Alaskan Community Ontology

A Book Made with Quarto

Liz Dobbins

Table of contents

1 Alaska Community Ontology

A Book Made with Quarto

1.1 Welcome

This repository aims to support needed conversations regarding a foundational element of Alaska Energy Data Gateway (AEDG): the creation of derived aggregates of socio-economic and energy data at upper levels such as administrative regions and grids.

To accomplish the computation of these aggregates in a reproducible way, a cross-walk between conceptual groupings of communities is required. This conceptual model will be applicable to diverse types of data. It will also guide the development of the algorythms that compute the aggregations and convert between them.

If possible, this cross-walk will be machine readable/actionable. But to establish with consensus between our various subject matter experts, we need to start with a single set of definitions and axioms (rules) expressed in a human readable form. This could happen in a wiki, but we should maintain a record of revisions so we can document discussions and final recommendations. The best format for this process is a GitHub repo.

1.2 Goal

To provide a basis for tranformation of data provided at different scales so that multiple groupings can be displayed on AEDG dashboards.

This is required because data will provided at different scales. For instance:

- Because of interties, electricity is not always reported by individual communities.
- Some ways of avoiding personally identifiable information (such as for building footprints) involve reporting by zip code.
- Existing analysis such as [somebody's] energy burden data is reported at census tract level.

Also, different audiences need data reported at different scales:

- Communities want a summary of their own data.
- Legislative Aides want summaries by legislative district.
- Agency leads might want to compare PCE communities with railbelt communities.

The assumption is that this needs to function **for the current time only**. Of course, boundaries change over time, and that these changes are of interest to researchers. However, the main users of AEDG have been defined as requiring current data to support future projects or to determine the success of existing projects. Researchers can download data and work out time-dependent issues for themselves. In a future time, we might want to support this with a time-dependent version of this system, but now is not that time.

1.3 Format

Development will begin with a series of web pages oriented around ontological concepts. The concepts are:

- Classes: Types of entities. The base Class is Community; dispite the fact that this might be the hardest Class to define, Community is the organizational principle of previous versions of AEDG, so we will start with that. Other Classes (Borough, AEARegion, ZipCode) will be defined in relation to Community.
  - Classes could be translated to tables in a relational database, if that makes sense for development.
  - Instances of a Class are nodes in a graph, if you think about things that way.
- Properties: These are attributes of the Classes.
  - Properties could take the form of columns in tables in a relational database.
- Relations: These can be heirarchical (taxonomy) or could form a network or graph (semantic web) of the various entities. If they are based on domain specific rules (Axioms), then that is what distinguishs an ontology. Relations enable AI to infer new information without people having to explicitly define everything.
  - Relations are the edges in a graph.
  - In a relational database, Relations can be encoded using a lookup table with a combination of foreign keys.
- Axioms: Domain-specific rules that encapsulate what we know about the Classes.
  - Axioms could form the basis of quality checks to entered data or the definition of the ontology. For instance, if the axioms include `x = y` and `y = z`, trying to enter an `x` that isn't equal to `z` would fail a test.

1.4 Anticipated Workflow

1. Liz will draft an initial version based on discussions at the Nov 11, 2024, meeting of AEDG personnel.
2. SMEs will comment on the draft by submitting GitHub Issues. Problems will be raised and solutions will be suggested.
3. When consensus has been reached (which might happen in an offline discussion), the resolution will be recorded in the issue, revisions made to this web book, and the issue closed.
4. After the most substantive discussions have been resolved, Liz will attempt to translate this ontology into the OWL ontology language.
5. If that does not work, we will try a relational database for consistency with other AEDG products.
6. If the concepts do not resolve into a relational database, we can try a graph database. We might want to leave this option open for visualization possibilities.

Whenever possible, we will rely on existing resources and resist defining our own ontology from scratch. Hopeful candidates are:

- Darwin Core
- OBO Relations Ontology

As of Nov. 2024, it is unclear how to make use of an OWL formated ontology with code. If this becomes problematic, we reserve the right to use a relational database instead based on our greater familiarity with that tech. Ontology specific work can be delayed to future development cycles.

