A Pattern for Modeling Causal Relations Between Events*

Cogan Shimizu¹⊠, Rui Zhu³, Gengchen Mai³, Mark Schildhauer², Krzysztof Janowicz³, and Pascal Hitzler¹

 $^{\rm 1}\,$ Data Semantics Laboratory, Kansas State University, USA

Abstract. Space and time are useful nexuses for integrating data. For instance, events affect the places in which they occur and the people that participate in them. By capturing the effects that they may have on a place, coupled with authoritative sources on possible causality between types of events, we can model causal relations between events. In this paper we present an ontology design pattern for modeling the causal relations between events, discuss the primary conceptual components, how they may be instantiated, and present overarching examples related to the domain of disaster risk management.

1 Introduction

Space and time are frequently useful nexuses for integrating data. For example, using common spatial or topological calculi (e.g., such as RCC5, RCC8, or DE-9IM) one can describe how spatial entities (e.g., events or records of events) interrelate. However, there are fewer resources for modeling how events may (or did) interact causally. That is, via time and space, in such a way that they affect or cause each other to occur. We note that this is distinct from different conceptualizations of an event, such as the ontology design pattern for a recurring event series [1]. We emphasize the importance of causation. Certain notions, such as seasonality, is not the same as causation. To this end, we have developed an ontology design pattern that provides a framework for modeling spatiotemporal data, in particular events, and capturing the nature of relationships between them, emphasizing causality, as declared by some a priori notion of causation, such as the IRDR's (Integrated Research on Disaster Risk) taxonomy [7]. More specifically, this pattern addresses a scenario that is concerned with three key questions.

- How are events connected to each other?
- Who asserts that they are connected?
- How do these events affect the places at which they occur?

² Center for Spatial Studies, University of California, Santa Barbara, USA

³ National Center for Ecological Analysis & Synthesis, Santa Barbara, USA

 $^{^\}star$ Copyright © 2021 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Answering these questions is valuable for understanding both the nature of events, but also for understanding places. More concretely, it may allow us to examine the causes (and subsequently consequences) of sociodemographic or geographic triggers in an area. This is valuable in its own right, but also in particular valuable to the domain of disaster risk management, whereby understanding the causal relations between events can result in saved lives. We briefly consider, at a top level, two scenarios within the domain of disaster risk management.

- Identifying possible locations of future subsidence
- Predicting resurgence of endemic disease

We briefly expand on these use-case scenarios in the next section. In Section 3 we discuss the details of this ontology design pattern and in Section 4 we conclude.

2 Use-case Scenarios

The following use-case scenarios are motivated by the KnowWhereGraph project ⁴ and its partners within the domain of disaster risk management. KnowWhereGraph aims at providing a densely interlinked knowledge graph for environmental intelligence applications for enriching the data of decision-makers and data scientists with pre-integrated data custom-tailored to their spatial area of interest, thereby reducing the time needed to address an emerging crisis or to gain situational awareness.

We have identified two such use-cases that demonstrate the usefulness of our pattern. In the following, we have included a selection of competency questions that were used to guide the development of this pattern.

2.1 Wildfire Scenario

In this scenario, we are interested in the consequences of a wildfire that lead to the pre-conditions of other natural hazards. In particular, we may start with the trigger event of the wildfire. According to the IRDR Programme's "Peril Classification and Hazard Taxonomy" this may be a lightning strike or a human action [7]. The resultant wildfire can drastically and problematically induce soil erosion (e.g., by removing the flora and root systems that hold soil together) [8]. Subsequently, a storm with heavy rainfall can cause a landslide, which in turn can degrade soil, damage infrastructure, and so on. Modeling this chain of events in such a way that goes beyond temporal ordering can help decision makers detect locations where possible landslides or other forms of subsidence are likely to occur based on which events have occurred in certain places.

