#WikiForHumanRights 2023

Right to a Healthy Environment

A Brief (or not so brief) introduction to Petscan

WIKIMEDIA FOUNDATION

We will be recording today!

We plan to record this presentation today.

If you say something aloud or are on camera, it will like be in the recording.

If you share something in the chat, that will not be shared in the recording

Today's **agenda**

- Reminder of why you might use Petscan
- Orientation to Simple Petscan
- Using it with other tools
- Orientation to Petscan with Wikidata

What is Petscan?

#WikiforHumanRights

WIKIMEDIA FOUNDATION

What is it exactly?

Petscan is a Magnus Manske tool that

- mixes different kinds of data to create topical lists
- succeeded several tools focused only on Categories
- Adds features like filtering with Wikidata, editing Wikidata, and editing Wikimedia Commons

Why it's useful for organizers:

- Allows for an auto-updating list based on the criteria you set (if something changes on the wikis, it stays up to date).
- Helps take a massive body of knowledge and focus it

WIKIMEDIA FOUNDATION

In the chat...

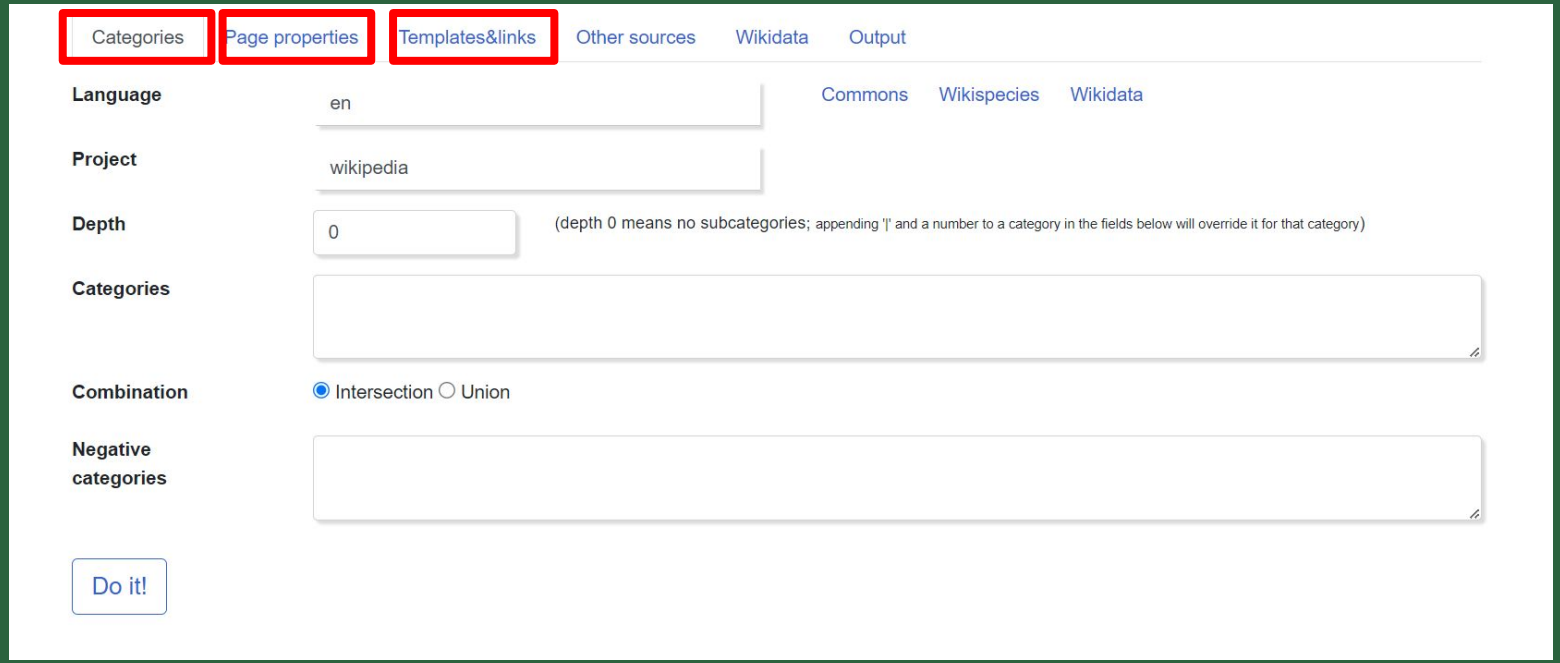
- Share examples of when you have seen Petscan in use?
- Why was it useful in that context?