1.5 This is a stoat

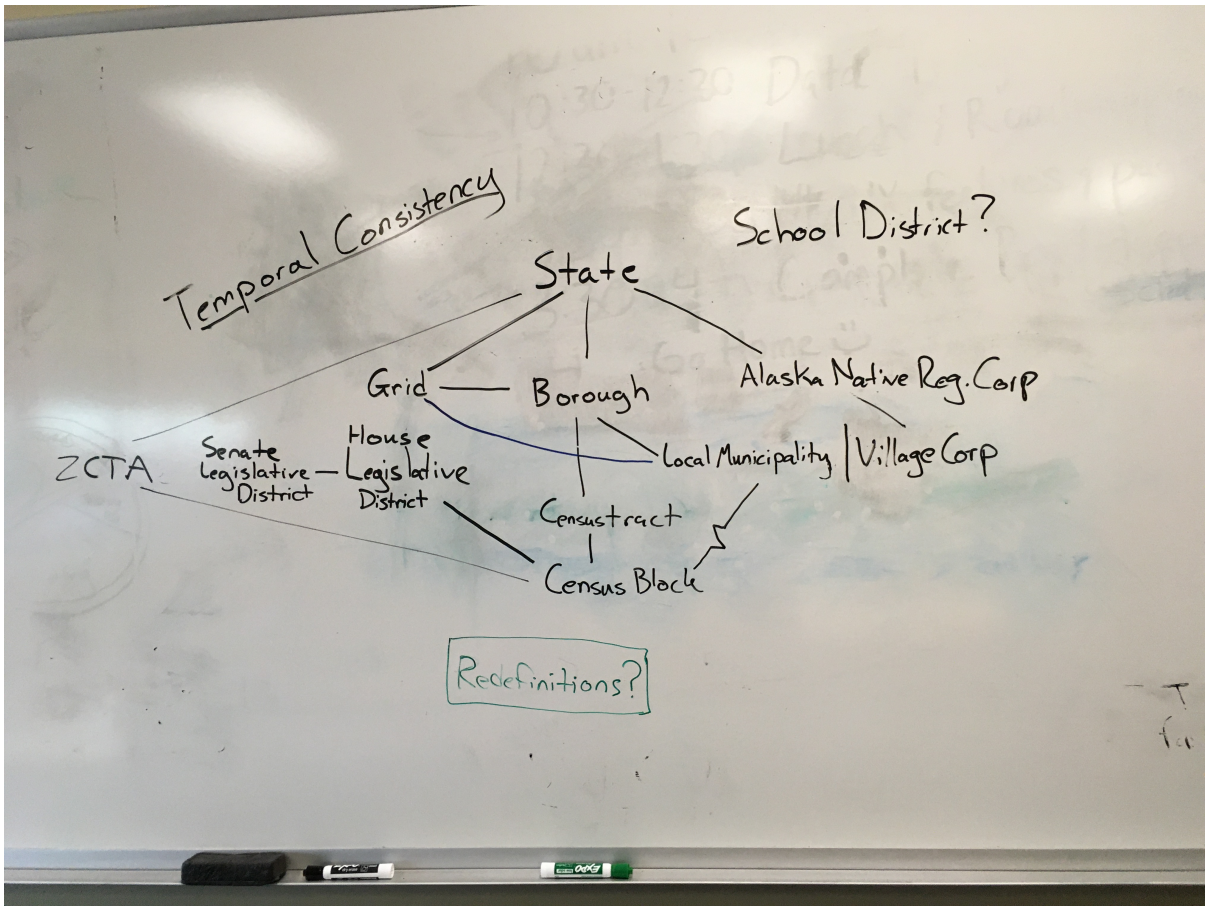

Figure 1.1: Initial Brainstorming.

2 Classes

Types of entities in the Alaska Community Ontology

2.1 Classes

Classes are types of entities. The base Class is Community; dispite the fact that this might be the hardest Class to define, Community is the organizational principle of previous versions of AEDG, so we will start with that. Other Classes (Borough, AEARegion, ZipCode) will be defined in relation to Community.

