KNOW WHERE GRAPH FOR SUSTAINABILITY

Or, “Context Makes Sustainable Places”

Cogan Shimizu
Wright State University

KnowWhereGraph?

KnowWhereGraph

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 one else.
- **Integrated thematic layers:** develops geographic regions that range from extreme events, soils, and crops to social vulnerability and critical infrastructure.

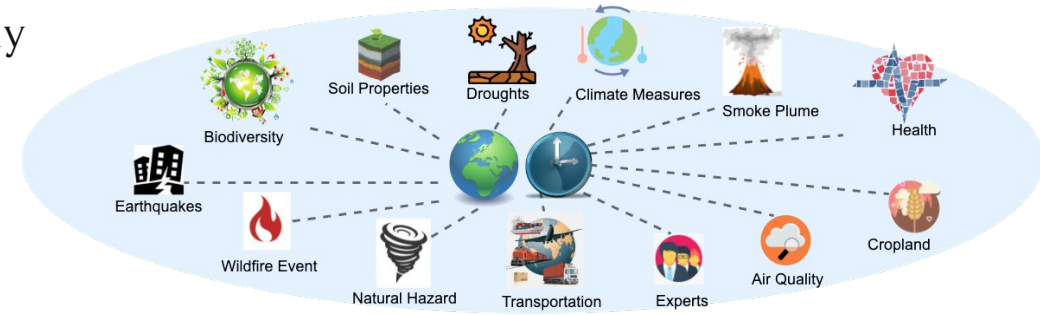
KnowWhereGraph

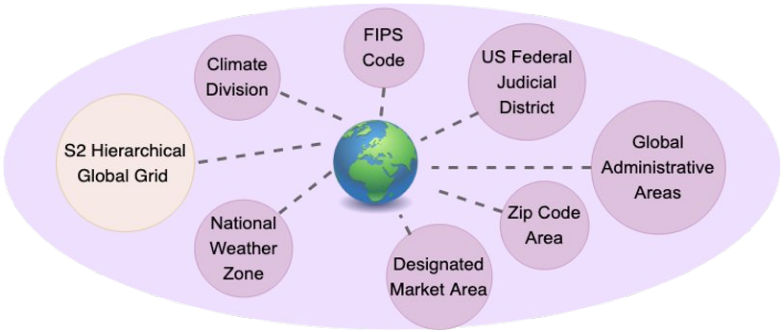
We make data-driven decision-making affordable and effective by providing data analysts with contextual information about their study area on-demand.

- **A hub for geo-enrichment:** allows any 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.

KnowWhereGraph: *Layers*

- 10 **region identifiers** + S2 cells for any shape on Earth
- 20+ **thematic layers** across domains
- Space (and time) as the nexus to connect them
- Results in 16B triples about the world





Datasets

- Data layers are primarily US-Centric and recent years... for now!

