README.md

Eduard Jurášek - Internet slang vocabulary (topic 5)

https://github.com/ejurasek00/isv

The proposed seminar project aims to implement graph database of the contemporary Internet slang vocabulary represented in RDF. By internet slang, I refer words, phrases or acronyms frequently used online and/or in casual speech as well.

These words have complex relationships and linguistic backstories. The graph database is planned not only as a "word" – "meaning" principle, but include: -Definition -Origin -Example in sentence -Community usage

This project aims to make a Urban Dictionary - like database represented as a graph, but with additional perks mentioned above. A real-world usage is to make a web UI layer and plot the results in an appealing form directly in a web browser (not yet implemented).

The development was made in a VSCode, Protege and Fuseki on a Kubuntu Linux 25.10 and 26.04 (Beta preview)

Tip: Install the "Turtle Language Server" extension in VSCode from https://marketplace.visualstudio.com/items?itemName=stardog-union.vscode-langserver-turtle to get the code highlighted and to get the syntax checked!

Important: Fuseki requires Java to be installed!

Legal Notice

Some terms and definitions used in the ontology are obtained from external sources for **scientific and educational purposes** within the meaning of Sections 31 and 34 of Copyright Act No. 121/2000 Coll. (Czech Republic). When further distributing or commercially using this ontology, please note that these specific text segments (primarily from Urban Dictionary, Know Your Meme, and similar commercial websites) may be subject to the copyright of the original sources.

Words

Around 120 words are covered in .ttl

Bait, Troll, b4, 2day, brb, OMG, cmng, idc, LOL, Only in Ohio, Floptropica, Wakanda, Covfefe, Rizz, Bussin, Cheugy, Tea, Blaccent, Finsta, Slaycation, Skibidi, AMA, HP, gg, glhf, DPS, opp, Karen, NPC, poggers, u, ngl, af, fw, ts, ur, rn, fr, tf, Cooked, pfp, Zaddy, Goated, Pookie, Twin, Honse, Car, Type beat, Scrumpt, Offed, tbr, atp, GTA, gd, ikr, Vibe, smh, yt, Poser, Gatekeeper, Based, mb, Rule 34, Ragebait, 6-7, 69, 420, Rule 621, 21, 911, cap, rly, /s, Gymmaxing, DYEL, op, Frogposter, AI slop, r/, omegalul, monkaS, Sadge, Sus, Amogus, Impostor, Copium, pls, Selfie, Gigachad, Sigma, UwU, OwO, :3, bnuy, Birb, Yeet, Jet2Holiday, Grubhub, Bing chilling, Clanker, W/L, FL, Mid, Peak, Flop, Mog, Soft launch, Stand, Simp, Delulu, Ratio, Main Character Energy, Fanum Tax, Gyatt, Mewing, Brainrot

The Definition and the Example in Sentence were used from https://www.urbandictionary.com/ , https://knowyourmeme.com/ , https://en.wiktionary.org/, https://www.dictionary.com/culture/slang/ and others. Some parts were edited or censored.

Using already existing ontologies

`@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .` Used for `is a shomething`, eg. `Bait a isv:SlangTerm`.

`@prefix skos: <http://www.w3.org/2004/02/skos/core#> .` This is a backbone of the project and helps to describe the concept, definition and example of usage for each SlangTerm.