2.2 Hurricane Scenario

Hurricanes are seasonally recurring events that often lead to disasters with strong primary and secondary humanitarian relief implications—from emergency medical considerations due to injuries and exposure from the storm event itself, to

⁴ See https://knowwheregraph.org/.

secondary and tertiary implications arising from disruption of food and water supply systems, and elevation in the incidence of specific diseases, such as cholera and dysentery. Regional variation in these factors also exist due to differences in robustness of their infrastructures, endemism of certain diseases, and so on. Representing associations among these factors based on past events can be used to forecast region-specific disaster relief needs, as well as better understand the efficacy of certain disaster relief actions.

2.3 Selected Competency Questions

We have included competency questions (CQ) that pertain to the use case described above and CQs that may be used with other event and more general use cases.

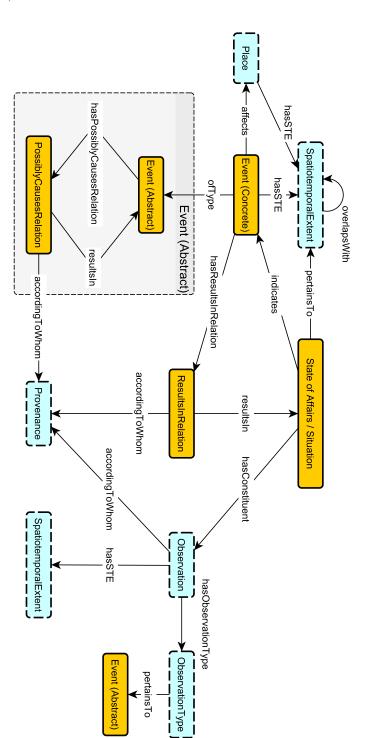
- CQ 1. Given a fire x, which regions will be effected by smoke exposure, given current wind direction projections?
- CQ 2. How were the 2019 Southern California fires affecting the tourism industry?
- CQ 3. Was the Cholera outbreak in Mozambique contributing to the food shortage in year x?
- CQ 4. What are the causalities of the wildfire?
- CQ 5. What factors can you find in a specific region that would help explain e.g. the stroke belt. Which contaminants of farms may be related from the health literature to strokes?
- CQ 6. What farmlands or vegetation covers have been mostly affected in the fire?
- CQ 7. What were the reasons for the landslide east of Santa Barbara in April 2017?
- CQ 8. What were wildfires affecting the Ventura area in the 2010s?
- CQ 9. Where are areas of increased heat exceedence and pollution, where migration is not driven by urbanization?
- CQ 10. Where are the places where heat is rising and (human) migration is occurring where there are no indicators of urbanization?
- CQ 11. Which farm has experienced disease?
- CQ 12. Which region affected by a transmissible disease is affected by a hurricane?
- CQ 13. Which region affected by the current hurricane just suffers from another natural disaster?
- CQ 14. Which regions affected by wildfires are expected to experience heavy rain?

3 The Causal Event Pattern

Overview

The Causal Event pattern has four main components: Event(Abstract), Event (Concrete), Provenance, and Place. That is not to say that the others are unimportant, but that these are the key conceptual components. We note that a

4



box indicates a module, which indicates a conceptual grouping of nodes in the schema diagram. external dependency (i.e., they are left unmodeled in the pattern); and edges with filled arrows are object properties. The large Gray Fig. 1. The schema diagram for the CausalEvent Pattern. Yellow boxes are classes; blue dashed boxes are also classes, but acknowledge

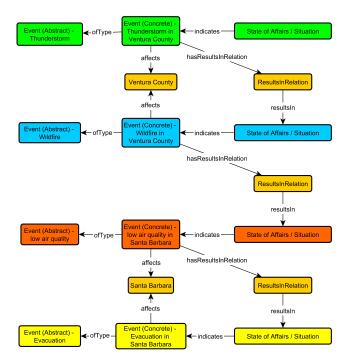


Fig. 2. This diagram shows naively constructed instance data that populates a portion of the pattern, emphasizing how it might connect to multiple instances of itself.

notion of time is also an incredibly important but that we do not commit to any specific conceptualization thereof. Furthermore, the distinction between the two different types of events (read: algorithm vs. execution thereof) is central. This allows us to make top-level statements about the type of an event, such as the environmental characteristics necessary for it to occur and the consequences that follow. Provenance is also important as it drives the trust of the overall knowledge graph – whose reputation is at stake for making these claims? Finally, the notion of Place is important for grounding these events in space (and time) in a human meaningful way.