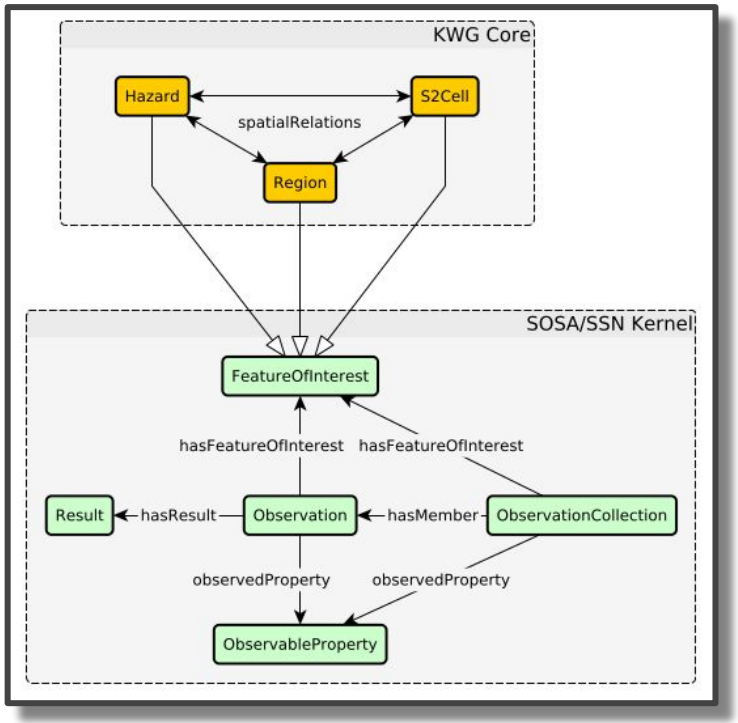
Ontologies

- SOSA/SSN, PROV-O, QUDT GeoSPARQL
- Expert/Expertise, HIP
- 3 ODPs

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			

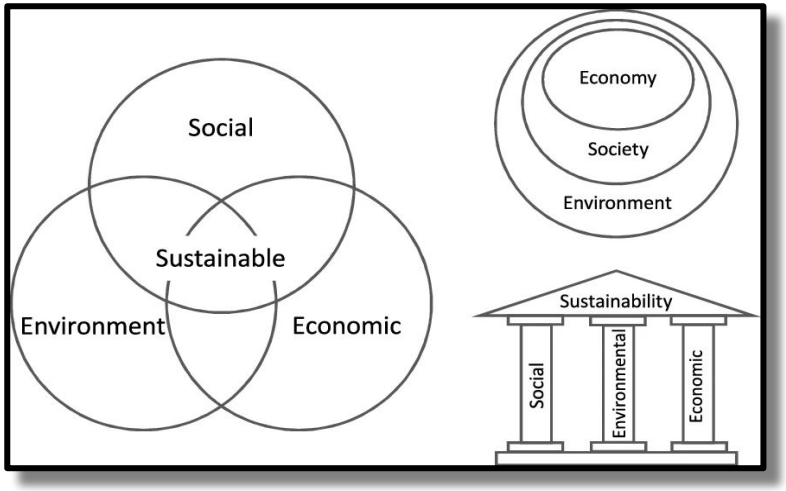
KnowWhereGraph: *Kernel*

- The reusable “core” of KWG
  - KWG Core
  - The SOSA/SSN Ontology and extracted pattern
  - Data about Places, Hazards, People are connected to regions, and to each other through these places



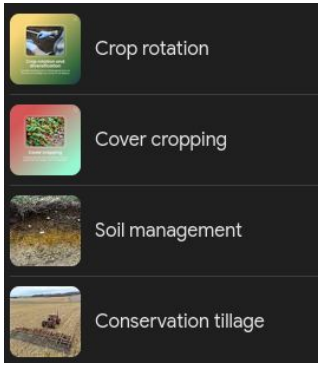
OK, but so what?

- Sustainability practices impact the environments in which they are implemented.



