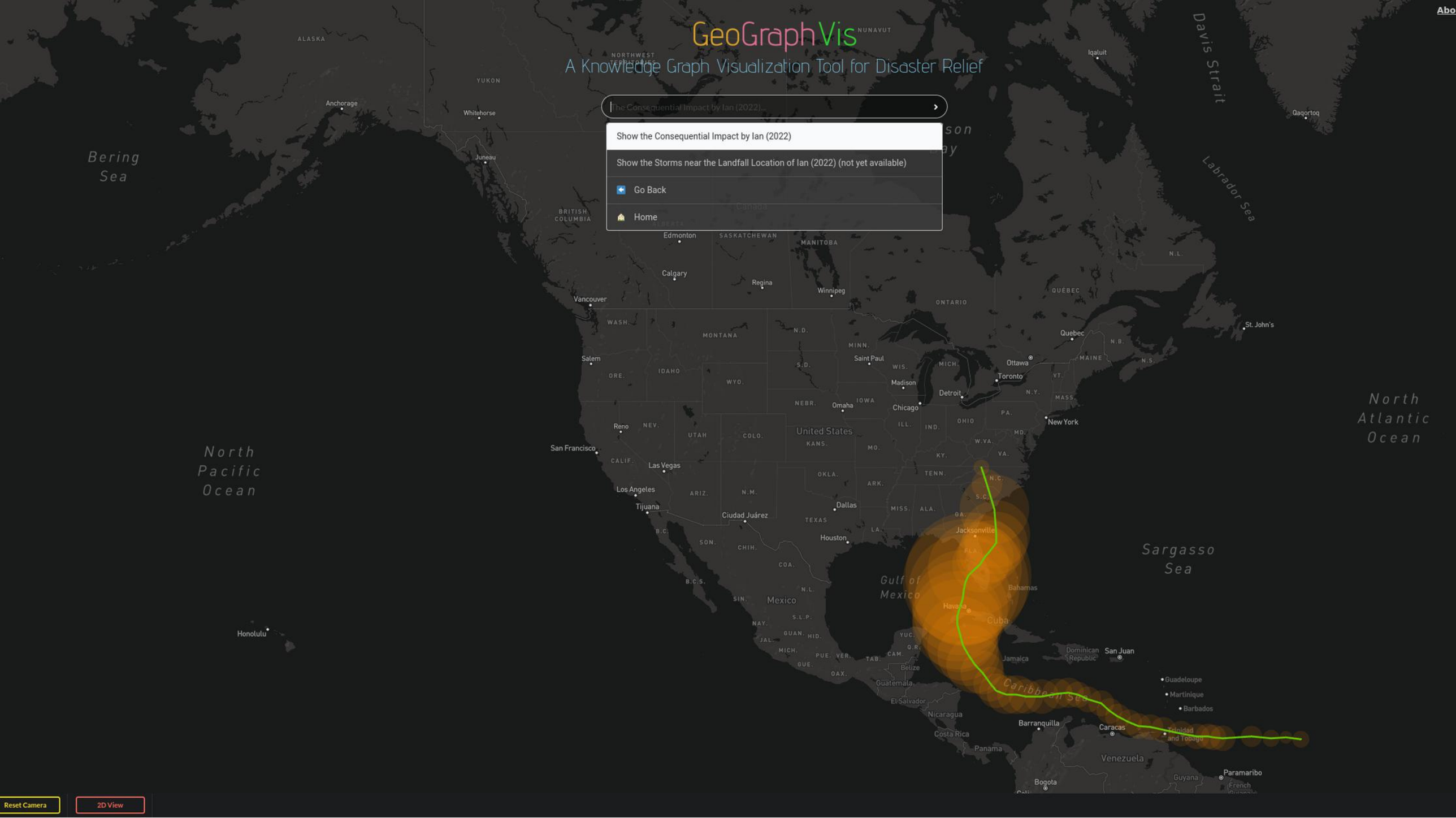
Community Lifeline System Backbone

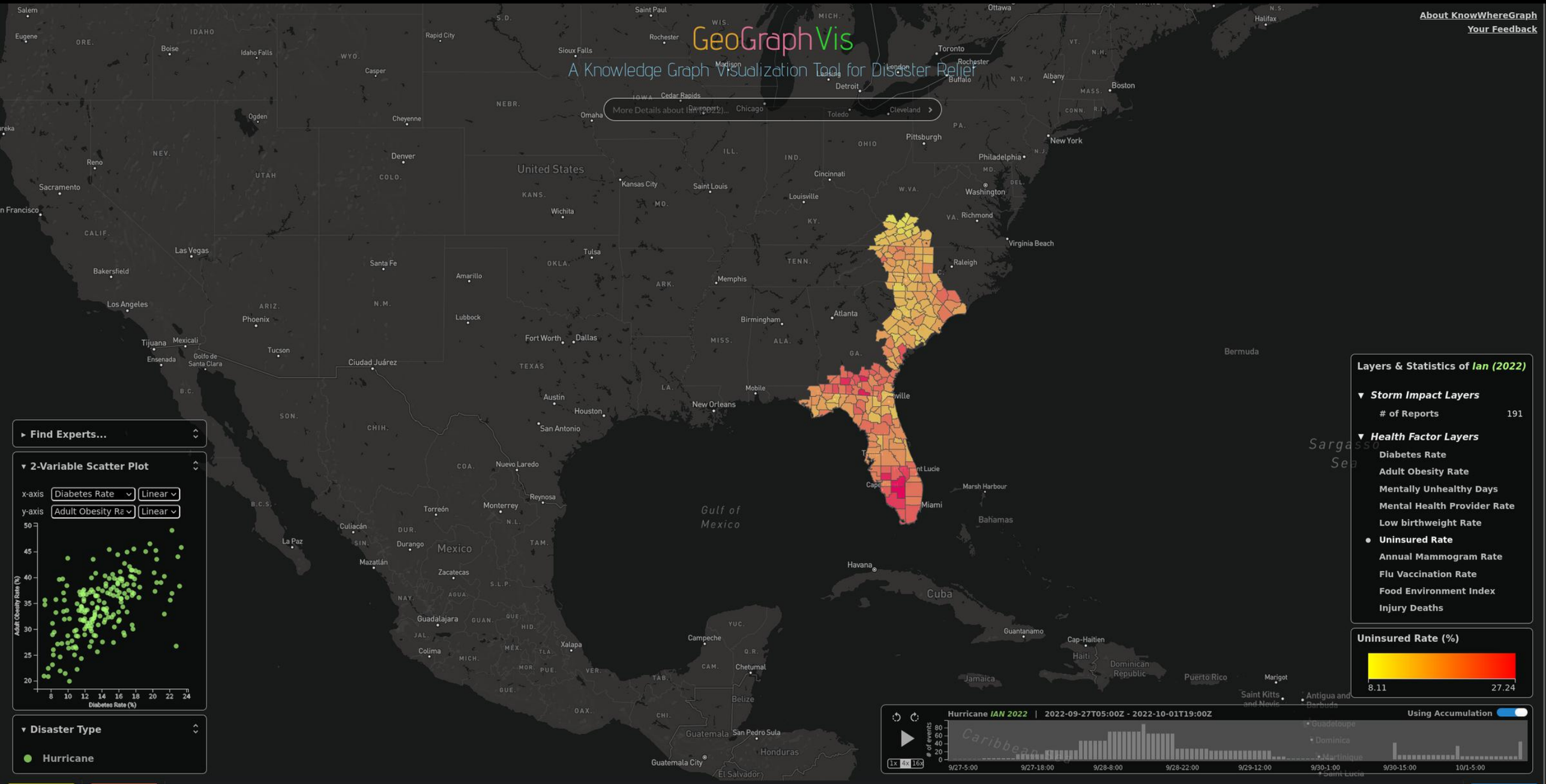
Provide a data library for time-critical situational awareness assessment by delivering data for an affected area with over 25 data layers in seconds

