Tyler Procko – *Semantic Science*
86130 Williams Avenue
Yulee, FL 32097
November 17th, 2023

Semantic Science

- Ontology Engineering & Research Publication -

www.semanticscience.us

General Workflow

I began with a survey of extant taxonomies describing related content. The best one I knew of was the DMOZ project from AOL, which started in 1999. It was the world's first attempt at humans organizing the web. It has a massively deep taxonomy that informed a lot of what I did here.

Arts Movies, Television, Music...	**Business** Jobs, Real Estate, Investing...	**Computers** Internet, Software, Hardware...
Games Video Games, RPGs, Gambling...	**Health** Fitness, Medicine, Alternative...	**Home** Family, Consumers, Cooking...
News Media, Newspapers, Weather...	**Recreation** Travel, Food, Outdoors, Humor...	**Reference** Maps, Education, Libraries...
Regional US, Canada, UK, Europe...	**Science** Biology, Psychology, Physics...	**Shopping** Clothing, Food, Gifts...
Society People, Religion, Issues...	**Sports** Baseball, Soccer, Basketball...	**Kids & Teens Directory** Arts, School Time, Teen Life...

I started with a basic mapping of all your sites' high-level (presumably determined to be the most *important*) categories. Obviously, some of the DMOZ categories, like science, are irrelevant.

Articles ⌄　　Health + Wellness　　Things To Do

40+ Best Things to Do　　Events　　Disney　　Eats　　Drinks　　Getaways

Things to Do　　Do　　Eats　　Drinks
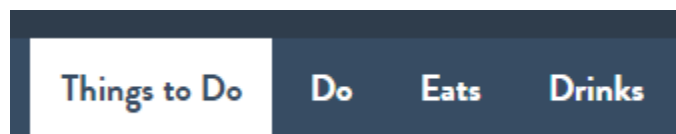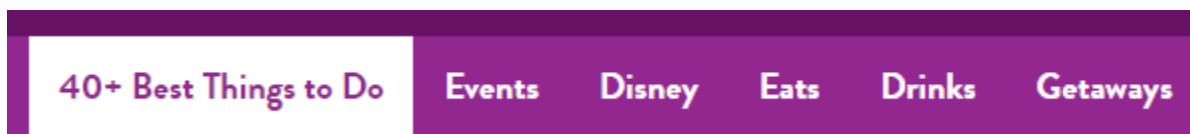
ARTICLES ⌄　　HOLIDAY FAMILY FUN　　THINGS TO DO ⌄

On both the Parenting websites, I notice you *do* allow the sitemap taxonomy to show, at least to a small degree.
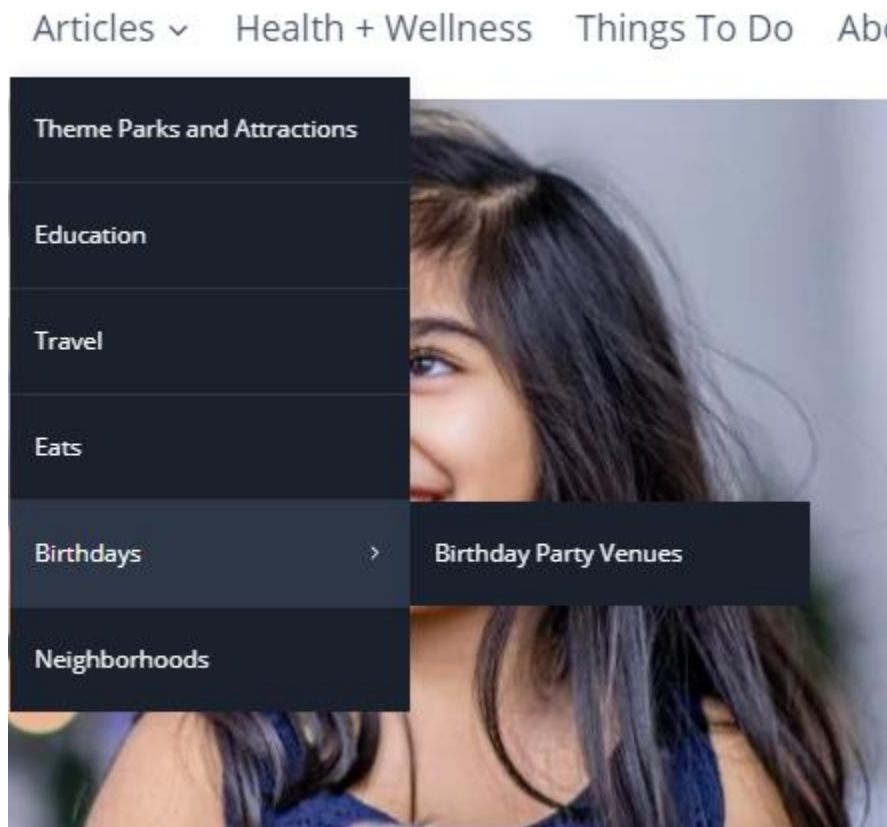
This is good, since I wanted to develop a deeper taxonomy. On your end, you can pick and choose which categories (classes in my lingo) to show, depending on business needs.

