KNOWHERE GRAPH

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

Presenting:
Cogan Shimizu, WSU


KNOW
WHERE GRAPH


esri
THE SCIENCE OF WHERE™


HYDRONOS LABS
OLIVER WYMAN
DirectRelief


FMI
THE FOOD
INDUSTRY
ASSOCIATION


CiBO
TECHNOLOGIES


ontotext


WRIGHT STATE
UNIVERSITY


USDA


USGS
science for a changing world


USC


UCSB


Climate
Hazards
Center
UC Santa Barbara


KANSAS STATE
UNIVERSITY


spatial@ucsb
CENTER FOR SPATIAL STUDIES


ASU


NCEAS
National Center for Ecological Analysis and Synthesis


MICHIGAN STATE UNIVERSITY

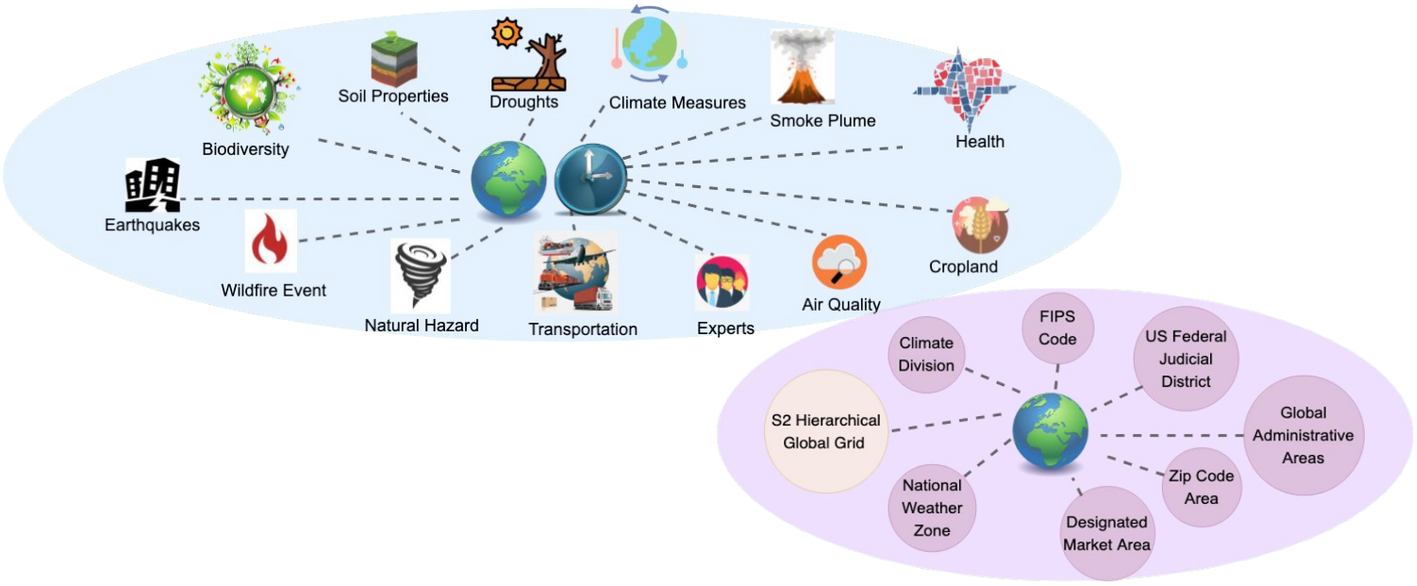
Project Overview

We make data-driven decision-making affordable and effective by providing data analysts with contextual information about their study area on-demand.

- **A gazetteer of gazetteers:** creates many different place type identifiers and S2 cells for generating essentially any possible geospatial region of interest. These KWG place and region identifiers can be used by any other team.
- **Integrated thematic layers:** develops geographic regions that range from extreme events, soils, and crops to social vulnerability and critical infrastructure.
- **A hub for geo-enrichment:** allows any OKN – or any other project – to access millions of facts about regions to enrich their own data.
- **Homogeneous data access:** KWG integrates and links the data to increase graph density through a uniform observation (and sensors) driven framework.