WIKIMEDIA FOUNDATION

You probably **want to use Petscan when:**

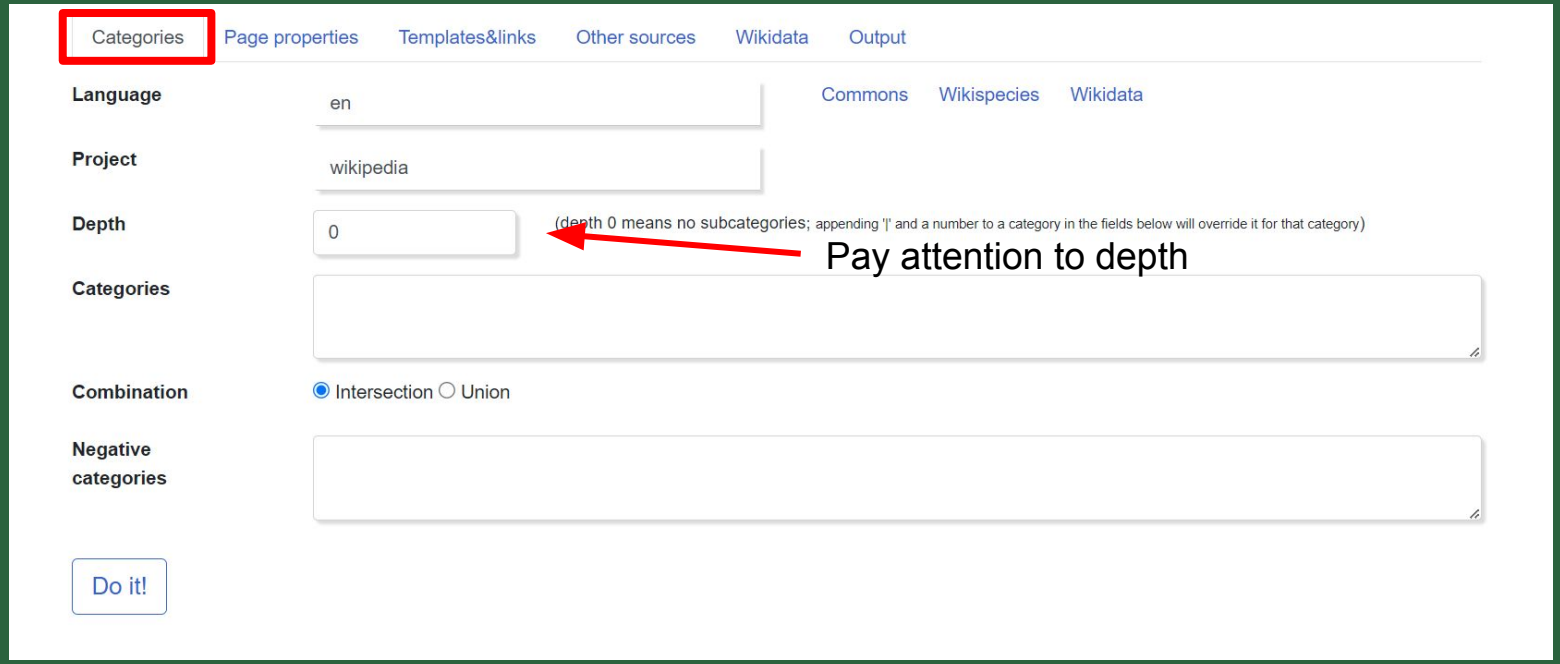
- You expect the topic list that you are building will change
- Other tools don’t meet your needs (see last week)
- You need to edit Wikidata, Commons or use other tools w/ Pagepile

Using PetScan simply...



