An Ontology and Geospatial Knowledge Graph for Reasoning About Cascading Failures

Torsten Hahmann
School of Computing and Information Science, University of Maine, Orono, ME, USA

David K. Kedrowski
School of Computing and Information Science, University of Maine, Orono, ME, USA

Abstract

During a natural disaster such as flooding, the failure of a single asset in the complex and interconnected web of critical urban infrastructure can trigger a cascade of failures within and across multiple systems with potentially life-threatening consequences. To help emergency management effectively and efficiently assess such failures, we design the *Utility Connection Ontology Design Pattern* to represent utility services and model connections within and across those services. The pattern is encoded as an OWL ontology and instantiated with utility data in a geospatial knowledge graph. We demonstrate how it facilitates reasoning to identify cascading service failures due to flooding for producing maps and other summaries for situational awareness.

2012 ACM Subject Classification Information systems → Geographic information systems; Computing methodologies → Ontology engineering; Computing methodologies → Spatial and physical reasoning

Keywords and phrases knowledge graph, ontology, OWL, spatial reasoning, cascading failures, urban infrastructure

Digital Object Identifier 10.4230/LIPIcs.COSIT.2024.21

Category Short Paper

Supplementary Material
Software: `https://github.com/UFOKN/Knowledge-Graph/tree/master/ontologies/v2.1` [10]
archived at `swh:1:dir:00a0f1fbcd022a6beed4d88b6430567ef1314872`
Software: `https://github.com/UFOKN/Knowledge-Graph/tree/master/ontologies/sparql`
archived at `swh:1:dir:8875f0fda487e4304638d4cd8e82a542453906c1`
Software: `https://github.com/UFOKN/Knowledge-Graph/tree/master/interactive-maps`
archived at `swh:1:dir:66f93d6c6f4866a156f7ba14cd71ed7753a58843`

Funding This material is based upon work supported by the National Science Foundation under Grant Numbers OIA-1937099 and ITE-2033607 to support the development of the Urban Flooding Open Knowledge Network (UF-OKN).

Acknowledgements We are grateful to the entire UF-OKN team for the collaboration on this project. The presentation of the work here benefited greatly from feedback and discussions of earlier versions of the presented research with members of the UF-OKN team, in particular Lilit Yeghiazarian, Doug Fils, Thomas Narock, Sadegh Riasi, Siddharth Saksena, and Adam Shepherd.
The authors also thank the three anonymous reviewers for their constructive comments.

1 Introduction

The complex web of critical infrastructure that sustains urban environments epitomizes interconnectedness. From the provisioning of essential utilities like electricity and drinking water to vital services such as healthcare and transportation, the smooth operation of each relies upon a network of physical assets interdependent on one another. For instance, the functionality of a city's drinking water system hinges on electricity to power pumps and filtration systems, while its public transportation network requires electricity and operational communication systems for traffic management and safety protocols. Similarly, healthcare facilities rely on a steady supply of electricity, water, and effective communication systems.

However, this interconnectedness exposes urban areas to significant vulnerability during natural disasters. A single failure in any infrastructure asset, triggered by hurricane-force winds or by accompanying flooding, can trigger a domino effect that results in *cascading failures* across multiple systems [3]. These cascading failures can extend far beyond the initial point of disruption and amplify the overall impact of the original failure. The repercussions of such failures reverberate throughout communities, affecting residents and businesses alike, even those seemingly untouched by the initial calamity.

1.1 Use Case Description

A flooding event may disable an electrical substation causing power outages that impact water supply assets (e.g., pumping stations), wastewater plants, or telecommunications infrastructure, which in turn may impact other infrastructure and essential services, such as medical or food services. Effective emergency management hinges on access to accurate information about the magnitude and ramifications of the flood. Emergency managers need to know which infrastructure assets, buildings and facilities have lost which services and for what direct or indirect reasons to effectively deploy limited resources where the greatest threats to human life and welfare are.

Likewise, utility providers will be interested in the hidden root causes of outages. A hospital may have lost water service but not electric service. However, the outage may be caused because water assets elsewhere, such as pumps, have lost their electric service. The root cause of the hospital's water issue is then an electrical one.

