

Krzysztof Janowicz, UCSB & UniVie (PI)

Pascal Hitzler, K-State (Co-PI)

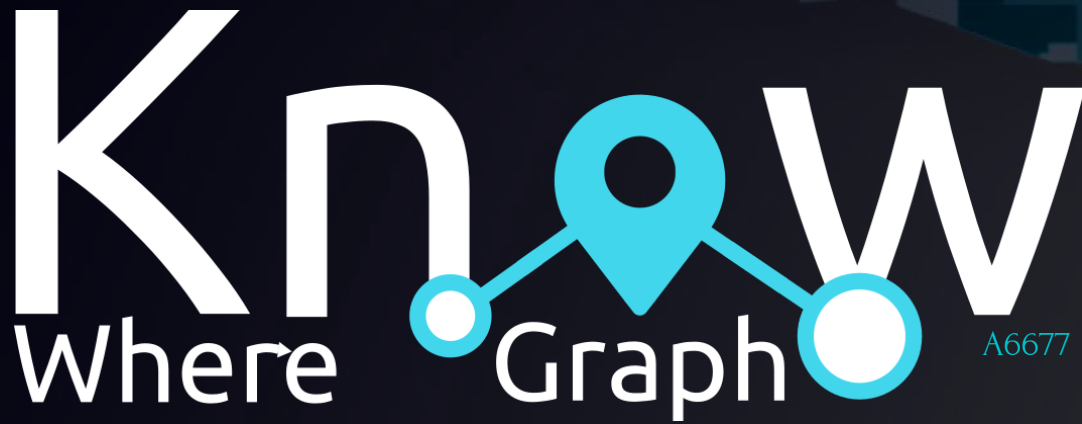
Wenwen Li, ASU (Co-PI)

Mark Schildhauer, NCEAS (Co-PI)

Dean Rehberger, MSU (Co-PI)

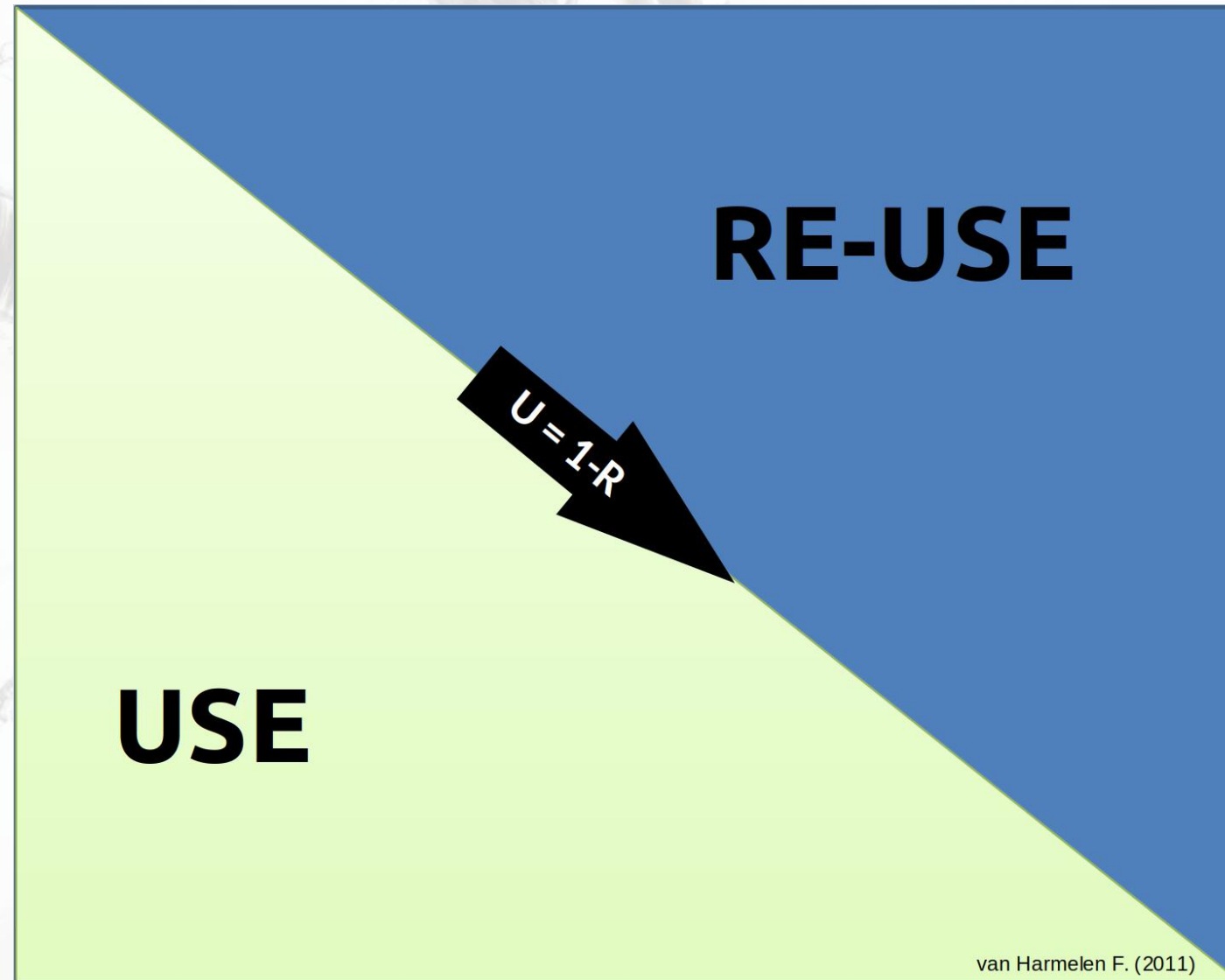
Knowledge Graphs and Disaster Mitigation

Ontology Summit 2022



We provide area briefings within seconds for any region on Earth

Can We Make Useful Data Re-usable?



Probably the most pressing **foundational challenge** in (**meta**) data management.



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

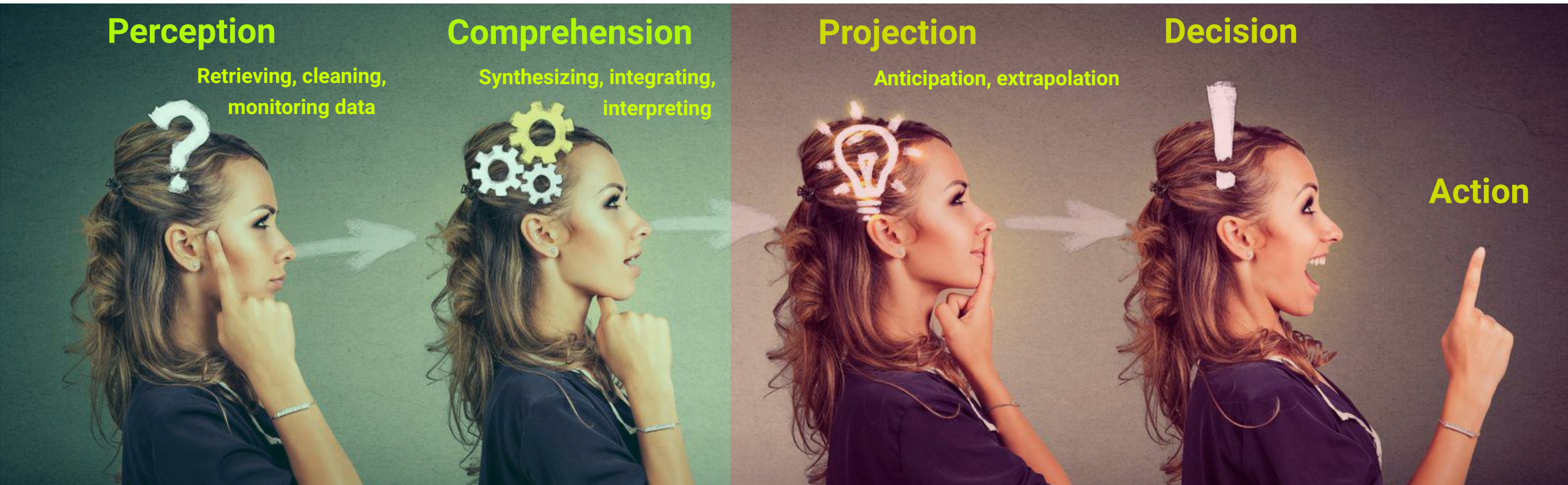
- ***“What is here?”***
- ***“What happened here before?”***
- *“Who knows more?”*
- *“How does it compare to other regions or previous events?”*

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.