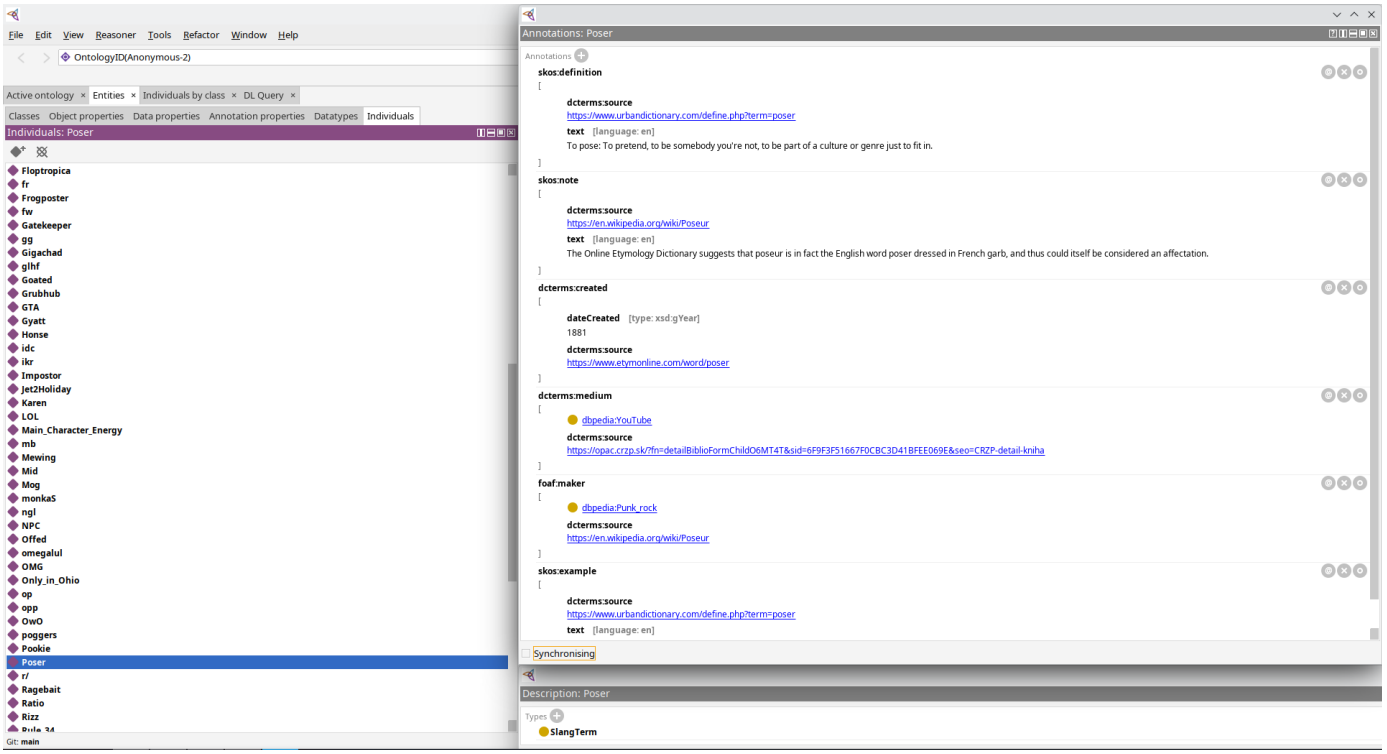
`@prefix foaf: <http://xmlns.com/foaf/0.1/> .` `foaf:maker` declares who is responsible for the origin of the SlangTerm.

`@prefix dcterms:<http://purl.org/dc/terms/> .` `dcterms:created` is used to state when was the SlangTermrd created. dcterms:medium states which communities or sicial media use this SlangTerm.

`@prefix dbr: <http://dbpedia.org/resource/> .` `dbr:something` connects the SlangTerm to a link on DBPedia where more information about the topic can be found if someone is not familiar in the field.

`@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .` `dcterms:created "YEAR"^^xsd:gYear`; assign year or creation of the SlangTerm as a date format.

`@prefix schema: <https://schema.org/> .` Used to delcare first occurance of the SlangTerm using `schema:dateCreated`, definition and example of the SlangTerm using `schema:text`.



Using own NON-ontology prefixes

These websites or documents were used as a source of information for this very project, alongside with other sources stated for each SlangTerm separately. In some cases, however, full IRIs had to be used because of special characters in the adress.

`@prefix kym: <https://knowyourmeme.com/memes/sites/>.` Know your meme.

`@prefix wtnr: <https://en.wiktionary.org/wiki/>.` Urban dictionary.

`@prefix riecky: <https://opac.crzp.sk/?fn=detailBiblioFormChildO6MT4T&sid=6F9F3F51667F0CBC3D41BFEE069E&seo=CRZP-detail-kniha>.` Master Thesis of Mgr. Jakub Riecky.

`@prefix ud: <https://www.urbandictionary.com/define.php?term=>.` Urban Dictionary.