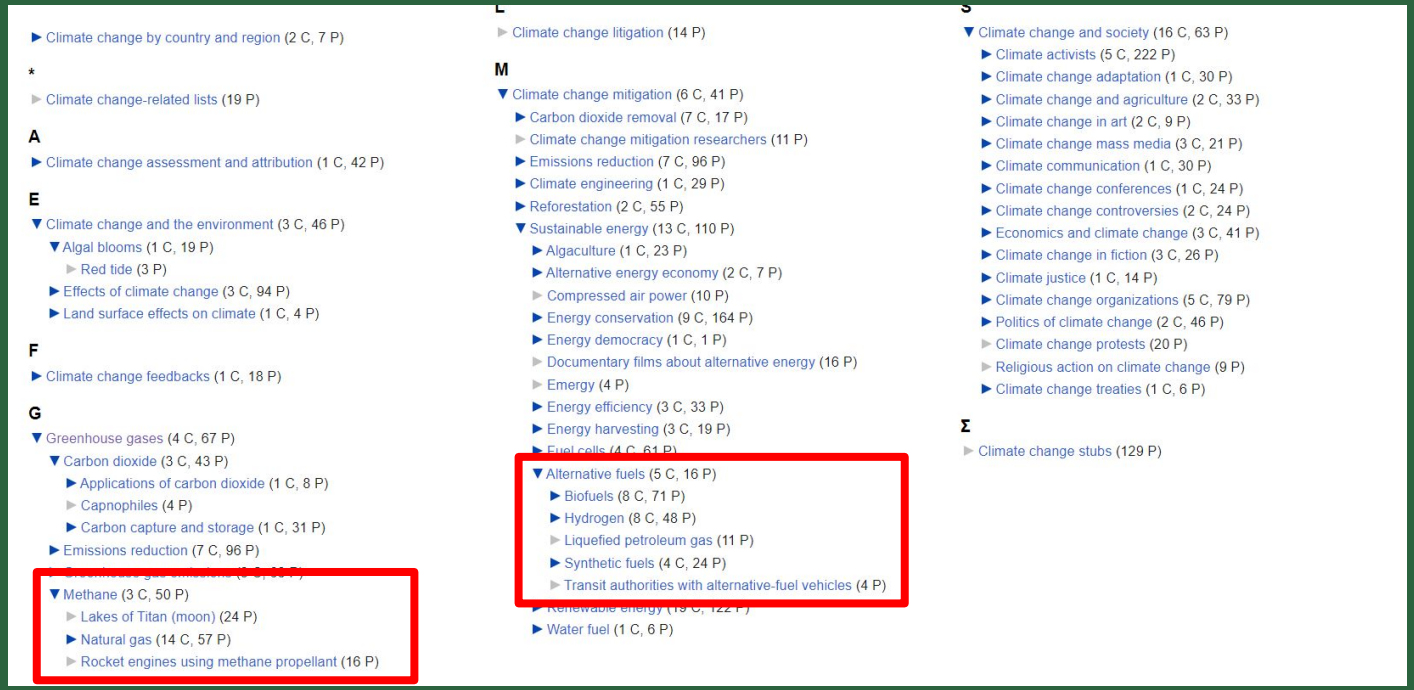
Let’s look at the first three tabs

#WikiforHumanRights



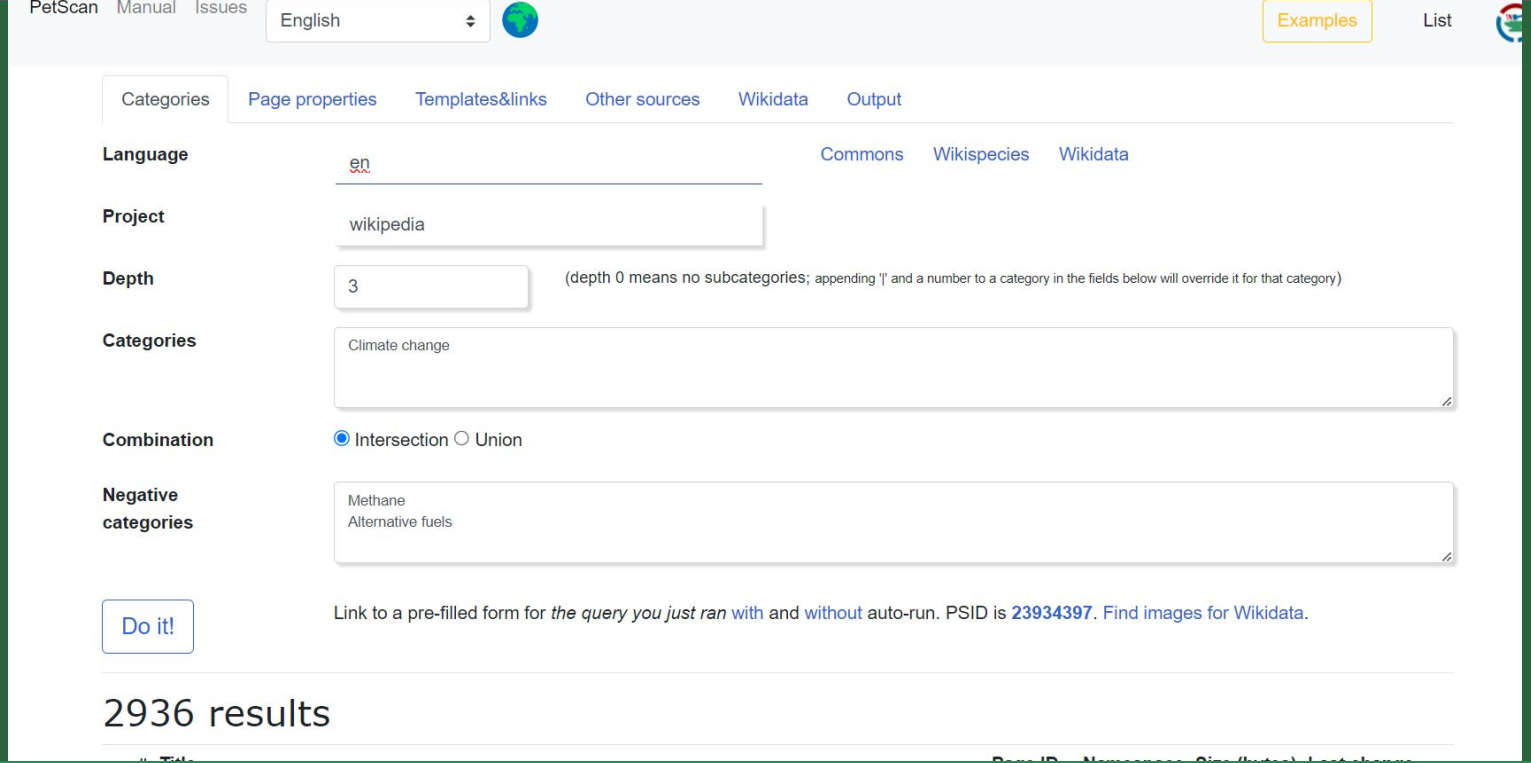
The first tab is focused on categories and is most intuitive

#WikiforHumanRights



For example, I probably want most of the articles in these categories, but... several subcategories introduce problems...

#WikiforHumanRights



The list is not bad.... https://petscan.wmflabs.org/?psid=23934397

#WikiforHumanRights

PetScan Manual Issues English Examples List

Categories | Page properties | Templates&links | Other sources | Wikidata | Output

Language en Commons Wikispecies Wikidata

Project wikipedia

Depth 3 (depth 0 means no subcategories; appending '|' and a number to a category in the fields below will override it for that category)

Categories Climate change

Combination Intersection Union

Negative categories Methane
Alternative fuels

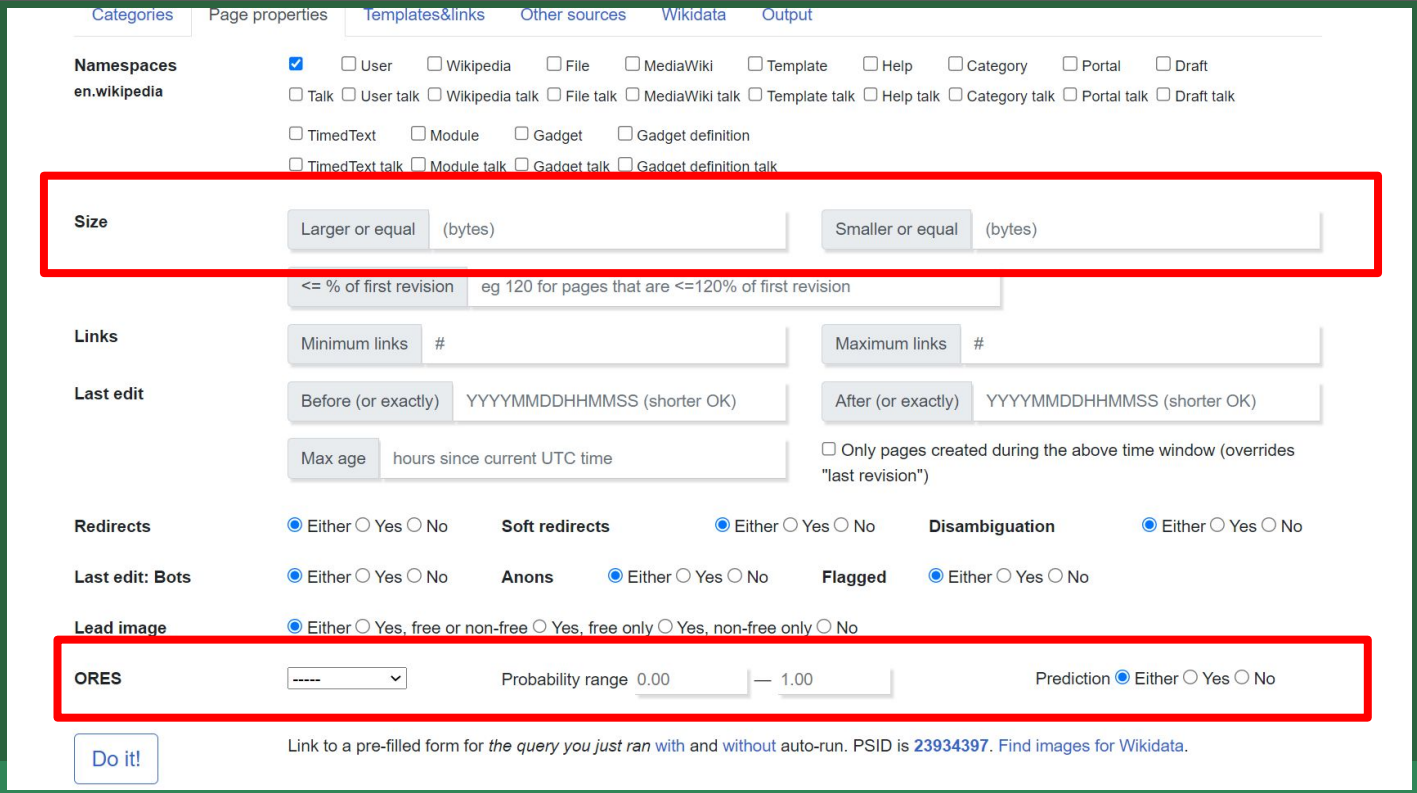
Do it!