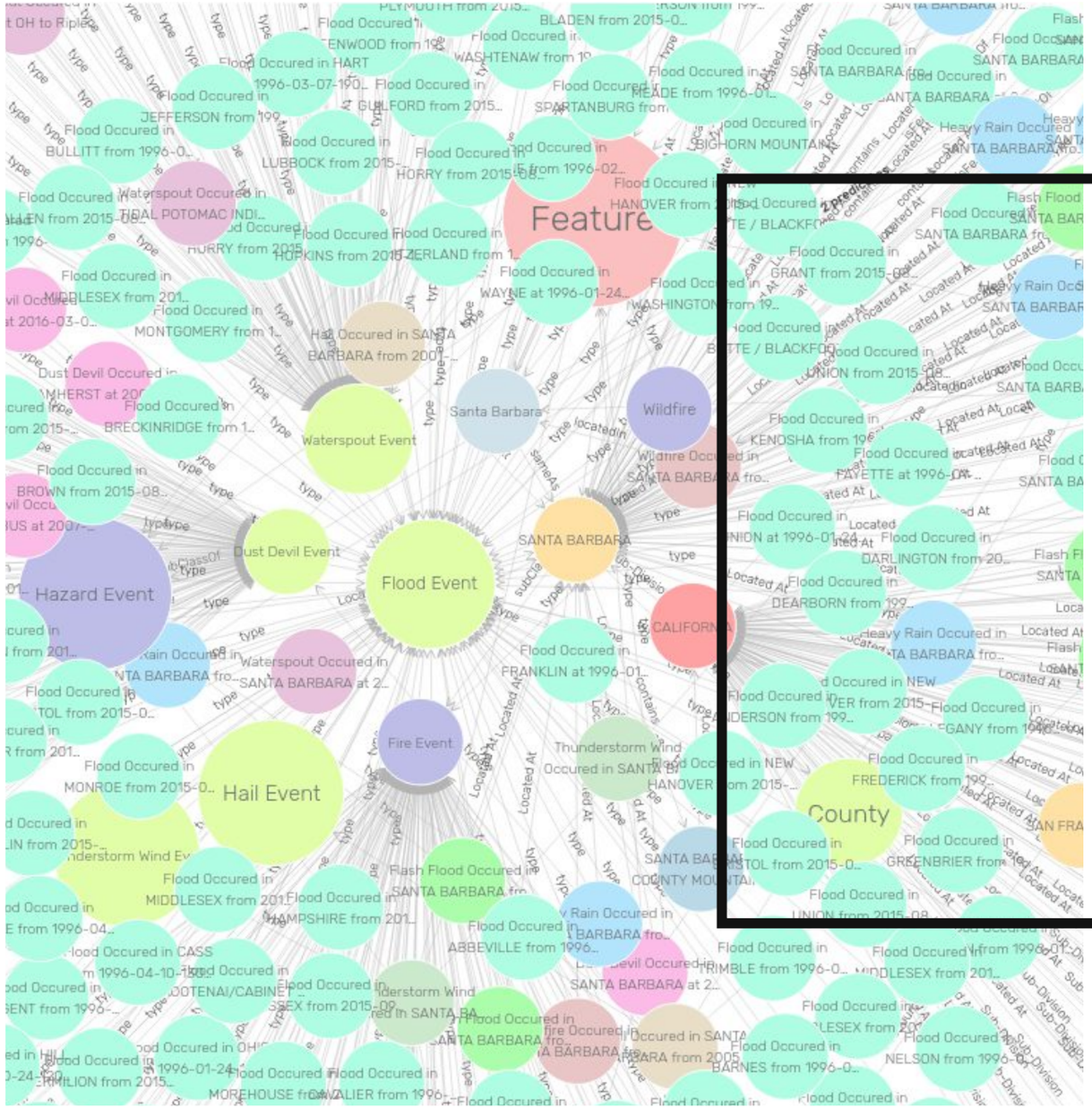
Humanitarian Aid

Apply our technologies to the humanitarian supply chain needs of *Direct Relief* during the COVID-19 crisis and help them to find experts.