The information that emergency and utility managers need to tailor their responses encompasses details such as (1) the locations where service disruptions occur, (2) the types of services affected, (3) the direct and indirect impacts on individuals and communities, and (4) the *root causes* behind specific facility or service outages. Armed with this comprehensive understanding, emergency responders can strategically allocate resources, prioritize rescue and relief efforts, and expedite repairs, focusing on restoring the most critical services and addressing the areas most severely affected by the flooding.

1.2 Objectives and Contributions

The work presented here is part of a larger effort to construct the *Urban Flooding Open Knowledge Network* (UF-OKN) [12], which includes a knowledge graph and decision support tools to aid emergency management before, during and after natural disasters. Our specific aim here is to develop and test an ontological model that (1) captures the dependencies between different types of utility assets and services and their users and that (2) enables semantic reasoning to quickly answer key questions about cascading failures and their impacts in a flood event to help direct and prioritize emergency response and restoration efforts with regards to service outages.

A central contribution is the *Utility Connection Pattern*, an Ontology Design Pattern (cf. [2,8,9]) for representing different types of "utility features" and other kinds of geospatial features in a network-like dependency structure. We implement the pattern as an OWL ontology and test it in a geospatial knowledge graph specifically for utilities and services essential to urban environments, such as water, energy, communication and medical services using data from Hamilton County, Ohio. We develop SPARQL graph query patterns to demonstrate how to identify and summarize cascading flooding impacts and root causes

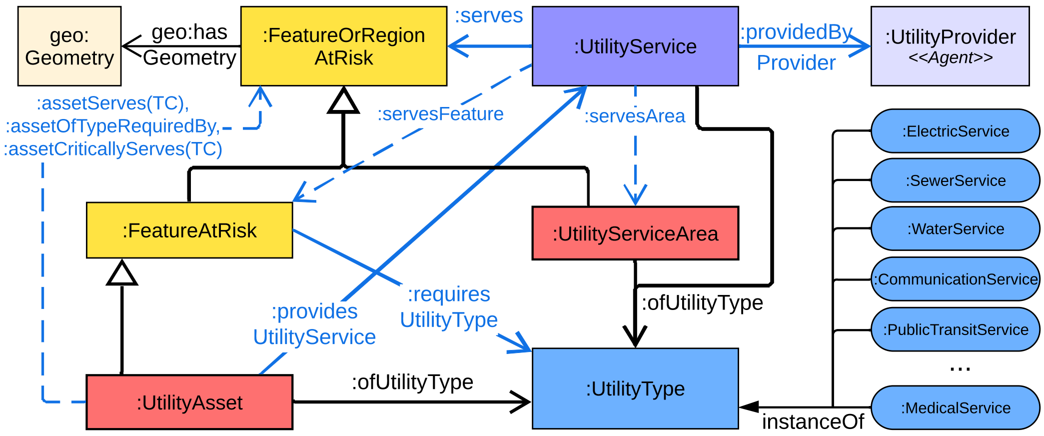


Figure 1 The core concepts of the Utility Connection Pattern and their relationships. Rectangles denote classes and rounded entities individuals. Triangular arrows depict subclass relations, for example *UtilityAsset* is a subclass of *FeatureAtRisk*, meaning that every *UtilityAsset* is also a *FeatureAtRisk*. All other arrows – for example *serves* between *UtilityService* and *FeatureOrRegionAtRisk* – denote object properties, which are used in OWL to encode binary relations. The dashed arrows are semantically inferred (defined) relations, e.g. the *servesFeature* object property is the specialization of the *serves* property where the object is a *FeatureAtRisk*.

from simple flood information – akin to a flood map – on demand using the ontology's semantics. The results and visualizations thereof can be recomputed in seconds when new flood information becomes available without changes to the ontology, graph, queries or algorithms.

2 Conceptualization: The Utility Connection Pattern

The proposed *Utility Connection Pattern*, shown in Figure 1, is built around three concepts: (1) geospatial features (*FeatureOrRegionAtRisk*) including infrastructure assets and their service areas as well as features that rely on them, (2) the types of utilities and other services of interest (*UtilityType*), and (3) *UtilityService* as the class that establishes connections between two or more features or regions. The Turtle encoding of the resulting OWL ontology is available from `https://github.com/UFOKN/Knowledge-Graph/tree/master/ontologies/v2.1`.

FeatureOrRegionAtRisk encompasses individual features (*FeatureAtRisk*), including physical utility assets (*UtilityAsset*) and features that rely on but do not provision services, and *UtilityServiceAreas* that represent entire spatial regions served by a particular asset.

