An Ontology Design Pattern for Modeling Pollution*

Saad Ahmad[0000−0002−3040−508X], Sudhir Attri[0000−0002−3667−9110], and Raghava Mutharaju[0000−0003−2421−3935]

Knowledgeable Computing and Reasoning (KRaCR) Lab,
IIIT-Delhi, New Delhi, India.
{saad18409,sudhir18267,raghava.mutharaju}@iiitd.ac.in

Abstract. Pollution has been identified as a significant risk to global ecosystems and living beings. However, the information about pollution is fragmented and there is no meaningful organization of information despite several ongoing efforts to monitor it. Organizing the information about the pollution, such as the pollutants, their observations at different spatio-temporal points and the carriers in the form of an ontology will be very helpful to the applications that work with the different heterogeneous pollution data sources. We propose an ontology design pattern (ODP) for pollution that captures its general characteristics and can be used as a building block for modeling specific categories of pollution such as air, water and soil. The *Pollution* ODP is available on the ODP portal at `http://ontologydesignpatterns.org/wiki/Submissions:Pollution`. It is also available for public comments at `https://github.com/kracr/aq-structured-platform/blob/main/Ontology/PollutionODP/PollutionODP.owl` with an Apache License 2.0.

Keywords: Pollution · Ontology Design Pattern · Pollutants · Air Pollution.

1 Introduction

The introduction of substances into the environment that are harmful to the living organisms is defined as pollution [11]. These harmful substances can be solids, liquids, or gases produced in higher than usual concentrations. These substances are referred to as pollutants. Pollution could affect many natural resources, such as the air, water sources and soil. The need to study these pollution types is of utmost importance. Air pollution is regarded as a global health emergency and the effect of bad air quality is deadly. It leads to asthma, other respiratory illnesses and heart disease. Air pollution is responsible for more deaths than many other risk factors, including malnutrition, alcohol use and physical inactivity [6]. Similarly, there is a pressing need to study water and soil pollution [1]. Different

heterogeneous data sources are made use of to build pollution monitoring applications [2,3]. An ontology design pattern(ODP) [8] that captures the abstract details of the pollution will be beneficial to these applications.

2 Related Work

Attempts to model pollution have focused on air pollution, probably because of the readily available data of pollutant concentrations. Claudine et al. [12] built an air quality ontology and used that along with the 3D models of a city. The ontology proposed in [13] links air pollutants with meteorological factors. EnvO[1] is a very broad ontology that describes the environment by focusing on biomes. Dalia et al. [4] built an air pollution ontology by focusing on species, sensors, pollutants and meteorological factors. They have designed one ontology for each of these factors. We studied these ontologies along with several data sources such as pollutant concentration[2], weather data[3] and wind trajectory[4] to design the *Pollution* ODP. We also capture the spatio-temporal aspect of pollution in the ODP, which has not been considered in some of the pollution ontologies. The *Pollution* ODP has been annotated to indicate the ODPs that were used and submitted for a review on the ODP community portal[5].

3 Pollution ODP Description

The pollution ODP makes use of the trajectory ODP [9], observation ODP [5] and the stub meta-pattern [10]. We use cpannotationschema[6] to describe the intent, scenarios, consequences and components of the ODP. The schema of the pollution ODP is given in Figure 1. We describe the concepts and the properties of the pollution ODP in the subsequent sections.

3.1 Concepts

- **Pollution**. It is the core concept in the ODP to represent pollution and is linked to the contributors of pollution. Some of the instances of `Pollution` could be air pollution, water pollution, soil pollution, space pollution, and sound pollution.
- **Observation**. The `Observation` concept is modeled from the Observation ODP. It represents a spatio-temporal observation. We use it to capture the concentration and prescribed standards for a particular pollutant. With respect to the Observation ODP, `Pollutant` and `PrescribedStandardForPollutant` is the situation and `TimeEntity`, `PlaceEntity` are the parameters of the observation.