Environmental Intelligence


GeoGraphVis
A Knowledge Graph Visualization Tool for Disaster Relief
The Consequential Impact by Ian (2022)
Show the Consequential Impact by Ian (2022)
Show the Storms near the Landfall Location of Ian (2022) (not yet available)
Go Back
Home
Abo
Reset Camera
2D View
ALASKA
NUNAVUT
NORTHWEST TERRITORIES
YUKON
Anchorage
Whitehorse
Juneau
Iqaluit
Davis Strait
Qaqortoq
Bering Sea
Labrador Sea
BRITISH COLUMBIA
Canada
ALBERTA
SASKATCHEWAN
MANITOBA
Edmonton
Calgary
Regina
Winnipeg
Vancouver
ONTARIO
QUÉBEC
N.L.
St. John's
Quebec
N.B.
MAINE
N.S.
WASH.
MONTANA
N.D.
MINN.
Salem
Saint Paul
Ottawa
Toronto
ORE.
IDAHO
S.D.
WIS.
MICH.
VT.
WYO.
Madison
Detroit
N.Y.
MASS.
NEBR.
Omaha
IOWA
Chicago
PA.
Reno
NEV.
UTAH
COLO.
United States
ILL.
IND.
OHIO
New York
KANS.
MD.
W.VA.
San Francisco
MO.
KY.
VA.
CALIF.
Las Vegas
OKLA.
TENN.
ARK.
N.C.
Los Angeles
ARIZ.
N.M.
S.C.
Tijuana
Dallas
MISS.
ALA.
GA.
Ciudad Juárez
TEXAS
Jacksonville
B.C.
LA.
Houston
SON.
CHIH.
FLA.
COA.
Sargasso Sea
North Atlantic Ocean
North Pacific Ocean
B.C.S.
Gulf of Mexico
Bahamas
N.L.
SIN.
Mexico
Havana
Cuba
S.L.P.
NAY.
GUAN.
HID.
JAL.
YUC.
Q.R.
MICH.
PUE.
VER.
TAB.
CAM.
Jamaica
Dominican Republic
San Juan
GUE.
OAX.
Belize
Guatemala
Caribbean Sea
Guadeloupe
Martinique
Barbados
El Salvador
Nicaragua
Barranquilla
Caracas
Costa Rica
Trinidad and Tobago
Panama
Venezuela
Paramaribo
Guyana
Bogota
French


GeoGraphVis
A Knowledge Graph Visualization Tool for Disaster Relief
About KnowWhereGraph
Your Feedback
More Details about IAN (2022)...
Layers & Statistics of Ian (2022)
Storm Impact Layers
# of Reports 191
Health Factor Layers
Diabetes Rate
Adult Obesity Rate
Mentally Unhealthy Days
Mental Health Provider Rate
Low birthweight Rate
Uninsured Rate
Annual Mammogram Rate
Flu Vaccination Rate
Food Environment Index
Injury Deaths
Uninsured Rate (%)
8.11
27.24
Find Experts...
2-Variable Scatter Plot
x-axis Diabetes Rate Linear
y-axis Adult Obesity Ra Linear
Adult Obesity Rate (%)
Diabetes Rate (%)
Disaster Type
Hurricane
Hurricane IAN 2022 | 2022-09-27T05:00Z - 2022-10-01T19:00Z
Using Accumulation
# of events
9/27-5:00
9/27-18:00
9/28-8:00
9/28-22:00
9/29-12:00
9/30-1:00
9/30-15:00
10/1-5:00
Reset Camera
2D View
Home


Feature
Flood Event
Hazard Event
Hail Event
Fire Event
Dust Devil Event
Waterspout Event
Wildfire
Santa Barbara
SANTA BARBARA
County