- Classes could be translated to tables in a relational database, if that makes sense for development.
- Instances of a Class are nodes in a graph, if you think about things that way.
- Classes are typically capitolized to differentiate them from instances or objects of that class, which are lower case.

This page also includes Properties, which are attributes of the Classes.

- Properties could take the form of columns in tables in a relational database.

2.2 Common Properties

These are attributes shared by all Classes

- id (int): All Classes should have an id column for inclusion in lookup tables.
  - How we determine these is a whole 'nother thing.
- label (string): Human readable name. This cannot be the key of the table becaues multiple Classes could have the same label (example: Fairbanks the Municipality and Fairbanks the Borough)

- classModificationDate (time): Classes will evolve as we evaluate their effectiveness, so we should record when the class definition changes to identify which Classes might need reevaluation.
- entityModificationDate (time): This is to allow evolution in time, which we do not plan to utilize in the first development cycle of AEDG.
- geometry (geom): As it is developing, each of these entities seems to have spatial component, though it might be a Point, Polygon, or Line.

2.3 List of Classes

2.3.1 Community

- **Definition**: Construct of AEDG that is the fundamental unit of displays and dashboards.
- **Properties**:
  - location (WKT Point): longitude, latitude of the Community for convenience of display. Perhaps the centroid of the Municipality or a commonly accepted "place".

2.3.2 Municipality

- **Definition**: Administrative entity for a town, city, or village.
- **Properties**:
  - boundary (WKT Polygon): longitude, latitude points that denote the boundary of the town

2.3.3 VillageCorp

- **Definition**: a type of Municipality?
- **Properties**:
  - boundary (WKT Polygon): longitude, latitude points that denote the boundary of the village (are these available?)
  - Additional properties should be determined by what is important to the people.

2.3.4 Borough

- **Definition**: Administrative collection of Municipalities analagous to a county in other places.
- **Properties**:
  - type (string): 1st class or 2nd class
  - boundary (WKT Polygon): longitude, latitude points that denote the boundary of the Borough

2.3.5 HouseDistrict

- **Definition**: Alaska state House districts as available from [link]
- **Properties**:
  - [wassit] (int): digit that denotes the district
  - boundary (WKT Polygon): longitude, latitude points that denote the boundary of the house district.

2.3.6 SenateDistrict

- **Definition**: Alaska state Senate districts as available from [link]
- **Properties**:
  - [wassit] (character): Capitol letter that denotes the district
  - boundary (WKT Polygon): longitude, latitude points that denote the boundary of the house district.

2.3.7 RegionalCorp

- **Definition**: An Alaska Native Regional Corporation established by the [wassit] as available from [link]
- **Properties**:
  - boundary (WKT Polygon): longitude, latitude points that denote the boundary of the Regional Corporation.
  - Additional properties should be determined by what is important to the people.

2.3.8 Region

- **Definition**: Definitions of regions that have been used to summarize energy trends in the Energy Statistics Workbooks or the Alaska Electricity Trends Report (2024). If new regions are defined, they can be included, but this is intended to be either AEA's 11 regions or ACEP's 3 regions (Coastal, PCE, or Railbelt).
- **Properties**:
  - organizaton (string): 'ACEP' or 'AEA' as the origanization that defined the region.
  - ~~boundary (WKT Polygon):~~ A spatial boundary does not make sense because these were used to group communities together, and in the case of ACEP's regions, they were intended to be discontinuous. Meaning, there are some PCE communities that are pretty near the coast because the designation is based on power costs and not geography.

2.3.9 CensusBlock

- **Definition**: TBA
- **Properties**:

2.3.10 CensusTract

- **Definition**: TBA
- **Properties**:

2.3.11 Grid

- **Definition**: An entity that provides electricity to a Community. Includes generation and distribution, and is typically categorized by capacity.
- **Properties**:
  - microgrid (boolean): True/False (need a definition for this)
  - capacity (int):

2.3.12 Utility Service Area

- **Definition**: The corporate entity that runs the Grid
- **Properties**:
  - coop (boolean): True/False is it organized as a member owned co-operative?
  - private (boolean): True/False is it a private company?