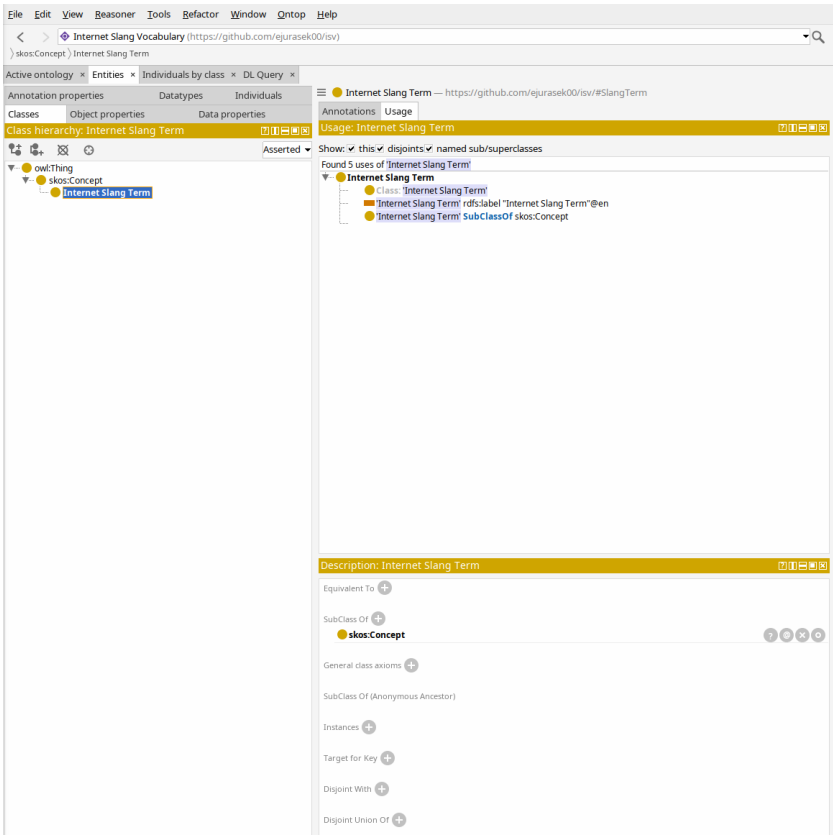
`@prefix ds: <https://www.dictionary.com/culture/slang/>.` Dictionary.com

Using own ontology

I named this ontology as ISV - Internet Slang Vocabulary.

`@prefix isv: <https://github.com/ejurasek00/isv/#> .` (Custom) – This ontology defines the specific class isv:SlangTerm and project metadata.

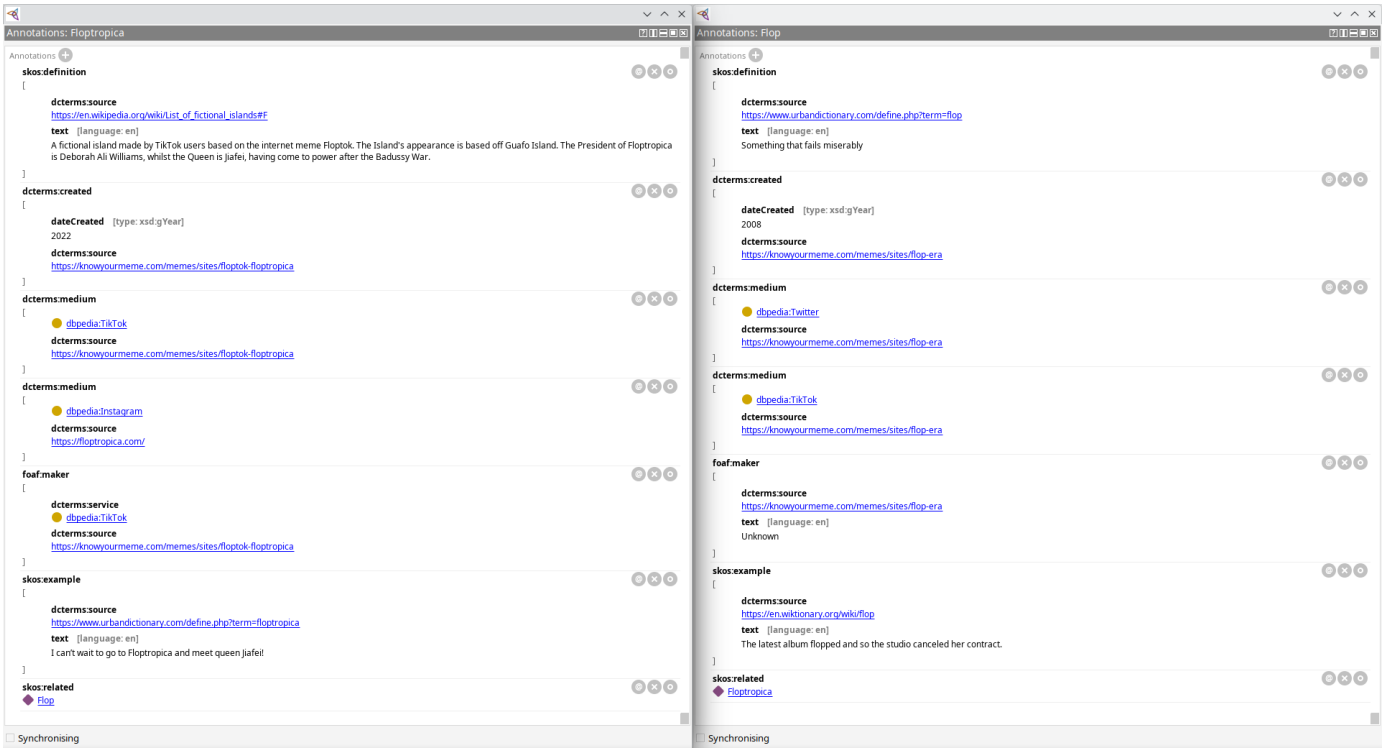
isv:SlangTerm is a subclass of skos:Concept which is a subclass of owl:Thing.