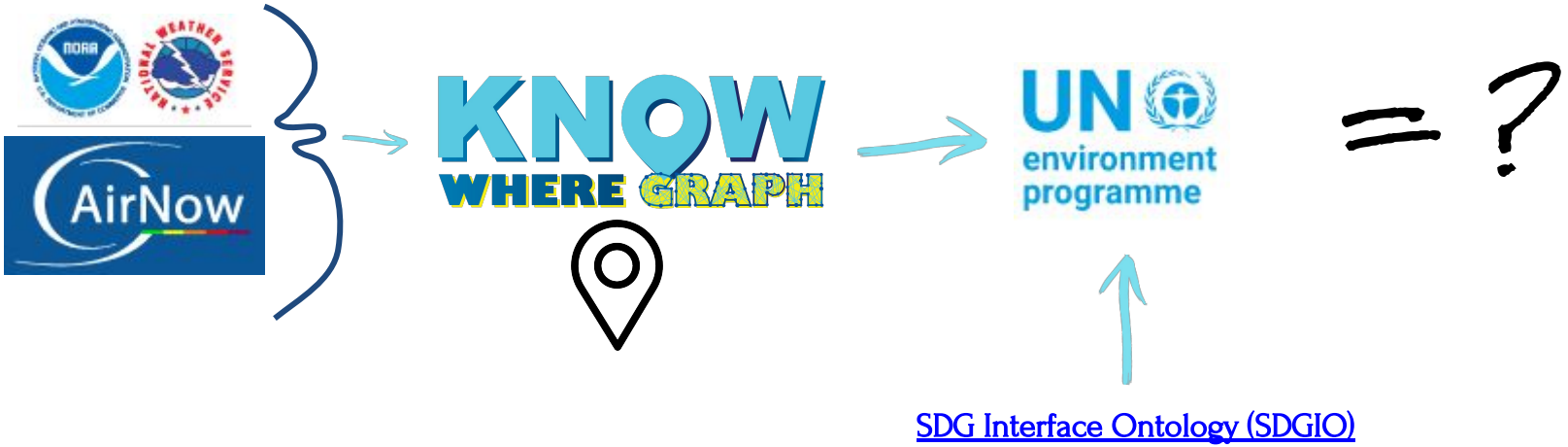
Taken from Wikipedia

Sustainable Agricultural Practices



The **Dust Bowl** was the result of a period of severe dust storms that greatly damaged the ecology and agriculture of the American and Canadian prairies during the 1930s. The phenomenon was caused by a combination of natural factors (severe drought) and human-made factors: a failure to apply dryland farming methods to prevent wind erosion, most notably the destruction of the natural topsoil by settlers in the region.[1][2] The drought came in three waves: 1934, 1936, and

Connecting to an Environment



SDG Interface Ontology (SDGIO)

Is a particular place hitting their Sustainability Goals related to air pollution?

Connecting to an Environment

- Does increasing EV presence *actually* worsen air pollution due to rubber particles?
  - What is the relative impact of reduced car pollution marginalized communities?

Environment – or Place?

- Place has sociocultural meaning or “A Slow Moving Event”
- Caring about “the environment” versus caring about a place (or your community, as the case may be)
- People have connections to la

'A great sadness': Venezuela is first Andean country to lose of all its glaciers (nbcnews.com)

submitted 15 hours ago by Free_Swimming to r/worldnews

303 comments share save hide report crosspost

3.3 Hypothesis 2: neighborhood and community-level flood experiences

Unlike the measure of personal damage experienced in models in Table 2 (and Online Resource Table 7), the measures of perception of community damage was found to be a significant positive predictor in three models. Perception of neighborhood damage was found to be significantly and positively related to climate change beliefs in Model 1 in Ta 2 (and Online Resource Table 7). The measure of zip-code level damage reported by FEM was found to be positively associated with *climate belief* in three of the models. *These res*

Albright, E.A., Crow, D. Beliefs about climate change in the aftermath of extreme flooding. *Climatic Change* **155**, 1–17 (2019). https://doi.org/10.1007/s10584-019-02461-2

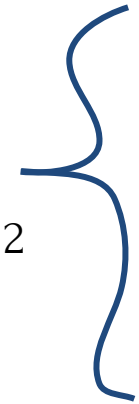
But the weather doesn't have to be that extreme to leave a mark. In a study published in the *Proceedings of the National Academy of Sciences* in December 2016, scientists have found that experiencing record high or low temperatures affects people's stated belief in climate change. The study was funded by the Robertson Foundation and the British Academy.

Spatial heterogeneity of climate change as an experiential basis for skepticism

Robert K. Kaufmann, Michael L. Mann, Sucharita Gopal, +4, and Michelle Gilmore Authors Info & Affiliations

Place & Sustainability

- Taking a look at the KG4S 2024 Proceedings
  - Uses Place: 2
    - `has -> Environment`
    - `:locatedAt ?location`
  - Doesn't (But Could!) Use Place: 5
  - Doesn't Use Place/Out of Scope: 2



- Location (and climate) matters for planting
- (Circular) Economy is tied to places or transport between them
- Water (throughput) is tied to region – or even sets of regions!
- Environment, Social, Governance...