UtilityAsset represents physical assets, such as entire buildings (e.g., hospitals, grocery stores), structures (e.g., electric substation, cell phone towers), or specific equipment (e.g., communication servers) used in the generation, storage and distribution of services essential to the functioning of urban systems. Each *UtilityAsset* is of one particular *UtilityType* as exemplified on the right of Fig. 1. The ontology allows modeling utility assets at much greater levels of detail – down to individual wires or pipes – where such information is available.

Utility Service represents the connection between an asset providing a particular kind of service, such as drinking water, wastewater, energy, telecommunication, medical, food or transportation services, and a set of users that use and depend on that service. Each instance links one or more provider assets, such as an electric substation or cellphone tower, to one or multiple, possibly thousands, of features that rely on it. The *providesUtilityService* and *serves* relations link the *UtilityService* instance to the providers and consumers, respectively. Note that users of particular services do not necessarily rely on those particular services, but rather on certain *types* of service. For example a residential building may require electric, water, telecommunication and sewer services, while an electric asset may require no or only telecommunication services. This is captured by the *requiresUtilityType* relationship between *FeatureAtRisk*s and *UtilityType*s.

Utility Providers represent the legal entities (e.g., companies) that operate services (linked to *UtilityService*s via *providedByProvider*) but are not further discussed here.

Utility Service Areas describe an entire spatial region served by a particular asset, thus aggregating a number of features and not having to link individual features to the *UtilityService* instance. The service area might be an entire city or a neighborhood or be defined by a distance from the asset, such as a cellphone tower, providing the service. The relation *serves* is specialized into *servesFeature* and *servesArea* as shown in Figure 1 to distinguish whether individual features or an area is served by a particular service.

Inferred Relationships. In addition to the relationships shown by solid lines in Fig. 1, additional relationships can be inferred to (1) calculate indirect dependence and, thus, cascading outages and to (2) distinguish between outages that are critical or not for the operation of assets and other features[1]. We add a direct, node-to-node (N2N) relationship *assetServes* as the composition of *providesUtilityService* with *serves*. *assetServesTC* is defined as its transitive version that captures also indirect service dependence, with *assetServes* treated as a specialized subproperty thereof. These two relations are automatically populated as instructed by the following axioms: first, *assetServes* is declared as a subproperty, i.e. a more specialized case, of the more general transitive version *assetServesTC*. Further, *assetServes* is declared as a subproperty of any chain of *providesUtilityService* and *serves*, meaning that if some asset A provides a utility service B and that service serves some *FeatureAtRisk* or *UtilityServiceArea* C, then the asset A *assetServes* C. Semantically, the property chain is also a subproperty of *assetServes*, but this direction is not needed and cannot be expressed in OWL for tractability reasons. These kinds of axioms permit leveraging OWL's semantic inferencing capabilities to simplify querying and reasoning about cascading impacts by end users.

```
ufokn_c:assetServes rdfs:subPropertyOf ufokn_c:assetServesTC ;
    owl:propertyChainAxiom (ufokn_c:providesUtilityService ufokn_c:serves) .
ufokn_c:assetServesTC rdf:type owl:TransitiveProperty .
```

However, the transitive propagation of outages can go too far when an asset serves an entire service area. Then, every feature inside that area would be presumed to be impacted as well. But not all features therein actually depend on all services: e.g., if there is a drinking water outage, it may not actually affect electric or communication assets because those do not rely on a drinking water supply. For that reason, we distinguish *assetCriticallyServes* as when an asset both serves (*assetServes*) and is required by another feature. This requirement is modeled using the relation *assetOfTypeRequiredBy*, which itself chains *assetOfUtilityType* – a specific case of *ofUtilityType* where the subject is a utility asset – and *utilityTypeRequiredBy*. We again introduce *assetCriticallyServesTC* as the transitive version of *assetCriticallyServes*.