Related / Interconnected SlangTerms

Some SlangTerms in isv: are related to each other. This is declared by `skos:related`.

Currently implemented: OwO <-> UwU <-> :3 AND Floptropica <-> Flop AND Sus <-> Impostor <-> Amogus.



dcterms:medium and foaf:maker. What is the difference here?

`foaf:maker` is the website, location, person,... that brought this SlangTerm to life. Long story short, whom can we say thank you for making this word exist.

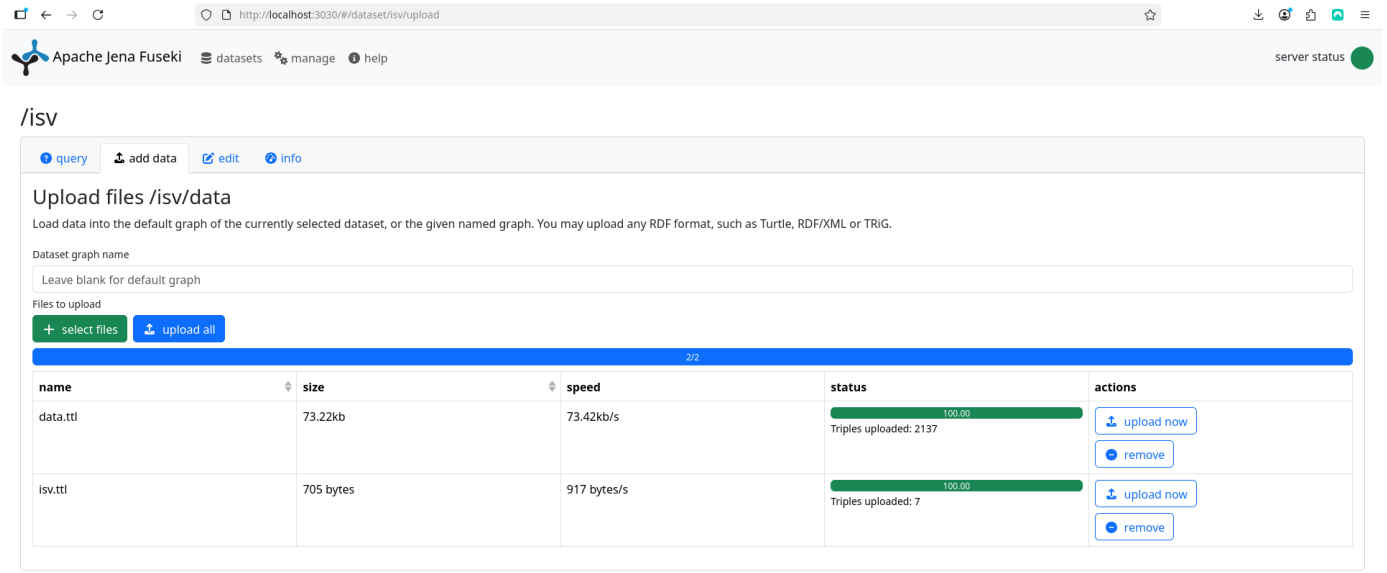
`dcterms:medium` represents the medium or group/community where can we come across this SlangTerm. In another words, who USES this SlangTerm in their lingo.