Link to a pre-filled form for *the query you just ran* with and without auto-run. PSID is **23934397**. Find images for Wikidata.

2936 results

But let's try filtering it....

#WikiforHumanRights

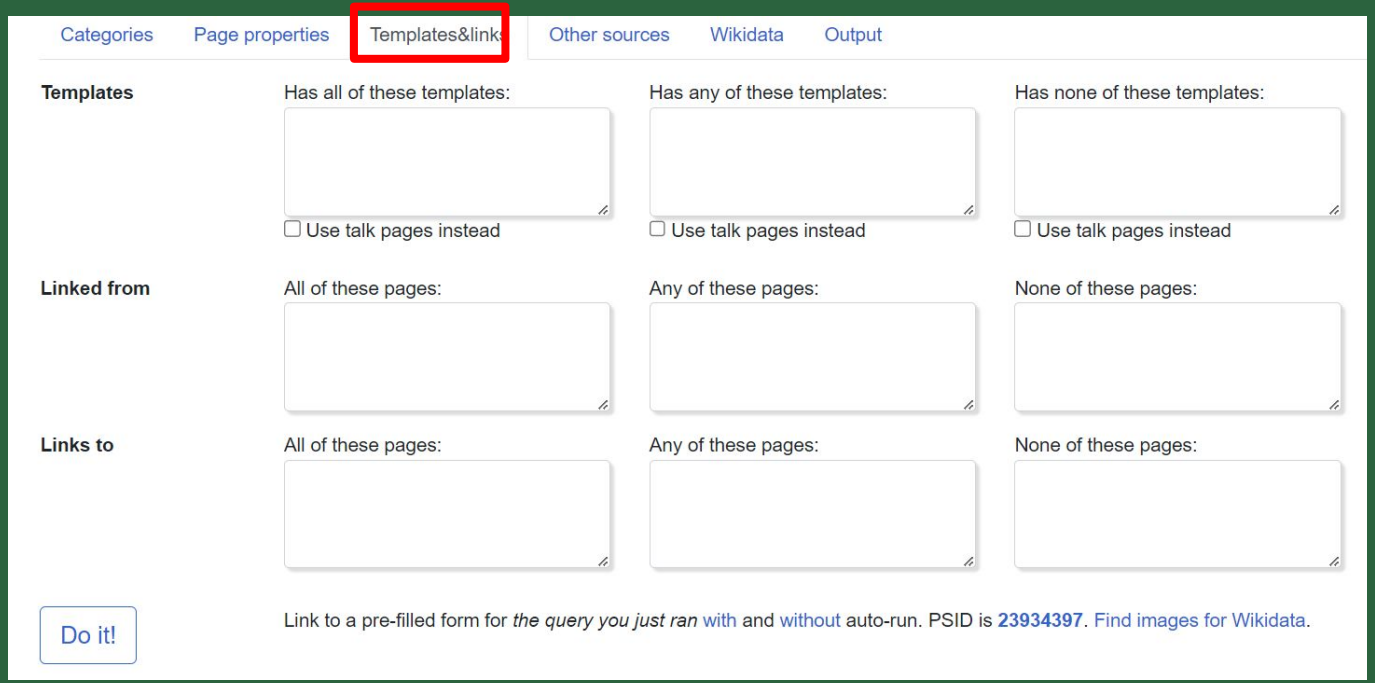


Bytes is a proxy for complexity and length of article -- lots of bytes implies more references

ORES uses other signals like references, other page features

Two important filters for organizers of writing events, others useful in other contexts....

#WikiforHumanRights



You should have a theory about what kinds of topics should be covered in the text.

For example, if I think all topics that talk about agriculture or climate smart agriculture should talk about deforestation....

Let's filter this query Templates/Links page https://petscan.wmflabs.org/?psid=23934397

#WikiforHumanRights

WIKIMEDIA FOUNDATION

Let me show you how I built another query about crops

https://petscan.wmflabs.org/?psid=23229381

Why did I want to build the list?