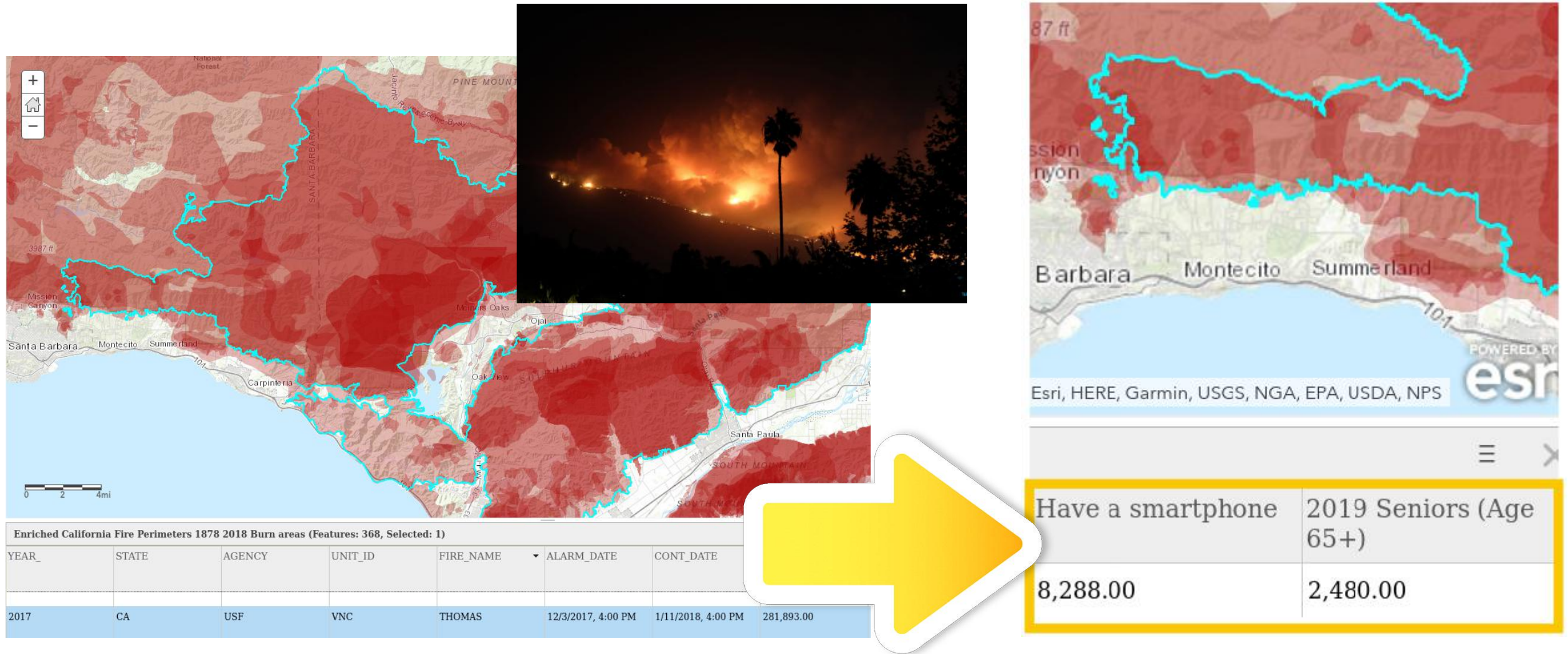
Our ultimate goal is to flip the 80/20 bottleneck on its head.

Situational Awareness

High situational awareness is a prerequisite to good decision-making. We aim at helping decision-makers and data scientists to fast-forward through the **perception** and **comprehension phases**.