WHAT

HOW

The 30,000 Foot View

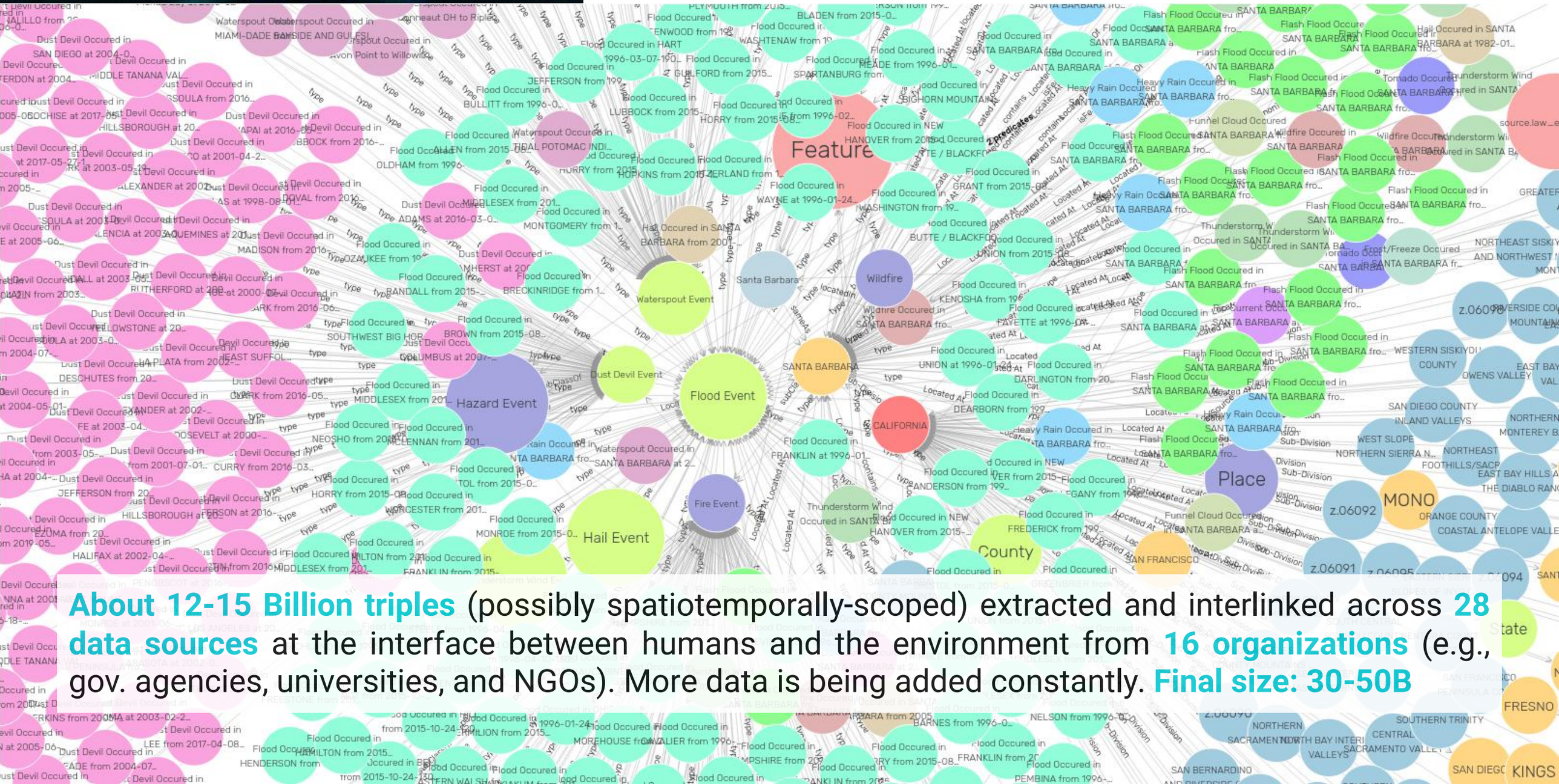
KnowWhereGraph serves 8+ different types of globally unique place and **region identifiers** and 10+ different thematic layers containing **millions of facts about these regions.**

- Discrete, Hierarchical Global Grid System
  S2 Cells (L9 globally, L11 USA, L14 CA, L16 for urban (*))
- Global Administrative Areas
- US Federal Judicial District (*)
- National Weather Zones
- FIPS Codes
- DMA (*)
- ZIP (*)
- Climate Division

Our region identifiers are linked to Wikidata/Linked Data Cloud.
A team/customer that uses our identifiers gets access to billions of triples outside in addition to ~1.2B from the current KWG.