[1] `http://environmentontology.org/`
[2] `https://cpcb.nic.in/`
[3] `https://weatherstack.com/`
[4] `https://www.ready.noaa.gov/HYSPLIT_disp.php`
[5] `http://ontologydesignpatterns.org/wiki/Main_Page`
[6] http://www.ontologydesignpatterns.org/schemas/cpannotationschema.owl

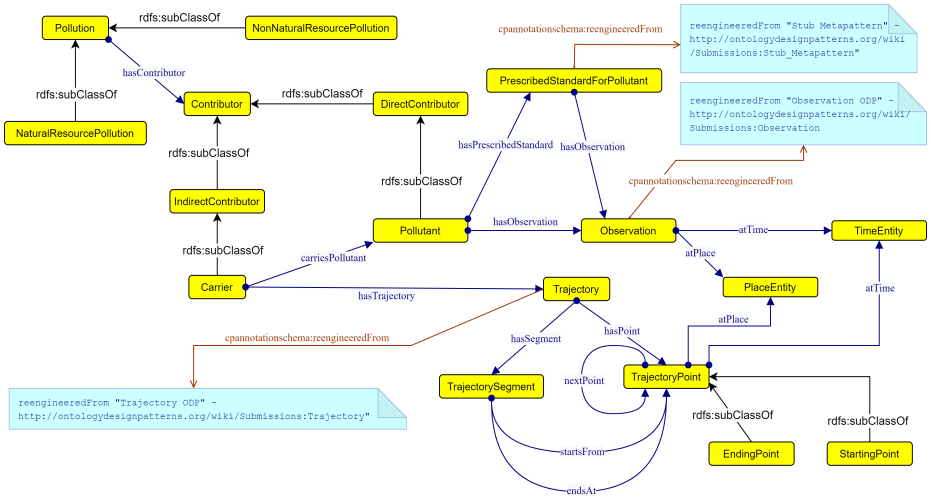

Fig. 1. Core concepts and properties of the Pollution ODP

- **DirectContributor**. The `DirectContributor` concept represents the contributors that directly affect the pollution. `Pollutant` is a subclass of `DirectContributor`. Some examples of pollutants include biological pollutants (viruses, pathogens, bacteria, etc.), chemical pollutants (toxic metal, radionuclides, organophosphorus compounds, gases, etc.) and physical pollutants (sound, thermal energy, space debris, etc.) present in a particular environment.
- **IndirectContributor**. The `IndirectContributor` concept represents concepts that indirectly contribute to the pollution at a particular spatio-temporal point. These include environmental factors like temperature, the air or water streams flowing into or out of a particular place, the socio-economic factors such as policies and demographics. We have modeled only the `Carrier` concept as the subclass of `IndirectContributor` since other indirect contributors are specific to certain kinds of pollution.
- **PlaceEntity and TimeEntity**. These concepts denote the place and time for a spatio-temporal observation. They are linked to the `Observation` and `TrajectoryPoint` concepts by the `atPlace` and `atTime` properties.
- **PrescribedStandardForPollutants**. This is a stub meta-pattern [10] that can be used to describe the prescribed ranges for pollutants. For a particular location, pollutants have a defined range of permissible concentrations that are specified by the various global authorities[7]. A standard for a pollutant may change with time and place. Hence the `PrescribedStandardForPollutants` concept is linked to the `Observation` concept by the `hasObservation` property.

[7] 2005 WHO guidelines prescribe the range for air pollutants such as particulate matter (PM), ozone (O3), nitrogen dioxide (NO2), etc., available at `http://whqlibdoc.who.int/hq/2006/WHO_SDE_PHE_OEH_06.02_eng.pdf?ua=1`.

- **Carrier**. This concept is a subclass of the `IndirectContributor` concept and represents the air, water, or other kinds of streams flowing into or out of a particular place. It is linked to the `Trajectory` concept through the `hasTrajectory` property. Carriers are generally observed to affect the concentration of pollutants at a particular place and are important in modeling pollutants. To represent the pollutants that might be carried through a trajectory, the `Carrier` concept is linked to the `Pollutant` concept by the `carriesPollutant` property. To specify the location of the source of pollutants in a carrier trajectory, `nearby` property can be used. This links the pollutant sources such as factories in the case of wind stream carrier or drains in the case of water stream carrier to the `TrajectoryPoint`. The applications that have such a requirement can make use of the `nearby` property, but we excluded it from the ODP because it is not sufficiently general.
- **Trajectory**. The `Trajectory` concept represents a set of ordered spatio-temporal points and has been directly adopted from [9]. It is linked to the `TrajectoryPoint` and `TrajectorySegment` through the `hasPoint` and `hasSegment` properties. The `nextPoint` property links trajectory points in the appropriate order within a trajectory. The segments in the trajectory are defined by a starting trajectory point $\{x_i, y_i, t_i\}$ and an ending trajectory point $\{x_j, y_j, t_j\}$ where t_i, t_j denote time points such that $t_i < t_j$. The `TrajectorySegment` concept represents this notion of a segment which is connected to two fixes through `startsFrom` and `endsAt` properties.