- Food choices are important for individual climate action
- Climate impacts are profound on food

So I needed to identify:

- Articles about food vulnerable to climate change (all of it)
- Remove any of the non-agriculture articles
- Make it easy to figure out what articles already talk about climate change

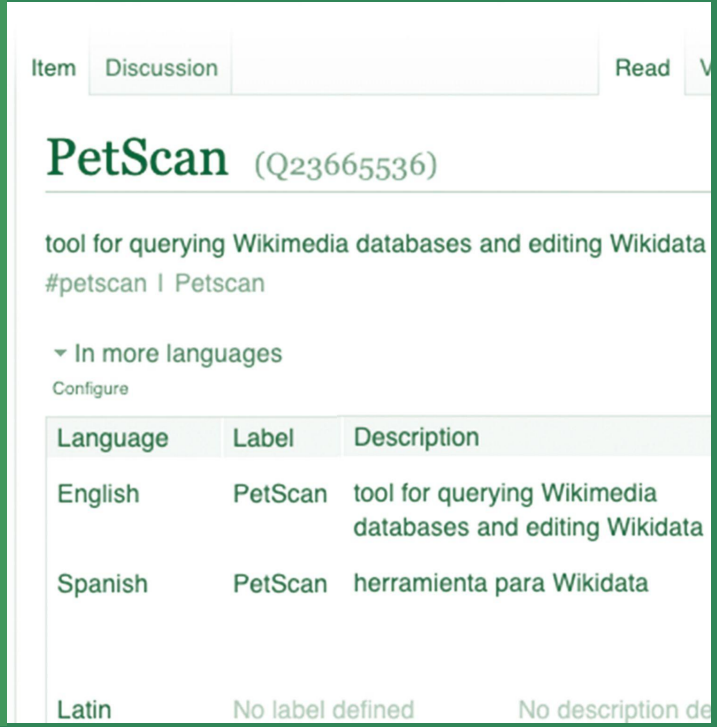
WIKIMEDIA FOUNDATION

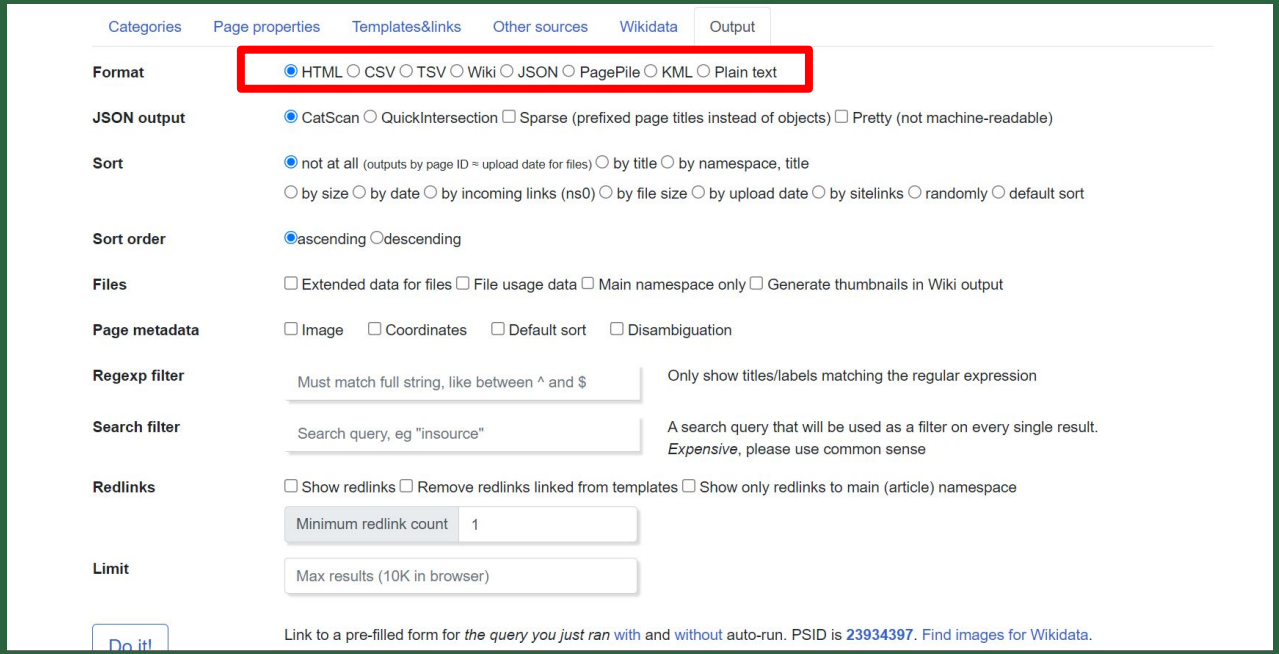
Want to clarify anything?

Now is the time to ask questions or share Comments about the first section of the presentation.

WIKIMEDIA FOUNDATION

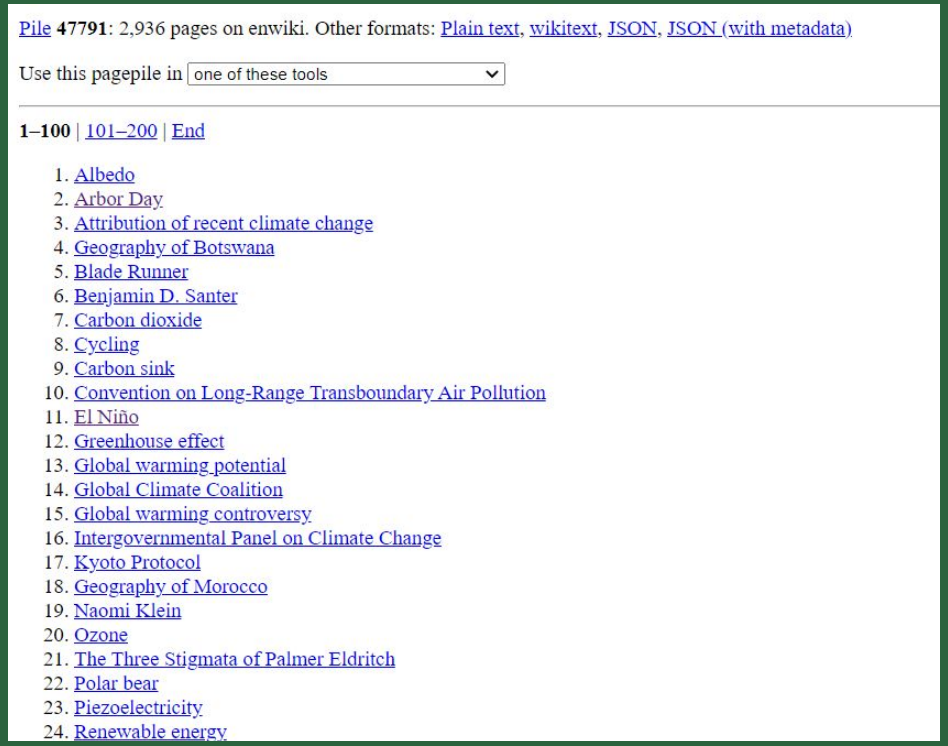
Outputs from Petscan





The 8 types of outputs are super useful, let me show you what they look like....