Apache Jena Fuseki

For the query part, Apache Jena Fuseki 5.6.0 on Kubuntu Linux 24.04.2 LTS was used. I used the Web UI in Firefox using the port 3030. http://localhost:3030/

While adding data to the dataset, both data.ttl and isv.ttl are needed to be uploaded.

2137 tripples were uploaded in data.ttl and 7 triples in isv.ttl



Queries

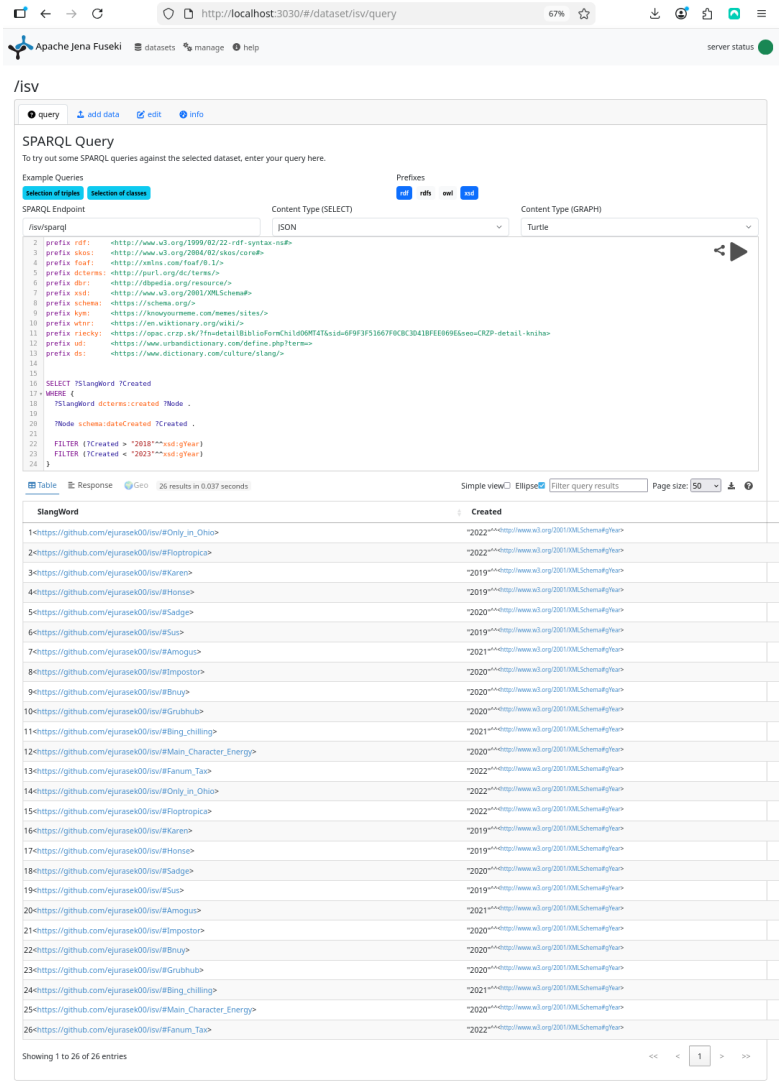
Two queries were executed in Apache Jena Fuseki.

Query 1 - Covid-19

In the first one, I was interested which of the Concepts were created during the Covid-19 Pandemic. To get the results, two filters were used: year of creation is greater than 2018 and smaller than 2023.