CQs for Place & Sustainability

- Place as Context for Sustainability
  - How will the practice impact a place?
  - What sustainability practices are appropriate for a specific place or environment?
  - Who will (locally or broadly) be effected by the practice?
  - Which (downstream) practices will impact *this* practice?

Icon from flaticon.com

What Next?

- Emphasize *Place* in approaches. It matters to a community!
- Get Involved! KWG is always looking to support new use cases and integrate more data.

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

https://proto-okn.net/

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

KANSAS STATE UNIVERSITY

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

WRIGHT STATE UNIVERSITY

Thanks!

Offline questions can be directed to:

cogan.shimizu@wright.edu

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

KnowWhereGraph: *Use Cases*

Accessing KWG

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

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%.

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

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

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

We are X-Ray for Places

A6677

Deep KG Schema Pros

- Easy to understand for an expert
- Versatile (different applications)
- Easily extendable
- Makes data integration easier
- Carries little ambiguity

Deep KG Schema Cons

- High initial modeling effort (cost)
- Results in complex schema
- Traditionally done via waterfall process

We are addressing these!

Modular Ontology Modeling for High-quality Schema

Traditional Waterfall Model

1

Versatile scope description

2

Divide (and conquer)

3

Design modules from patterns

4

Connect and document

5

Create data artifact

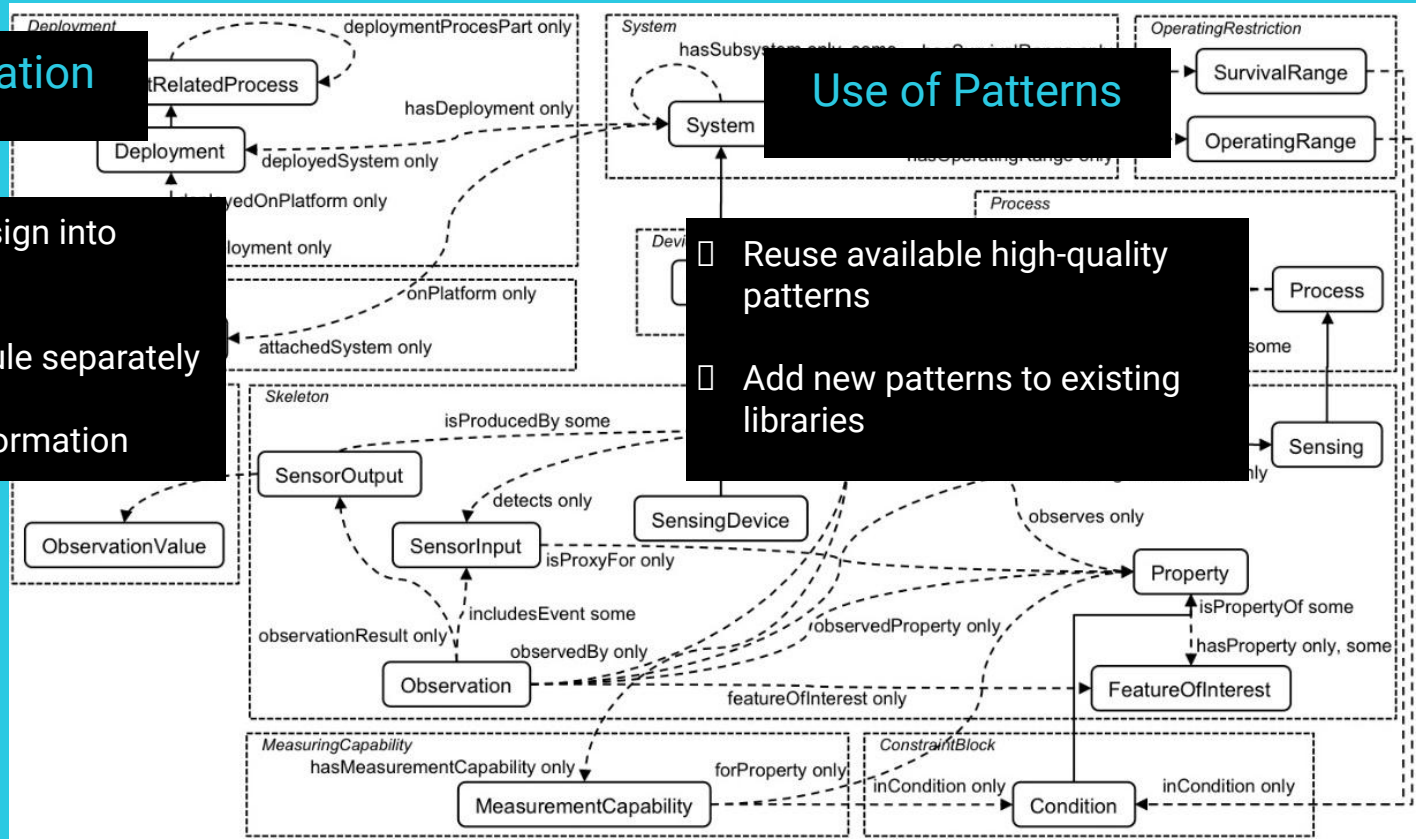
Addressing: High Initial Modeling Effort