A lot of the tags your content writers create are incredibly specific and often used for a single post, to be discarded and forgotten.

https://www.orlandodatenightguide.com/tag/pechakucha/
https://www.orlandodatenightguide.com/tag/pizza/
https://www.orlandodatenightguide.com/tag/plaza/
https://www.orlandodatenightguide.com/tag/plaza-live/
https://www.orlandodatenightguide.com/tag/poetry-slam/
https://www.orlandodatenightguide.com/tag/princess-ball/
https://www.orlandodatenightguide.com/tag/psychedelic-rock/

So, the general idea was to create the most general taxonomy possible, allowing all four sites (and future ones!) to re-use from without re-inventing the wheel. My operating assumption was that you should enforce a strict, controlled taxonomy of categories to place posts under (what I have developed so far), only allowing authors to create new tags as needed. In other words: you should *own* the taxonomy as IP, ensuring your authors use it to classify their posts.

I draw the line, philosophically, between a *category* and a *tag*, at the number of uses and its specificity. If something is only used once, or is very specific, it is most likely a tag. If something is used more than once, or is rather general, it is a category. For example: the category *Bars* is indeed a category, whereas *Willie's Brewskis on 5th Avenue* is a tag. See the Conclusion section for more on this.

The process of categorization/tagging I propose:

1. Authors write their content
2. They traverse the taxonomy from top to bottom and pick ~3-5 (your choice!) categories they like
3. (TBD) They check the tag "bins" and re-use or make some new ones for others to have

This way, you definitively control your categories, but allow some freedom of expression in your tags.

An additional thought I had was to use GPT-4, specifically by "training" a domain-specific GPT, to automatically categorize your posts, so there isn't any author deviance. I set a model up pretty quickly and asked it to give me the 5 best categories for this post: https://orlando-parenting.com/free-christmas-events-holidays-orlando-kids/. Its response was the following:
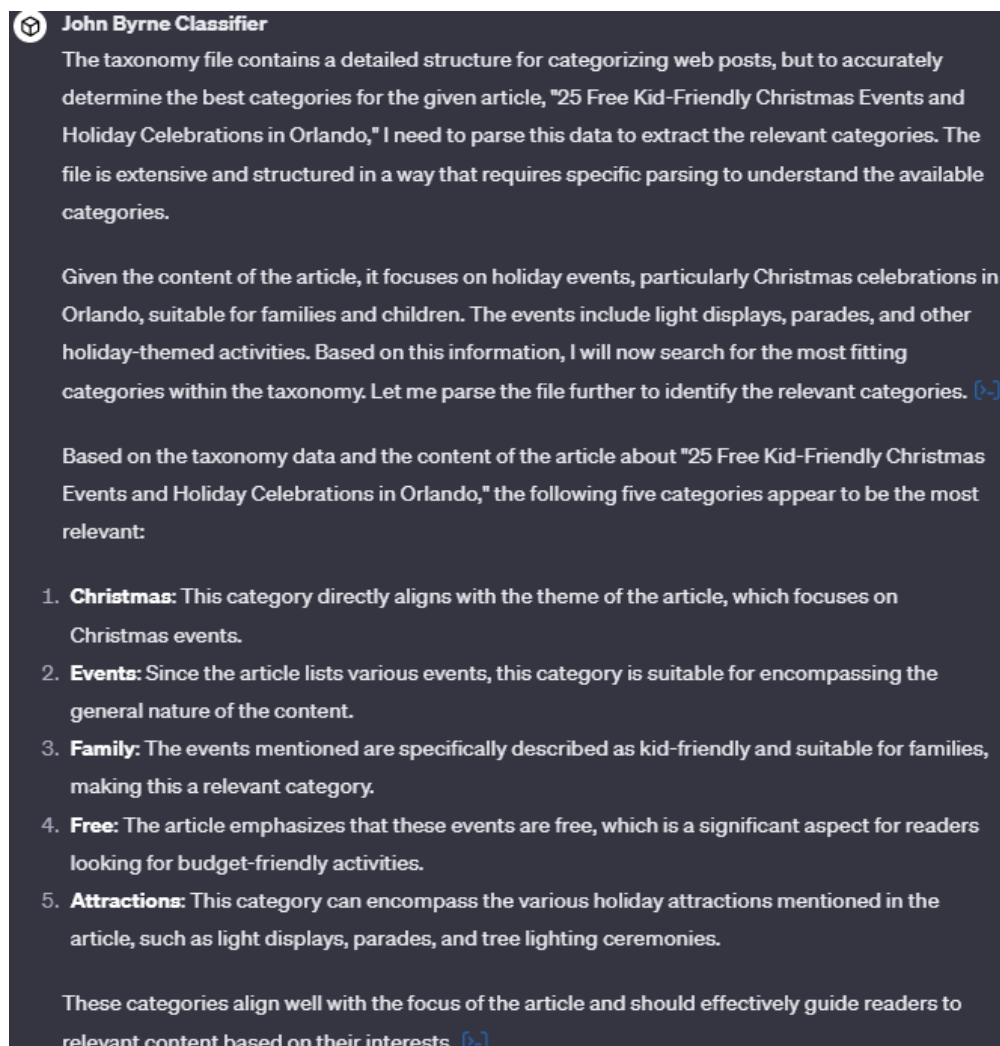
I can pretty easily set this up as an API call so it can be used by anyone. The same idea can be done for tagging.