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



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

GeoEnrichment Limitations



Pros

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



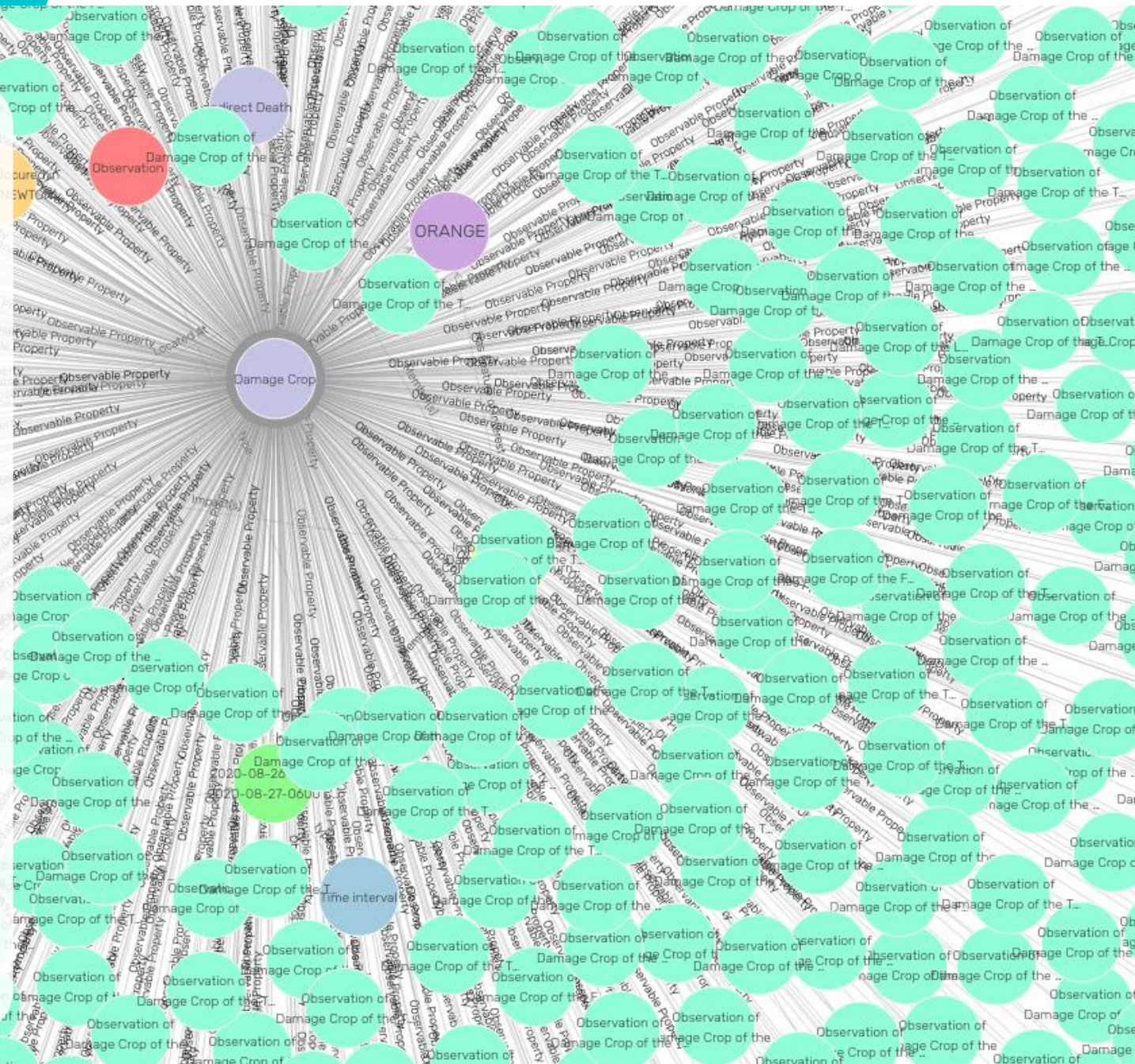
Cons

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

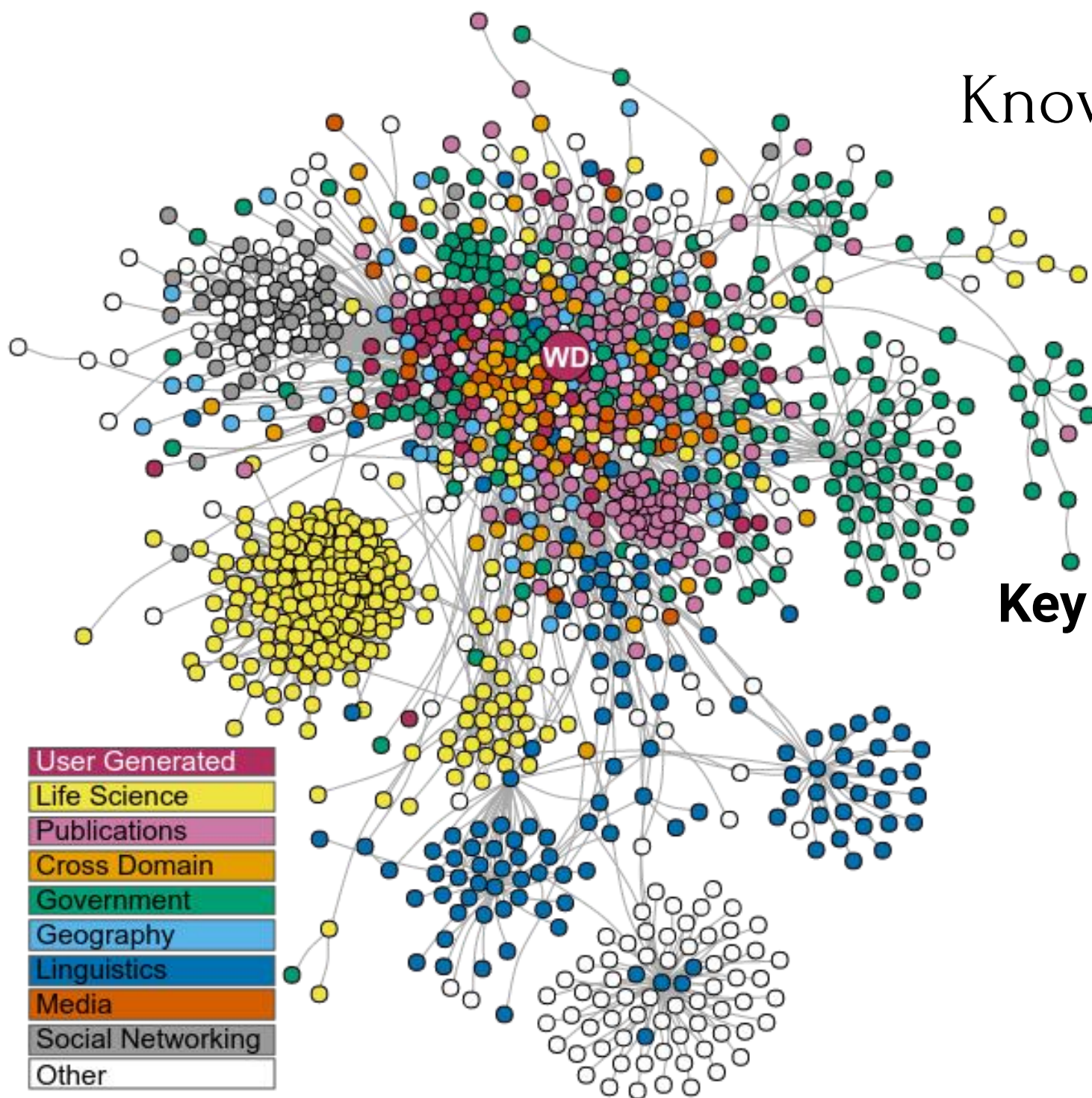
Knowledge Graphs Connect Data Across Themes