3.2 Properties

- **atPlace, atTime**. They connect `Observation` and `TrajectoryPoint` concepts to the `PlaceEntity` and `TimeEntity` respectively. Since an instance of `Observation` or `TrajectoryPoint` can be associated with at most one timestamp, the `atPlace` and `atTime` properties are functional.
- **startsFrom, endsAt**. These functional properties connect a `TrajectorySegment` to starting and ending `TrajectoryPoint` representing the starting and ending point of a segment.
- **carriesPollutant**. This property represents the pollutants that a `Carrier` can carry. It connects `Carrier` to the `Pollutant` concept and represents the pollutants being carried away by a carrier. The trajectory of the carrier dictates the displacement of pollutants.
- **hasTrajectory**. This property connects the `Carrier` concept to the `Trajectory`.
- **hasContributor**. This property connects the `Pollution` concept to the `Contributor` concept.
- **hasPoint**. This property connects the `Trajectory` with the `TrajectoryPoint`.
- **hasPrescribedStandards**. This property connects the `Pollutant` concept to the stub `PrescribedStandardForPollutant` concept.
- **hasObservation**. This property connects `Pollutant` and `PrescribedStandardForPollutant` concept to the `Observation` concept. It can be used to capture the concentration of pollutants or a prescribed standard for a particular pollutant with the `TimeEntity` and `PlaceEntity` as the parameters of the observation.

- **hasSegment**. This property connects the `Trajectory` to the `TrajectorySegment`.
- **nextPoint**. This property links each point represented by the `TrajectoryPoint` to the next point forming a chain of ordered points.

3.3 Axioms of the Pollution ODP

The axioms that are part of the Pollution ODP are discussed here.

$$\texttt{Carrier} \sqsubseteq \forall \texttt{carriesPollutant.Pollutant} \tag{1}$$

$$\texttt{Carrier} \sqsubseteq \exists \texttt{hasTrajectory.Trajectory} \tag{2}$$

$$\texttt{Pollution} \sqsubseteq \exists \texttt{hasContributor.Contributor} \tag{3}$$

$$\texttt{Observation} \sqsubseteq \exists \texttt{atPlace.PlaceEntity} \tag{4}$$

$$\texttt{Observation} \sqsubseteq \exists \texttt{atTime.TimeEntity} \tag{5}$$

$$\texttt{TrajectoryPoint} \sqsubseteq \exists \texttt{hasPoint}^{-}\texttt{.Trajectory} \tag{6}$$

$$\texttt{Trajectory} \sqsubseteq \exists \texttt{hasSegment.TrajectorySegment} \tag{7}$$

$$\texttt{hasSegment} \circ \texttt{startsFrom} \sqsubseteq \texttt{hasPoint} \tag{8}$$

$$\texttt{hasSegment} \circ \texttt{endsAt} \sqsubseteq \texttt{hasPoint} \tag{9}$$

Axioms 6, 7, 8 and 9 capture the relation between the `Trajectory`, `TrajectoryPoint` and the `TrajectorySegment` concepts.

3.4 Competency Questions

The *Pollution* ODP answers the following competency questions.

1. What are the contributors of the pollution?
2. What is the pollutant concentration at a particular time and place?
3. What are the carriers that contributed to the pollution?
4. What are the pollutants carried by a carrier?
5. What are the prescribed standards for a particular pollutant?
6. What is the trajectory of a carrier for a pollutant?

4 Use Cases

The *Pollution* ODP can be used as a building block to model various pollution types such as air, water and soil. They can be added as subclasses of the `Pollution` concept. An example use case of air pollution can describe the concentration of pollutants in the air at a particular spatio-temporal point through the `Observation` and `Pollutant` concepts. Similarly, the pollutants carried by the air stream can be captured by the `Carrier` concept. The `Weather` concept can be added as a subclass of `IndirectContributor` concept in the concrete implementation of this ODP to represent the weather related factors that contribute to air pollution. These concepts can be extrapolated to water, soil and other types of pollution as well.