Taxonomy

A full, interactive visualization of the taxonomy can be found here. It is reproduced below (hard to see):

It is given as a bulleted list below:
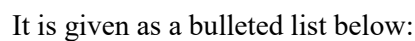
- People
  - InterestingPeople
    - Kids
    - Moms
- ThingsToDo
  - Dancing
  - Dining
    - Bars
    - Breakfast
    - Breweries
    - Brunch
    - Desserts
    - Dinner
    - Drinks
      - Beer
        - CraftBeer
      - Coffee
      - Wine
    - Lunch
    - Restaurants
  - Events
    - Camping
      - Glamping
    - DateNight
    - DayTrip
    - FamilyFun
    - GirlsNightOut
    - Parade
    - RoadTrip
  - Games
  - Holidays
    - CulturalHolidays
      - CincoDeMayo
      - FathersDay
      - Halloween
        - TrickOrTreating
      - July4th
        - July4thFireworks
      - Kwanzaa
      - LaborDay
      - MLKDay
      - MemorialDay
      - MothersDay
      - NationalChocolateDay
      - NewYears
      - NewYearsEve
      - PresidentsDay
      - SpringBreak
      - StPatricksDay
      - Thanksgiving
      - ValentinesDay
      - VeteransDay
    - ReligiousHolidays
      - Christmas

- - - ChristmasLight
    - ChristmasTree
  - Easter
    - - EasterEggHunting
  - Hanukkah
- Scene
  - Indoor
  - Outdoor
- Seasons
  - Fall
    - FallFestivals
      - PumpkinPatch
  - Spring
  - Summer
  - Winter
- Shopping
  - Deals
- Sports
  - Biking
  - Hiking
  - Running
    - Marathon
    - 5k
  - Swimming
- Travel
  - Ballroom
  - BirthdayPartyVenue
  - Cruise
  - Getaways
  - Locations
    - Accommodations
      - Hotel
      - Inn
      - Wedding
    - Attractions
      - Aquarium
      - Beaches
      - Concert
      - Contest
      - Festival
      - Museums
        - ArtMuseum
        - HistoryMuseum
      - Park
        - StatePark
        - ThemePark
      - Zoo
        - PettingZoo
    - Camps
      - SummerCamp
    - DanceStudio
    - FarmersMarket
    - Farms
      - UPickFarm

- - - Libraries
    - Neighborhoods
    - Playgrounds
  - Volunteer
- Health
  - Fitness
  - MentalHealth
- Animals
  - Pets
    - Cats
    - Dogs
- Regions
  - State
    - Florida
      - Clearwater
      - Orlando
      - StPetersburg
      - Tampa
        - TampaBay
      - WinterGarden
- Home
  - Beauty
  - Birthday
  - Craft
  - Education
    - BackToSchool
  - Kitchen
    - Recipes
  - Parenting
    - ElementaryYears
    - HighschoolYears
    - Infancy
    - MiddleSchoolYears
    - PrePregnancy
    - Pregnancy
    - ToddlerYears
    - YoungAdulthood
  - Style

The top-level categories are:

- Health
- Home
- Animals
- News
- People
- Regions
- Things-To-Do

There is room to expand and room for debate. I have covered most (~600) of the categories/tags from your spreadsheets, but some remain to be finalized. Feedback is key!