Modularization

- Break schema design into sub-problems
- Design each module separately
- Retain module information

Use of Patterns

- Reuse available high-quality patterns
- Add new patterns to existing libraries



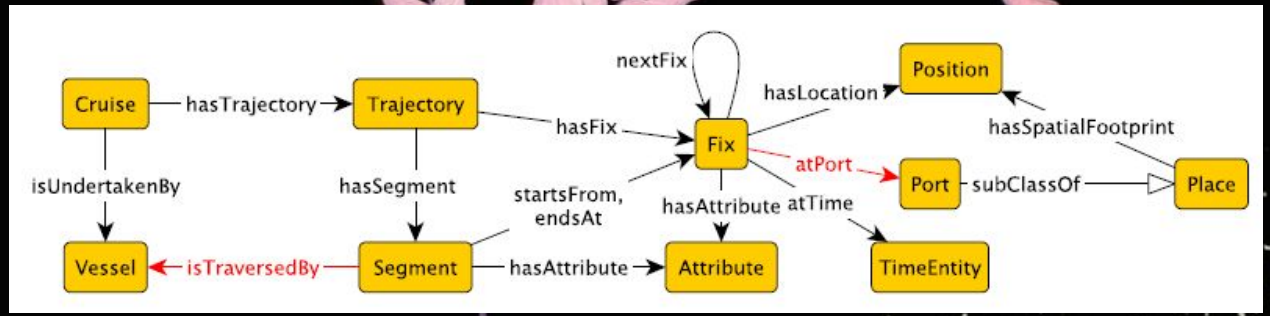
Addressing: Dealing with Complex Schemas

Systematic use of diagrams

- For design discussion with domain experts
- For drafting
- For documentation

Shortcuts and views

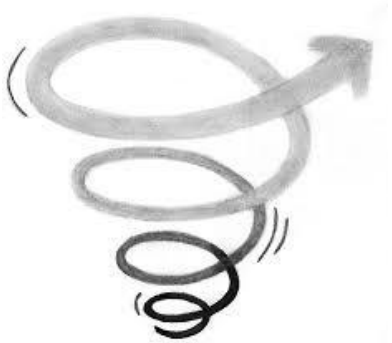
- Complex schema for integration
- Simplified version for interfaces
- Mappings between them