5 Conclusion and Future Work

Pollution exists in many forms in the environment and affects individuals as well as ecosystems. An ontology design pattern for modeling various sources and characteristics of pollution can be used effectively as a building block by multiple applications that work with pollution data. We introduce an ODP for modeling pollution and discuss its competency questions, concepts and properties. We also describe some use cases of the ODP. The ODP and its documentation are publicly available at `http://ontologydesignpatterns.org/wiki/Submissions:Pollution` and at `https://github.com/kracr/aq-structured-platform/blob/main/Ontology/PollutionODP/PollutionODP.owl` with an Apache License 2.0.

We plan to use the *Pollution* ODP to model air pollution by considering the pollutant concentration, weather and wind trajectory data sources. The air pollution ontology will be populated using these data sources by converting the semi-structured data (csv, json, etc.) into an RDF[8] graph using YARRRML [7].

Acknowledgements. This work is partially supported by the Infosys Center for Artificial Intelligence (CAI) at IIIT-Delhi, India.

References

1. Air and Water Pollution: An Important Nexus of Transportation and Health. In: Meyer, M., Elrahman, O. (eds.) Transportation and Public Health, pp. 65–106. Elsevier (2019). https://doi.org/10.1016/B978-0-12-816774-8.00004-9
2. Bellas, J., Hylland, K., Burgeot, T.: Editorial: New Challenges in Marine Pollution Monitoring. Frontiers in Marine Science **6**, 820 (2020). https://doi.org/10.3389/fmars.2019.00820
3. Cromar, K.R., Duncan, B.N., Bartonova, A., Benedict, K., Brauer, M., Habre, R., Hagler, G.S.W., Haynes, J.A., Khan, S., Kilaru, V., et al.: Air Pollution Monitoring for Health Research and Patient Care. An Official American Thoracic Society Workshop Report. Annals of the American Thoracic Society **16**(10), 1207–1214 (Oct 2019)
4. Dalia Dahleh, M.S.F.: An Environmental Ontology for Global City Indicators (ISO 37120) (09 2016)
5. Eva Blomqvist: Submissions:Observation - Odp. `http://ontologydesignpatterns.org/wiki/Submissions:Observation`, (Accessed on 15/07/2021)
6. Goldsmith, J.: Effects of Air Pollution on Human Health. In: STERN, A.C. (ed.) Air Pollution and its Effects (Second Edition), pp. 547–615. Academic Press, second edition edn. (1968). https://doi.org/10.1016/B978-0-12-666551-2.50021-2
7. Heyvaert, P., De Meester, B., Dimou, A., Verborgh, R.: Declarative Rules for Linked Data Generation at your Fingertips! In: Proceedings of the 15th ESWC: Posters and Demos (2018)
8. Hitzler, P., Gangemi, A., Janowicz, K., Krisnadhi, A.A., Presutti, V.: Ontology Engineering with Ontology Design Patterns: Foundations and Applications. IOS Press, NLD (2016)

[8] `https://www.w3.org/TR/rdf11-primer/`

9. Hu, Y., Janowicz, K., Carral, D., Scheider, S., Kuhn, W., Berg-Cross, G., Hitzler, P., Dean, M., Kolas, D.: A Geo-ontology Design Pattern for Semantic Trajectories. In: Tenbrink, T., Stell, J., Galton, A., Wood, Z. (eds.) Spatial Information Theory. pp. 438–456. Springer International Publishing, Cham (2013)
10. Krisnadhi, A., Hitzler, P.: The stub metapattern, pp. 39–45 (12 2017)
11. Manisalidis, I., Stavropoulou, E., Stavropoulos, A., Bezirtzoglou, E.: Environmental and Health Impacts of Air Pollution: A Review. Frontiers in Public Health **8**, 14 (2020). https://doi.org/10.3389/fpubh.2020.00014
12. Métral, C., Falquet, G., Karatzas, K.: Ontologies for the Integration of Air Quality Models and 3D City Models (2012)
13. Oprea, M.M.: AIR_POLLUTION_Onto: an Ontology for Air Pollution Analysis and Control. In: Iliadis, Maglogiann, Tsoumakasis, Vlahavas, Bramer (eds.) Artificial Intelligence Applications and Innovations III. pp. 135–143. Springer US, Boston, MA (2009)