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



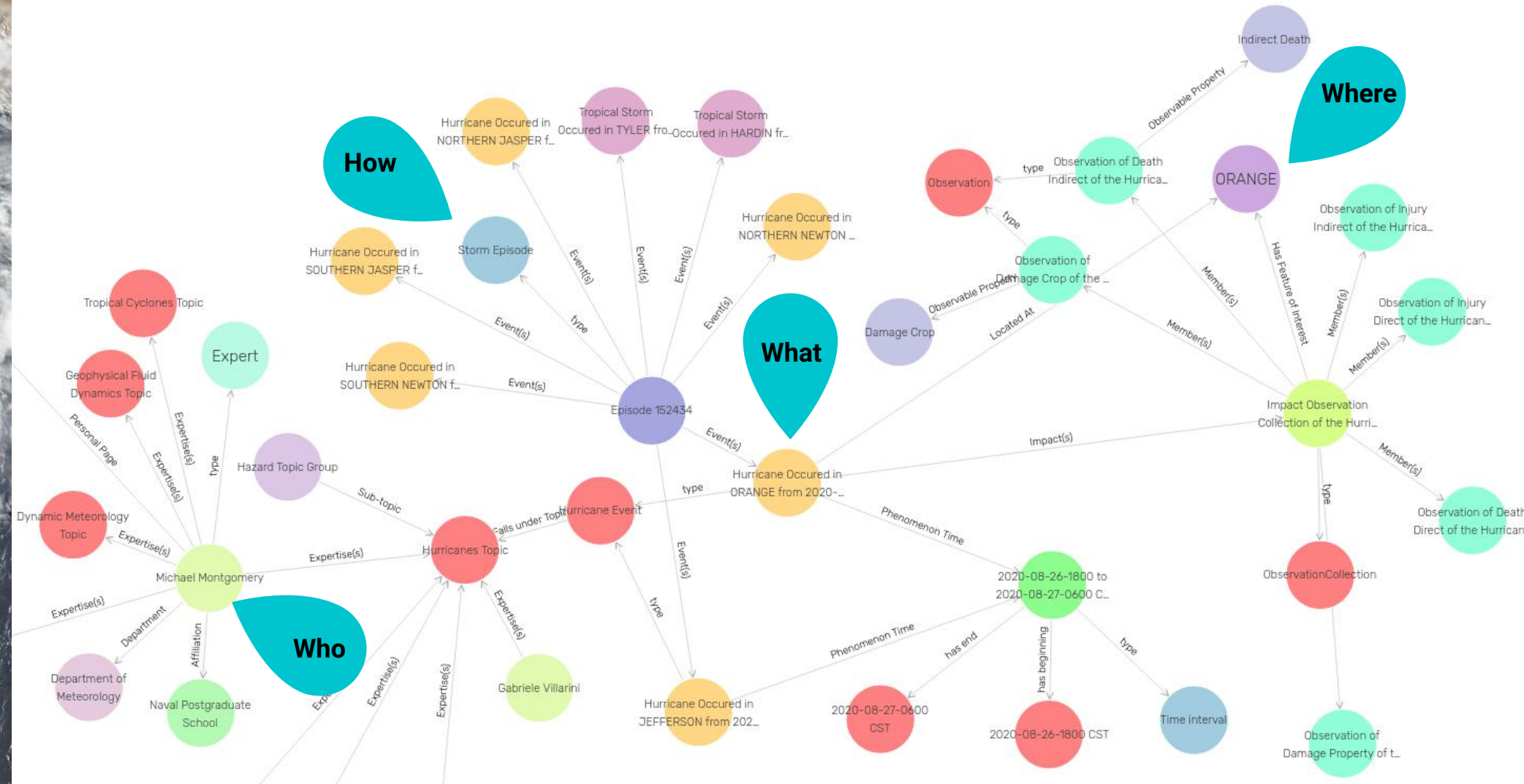
Knowledge Graph Ecosystem



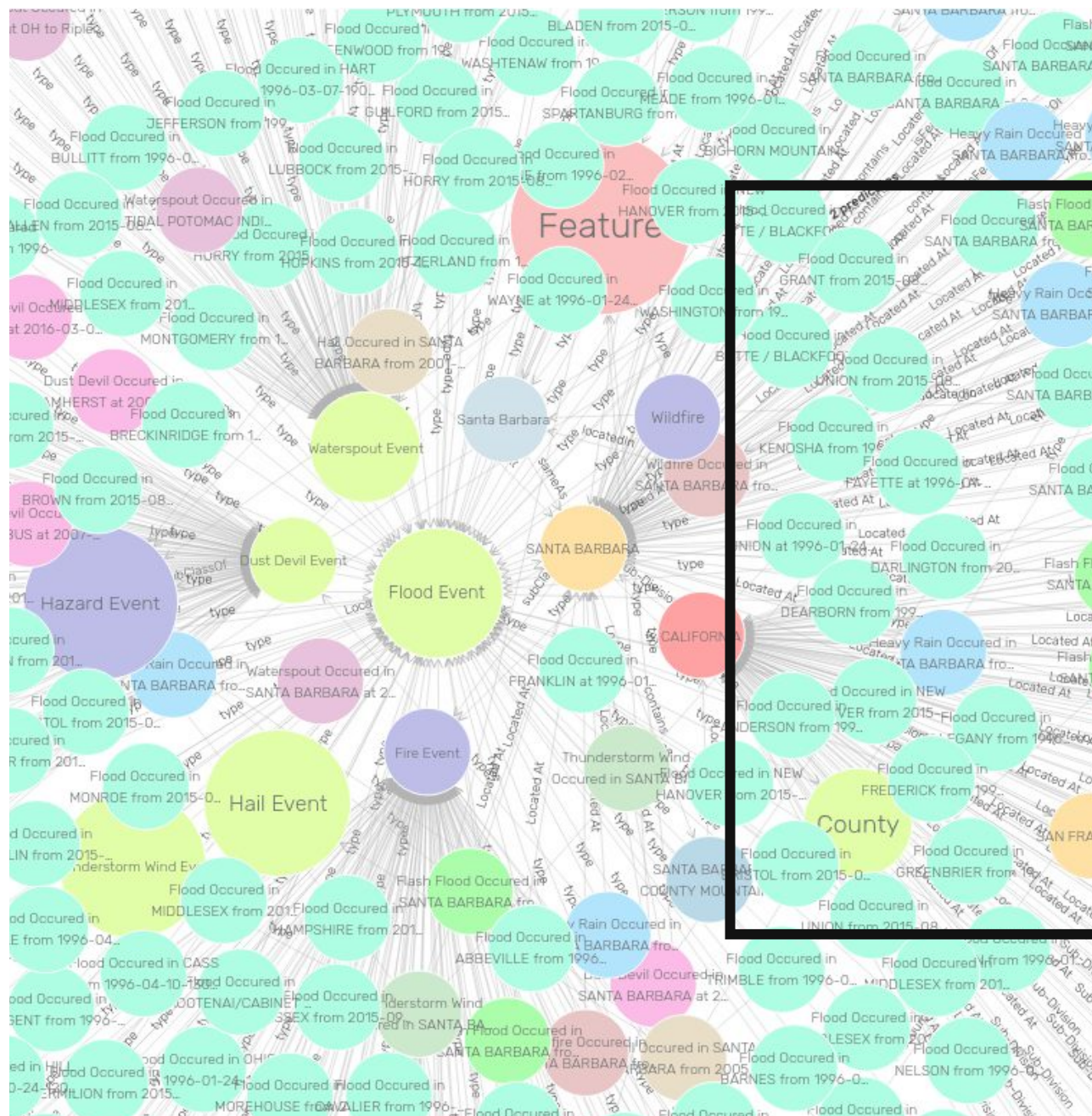
Key Ideas

- Relating (sub)**classes** and relations (*Disaster* <?> *Hazard*)
- Relating **entities** (*Vienna* in graph1 is the **same** as *Wien* in graph2)
- **What about the relation** between *Disaster* in G1 and *Disaster* in G2?
- **Gaining a more holistic understanding!**

KnowWhereGraph = Knowledge Graph + GeoEnrichment

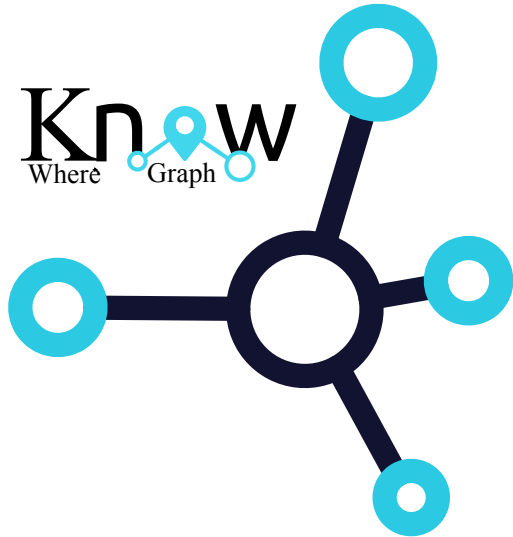


We are the first to **combine geoenrichment with knowledge graphs** to power **environmental intelligence** applications



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.

Hurricanes and COVID-19

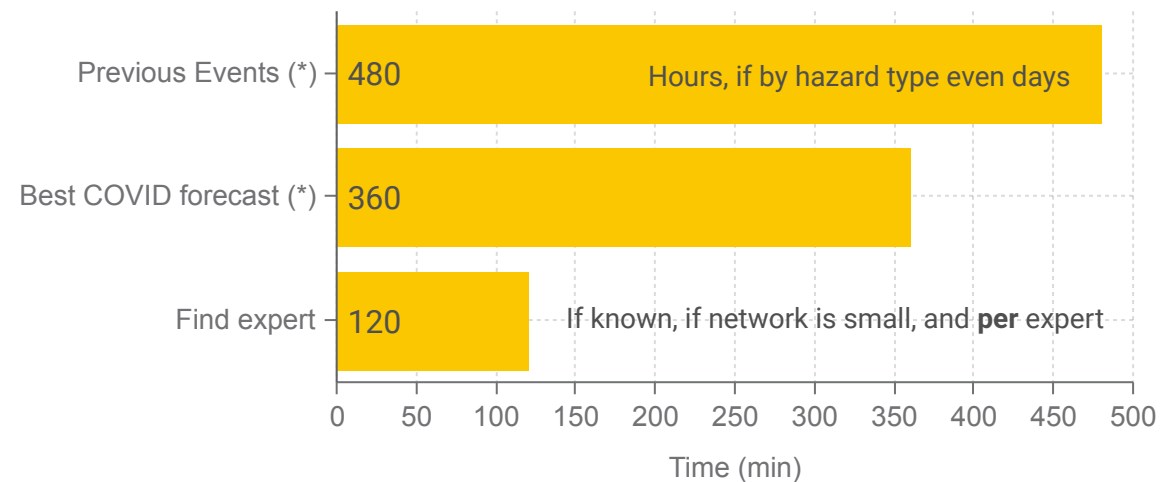
"Many Displaced from Hurricane Laura Are Now in Path of Hurricane Delta"