The rest of the concepts play a supporting role: StateOfAffairs, Observation, and ObservationTypes allow us to record the empirical data that indicates the presence of an event, or model the conditions in a location according as they set-up, or have been impacted by, events. The PossiblyCausesRelation and ResultsInRelations are reifications of simpler properties that allow us to more appropriately, and directly, capture provenance.

A schema diagram for this pattern is shown in Figure 1 and an example of it in a naive population is shown in Figure 2. For each concept in the pattern, we provide the formalization as well as further discussion regarding its role in the pattern and how it may be used and instantiated.⁵ Formalization was conducted according to the "Systematic Axiomatization" Step in the Modular Ontology Modeling paradigm, utilizing the axiom patterns (and labels) as found in [3].

Each axiom appears only once in this section and appears in the section corresponding to the "source" of the arrow representing the relation. Throughout, we use initialisms for formatting purposes (e.g., STE in place of SpatiotemporalExtent, RIR in place of ResultsInRelation, and PCR in place of PossiblyCauses-Relation).

The OWL file can be found online ⁶ and is annotated with extended OPLa [6,5]. This pattern has already been integrated into the internal MODL for Co-ModIDE [11]. Scoped Domain and Scoped Range axioms restrict the domains and ranges based on fillers; this is a strict axiom that intends to limit the overall impact of the axiom on the rest of the ontology.

Finally, we note that the names for the classes and properties can be contentious. However, this pattern is meant to be used as a template (i.e., turned into a module through template-based instantiation [4]); in template-based instantiation, the structure of the pattern is re-used, where the names provide guidance for the initial conceptualization in an ontology engineering workflow. Thus, the names generally change in this process and are no longer a concern.

SpatiotemporalExtent

$$STE \sqsubseteq \forall overlapsWith.STE$$
 (Scoped Range) (1)

⁵ By this we mean *template-based instantiation* which is the method by which a pattern is adapted to a use-case [4], particularly in the MOMo setting [12].

 $^{^6}$ See https://github.com/KnowWhereGraph/causal-events-pattern.

$$\exists overlapsWith.STE \sqsubseteq STE$$
 (Scoped Domain) (2)

SpatiotemporalExtent is left unmodeled in this pattern and is instead left as a "hook" for potentially more complex modeling depending on specific needs. In the past we have *instantiated* this concept as a pair of data properties connecting latitude and longitude (and ignoring the temporal component). Alternatively, we have utilized concepts from the commonly accepted standards from the Open Geospatial Consortium and W3C GeoSPARQL [9] and owl:Time [14], respectively.

Place

$$\begin{array}{ll} \mathsf{Place} \sqsubseteq \forall \mathsf{hasSTE.STE} & (\mathsf{Scoped\ Range}) & (3) \\ \exists \mathsf{hasSTE.Place} \sqsubseteq \mathsf{STE} & (\mathsf{Scoped\ Domain}) & (4) \\ \end{array}$$

Place refers to a conceptual location that goes beyond mere coordinates. These might be very well defined, such as the boundaries of a voting district, or vague regions, such as "Southern California." In the case of the latter, the ontology engineer may opt to remove the hasSTE property and perhaps utilize a locatedIn property that points back at Place. One way of instantiating this node would be through a gazetteer.

Event(Concrete)

Here, we use Event(C) in place of Event(Concrete).

$Event(C) \sqsubseteq \forall hasSTE.STE$	(Scoped Range)	(5)
$\exists hasSTE.Event(C) \sqsubseteq STE$	(Scoped Domain)	(6)
$Event(C) \sqsubseteq \exists hasSTE.Event(C)$	(Existential)	(7)
$Event(C) \sqsubseteq \leq 1hasSTE.Event(C)$	(Functionality)	(8)
$Event(C) \sqsubseteq \forall affects.Place$	(Scoped Range)	(9)
\exists affects.Event(C) \sqsubseteq Place	(Scoped Domain)	(10)
$Event(C) \sqsubseteq \forall ofType.Event(Abstract)$	(Scoped Range)	(11)
$\exists ofType.Event(C) \sqsubseteq Event(Abstract)$	(Scoped Domain)	(12)
$Event(C) \sqsubseteq \forall hasRIR.RIR$	(Scoped Range)	(13)
$\exists hasRIR.Event(C) \sqsubseteq RIR$	(Scoped Domain)	(14)