Project Overview

- 10 **region identifiers** + S2 cells for any shape on Earth
- 20+ **thematic layers** across domains
- Space (and time) as the nexus to connect them
- Adheres to standards by W3C and OGC
  - Semantic Sensor Network Ontology (SSN)
  - Time Ontology in OWL (OWL-Time)
  - GeoSPARQL



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			

Achievement of Project Objectives

- **Three Major Open Graph Releases:**
  - Vienna, Manhattan, Santa Barbara (forthcoming)
  - 20B+ triples in the graph – 1000% growth over the project
  - Expansive ontology covering 150 classes
- **Repositories:** 40 resources containing code, tools, scripts, resources, data, schema/ontologies
  - Many open, more to be opened
- **Broad:** Knowledge Explorer, Faceted Search, KWG API, and more
- **Pilots:** GeoGraphViz (Direct Relief), Cropland Impacts Tool (Food Industry Association), Land Valuation Tool (Farm Credit Association)
- **Cross-track Integration:** SPOKE and UFOKN

- **Re-assertion of the primacy of place:**
  - Even with all the integrative power of KGs, it's still easy to lose the context of a place
  - In many ways, **we have have succeeded** in recontextualizing and rehumanizing what a place or event *is*.
- **Leveraging Topological Linking:** Our work flips the 80/20 ratio from same-as or equivalency links to reuse of clean, curated, contextualized identifiers

Success Stories

- **KWG KnowHow's Growing List of**
  - Inquiries for lessons learned;
  - Research outcomes used by key players in Silicon Valley in their products now (e.g., Spatially-explicit ML);
  - Close collaborations with industry partners.
- **Proliferation of KWG Fundamentals:**
  - Follow-on Projects;
  - Institute for Digital Agriculture and Advanced Analytics (ID3A) at Kansas State University;
  - Transfers to other (emerging) domains such as Conflict & Hazard graphs for Ukraine, Israel, etc.;
  - **89 publications** over the entire project lifetime.
- **Outstanding Placement for Team alumni:**
  - 3 Tenure-track professor positions;
  - Major industry players including Google, IBM, TigerGraph, Walmart, TBL's new SOLID Inrupt, and others on graph related work.
- **Cross-Track and Program Integration:**
  - SPOKE & UFOKN Integration;
  - Proto-OKN Collaboration –
    - DreamKG, Neighborhood Safety, IJP
    - Geospatial Working Group

Project Archiving

&

&

KnowWhereGraph hosted at **National Center for Ecological Analysis and Synthesis (NCEAS)** for longer term archiving

Software & Tools is hosted under open and permissive licenses in a family of **GitHub repositories** under the KnowWhereGraph organization

Institutional Mirrors will be hosted at Wright State University and Kansas State University

Future Plans

KWG API

Enable broader sets of tools through programmatic access to KnowWhereGraph without any knowledge of SPARQL or semantics.

https://github.com/KnowWhereGraph/kwg-api

Continuous Integration

Extend the framework by which we keep the graph current, including the ability to add new data layers.

Commercialization

Exploring options for translating academic success to sustained commercial outcomes

Future Plans

Proto-OKN Geospatial WG

KnowWhereGraph will provide **hard-won expertise** in geosemantics and KG-based geospatial data representation, analysis, to the Proto-OKN cohort

Nexus for Topological Place Linking

KnowWhereGraph provides a many different ways for all other OKN to link into the graph thereby also linking to each other indirectly and geo-enriching all their data

Follow-on Projects

Towards a Global Food Systems Data Hub: Seeding the Center for Sustainable Wheat Production
Kansas State University

KWG-Ohio: Building a Digital Twin of Ohio
Wright State University

WRIGHT STATE UNIVERSITY

KANSAS STATE UNIVERSITY

Post-Project Recommendations

- Community building really difficult
  - It doesn't happen by accident, focus on a principled approach from the beginning
  - Corollary: Avoid Pandemics

Icon from flaticon.com

- Strict timeline may sometimes be difficult to handle
  - Especially in challenging times: again, avoid pandemics ;)
  - Research or exploratory outcomes "on-demand" are hard to predict, which means an optimized team is very important (see next point)

- When mixing foundational research with production-ready software development, devs are key but they are frequently hired away the moment they show success, and the competition pays 300%.

Thanks!

Offline questions can be directed to:

cogan.shimizu@wright.edu