Addressing: Waterfall Process

Have both top-down and bottom-up design

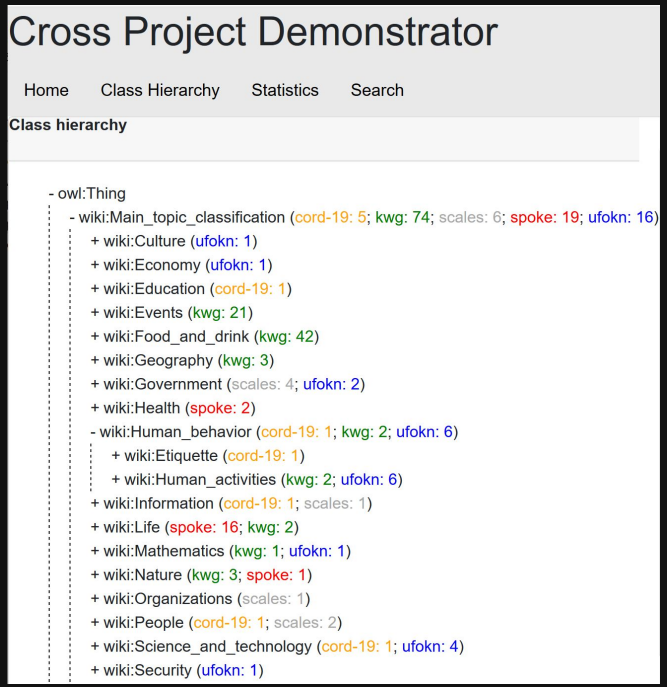
- Top-down
  - I.e. careful design first
  - Focus on quality, reusability, versatility
- Bottom-up
  - I.e. ad hoc, data-driven
  - Focus on speed and data efficiency



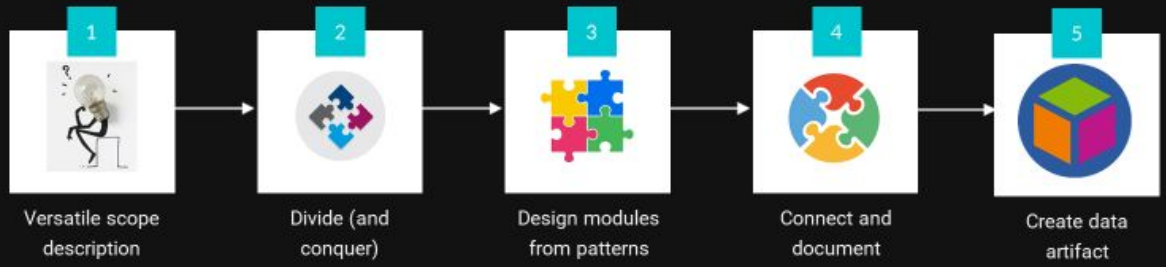
Continuous consolidation and iteration

- Same group of people doing both!
- Goal:
  - Top-down effort to result in the high-quality complex schema
  - Bottom-up effort to result in the compatible shortcut/view version

Cross-project Integration Effort (Track A)



https://cpd.cs.ksu.edu/hierarchy



- Follow essentially the same process
- Do not impose changes on source graphs
- Status:
  - Mutual understanding of sources
  - General agreement on approach
  - First initial scope descriptions