ns1:loc						done
https://tampabaydatenightguide.com/category/uncategorized/	🟥					maybe done
https://tampabaydatenightguide.com/category/do/	🟩					ignored
https://tampabaydatenightguide.com/category/events/	🟩					tag
https://tampabaydatenightguide.com/category/cheapfree/	🟦					location
https://tampabaydatenightguide.com/category/eats/	🟩					
https://tampabaydatenightguide.com/category/drinks/	🟨					
https://tampabaydatenightguide.com/category/getaways/	🟨					
https://tampabaydatenightguide.com/category/girls-night-out/	🟨					
https://tampabaydatenightguide.com/category/deals/	🟩					
https://tampabaydatenightguide.com/category/holidays/	🟩					
https://tampabaydatenightguide.com/category/holidays/christmas-lights/	🟩					
https://tampabaydatenightguide.com/category/celebrate-fall/	🟩					

I think that many of the very specific tags your authors have used can be represented by selecting 2 or more categories. For instance, the tag "thanksgiving-fall-food-dinner" is a bit silly, and can be better represented by selecting the categories: *Thanksgiving*, *Fall*, *Dining*, *Dinner*.

<u>Sitemaps</u>

One thing you mentioned to me was the difficulty in semantically representing what a Pumpkin Patch is. I believe fringe things like this can now be adequately represented:

So, the sitemap URL would look like:

www.yoursite.com/category/things-to-do/seasons/fall/fall-festivals/pumpkin-patch

When a search engine sees your post under this directory in the sitemap, it has greater semantic depth than if it were just tagged with *pumpkin-patch*, i.e., *www.yoursite/com/tag/pumpkin-patch*

Further, if you categorize such a post with more than just this one category, by also categorizing it under say, *Family Fun* and *Attractions*, your post has 3 degrees of improved sitemap semantics.

LODE Taxonomy

I serialized this work in RDF/RDFs using Protégé. Do not worry about format; transcribing it to something else is trivial work, as I am a software engineer by trade. In any case, we can create a legitimate static HTML site about the taxonomy using PyLODE.

This static site could be hosted as its own URL or under your parent site. This doesn't buy you anything directly save for scientific recognition. You would have to manifest this taxonomy into your Wordpress hierarchy, essentially maintaining parity between the RDF/RDFS "canonical" serialization and what your authors see.

This taxonomy can be updated and versioned in future iterations with GitHub, which I can show you how to work with.

We aren't yet an ontology in full, as we have no relations between the categories. But, some things are already possible, such as "related category", as, for example, the *Dining* category, under *Things to Do,* is often paired with the *Recipes* category, under *Home*.

Considerations

The process for categorizing a post should be a tree traversal, from top to bottom. Keep the taxonomy in mind, or pull up the image. For some post, say https://tampabayparenting.com/summer-camp-lunch-ideas/, the taxonomy should be run down in order until all categories have been exhausted, and 3-5 good ones picked. For example:

- Animals category, pets, irrelevant
- Health, possibly relevant (this category needs some more expansion)
- Home, kitchen, recipes, GOOD
- Home, parenting, elementary years, GOOD
- Things to do, dining, lunch, GOOD
- Things to do, travel, locations, camps, summer camp, GOOD
- And so on…

I have future plans on enforcing a few tag "bins" to keep things slightly more organized and re-usable. For instance, there are many, many tags about particular types of food, e.g., ribs, steaks, kombucha, mimosas, chocolate, etc., that are too infrequently used and too specific to consider categories; *but*, it is sensible to "bin" these tags under the category, *Dining*, i.e., *Food*. These tag "bins" can be pulled from the taxonomy I've created as needed. So, instead of having tags like:

- www.yoursite.com/tag/chocolate
- www.yoursite.com/tag/hot-yoga
- www.yoursite.com/tag/asian-noodles

Your sitemaps would retain some organization, and your authors would be less likely to create new tags, with:

- www.yoursite.com/tag/food/chocolate
- www.yoursite.com/tag/food/asian-noodles
- www.yoursite.com/tag/fitness/hot-yoga

It remains to be seen, but I think keeping the tag "bins" just 1-deep, and not a full taxonomy like your categories, makes more sense. In any case, determining what "bins" are needed will come from the taxonomy itself and the precedence of your posts. Food is definitely one, but there are others.

Conclusion

Summarizing, this was a very successful preliminary investigation into the project. In reality, despite the seeming diversity of the domain, it is rather restricted, and easily extended, given the taxonomy I have developed. I still intend to finish integrating the remaining tags/categories from your spreadsheets with the taxonomy, and also working to make the tag "bins" for the too-specific terms. I also want to cross-reference/integrate with your competitors tags, ThatsSoTampa and TampaMagazine. That might be a worthwhile investigation.

At your service,

Tyler

Semantic Science

- Ontology Engineering & Research Publication -

www.semanticscience.us