KnowWhere Graph (*)

(*) and KnowWhen



About 12-15 Billion triples (possibly spatiotemporally-scoped) extracted and interlinked across **28 data sources** at the interface between humans and the environment from **16 organizations** (e.g., gov. agencies, universities, and NGOs). More data is being added constantly. **Final size: 30-50B**

KnowWhere Graph

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			
Expert - Covid-19 Mobility	Direct Relief (DR)	name, affiliation, expertise	Global	2021	US Federal Judicial District	DoJ, ESRI	US
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	CDC, US Census	below poverty level percent, diabetes age adjusted 20 plus percent, obesity age adjusted 20 plus percent	US	2017			
Social Vulnerability	CDC/ATSDR	social vulnerability index	US	2018			
Hurricane Tracks	NOAA	max wind speed, min pressure	US	1851-2020			

https://www.knowwheregraph.org/graph/

[This overview is always outdated]

Classical Knowlege Representation and Reasoning

F. Baader, W. Nutt

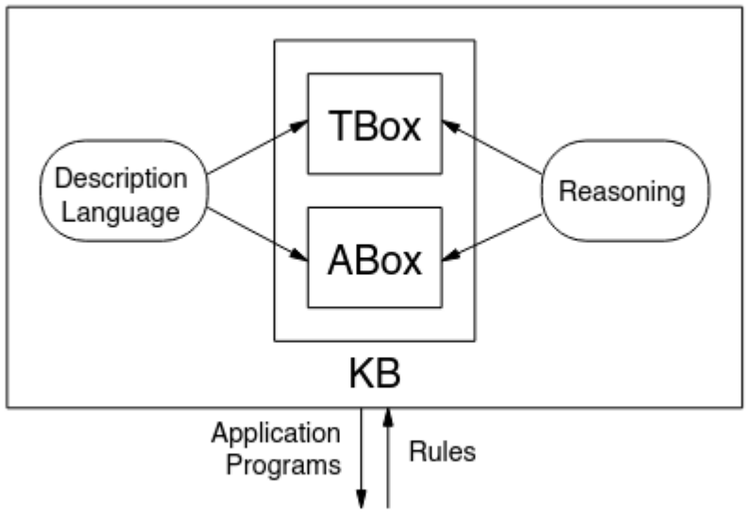
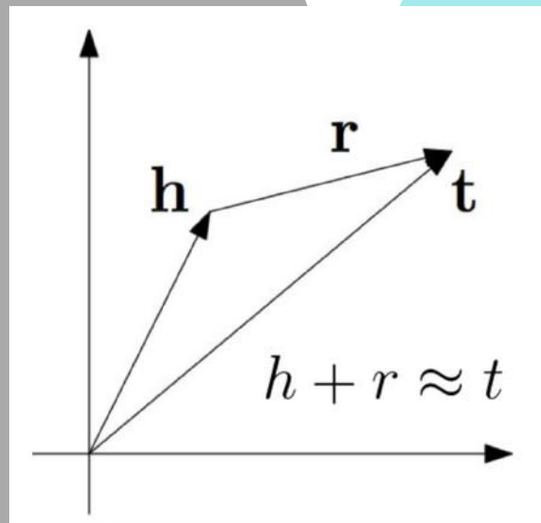


Fig. 2.1. Architecture of a knowledge representation system based on Description Logics.