Patterns & Schemas

Reusing (visual) knowledge fragments

Spatially Explicit Design Patterns

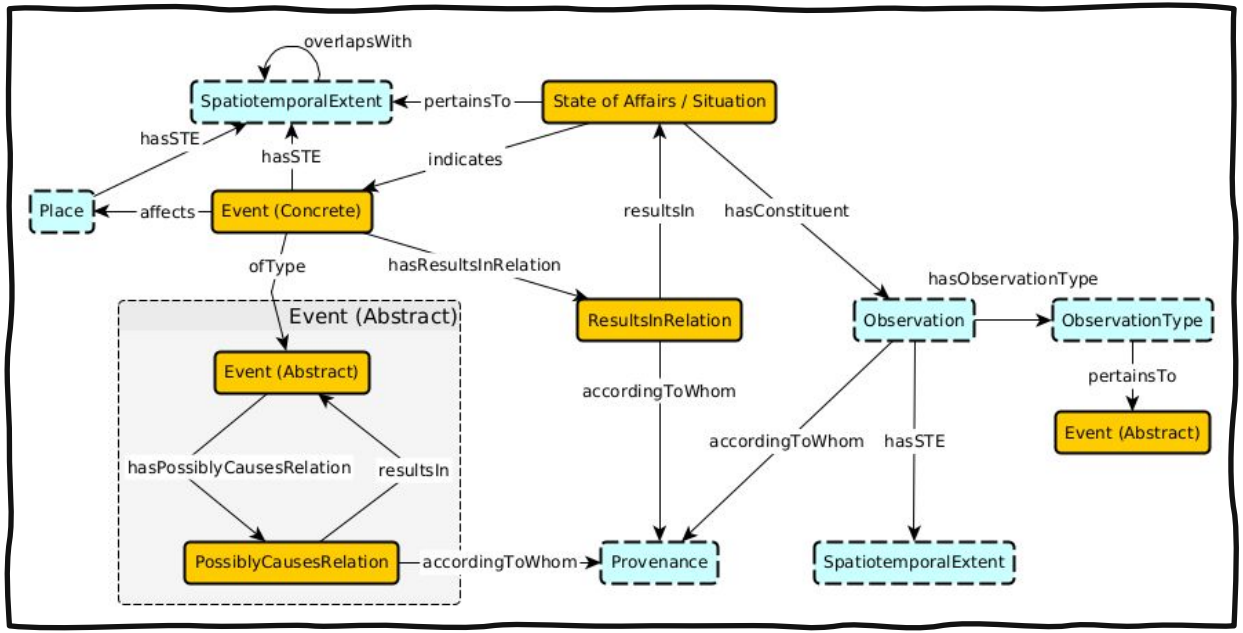
- Interoperablity starts in the schema.
- Reusability is both technical and a mindset.
- To help, we've developed spatially explicit design patterns that can, out of the box, help model data with strong spatiotemporal components.

A pattern for the causal relations:

How are events connected?
Who even says that they are?
How are these events connected to places?

A pattern for features of cells in a hierarchical grid:

What is here?
What *else* is here?

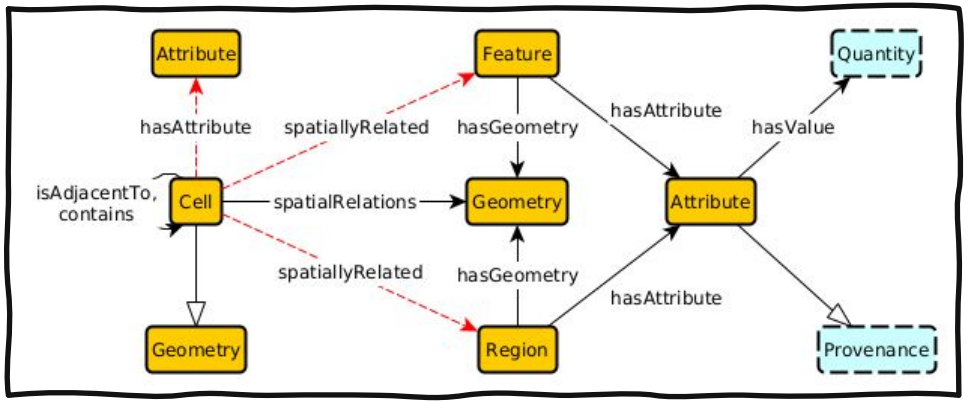


Causal Relationships Between Events

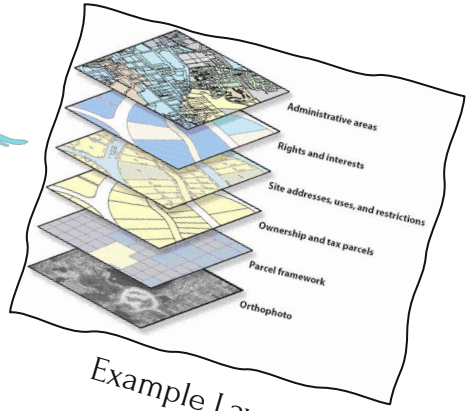
- Distinguishing between the “Concrete” and the “Abstract” notion of an event
- Provenance is a central aspect, which asserts the causality
- Obvious entanglement or alignment to W3C standards