  - ~~Boundary (WKT Polygon)~~: I feel like a spatial version of a corporate entity does not make sense. Especially for AVEC, whose region is discontinuous.

2.3.13 Intertie

- **Definition**: The transmission lines that create a Grid that includes multiple Communities
- **Properties**:
  - path (WKT Line): longitude, latitude points that trace an acceptible, if not accurate, route of the lines.
  - owner (string): Who owns it?
  - maintainer (string): Who maintains it? Is this the same as Owner?

2.4 Optional

2.4.1 ZipCode

- **Definition**: Postal zip codes as available from [link] (ZCTA)
- **Properties**:
  - boundary (WKT Polygon): longitude, latitude points that denote the boundary of the house district.

2.4.2 SchoolDistrict

- **Definition**: Community based school districts. This might be out of scope or better handled as research.
- **Properties**:
  - boundary (WKT Polygon): longitude, latitude points that denote the boundary of the school district. This is too much to expect, probably.

2.4.3 PowerPlant

- **Definition**: Entities that generate electricity for grids. This might be so complex that is it better handled by a different system.
- **Properties**:
  - fuel (string): OK. now I know this belongs elsewhere, because fuel type should be a lookup table and that doesn't belong with communities. Perhaps already handled by the Open Energy Ontology.

2.5 From Initial Diagram

House Legislative District Senate Legislative District Borough Alaska Native Regional Corporation Local Municipality Village Corporation (same as Local Municipality?) Census Tract Census Block Grid Utility Service Area Intertie (same as Grid?) School District

3 Relations

Domain-defined ways that entities relate to each other in the Alaska Community Ontology

Relations can be heirarchical (taxonomy) or could form a network or graph (semantic web) of the various entities. These are based on domain specific rules (Axioms), and that defines an ontology. Relations enable AI to infer new information without people having to explicitly define everything.

- Relations are the edges in a graph.
- In a relational database, Relations can be encoded using a lookup table with a combination of foreign keys.

4 Axioms

Domain-specific rules that describe entity relations in the Alaska Community Ontology

Axioms are domain-specific rules that encapsulate what we know about the Classes.

- Axioms could form the basis of quality checks to entered data or the definition of the ontology. For instance, if the axioms include `x = y` and `y = z`, trying to enter an `x` that isn't equal to `z` would fail a test.
- An Axiom should describe only one quality because it will eventually have a code-based test, and we want to keep those independent of each other.

4.1 Community

- Communities are primarily determined by the expectations of the AEDG audience.
- Communities should be equal to a single Municipalities or VillageCorps (preferred))
- Communities could be collections of Municipalities or VillageCorps (example: Fairbanks should include College and Farmers Loop, which are separate places in the [wassit])

4.2 State

- The State is defined to be Alaska
- All Communities are in the State

4.3 Districts

- HouseDistrict is a subset of SenateDistrict (is this true?)
- HouseDistricts can contain multiple Communities.
- A single Community might intersect several HouseDistricts (example: Fairbanks)
  - Therefore: HouseDistrict population cannot be determined from a sum of Community populations.

4.4 Boroughs

- Communities can be unencorporated, meaning they are not in a Borough.
- If encorporated, a Community can only be in a single Borough

4.5 ## Census

4.6 Energy-Related

- Communities can be isolated (microgrids)
- Communities can be connected via Interties

4.7 Zip Codes (Optional)

- A ZipCode can contain multiple Communities, especially in rural areas.
- A single Community can contain multiple ZipCodes, especially in urban areas.

A About

This respository exists to support development of the Alaska Energy Data Gateway by the Alaska Center for Energy and Power (ACEP), UAF, and the Institute for Social and Economic Research (ISER), UAA. It does not exist to be its own thing. If it ceases to be applicable to the AEDG project, it should be archived.