[1] Following common practice, all relationships are modeled in the ontology as what are called *object properties* in the OWL language: https://www.w3.org/TR/owl-overview.

```
ufokn_c:assetCriticallyServes rdfs:subPropertyOf ufokn_c:assetServes ,
            ufokn_c:assetOfTypeRequiredBy ,
            ufokn_c:assetCriticallyServesTC .
ufokn_c:assetCriticallyServesTC rdf:type owl:TransitiveProperty .
```

Because OWL does not provide a construct for defining a relation to be precisely the intersection of two (or more) relations, which is a stronger constraint than the relation being declared as a subproperty of each of the two relations above, we instruct our graph database (GraphDB) to infer *assetCriticallyServes* relations using an inference rule:

```
 x  <ufokn_c:assetServes>  y
 x  <ufokn_c:assetOfTypeRequiredBy>  y
-------------------------------
 x  <ufokn_c:assetCriticallyServes>  y
```

3 Implementation: Graph Construction

We test our ontology and graph using data about Hamilton County, Ohio, which includes the City of Cincinnati and is bordered by the Ohio River to the south. Many of its tributaries flow through the county and it has experienced frequent flooding in the past. All geospatial data comes from open sources like OpenStreetMaps (OSM: `www.openstreetmap.org`), municipal government websites, and the Federal Emergency Management Agency (FEMA).

Features. We use key:value pairs to extract features for three utility types from OSM: Power stations are used as electric utility assets, water towers as water utility assets, and hospitals as medical utility assets. We use schools and similar facilities (e.g., *childcare*, *school*, and *college*) as our non-utility end users[2]. Overall, our test data contains ∼150 utility assets and ∼600 non-utility features. Electric and water service areas are modeled using Voronoi polygons, while medical service areas are modeled using municipal divisions (outside Cincinnati proper) and Cincinnati Statistical Neighborhood Approximations[3].

After creating an initial knowledge graph from the utility and feature data, it is necessary to run a pre-processing query that converts utility service areas to node-to-node (N2N) service connections because the knowledge graph system does not support property chains or rules that involve GeoSPARQL relations such as `sfWithin`.

Flood data. Our sample flood areas are derived from FEMA National Flood Hazard Layer (NFHL) data[4]. The FEMA flood plain was spatially buffered to ensure a sufficient number of assets were actually affected in our testing. The flood data is dynamic in that it includes minimum and maximum flood levels and associated times. If the flood level at an asset exceeds its critical depth, an outage of the associated utility service is presumed. Flood data can be added and removed from the knowledge graph as needed. When data is added, a pre-processing query determines which assets have failed due to the flood level at their location; i.e., are root cause failures.

[2] The choice is based on the fact that such facilities often serve as emergency shelters.
[3] `https://www.cincinnati-oh.gov/planning/maps-and-data/frequently-requested-maps/`
[4] `https://www.fema.gov/flood-maps/national-flood-hazard-layer`

4 Results: Graph Querying and Visualization

As previously noted, emergency and utility managers need to know *where* service disruptions occur, *which* services are affected, *how* individuals and communities are impacted, and *why* services are out – the *root causes*. To produce suitable maps and other summaries we use simple yet flexible graph queries that rely on the transitivity and the composition of the dependencies between assets and impacted features as outlined in Listing 1[5]: Each flooded *UtilityAsset* is queried for all features it critically serves and thus are also impacted. It can be refined to focus on outages among specific classes of assets (e.g., *ElectricUtilityAsset*).

Listing 1 The core of our SPARQL queries (namespace declarations are omitted for brevity) for identifying root cause failures (`?rootAsset`: a *UtilityAsset* with a *CriticalFloodObservation*) and any (transitively) impacted assets (`?utilityAsset`).

```
SELECT * WHERE {
 ?rootAsset  rdf:type ufokn_c:UtilityAsset ;
             ufokn_fl:hasCriticalFloodObservation ?critFloodObs ;
             ufokn_c:assetCriticallyServesTC ?utilityAsset . }
```