Event(Concrete) is an event that occurs in space and time. This concept is complementary, and disjoint with, the Event(Abstract) class. Essentially, the difference is that the Event(Abstract) is the prototypical or archetypal notion of a type of event. For example, a Hurricane (the scientific topic) can cause flooding after landfall. This is not about any specific hurricane, but hurricanes in general. Hurricane Katrina, for example, did cause flooding and we can leverage this

connection between abstract and concrete for a high fidelity model. Nearly any ontology for events can be used here.

We note, in Axioms 7 and 8, that an event must occur in space and time; that is, it has a spatiotemporal extent.

Event(Abstract)

Event(Abstract) is the abstract notion of an event. For instance, an expert may study hurricanes or wildfires.

StateOfAffairs

Here, we use Obs in place of Observation.

$StateOfAffairs \sqsubseteq \forall pertainsTo.STE$	(Scoped Range)	(19)
$\exists pertainsTo.StateOfAffairs \sqsubseteq STE$	(Scoped Domain)	(20)
$StateOfAffairs \sqsubseteq \forall indicates.Event(C)$	(Scoped Range)	(21)
$\exists indicates.StateOfAffairs \sqsubseteq Event(C)$	(Scoped Domain)	(22)
${\sf StateOfAffairs} \sqsubseteq \forall {\sf hasConstituent.Obs}$	(Scoped Range)	(23)
\exists hasConstituent.StateOfAffairs \sqsubseteq Obs	(Scoped Domain)	(24)

State Of Affairs is a collection of conceptually linked observations. A straightforward choice for instantiating this node may be the Observation Collection from the extended SOSA/SSN ontology [2]. In SOSA/SSN members of such collections all share at least one attribute, such as the time they occur, or their feature of interest. In this case, the constituent observations would share a temporal entity that is strictly after an Event.

ResultsInRelation

Here, we use according TW in place of according ToWhom.

$RIR \sqsubseteq \exists hasRIR^RIR$	(Existential)	(25)
$RIR \sqsubseteq \forall resultsIn.StateOfAffairs$	(Scoped Range)	(26)
$\exists resultsIn.RIR \sqsubseteq StateOfAffairs$	(Scoped Domain)	(27)
$RIR \sqsubseteq \exists resultsIn.RIR$	(Existential)	(28)

$$\mathsf{RIR} \sqsubseteq \forall \mathsf{accordingTW.Provenance} \qquad \mathsf{(Scoped\ Range)} \qquad (29)$$

$$\exists \mathsf{accordingTW.RIR} \sqsubseteq \mathsf{Provenance} \qquad \mathsf{(Scoped\ Domain)} \qquad (30)$$

ResultsInRelation is a reification of the resultsIn property. This is used to attach provenance. We note in Axiom 26 that the inverse filler of hasResultsInRelation must exist and must be a ResultsInRelation.

Provenance

Provenance is left unmodeled and is instead left as a "hook" for potentially more complex modeling depending on specific needs. Generally, we suggest to utilize the EntityWithProvenance pattern included in MODL [13], which itself is based on the PROV Ontology [10].

Observation

Here, we use according TW in place of according ToWhom, Obs in place of Observation, and has OT in place of has Observation Type.

$Obs \sqsubseteq \forall according TW. Provenance$	(Scoped Range)	(31)
$\exists according TW. Obs \sqsubseteq Provenance$	(Scoped Domain)	(32)
$Obs \sqsubseteq \forall hasSTE.STE$	(Scoped Range)	(33)
$\exists hasSTE.Obs \sqsubseteq STE$	(Scoped Domain)	(34)
$Obs \sqsubseteq \forall hasOT. ObservationType$	(Scoped Range)	(35)
\exists hasOT.Obs \sqsubseteq ObservationType	(Scoped Domain)	(36)

Observation is some records of fact about a place in space and time. In the same manner as StateOfAffairs, the straightforward instantiation is also from SOSA/SSN with the eponymous sosa:Observation.

