**Common Action 2021 Quantum Hackathon Challenge**

**Impact Demo Ontology**

**Demo Ontology Details**

The backing ontology for the Common Action Hackathon data organizes information based on the “6Ws” - Who, What, Where, When, Why and hoW:

* Happenings/occurrences/states and conditions (what and how)
  + Examples are the generation of pollutants or electrical energy, and the perpetuation of conditions such as food insecurity or lack of sanitation in a country
  + Note that the various types of pollutants define how the environment is impacted, how electricity is produced, or how the people of a country live (e.g., in what conditions)
  + To codify the details of these occurrences and conditions, data is reported as types of Measurements and Assessments
* Agents/actors/organizations (who) and Locations (where)
  + Specifically, the data examines the environmental impact of over 1900 organizations, across 72 countries (where the organizations are headquartered)
  + In addition, the industries of the organizations are captured (classified by the Global Industry Classification Standard, GICS[[1]](#footnote-1))
* Time/sequences (when)
  + The data only includes annual/yearly information, so that is specifically reported in the KG
  + Impact, economic and environmental data is reported for multiple years for both organizations and countries
  + Measurements indicate both a :reported\_value and a :reported\_year
* Goal/intent (why)
  + The impact of organizations on several of the UN Sustainable Development Goals is available
* Causation, precondition and prevention (why)
  + Although the ontology can support cause/effect, enablement/prevention of events and conditions, and similar relationships, there is no supporting data that utilizes these associations, at the present time

The following figure shows the main concepts of the backing ontology and how they are related. These concepts define the basic terminology and structure by which the Hackathon data is organized. They represent the core semantics and the vocabulary for asking and answering queries related to sustainability, regardless of the various property names, foreign keys or web page formats used by the different data sources.

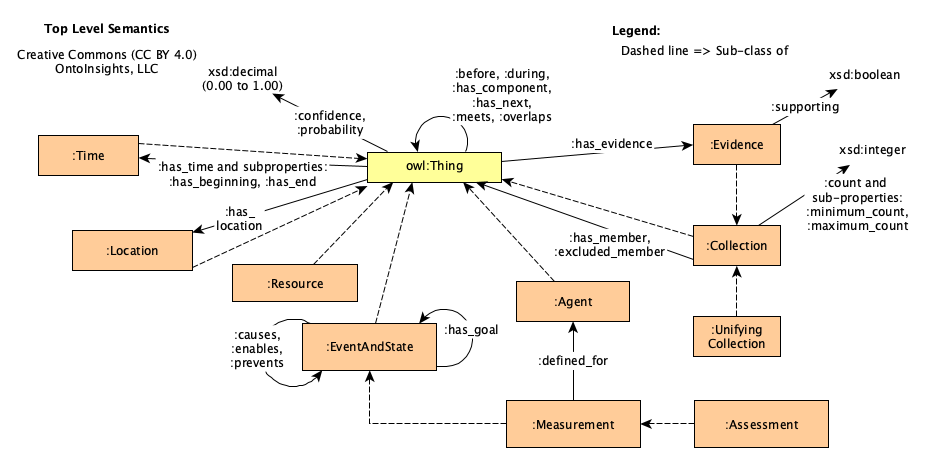


Figure 1. Top-Level Ontological Concepts

**Extending the Backing Ontology for the Demo**

An ontology’s main value proposition is to define the concepts and relationships of a domain, in order to organize the domain’s data/knowledge. All data from the Hackathon sources (the Harvard study[[2]](#footnote-2) and the CIA World Factbook[[3]](#footnote-3)) is translated and harmonized to the concepts, data properties and relationships of the Deep Narrative Analysis (DNA) backing ontology[[4]](#footnote-4), fusing the data from the disparate sources and making it consistent.

All the concepts from the DNA ontology are condensed from the various modules in the GitHub ontologies directory (referenced in footnote 2), to the file, background-ontology.ttl[[5]](#footnote-5). It was not necessary to do this condensing, but it simplified the resulting HTML documentation tree that was generated (and is discussed in a bit more detail below).

For the Hackathon, it was necessary to extend the DNA ontology to define the specific Measurements and Assessments that were captured from the data sources. The new Measurement and Assessment sub-classes are specified in the hackathon-esg-ontology.ttl file in GitHub[[6]](#footnote-6).