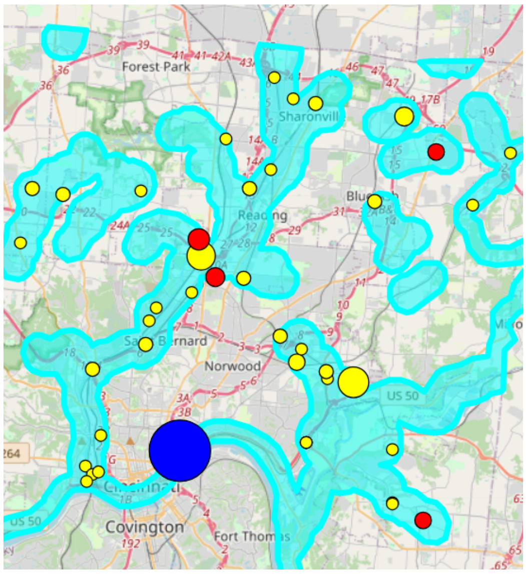
```
Summary of Non-Utility Feature Impacts
   218 non-utility features with no electric service
      35.9% of 607 non-utility features
      45.6% of 478 affected non-utility features
   252 non-utility features with no water service
      41.5% of 607 non-utility features
      52.7% of 478 affected non-utility features
   396 non-utility features with no medical service
      65.2% of 607 non-utility features
      82.8% of 478 affected non-utility features

Largest Cascades by Service Type
   WaterAsset-dngyt3h4cy37
      182 features affected (72.2% of water outages)
      1 water assets affected (including itself)
      28 medical assets affected

   ElectricAsset-dngyxwtxvvd4
      81 features affected (37.2% of electric outages)
      1 electric assets affected (including itself)
      1 water assets affected
      4 medical assets affected

   MedicalAsset-dngzjdkhb2p2
      43 features affected (10.9% of medical outages)
      1 medical assets affected (including itself)
```

(a) Summary Data.



(b) Root Cause Map.

Figure 2 Summarization of root causes: (a) top: total impact in the area; bottom: summary of the root cause assets of each service type with the largest number of affected features; (b) all flooded assets that are the root causes of all outages. Electric assets are yellow, water assets blue, and medical services red. The marker size is proportional to the number of features affected by cascades originating there.

More elaborate queries can follow all features that the impacted assets serve, aggregate and order them as desired, and retrieve additional geospatial and other attributes to produce summaries and maps like those shown in Figures 2 and 3.

For example, the outages can be aggregated by their root causes to produce the map in Figure 2b. Each of these root causes are assets that are flooded and thus at the beginning of any failure cascades. They explain why other assets and features have lost services. The

[5] This is a drastically simplified version; our full queries are available from: `https://github.com/UFOKN/Knowledge-Graph/tree/master/ontologies/sparql`

size of the dots visualizes the urgency of each root outage in terms of the total number of critically impacted features. The summary in Figure 2a provides even more detail. The top gives a sense of the scale of the overall impact on communities and people (e.g., what percentage of residences are impacted). The bottom summarizes for each utility type the cascade with the largest impact. Another query is used to display an individual cascade as shown in Figure 3 to provide a more detailed view of its spatial propagation. The requisite queries simultaneously retrieve additional data including identifiers, flood data, and locations.

Our figures are screenshots of interactive maps automatically produced from the knowledge graph using Python scripts and the SPARQLWrapper and Folium libraries. They can be dynamically updated by rerunning the graph queries. The maps contain multiple layers (e.g., one for each cascade) that users can turn on and off. They provide additional details via tooltips and popups that could incorporate contextual data, including demographic or socioeconomic data, to offer deeper insights. The maps are shareable as html pages (see `https://github.com/UFOKN/Knowledge-Graph/tree/master/interactive-maps`).

5 Related Work and Discussion

Our work leverages prior work on interdependencies within urban systems and on geospatial knowledge graphs. The various interdependencies between different kinds of urban systems have been analyzed and categorized comprehensively in, e.g., [3, 4, 13]. In our work, we do not try to incorporate and model all nuances and complexities of these interdependencies but rather choose a simpler model that focuses on what [4] describes as *necessity* (i.e., critical dependence) between *physical assets* to test and demonstrate reasoning over dependencies more generally as a means to identify the cascading impacts in an emergency. In the past, modeling and simulation [4, 13] were the main computational methods for analyzing and tracing infrastructure dependencies while using explicit semantics in the form of an ontology is new. Du et al.'s work [7] is most similar in that they also employ an ontology for analysis of the propagation of damages in urban environments. However, their approach differs in multiple ways: it relies on a probabilistic model to infer, weigh, and propagate factors; it focuses on environmental factors and models dependencies between them rather than between geospatial features; and thus, it does not spatially propagate cascading impacts across locations.