The query is listed below:

```
prefix isv:     <https://github.com/ejurasek00/isv/#>
prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix skos:    <http://www.w3.org/2004/02/skos/core#>
prefix foaf:    <http://xmlns.com/foaf/0.1/>
prefix dcterms: <http://purl.org/dc/terms/>
prefix dbr:     <http://dbpedia.org/resource/>
prefix xsd:     <http://www.w3.org/2001/XMLSchema#>
prefix schema:  <https://schema.org/>
prefix kym:     <https://knowyourmeme.com/memes/sites/>
prefix wtnr:    <https://en.wiktionary.org/wiki/>
prefix riecky:  <https://opac.crzp.sk/?fn=detailBiblioFormChildO6MT4T&sid=6F9F3F51667F0CBC3D41BFEE069E&seo=CRZP-detail-kniha>
prefix ud:      <https://www.urbandictionary.com/define.php?term=>
prefix ds:      <https://www.dictionary.com/culture/slang/>


SELECT ?SlangWord ?Created
WHERE {
  ?SlangWord dcterms:created ?Node .

  ?Node schema:dateCreated ?Created .

  FILTER (?Created > "2018"^^xsd:gYear)
  FILTER (?Created < "2023"^^xsd:gYear)
}
```



26 results in 0.037 seconds

Query 2 - Makers

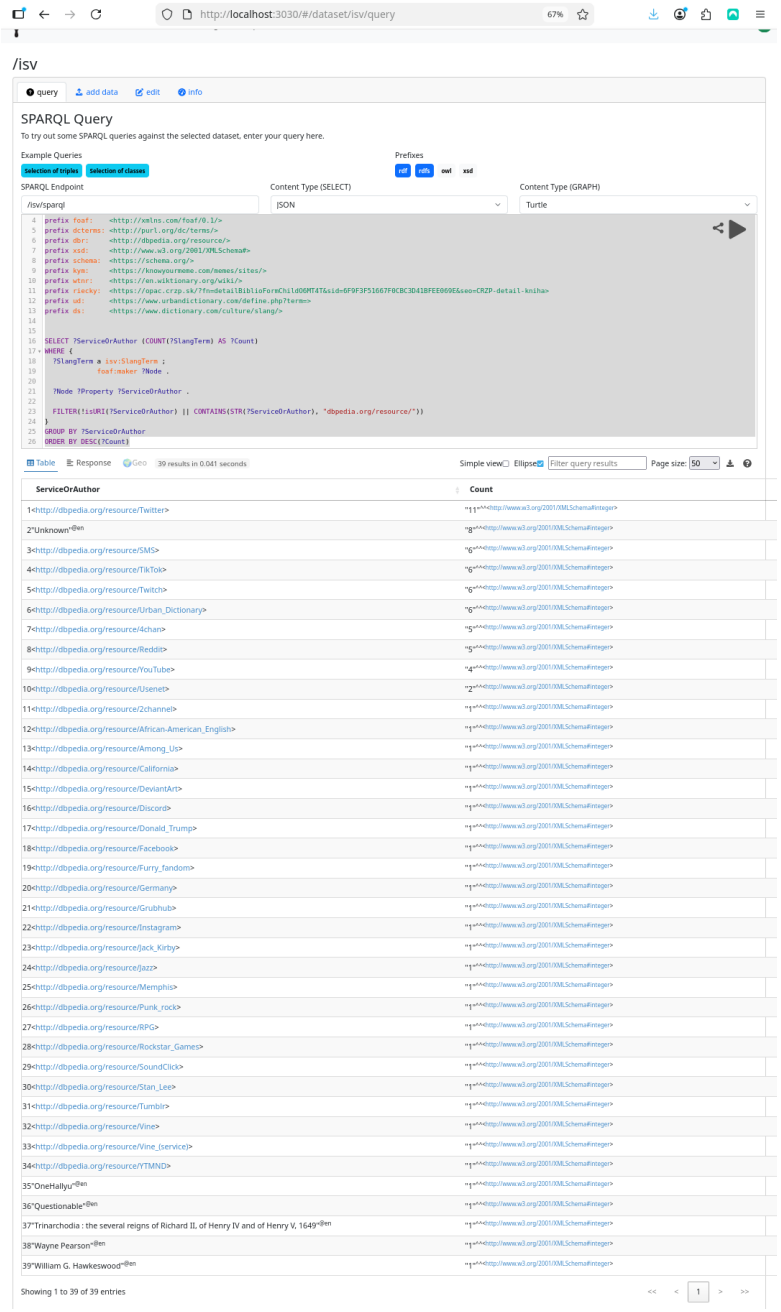
In the second one one, I was interested in how many SlangTerms each maker has. Another words, I wanted to see the biggest makers first. For this, the query counted amount of occurances for each maker.