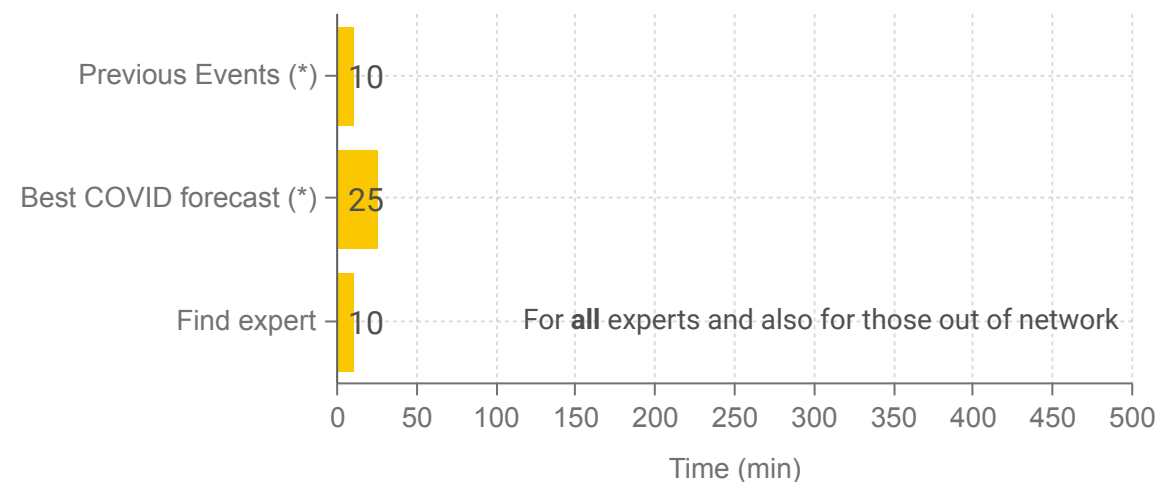
Tasks

- Find predicted path of the hurricane
- Find previous events along the path
- Understand COVID situation in x days
- Find experts for identified perils

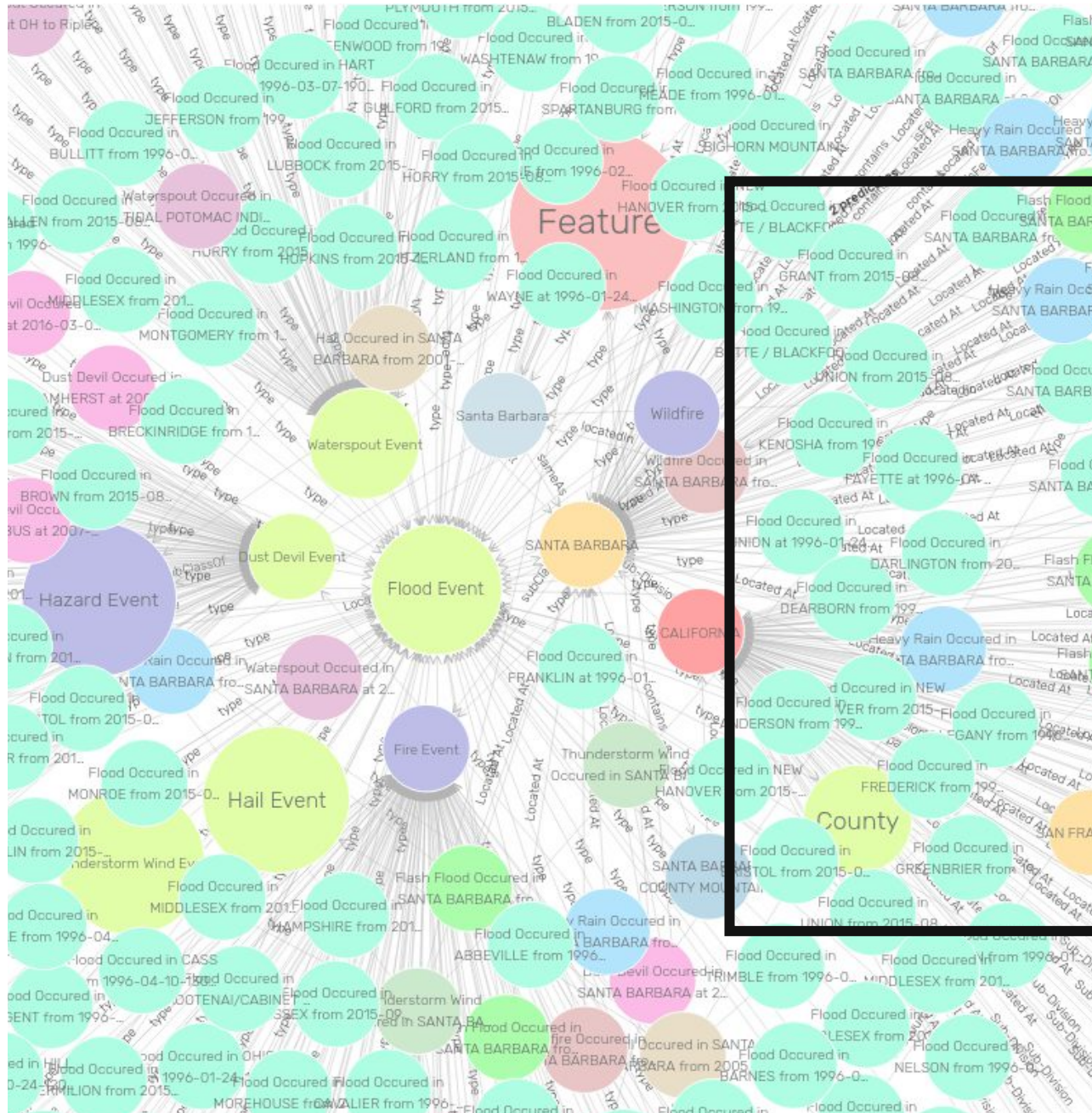
Without KnowWhereGraph GeoEnrichment



With KnowWhereGraph GeoEnrichment



(*) per region, e.g., state or 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

About 12-15 Billion triples (possibly
data sources) at the interface between
gov. agencies, universities, and NGOs.

[illegible]

About **12-15 Billion triples** (possibly spatiotemporally-scoped) extracted and interlinked across **27 data sources** at the interface between humans and the environment from **12 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 Areas	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	1950–2021			
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	1980–2021	Designated Market Area	Nielen	US
Smoke Plumes	NOAA	daily smoke plumes extent	US	2010-2021	ZIP	ZCTA	US
Climate Observations	NOAA	temperature, precipitation, PDSI, PHSI	US	1950 - 2020	Climate Division	NOAA	US
Disaster Declaration	FEMA	DesignatedArea, Program, amountApproved, ProgramDesignatedDate	US	1953 - 2021	Census Metropolitan Area	US Census	US
Smoke Plume Extents	NOAA	Smoke extent	US	2017 - 2021	[This overview is outdated]		
BlueSky Forecasts	Bluesky	PM10, PM5	US	2021-12-08			