ObservationType

Observation Type determines the aspect of reality that the Observation is recording. This is an explicit typing mechanism, but can also be instantiated instead as an Observable Property from SOSA/SSN.

Possibly Causes Relation

$$PCR \sqsubseteq \forall \mathsf{resultsIn.Event}(\mathsf{Abstract}) \qquad \mathsf{(Scoped Range)} \qquad (39) \\ \exists \mathsf{resultsIn.PCR} \sqsubseteq \mathsf{Event}(\mathsf{Abstract}) \qquad \mathsf{(Scoped Domain)} \qquad (40)$$

PossiblyCausesRelation is a reification of the possiblyCauses property, so that provenance may be attached.

4 Conclusion

Modeling the causal relationships between events is an important step in understanding places. By understanding what has happened in a location and, to some extent, why those events occurred, one can gain deep insight into the nature of a particular place, and possibly, what events can be expected to occur. As such, we have developed this pattern as a first step in understanding the nature of causation between complex events. To do this, we distinguish between the abstract and concrete notions of events. For example, consider the difference between a popular pizza recipe and the actual pizza that is produced. A recipe, when reasonably followed, produces some (hopefully) tight variation of the expected output. We consider the notion of a "hurricane" to be similarly useful. Thus by understanding the generalities of the abstract hurricane, we may reason more correctly about instances of a hurricane, such as "Hurricane Katrina." The pattern currently assumes that the ontology engineer has a priori knowledge of causal relations, such as using taxonomies from IRDR Programme or the United Nations. However, one could consider that the PossiblyCausesRelation to be generated by some KG mining algorithm detecting spatiotemporal overlap and indicating possible causation. Additionally, we have demonstrated two use-case scenarios, particularly within the domain of disaster risk management, where modeling such notions will have a high impact. Additionally, we provided a basic graphical example that easily maps to triple output.

Future work entails expanding the Event(Abstract) module for more sophisticated modeling, as well as including shortcuts to simplify the population of the pattern.

Acknowledgements. The authors acknowledge support by the National Science Foundation under Grant 2033521 A1: KnowWhereGraph: Enriching and Linking Cross-Domain Knowledge Graphs using Spatially-Explicit AI Technologies. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

 Carriero, V.A., Gangemi, A., Nuzzolese, A.G., Presutti, V.: An ontology design pattern for representing recurrent situations. CoRR abs/2101.00286 (2021), https://arxiv.org/abs/2101.00286