#WikiforHumanRights



Pile **47791**: 2,936 pages on enwiki. Other formats: Plain text, wikitext, JSON, JSON (with metadata)

Use this pagepile in one of these tools

1–100 | 101–200 | End

1. Albedo
2. Arbor Day
3. Attribution of recent climate change
4. Geography of Botswana
5. Blade Runner
6. Benjamin D. Santer
7. Carbon dioxide
8. Cycling
9. Carbon sink
10. Convention on Long-Range Transboundary Air Pollution
11. El Niño
12. Greenhouse effect
13. Global warming potential
14. Global Climate Coalition
15. Global warming controversy
16. Intergovernmental Panel on Climate Change
17. Kyoto Protocol
18. Geography of Morocco
19. Naomi Klein
20. Ozone
21. The Three Stigmata of Palmer Eldritch
22. Polar bear
23. Piezoelectricity
24. Renewable energy

For example, we can take the list to Massviews and learn which articles have the most pageviews (link)

Or can take it too tools like WDFIST to enrich information on Wikidata

Or if you convert the output to add Wikidata, edits via Tabernacle

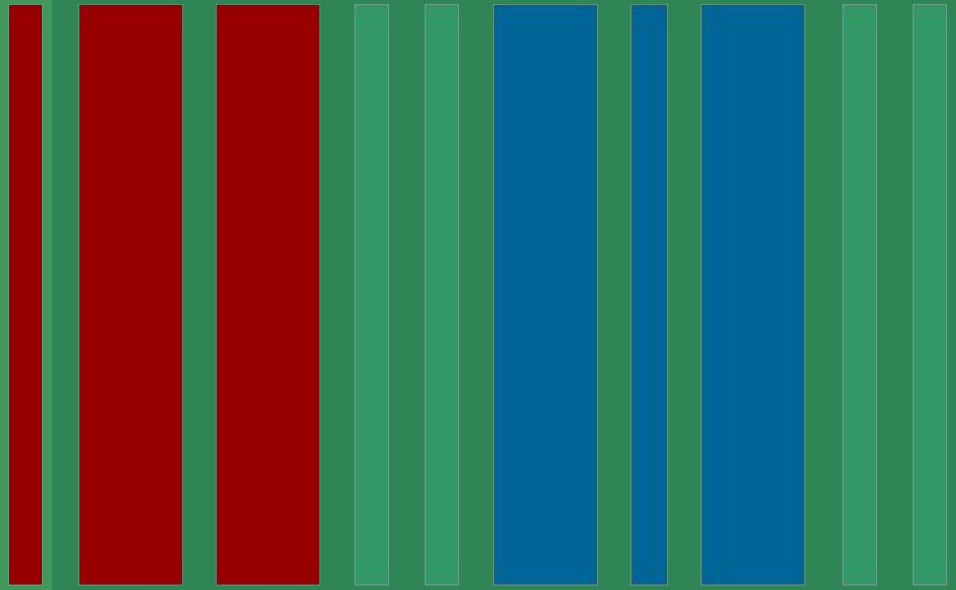
Particularly pagepile, which makes the list portable....

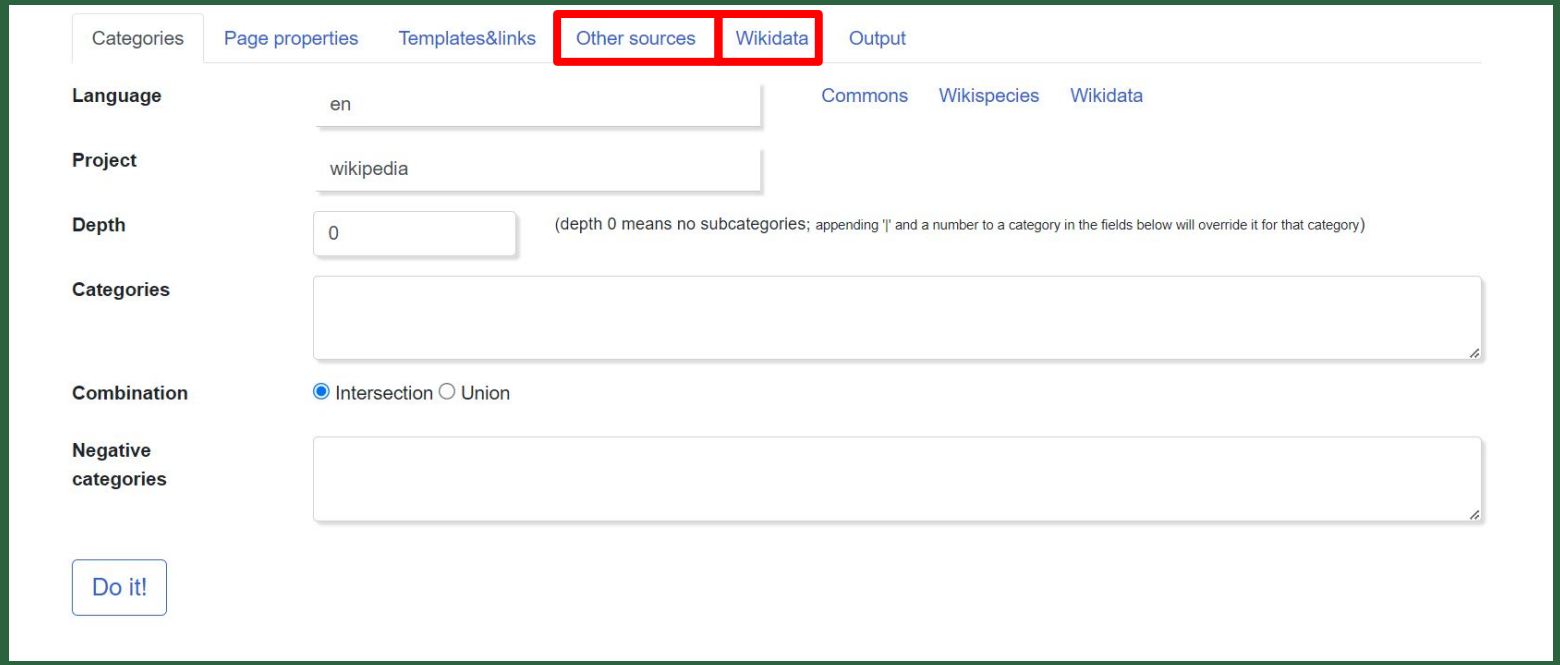
https://pagepile.toolforge.org/api.php?action=get_data&id=47791