Classical Knowledge Representation and Reasoning

F. Baader, W. Nutt

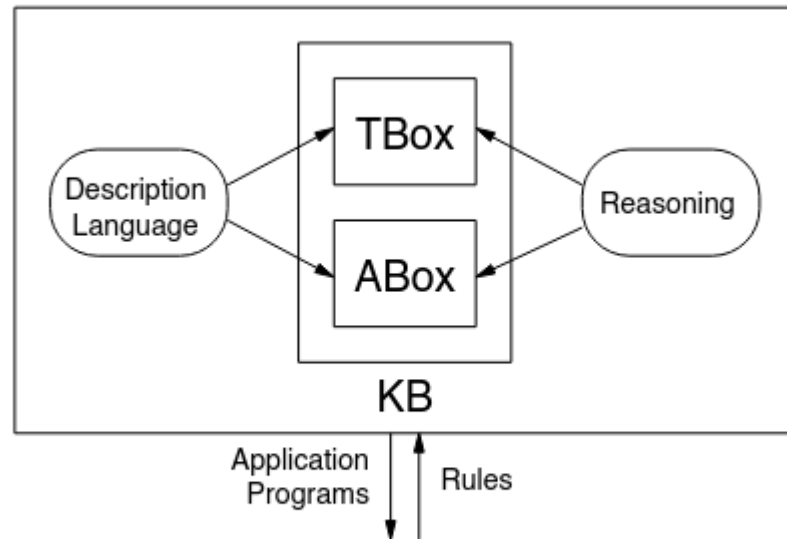
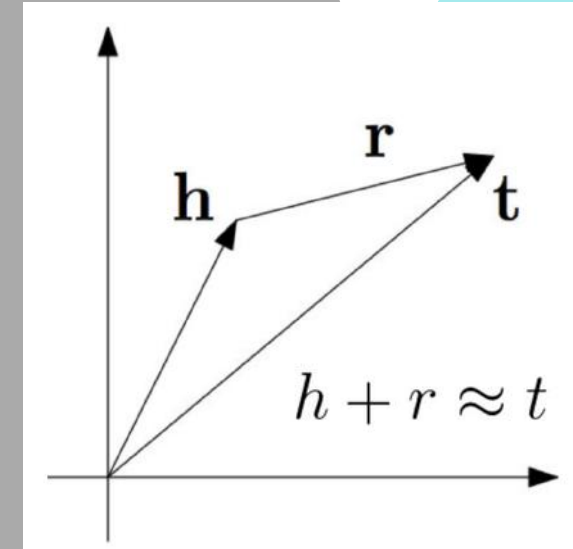


Fig. 2.1. Architecture of a knowledge representation system based on Description Logics.

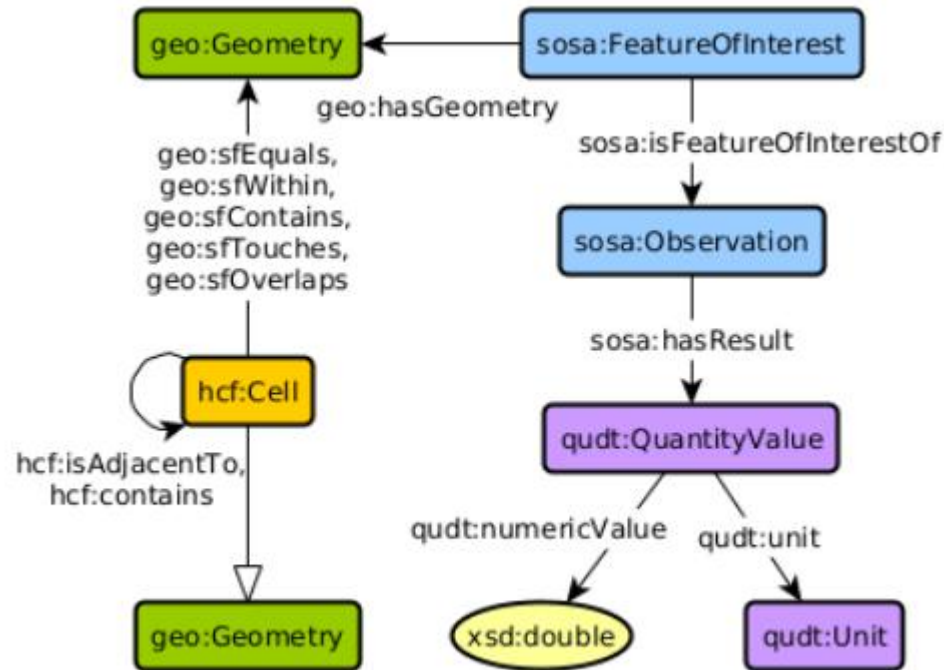
- Family of **satisfiability**-based reasoning services
 - **(Concept) subsumption**: $KB \models C \sqsubseteq D$
 - **Instance checking**: $KB \models C(a)$
- Requires (almost) **no data**
- **No uncertainty** at the cost of the **inability to handle noise**

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**

Classical Knowledge Representation and Reasoning



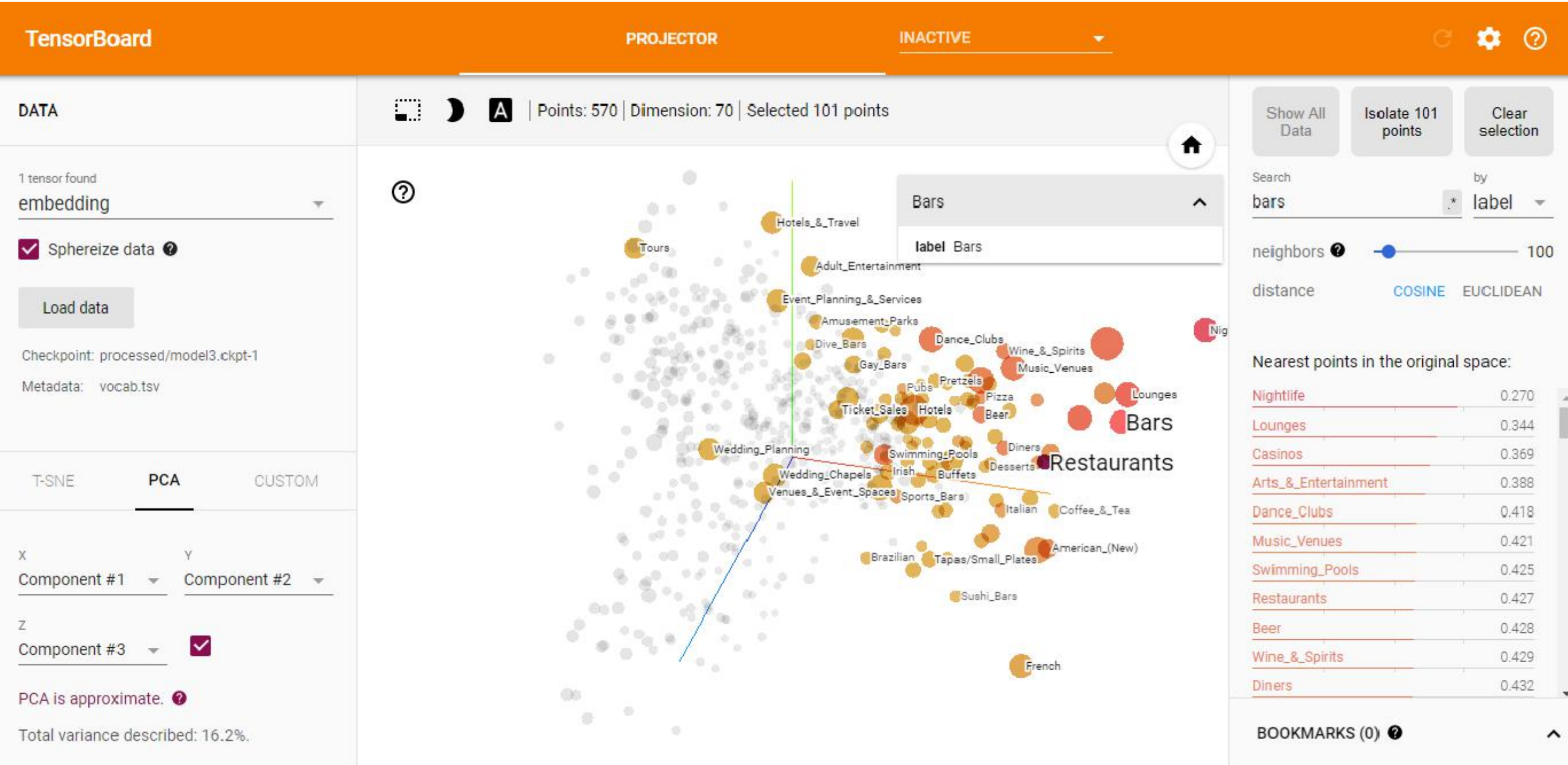
Some Simple Intuitions

- contains(CA,SB) and spatiallyRelated(SB,TF)
then spatiallyRelated(CA,TF)
- How do you **axiomatise** concepts such as *Democracy*? And **who** gets to write the axioms?