Features of Cells On A Hierarchical Grid

The Cell Feature Pattern



- Commonly encountered conceptualization
- Allows for a common underpinning
- Spatial Axiomatizations (RCC8) come attached



Example Layers

Using Patterns & Schema Diagrams

Discovery & Navigation

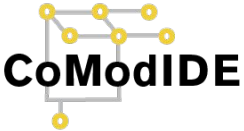
Using a schema diagram provides contextualizes local (or “node”) views of knowledge graphs, enhancing follow-your-nose searching and exploring.

Improved Interoperability

Using patterns improves base interoperability between knowledge graphs, promoting convergence and widespread reusability.

Rapid Iterative Visual Modeling

Patterns, and their schema diagrams, can be organized into libraries, based on coverage or domain, and then subsequently used to enhance modeling workflows.

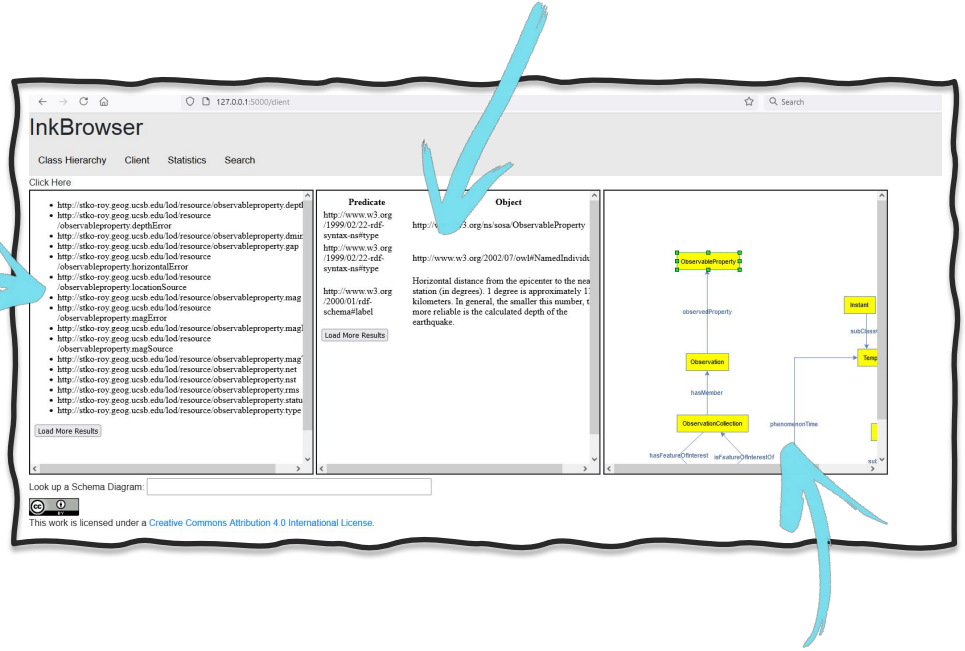


InK Browser

The Interactive Knowledge Browser

"Traditional" view, which lists the attributes of the particular chosen node.

The instance view that shows examples of other nodes of the same type in the graph.



The Schema diagram view that contextualizes other views with a broader view of the knowledge graph. Clicking on the nodes allows for class/type level navigation.

Growing & Year 2

- 16 papers published or submitted to relevant venues on ontology engineering topics alone
- Industry Expo with invited partners, including from Google, Bosch, United Nations, NIST, and AFRL
- more than a dozen keynotes, invited talks, tutorials for different communities

- Extension of schemas for KWG growth and interoperability
- Continued refinement of schema
- Compile spatio-temporal patterns library
- Expand Cross-Project Demonstrator
- Refine Software
- Evaluate software