The query is listed below:

```
prefix isv:     <https://github.com/ejurasek00/isv/#>
prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix skos:    <http://www.w3.org/2004/02/skos/core#>
prefix foaf:    <http://xmlns.com/foaf/0.1/>
prefix dcterms: <http://purl.org/dc/terms/>
prefix dbr:     <http://dbpedia.org/resource/>
prefix xsd:     <http://www.w3.org/2001/XMLSchema#>
prefix schema:  <https://schema.org/>
prefix kym:     <https://knowyourmeme.com/memes/sites/>
prefix wtnr:    <https://en.wiktionary.org/wiki/>
prefix riecky:  <https://opac.crzp.sk/?fn=detailBiblioFormChildO6MT4T&sid=6F9F3F51667F0CBC3D41BFEE069E&seo=CRZP-detail-kniha>
prefix ud:      <https://www.urbandictionary.com/define.php?term=>
prefix ds:      <https://www.dictionary.com/culture/slang/>


SELECT ?ServiceOrAuthor (COUNT(?SlangTerm) AS ?Count)
WHERE {
  ?SlangTerm a isv:SlangTerm ;
             foaf:maker ?Node .

  ?Node ?Property ?ServiceOrAuthor .

  FILTER(!isURI(?ServiceOrAuthor) || CONTAINS(STR(?ServiceOrAuthor), "dbpedia.org/resource/"))
}
GROUP BY ?ServiceOrAuthor
ORDER BY DESC(?Count)
```

This query found 39 results in 0.041 seconds

The biggest SlangTerm maker is Twitter (11x), followed by Unknown (8x), TikTok (6x), SMS (6x), UD (6x), 4cahn (5x), Reddit (5x), YouTube (4x), Usenet (2x) and the remaining ones.



Methodology

The selection of the SlangTerms was heavily inspired by the Internet vernacular - a study of neologisms in online discourse of cybercultures (Riecky, 2025). Additionally, I added some new SlangTerms and my personal favourites to make their count above 100.

The Definitions and Examples were primarily derived from UrbanDictionary. Origin and Creator was derived primarily from Know Your Meme.

However, not all words are covered by these “major” websites. Some of them are new, have unknown origin or are so niche that are difficult to find.

So, searching looked like “which social media use “offed” as slang” or “ikr acronym origin”

When no relevant results could be found, I used Perplexity, Gemini Pro and ChatGPT in a Deep research mode to find out the etymology and origin.

Notice

In some cases, the definition may not help to understand the meaning of the SlangWord. This may be caused by various factors, such as absurdity in usage or that the SlangTerm is used in a specific circumstances by specific groups.

In abbreviations, sometimes example can be more helpful to comprehend the meaning; however all of them are written as their long forms too and were explained what they mean.

Related Wors

The project was inspired by a masters thesis of a friend of mine. Thank you so much, Jakub!

Riecky, J. (2025). Internet vernacular - a study of neologisms in online discourse of cybercultures [Masters Thehis, Univerzita sv. Cyrila a Metoda v Trnave FF UCM KAA]. Centrálny register záverečných a kvalifikačných prác OPAC CRZP. https://opac.crzp.sk/?fn=detailBiblioFormChildM8GSN&sid=6F9F3F51667F0CBC3D41BFEE069E&seo=CRZP-detail-kniha

---

Support the Project

If you find this vocabulary helpful, consider buying me a coffee (or a beer, or a sauna entrance). I’d really appreciate the courtesy!

This project costed me:

- Lots of time during research

- Horse-doses of caffeine
- Some beers
- A USB drive purchase (thanks, *Planeo*, for being slow AF and making the staff stare at me in my tracksuit 3 minutes before closure in -3°C!).
- A full system reinstall (the reason I needed that USB in the first place)
- A train ticket
- Sauna sessions to recover my nerves
- Taking care of the cat

Thank you! <3

Remarks

Having any questions or suggestions on how to make this project better, do not hesitate to contact me via E-mail jure01@vse.cz