#WikiforHumanRights

WIKIMEDIA FOUNDATION

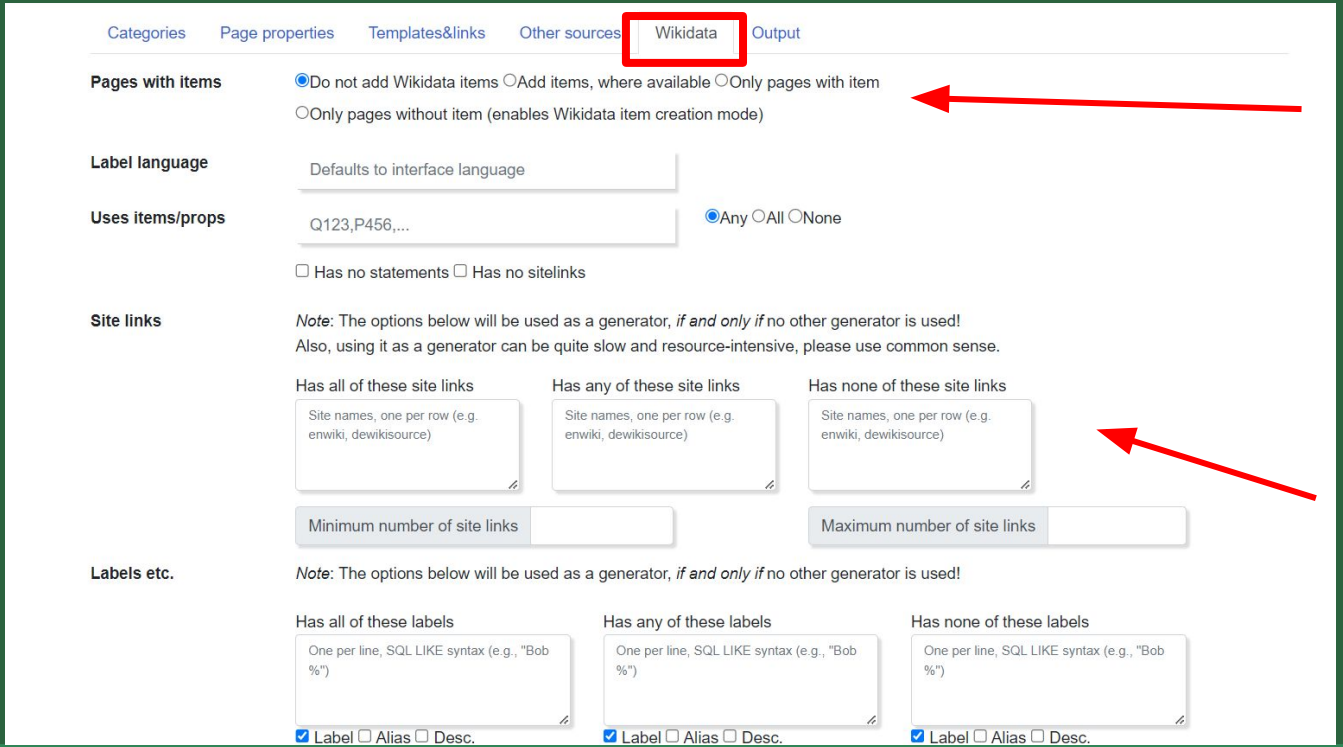
Using Petscan w/ Wikidata or more complex actions