In addition to the new Measurement classes, it was also necessary to define a few new properties. The new properties are listed below along with the reasons for their definition:

* about\_industy – Disambiguates the multiple “topics” of the AvgSalesByIndustry Measurement. This relationship identifies the particular industry for which average sales are reported, in the Country for which the Measurement is defined.
* environmental\_issues – Captures the unstructured text from the CIA World Factbook describing the general environmental conditions for a Country. The text is associated with the Country using this property.
* has\_headquarters – Clarifies the has\_location property to indicate that this Location/Country is where an Organization is headquartered.
* is\_in\_industry – Is semantically identical to the has\_line\_of\_business predicate from the backing ontology, but is renamed to align with the vocabulary of GICS (the Global Industry Classification System)
* land\_area\_sq\_kms – Defines an additional property for a Country indicating its total land area (from the CIA World Factbook). Note that the backing ontology only defines a general area\_sq\_kms property.
* localized – When true, indicates that a FoodInsecurityAssessment (of “High” or “Very High”) is not a general condition, but localized. This information is obtained from the CIA World Factbook.
* reported\_year – Simplifies the backing ontology’s has\_time property to avoid creating a time-related instance, just to indicate the year for which a Measurement is reported.
  + The more complex semantics of the backing ontology are needed for more general time declarations (and indeed can be reconstructed from this simplification)

Lastly, reviewing all the files in the hackathon-extensions directory of the DNA ontology on GitHub[[7]](#footnote-7), one finds a few more files beyond the background-ontology and hackathon-esg-ontology modules discussed above. These are:

* geonames\_countries.ttl – This file was directly copied from the DNA ontologies directory. It provides information on Countries, their currency and neighboring Countries, and the Continents that contain them.
* lob\_industry\_extensions.ttl – This module extends the LineOfBusiness classes defined in DNA’s agent.ttl. The extensions address the industries from GICS that are referenced in the Harvard study.

**Data Fusion from the Harvard Study and CIA Factbook**

The following figures show how the data from the Harvard study, the CIA Factbook and GeoNames[[8]](#footnote-8) are rendered in the Turtle syntax. (These images are also loaded in the hackathon-docs subdirectory on GitHub[[9]](#footnote-9).)

****

Figure 1. Company Data from Harvard Study and Country GeoNames Data

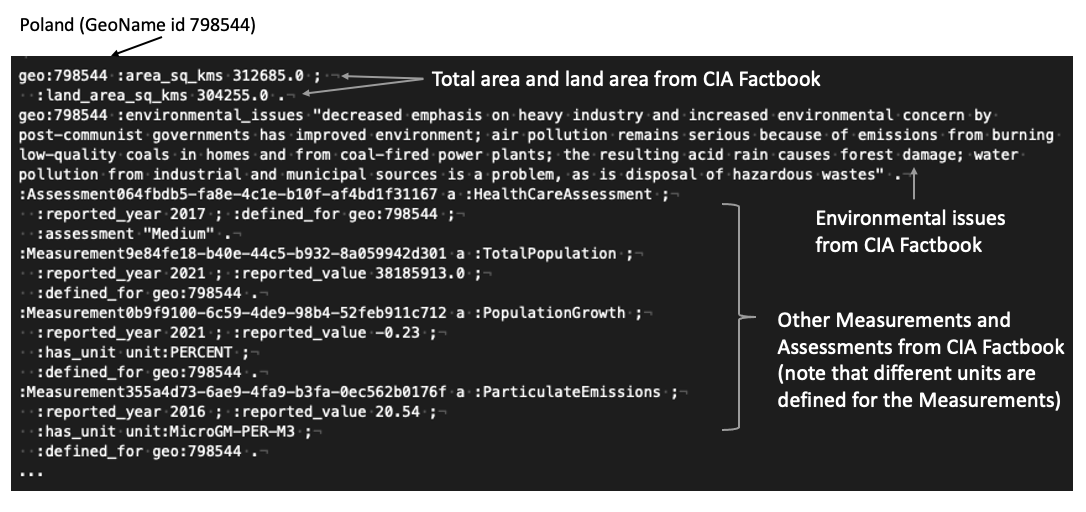


Figure 2. Country Data from the CIA World Factbook

**Reviewing the Impact Demo Ontology**

In order to better enable understanding of the ontology used in the Common Action Impact Demo, several resources are provided in the hackathon-docs sub-directory on GitHub (location footnoted on the previous page). These resources include:

* This document and all other “CA Quantum Hackathon” documents
* All PNG images embedded in the above documents
* Measurements.xlsx, which provides the complete definitional text for all Company and Country measurements and assessments
* Hackathon-ontology-tree.html, which is a searchable tree view of all the concepts (classes and properties) in the background-ontology and hackathon-esg-ontology modules
  + The HTML file can be downloaded and loaded to any browser
  + Its tree view shows the inheritance (generalization-specialization) hierarchy of the classes and properties (an example is shown in the figure below)
* Merged-hackathon-ontology.ttl, which combines the background-ontology and hackathon-esg-ontology modules into a single file that can be loaded into the Protégé ontology editor[[10]](#footnote-10)

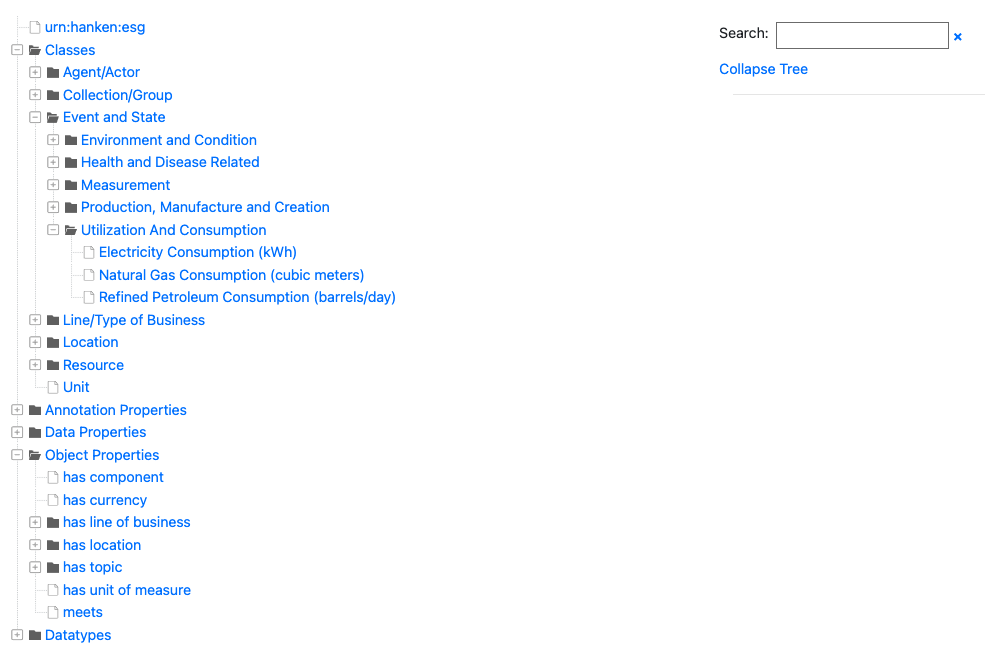


Figure 3. Example of the Hackathon Demo Ontology HTML Tree View

1. <https://www.msci.com/our-solutions/indexes/gics> [↑](#footnote-ref-1)
2. Produced by “Freiberg, David and Park, DG and Serafeim, George and Zochowski, Rob. 2020. Corporate Environmental Impact: Measurement, Data and Information. Harvard Business School, Impact-Weighted Accounts Project report.” [↑](#footnote-ref-2)
3. <https://www.cia.gov/the-world-factbook/> [↑](#footnote-ref-3)
4. <https://github.com/ontoinsights/deep_narrative_analysis/tree/master/ontologies> [↑](#footnote-ref-4)
5. <https://github.com/ontoinsights/deep_narrative_analysis/blob/master/ontologies/hackathon-extensions/background-ontology.ttl> [↑](#footnote-ref-5)
6. <https://github.com/ontoinsights/deep_narrative_analysis/blob/master/ontologies/hackathon-extensions/hackathon-esg-ontology.ttl> [↑](#footnote-ref-6)
7. <https://github.com/ontoinsights/deep_narrative_analysis/tree/master/ontologies/hackathon-extensions> [↑](#footnote-ref-7)
8. <https://www.geonames.org/> [↑](#footnote-ref-8)
9. <https://github.com/ontoinsights/deep_narrative_analysis/tree/master/ontologies/hackathon-extensions/hackathon-docs> [↑](#footnote-ref-9)
10. <https://protege.stanford.edu/> [↑](#footnote-ref-10)