- Family of **satisfiability**-based reasoning services
  - (**Concept**) subsumption: $KB \models C \sqsubseteq D$
  - **Instance** checking: $KB \models C(a)$
- Requires (almost) **no data**
- **No uncertainty** at the cost of the **inability to handle noise**

Representation Learning and Its Downstream Tasks



- E.g., learning **embeddings** as low(er)-dimensional **vector** (space) representations of entities and their relations
  - **Link prediction**: predicting unknown/missing statements between (two) entities.
  - **Node classification**: predict the class membership of unlabeled entities
- Requires a lot of (unbiased/representative) data
- **Great ability to handle uncertainty** and **noise**

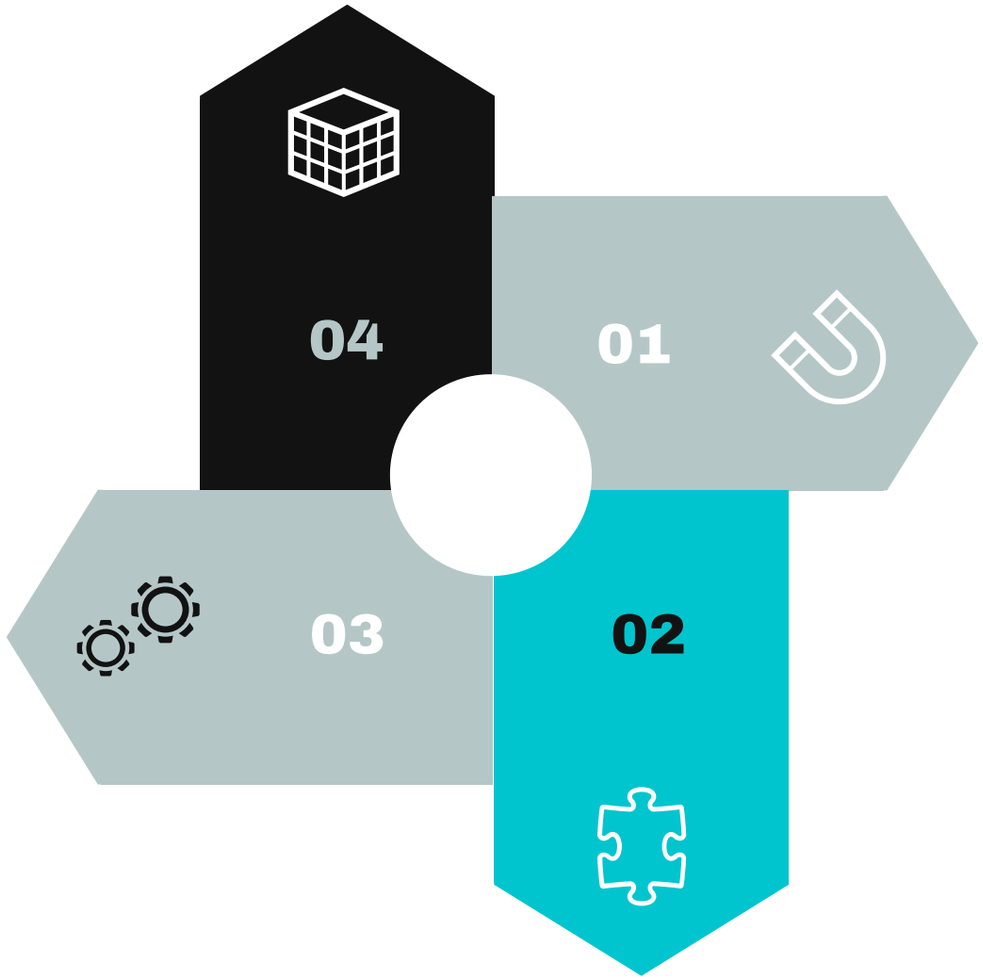
KWG and Epidemic Intelligence from Open Sources (EIOS)?

AI/ML Ready

KWG data (down to the level of individual observations) is directly ready for usage as features in AI/ML applications.

Not Another Portal

Instead of being the next software-based portal (remember Flash?), we serve smart data ready to use with off-the-shelf software.



Finding the right people

We use ML technology to infer expertize and recommend experts across topics and geographies.