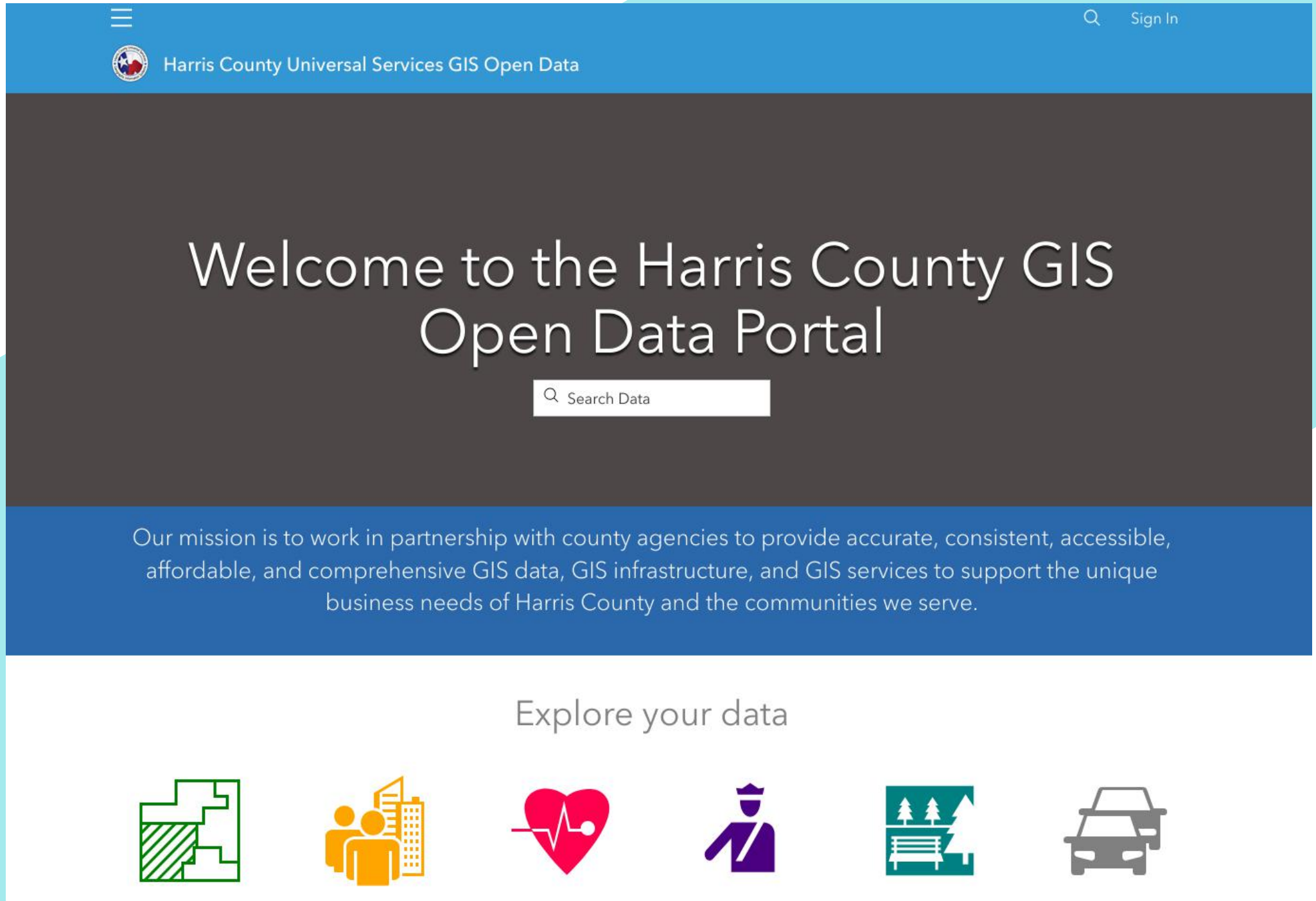
Cell

- $$\begin{aligned} \text{Cell} &\sqsubseteq \text{Geometry} & (1) \\ \text{Cell} &\sqsubseteq \forall \text{spatialRelations}.\text{Geometry} & (2) \\ \text{Cell} &\sqsubseteq \geq 0 \text{spatialRelations}.\text{Cell} & (3) \\ \text{Cell} &\sqsubseteq \geq 0 \text{isAdjacentTo}.\text{Cell} & (4) \\ \text{Cell} &\sqsubseteq \forall \text{isAdjacentTo}.\text{Cell} & (5) \\ \exists \text{isAdjacentTo}.\text{Cell} &\sqsubseteq \text{Cell} & (6) \\ \text{Cell} &\sqsubseteq \geq 0 \text{contains}.\text{Cell} & (7) \\ \text{Cell} &\sqsubseteq \forall \text{contains}.\text{Cell} & (8) \\ \exists \text{contains}.\text{Cell} &\sqsubseteq \text{Cell} & (9) \\ \\ \text{isFullyContainedIn} &\sqsubseteq \text{contains}^{-} & (10) \\ \text{contains} \circ \text{contains} &\sqsubseteq \text{contains} & (11) \\ \text{isFullyContainedIn} \circ \text{isFullyContainedIn} &\sqsubseteq \text{isFullyContainedIn} & (12) \\ \\ \text{spatialRelations} \circ \text{hasGeometry}^{-} &\sqsubseteq \text{spatiallyRelated} & (20) \\ \text{spatiallyRelated} \circ \text{hasAttribute} &\sqsubseteq \text{hasAttribute} & (21) \\ \text{contains} \circ \text{spatiallyRelated} &\sqsubseteq \text{spatiallyRelated} & (28) \end{aligned}$$

Learning Representations (for Types of Places)



The Issue With Portals



The Issue With Portals

data.houstontx.gov/group/flood-hazards

Debian.org Latest News Help

Log In Register

City of Houston Open Data Portal

Datasets About Search

Groups / Flood Hazards



Flood Hazards

Followers **0** Datasets **1**

Organizations

City of Houston Ope... **1**

Groups

Flood Hazards **1**

Tags

There are no Tags that match this search

Formats

Search datasets...