- Compton, M., Barnaghi, P.M., Bermudez, L., Garcia-Castro, R., Corcho, Ó., Cox, S.J.D., Graybeal, J., Hauswirth, M., Henson, C.A., Herzog, A., Huang, V.A., Janowicz, K., Kelsey, W.D., Phuoc, D.L., Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K.R., Passant, A., Sheth, A.P., Taylor, K.: The SSN ontology of the W3C semantic sensor network incubator group. J. Web Semant. 17, 25–32 (2012). https://doi.org/10.1016/j.websem.2012.05.003, https://doi.org/10.1016/j.websem.2012.05.003
- Eberhart, A., Shimizu, C., Chowdhury, S., Sarker, M.K., Hitzler, P.: Expressibility of OWL axioms with patterns. In: Verborgh, R., Hose, K., Paulheim, H., Champin, P., Maleshkova, M., Corcho, Ó., Ristoski, P., Alam, M. (eds.) The Semantic Web 18th International Conference, ESWC 2021, Virtual Event, June 6-10, 2021, Proceedings. Lecture Notes in Computer Science, vol. 12731, pp. 230-245. Springer (2021). https://doi.org/10.1007/978-3-030-77385-4_14, https://doi.org/10.1007/978-3-030-77385-4_14
- 4. Hammar, K., Presutti, V.: Template-based content ODP instantiation. In: Hammar, K., Hitzler, P., Krisnadhi, A., Lawrynowicz, A., Nuzzolese, A.G., Solanki, M. (eds.) Advances in Ontology Design and Patterns [revised and extended versions of the papers presented at the 7th edition of the Workshop on Ontology and Semantic Web Patterns, WOP@ISWC 2016, Kobe, Japan, 18th October 2016]. Studies on the Semantic Web, vol. 32, pp. 1–13. IOS Press (2016). https://doi.org/10.3233/978-1-61499-826-6-1, https://doi.org/10.3233/978-1-61499-826-6-1
- 5. Hirt, Q., Shimizu, C., Hitzler, P.: Extensions to the ontology design pattern representation language. In: Janowicz, K., Krisnadhi, A.A., Poveda-Villalón, M., Hammar, K., Shimizu, C. (eds.) Proceedings of the 10th Workshop on Ontology Design and Patterns (WOP 2019) co-located with 18th International Semantic Web Conference (ISWC 2019), Auckland, New Zealand, October 27, 2019. CEUR Workshop Proceedings, vol. 2459, pp. 76–75. CEUR-WS.org (2019), http://ceur-ws.org/Vol-2459/short2.pdf
- 6. Hitzler, P., Gangemi, A., Janowicz, K., Krisnadhi, A.A., Presutti, V.: Towards a simple but useful ontology design pattern representation language. In: Blomqvist, E., Corcho, Ó., Horridge, M., Carral, D., Hoekstra, R. (eds.) Proceedings of the 8th Workshop on Ontology Design and Patterns (WOP 2017) co-located with the 16th International Semantic Web Conference (ISWC 2017), Vienna, Austria, October 21, 2017. CEUR Workshop Proceedings, vol. 2043. CEUR-WS.org (2017), http://ceur-ws.org/Vol-2043/paper-09.pdf
- Irdr peril classification and hazard glossary, https://www.irdrinternational.org/knowledge_pool/publications/173
- Neary, D.G., Leonard, J.M.: Effects of fire on grassland soils and water: A review. In: Kindomihou, V.M. (ed.) Grasses and Grassland Aspects, chap. 3. IntechOpen, Rijeka (2020). https://doi.org/10.5772/intechopen.90747, https://doi.org/10.5772/intechopen.90747
- Geosparql a geographic query language for rdf data, https://www.ogc.org/ standards/geosparql
- 10. Sahoo, S., McGuinness, D., Lebo, T.: PROV-o: The PROV ontology. W3C recommendation, W3C (Apr 2013), http://www.w3.org/TR/2013/REC-prov-o-20130430/
- 11. Shimizu, C., Hammar, K., Hitzler, P.: Modular graphical ontology engineering evaluated. In: Harth, A., Kirrane, S., Ngomo, A.N., Paulheim, H., Rula, A., Gentile, A.L., Haase, P., Cochez, M. (eds.) The Semantic Web 17th International Conference, ESWC 2020, Heraklion, Crete, Greece, May 31-June 4, 2020, Pro-

- ceedings. Lecture Notes in Computer Science, vol. 12123, pp. 20–35. Springer (2020). https://doi.org/10.1007/978-3-030-49461-2_2, https://doi.org/10.1007/978-3-030-49461-2_2
- 12. Shimizu, C., Hammar, K., Hitzler, P.: Modular ontology modeling. Tech. rep., Kansas State University (2021), https://daselab.cs.ksu.edu/publications/modular-ontology-modeling, under Review.
- Shimizu, C., Hirt, Q., Hitzler, P.: MODL: A modular ontology design library. In: Janowicz, K., Krisnadhi, A.A., Poveda-Villalón, M., Hammar, K., Shimizu, C. (eds.) Proceedings of the 10th Workshop on Ontology Design and Patterns (WOP 2019) co-located with 18th International Semantic Web Conference (ISWC 2019), Auckland, New Zealand, October 27, 2019. CEUR Workshop Proceedings, vol. 2459, pp. 47–58. CEUR-WS.org (2019), http://ceur-ws.org/Vol-2459/paper4.pdf
- 14. Time ontology in owl, https://www.w3.org/TR/owl-time/