Cross Domain

Querying from KWG into SPOKE allows crosswalks from the place, to its population, local exposure, down to the pathogen and treatment.

Team and Partnership

Team

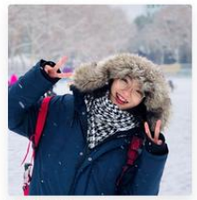
Randall Barker
Senior Personnel
IN10T

Bruno Basso
Senior Personnel
MSU

Ling Cai
Student at STKO Lab
Department of Geography
UC Santa Barbara

Ty Fitzpatrick
Senior Personnel
Esri

Catherine Foley
Senior Personnel
MSU



Zhining Gu
PhD Student
Arizona State University

Tim Murphy
Senior Personnel
Esri

Paulina Oliva
Senior Personnel
USC

Andrew Schroeder
VP of Research and Analysis
Direct Relief

Yuanyuan Tian
PhD Student
Geography
School of Geographical Sciences and Urban Planning
Arizona State University

Austin Truchan
Designer
MSU

Dalia Varanka
Research Scientist
U.S. Geological Survey

Kitty Currier
Postdoc
UCSB

Frank Davenport
Research Scientist
UCSB Climate Hazards Center

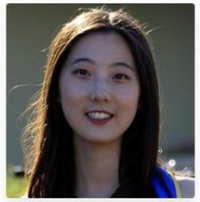
Anthony D'Onofrio
Developer
MSU

Tony Howser
Senior Personnel
UCSB

Matt Jones
Senior Personnel
UCSB

Zilong Liu
Ph.D. Student
STKO Lab
Department of Geography
UC Santa Barbara

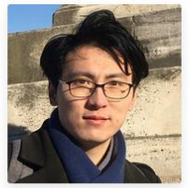
Alica Sheill
Center Manager
Matrix
Michigan State Unviersity



Meilin Shi
PhD Student
UCSB

Cogan Shimizu
Postdoc
K-State

Sizhe Wang
PhD Student
The School of Computing, Informatics, and Decision Systems Engineering
Arizona State University



Zhangyu Wang
Graduate Student
UCSB

Dawn Wright
Chief Scientist
Esri and Professor of Geography and Oceanography
Oregon State University

Karen Doehner
UCSB

Andrew Eells
Research Assistant
DaSe Lab
Kansas State University

Colby Fisher
Senior Personnel
Hydronos Labs, Oliver Wyman

Anna Lopez-Carr
Monitoring and Evaluation Specialist
Research and Analysis Group
Direct Relief

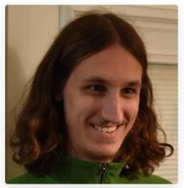
Myles McHugh
Senior Personnel
Hydronos Labs

Bryce Mecum
Science Software Engineer
NCEAS, UCSB

David Smith
Soil Scientist
Hydronos Labs, Oliver Wyman
New York

Shirly Stephen
Postdoc
UCSB

Thomas Thelen
Software Engineer
NCEAS, UCSB



Joseph Zalewski
Ph.D. Student
DaSe Lab
Kansas State University

Rui Zhu
Postdoc
UCSB

Mike Matheis
Oliver Wyman

PIs

Krzysztof Janowicz
Principal Investigator
Geoinformatics
UC Santa Barbara, USA
University of Vienna, Austria

Pascal Hitzler
Computer Science
Kansas State University

Mark Schildhauer
Environmental Sciences
NCEAS, UCSB

Wenwen Li
Urban Planning and Geoinformatics
Arizona State University

Dean Rehberger
Humanities
Michigan State University

Former Team

Seila Gonzalez
Head of Software Development
Matrix
Michigan State University

Gengchen Mai
Ph.D. Candidate
Space and Time for Knowledge Organization Lab
Department of Geography
UC Santa Barbara

Scott Robinson
Partner
Oliver Wyman's Commodity and Risk Practices

Lynn Usery
Senior Scientist
U.S. Geological Survey

Lu Zhou
Postdoc
K-State