We skipped two tabs... because they are mostly for Wikidata and Commons

#WikiforHumanRights

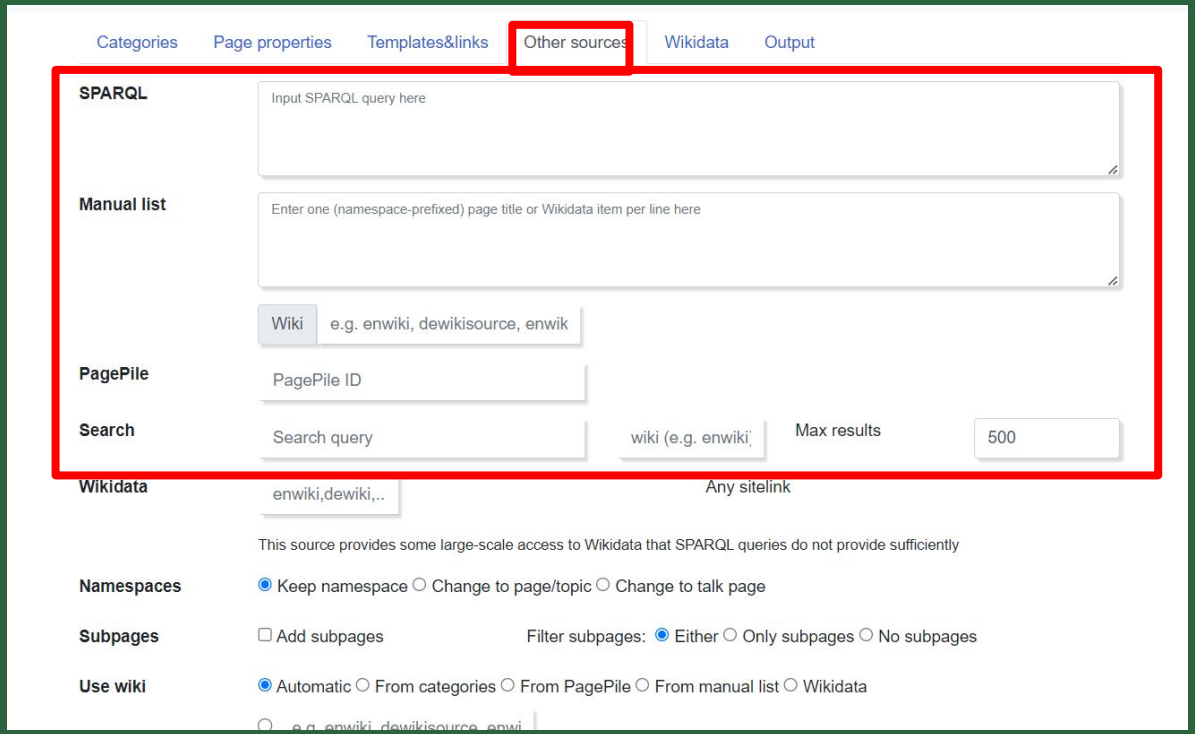


This first widget allows you to add the Wikidata output to a query -- important for most Wikidata based tools.

This section allows you to find out which wikis are missing for translation

Let's look at the Wikidata tab first

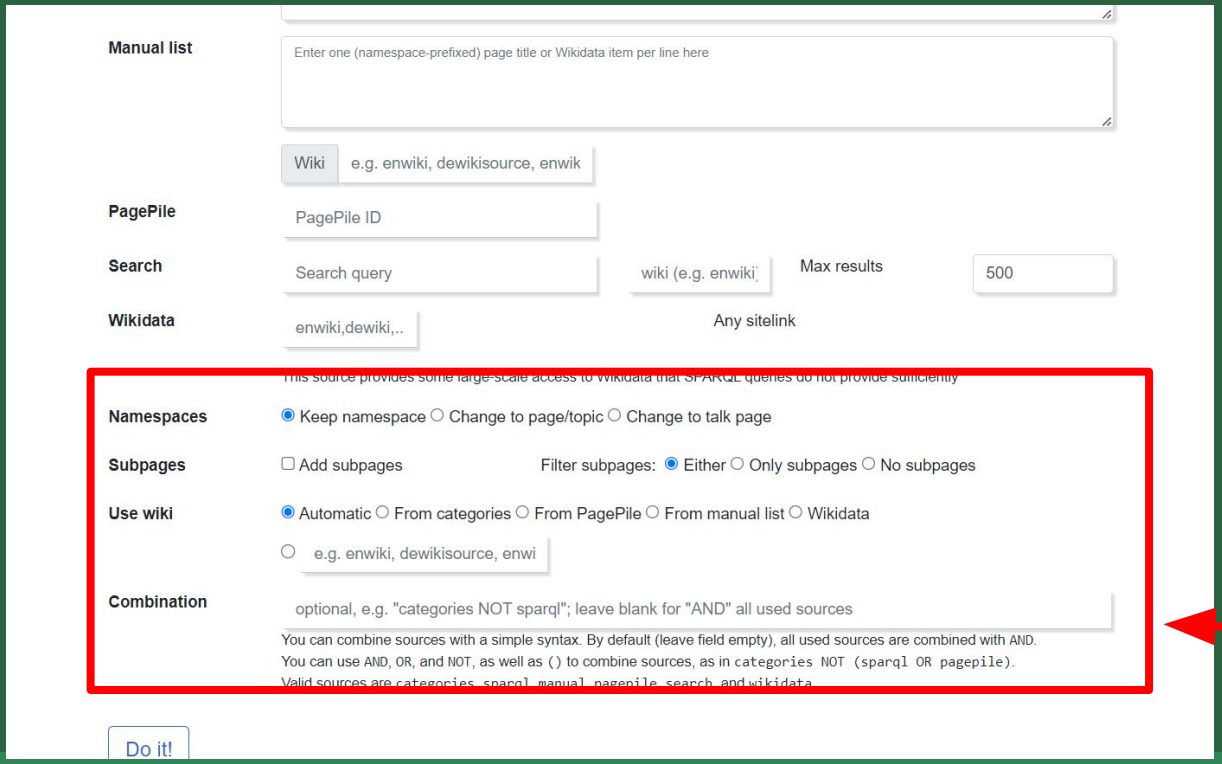
#WikiforHumanRights



This first section allows you to grab results from other queries (i.e. pagepile, search, the query service, etc). In my experience, I use SPARQL most, but all have been helpful.

The Other Sources tab has two types of functions

#WikiforHumanRights



This section allows you to change the output from lists

Notice that this combination feature allows you to use machine logic to combine results. I find this feature breaks a lot of my queries...

The Other Sources tab has two types of functions

#WikiforHumanRights

WIKIMEDIA FOUNDATION

Locations in Nigeria that aren't stubs without climate change links

Scenario:

- I am running editathons in Nigeria about impacts of climate change
- I want editors to write about climate change in city pages while improving articles that already exist

Here are the results:

- Main Query: https://petscan.wmflabs.org/?psid=23935851
- Minimum size of article 5000 bytes: https://petscan.wmflabs.org/?psid=23935901
- Pagepile
- Massviews of articles bigger than 5000 bytes

WIKIMEDIA FOUNDATION

Want me to clarify anything?

Now is the time to ask Questions or share Comments.

WIKIMEDIA FOUNDATION

We can use the rest of the time to try the tool and give you feedback/answer questions.