Geospatial knowledge graphs, such as [1,5,6,11,14], have gained popularity for sharing and linking geospatial information using knowledge graph technology. However, the focus has been on joint retrieval across datasets, which can be accomplished using shallow semantics, such as unique identifiers (IRIs) and uniform class and property names, following Jim Hendler's motto "a little semantics goes a long way". But these graphs are treated primarily as (graph) databases with semantic annotation in that the graphs often do not add much non-trivial inferred knowledge obtained via semantic reasoning. Thus, their utility depends very much on the users' ability to construct complex queries. In contrast, we have harnessed OWL capabilities more fully to facilitate complex and intricate reasoning (using the OWL-RL profile) that may remain hidden and still benefits users who may not be experts in ontologies without them having to tinker with it. Our ontology's advanced semantics enable the graph to perform cascading reasoning on dependency networks that transcends conventional spatial operations (e.g., intersections or distance searches) while also simplifying the queries. We have demonstrated that this requires remarkably little extra but powerful semantics typically underutilized in ontologies. Through a use case implementation, we have validated that this is not merely a theoretical option but that it tangibly enhances and simplifies decision

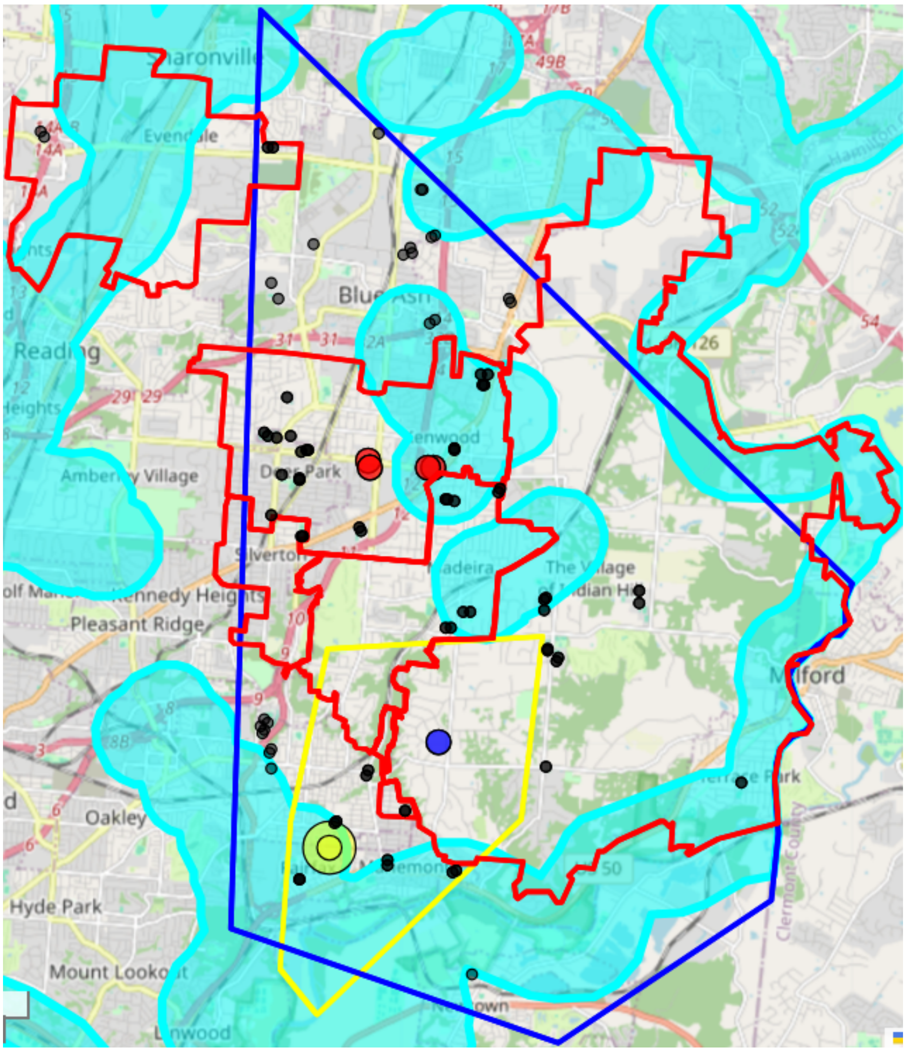


Figure 3 Visualization of a single cascade: A flooded electric asset is the root cause of its service area outage (shown as a yellow polygon). A water asset (blue dot) in that service area is impacted as well, which in turn affects its service area (blue polygon). That water outage impacts medical facilities and their service areas (shown as red dots and polygons) as well as end-use features, such as schools (black dots), that have lost at least one type but possibly multiple types of services. The indirectly impacted features outside the impacted electric service area do not experience a power outage but may have lost water or medical service.

support tools. Importantly, the idea of dependencies and cascades is not restricted to utilities but is much more broadly applicable to other types of services and supply chains. Also, we have only scratched the surface with respect to the kind of questions that can leverage the semantic connections.