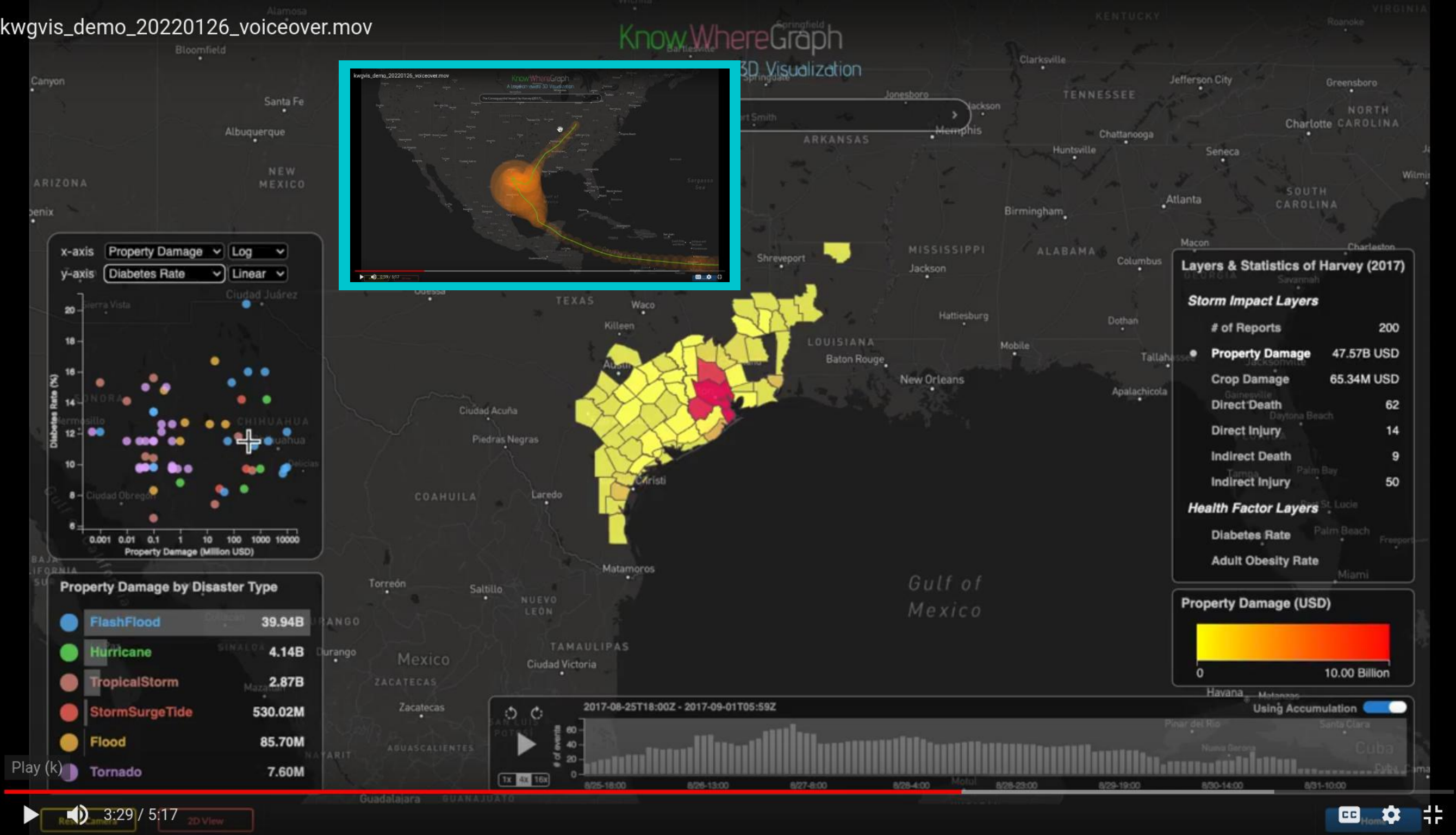
1 dataset found

Order by: Relevance

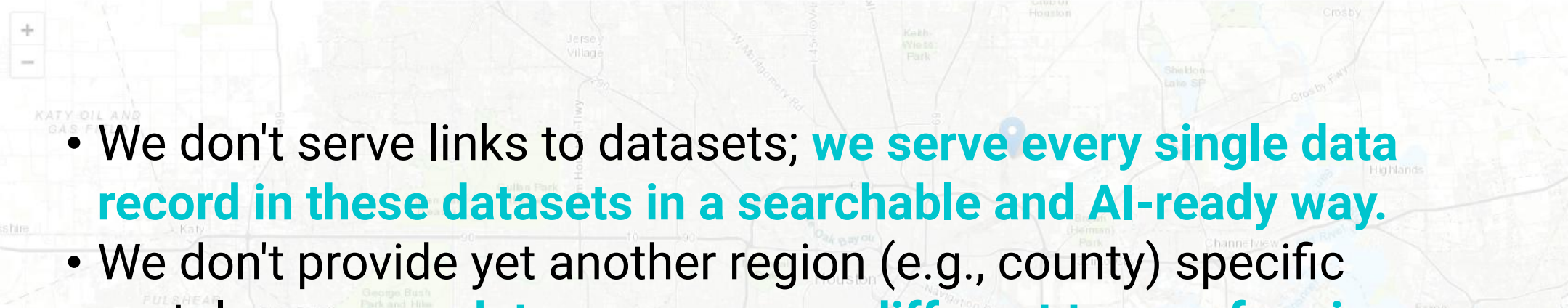
Other County Flood Zones

Flood Zones for outside of Harris County

ZIP



The Difference We Make

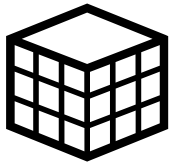
 "EPA Air Quality Site with Site ID 482010046 and Site Name Houston North Wayside"

- We don't serve links to datasets; **we serve every single data record in these datasets in a searchable and AI-ready way.**
- We don't provide yet another region (e.g., county) specific portal; **we serve data across many different types of region identifiers.**
- We don't make you download, overlay, and analyze different datasets on various topics relevant to your information needs; **we answer your questions directly and based on pre-integrated data.**
- We don't go offline with the next technology change, e.g., Flash, Silverlight; **we rely on a distributed, internationally standardized open-source technology paradigm.**

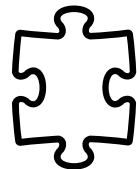
kgw-ont:siteName → "Houston North Wayside"

Key Research Questions

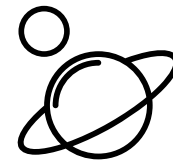
Spatiotemporally-explicit KG embeddings that are **invariant** under syntactic changes



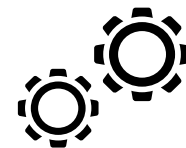
Learning geographic knowledge **graph summaries** for apportioned geo-enrichment



Quantifying and representing **regional differences** in the conceptualization of geographic space



How to **detect and mitigate bias** in geographic knowledge graphs



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



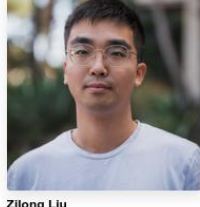
Anthony D'Onofrio
Developer
MSU



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 University



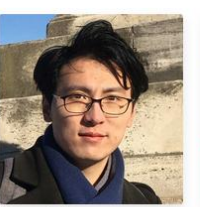
Meilin Shi
PhD Student
UCSB



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 Eelis
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



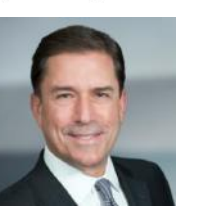
Thomas Thelen
Software Engineer
NCEAS, UCSB



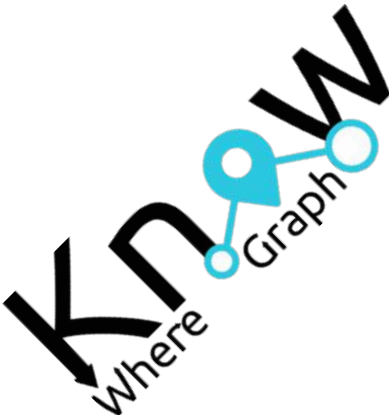
Joseph Zlewski
Ph.D. Student
DaSe Lab
Kansas State University



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
Humanities
Environmental Sciences
NCEAS, UCSB



Wenwen Li
Urban Planning and Geoinformatics
Arizona State University



Dean Rehberger
Humanities
Michigan State University

Former Team



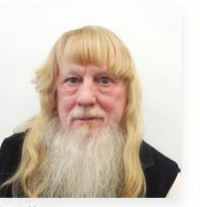
Sella Gonzalez
Head of Software Development
Matrix
Michigan State University



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