The presented cascades are powerful but still simplify the ground truth. For that reason, issues such as temporal delays in an outage (e.g., the availability of a battery backup), redundancy (e.g., two electric substations provisioning power to an area or specific facility or the availability of backup options such as a generator), alternatives (e.g., a functioning urgent care facility instead of a hospital with limited services), and partial outages (e.g., reduced service levels such as a cell phone tower operating on backup power being limited in the number of connections it can support) have been left for future work.

References

1 Sören Auer, Jens Lehmann, and Sebastian Hellmann. Linkedgeodata: Adding a spatial dimension to the web of data. In *Intern. Semantic Web Conference (ISWC)*, pages 731–746. Springer, 2009.

2 Eva Blomqvist, Torsten Hahmann, Karl Hammar, Pascal Hitzler, Rinke Hoekstra, Raghava Mutharaju, et al. *Advances in pattern-based ontology engineering.* IOS Press, 2021.

3 Sergey V. Buldyrev, Roni Parshani, Gerald Paul, H. Eugene Stanley, and Shlomo Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 104:1025–1029, 2010.

4 Neil Carhart and Ges Rosenberg. A framework for characterising infrastructure interdependencies. international journal of complexity in applied science and technology. *Intern. J. of Complexity in Applied Science and Technology*, 1:35–60, 2016.

5 Arkadiusz Chadzynski, Nenad Krdzavac, Feroz Farazi, Mei Qi Lim, Shiying Li, Ayda Grisiute, Pieter Herthogs, Aurel von Richthofen, Stephen Cairns, and Markus Kraft. Semantic 3D city database — an enabler for a dynamic geospatial knowledge graph. *Energy and AI*, 6, 2021.

6 Alishiba Dsouza, Nicolas Tempelmeier, Ran Yu, Simon Gottschalk, and Elena Demidova. WorldKG: A world-scale geographic knowledge graph. In *Intern. Conf. on Inf. & Knowledge Mgmt (CIKM)*, pages 4475–4484. ACM, 2021.

7 Heshan Du, Lijun Wei, Vania Dimitrova, Derek Magee, Barry Clarke, Richard Collins, David Entwisle, Mehran Eskandari, Giulio Curioni, Ross Stirling, Helen Reeves, and Anthony G. Cohn. City infrastructure ontologies. *Comput. Environ. Urban Syst.*, 104, 2023.

8 Aldo Gangemi. Ontology design patterns for semantic web content. In *4th Int. Semantic Web Conf.*, 2005.

9 Torsten Hahmann. Ontology repositories: A treasure trove for content ontology design patterns. In *Proc. of the Workshop on Modular Ontologies (WoMO 2014)*, CEUR Workshop Proceedings, Vol. 1248, 2014.

10 Torsten Hahmann and David K. Kedrowski. OWL Ontologies for the Urban Flooding Open Knowledge Network (UF-OKN). Software, version 2.1., swhId: `swh:1:dir:00a0f1fbcd022a6beed4d88b6430567ef1314872` (visited on 2024-08-28). URL: `https://github.com/UFOKN/Knowledge-Graph/tree/master/ontologies/v2.1`.

11 Krzysztof Janowicz, Pascal Hitzler, Wenwen Li, Dean Rehberger, Mark Schildhauer, Rui Zhu, et al. Know, Know Where, KnowWhereGraph: A densely connected, cross-domain knowledge graph and geo-enrichment service stack for applications in environmental intelligence. *AI Magazine*, 43(1):30–39, 2022.

12 J Michael Johnson, Tom Narock, Justin Singh-Mohudpur, Doug Fils, Keith Clarke, Siddharth Saksena, Adam Shepherd, Sankar Arumugam, and Lilit Yeghiazarian. Knowledge graphs to support real-time flood impact evaluation. *AI Magazine*, 43(1):40–45, 2022.

13 Min Ouyang. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability Engineering & System Safety*, 121:43–60, 2014.

14 Basel Shbita, Craig A. Knoblock, Weiwei Duan, Yao-Yi Chiang, Johannes H. Uhl, and Stefan Leyk. Building spatio-temporal knowledge graphs from vectorized topographic historical maps. *Semantic Web*, 2023.