Please go to the app: https://purl.org/earthsystemdatasetdiscovery
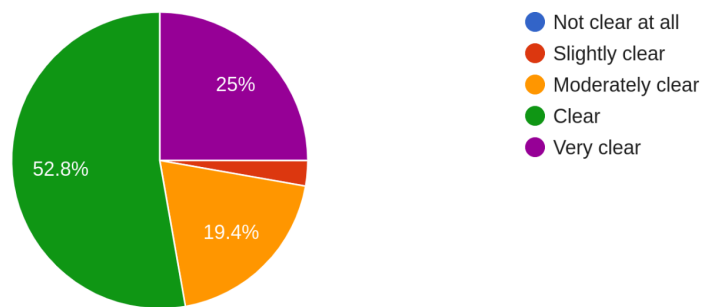
Step 1. **Enter** the keyword "Temperature" as the search term and click "retrieve datasets".

→ the retrieved datasets appear in the list at the bottom.
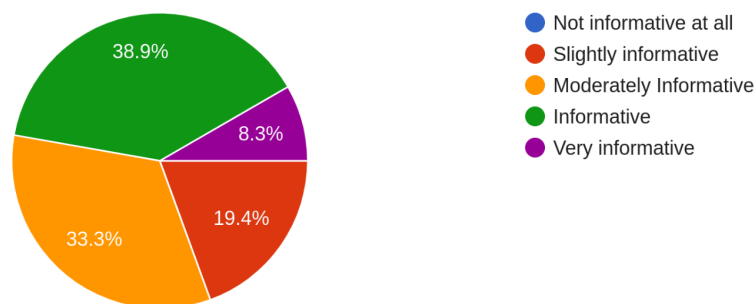
Questions about the retrieved datasets:

Q1. How clear do you find the provenance information (data hub) of retrieved datasets?
36 responses

- Not clear at all
- Slightly clear
- Moderately clear
- Clear
- Very clear

52.8%
25%
19.4%

Q2. How informative do you find the retrieved dataset list?
36 responses

- Not informative at all
- Slightly informative
- Moderately Informative
- Informative
- Very informative

38.9%
8.3%
19.4%
33.3%

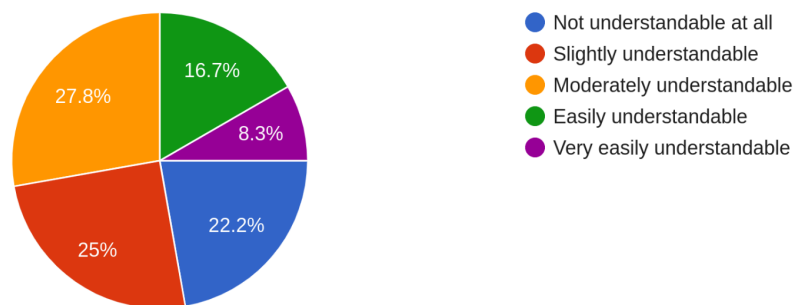Step2. **Click on** a specific dataset to have more details.

→ Relationships linking the search term to the retrieved dataset are shown as a **explanatory graph**.

→ Details of the selected dataset are displayed on the right.

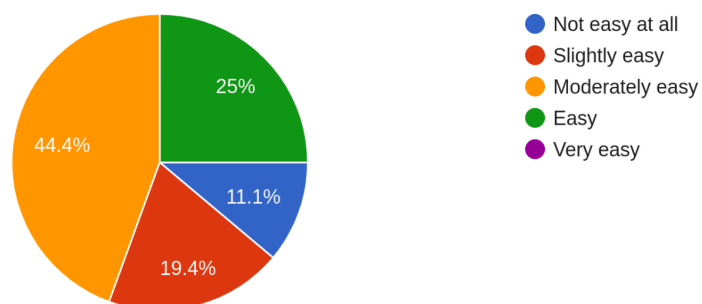**Questions about the explanatory graph:**

Q3. How understandable do you find the relationship between the search term and the retrieved datasets using the graph?

36 responses

- Not understandable at all
- Slightly understandable
- Moderately understandable
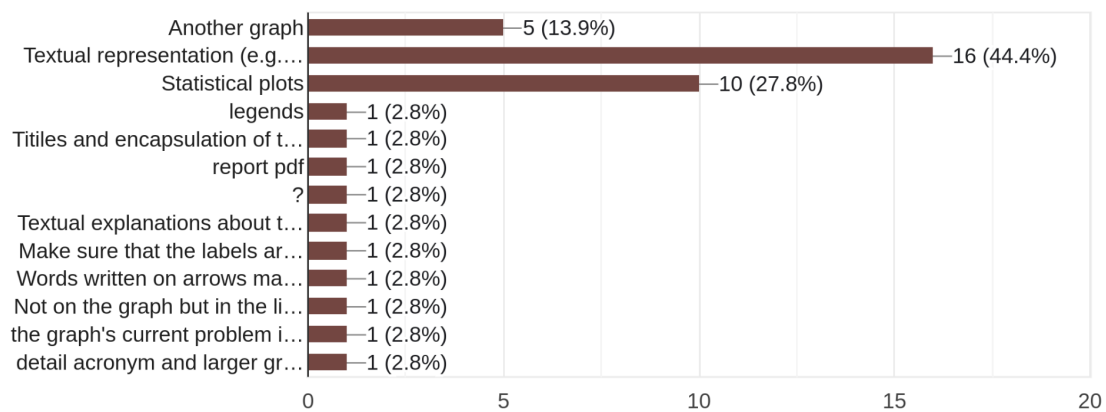- Easily understandable
- Very easily understandable

16.7%
27.8%
8.3%
25%
22.2%

Q4. How easy is it to assess the relevance of the retrieved datasets according to the search term?

36 responses

- Not easy at all
- Slightly easy
- Moderately easy
- Easy
- Very easy

25%
44.4%
11.1%
19.4%

Q5. Would you include an additional explanation for the graph? (multiple choice)

36 responses

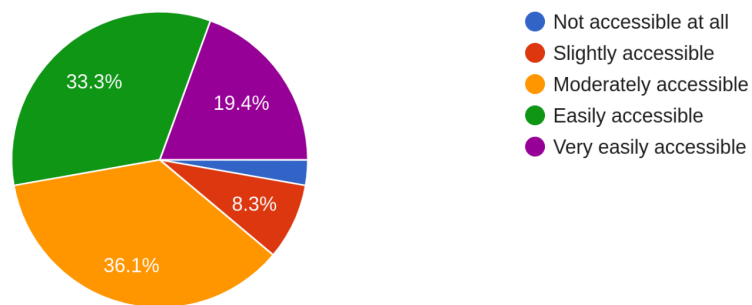| Response	Count
Another graph	5 (13.9%)
Textual representation (e.g….	16 (44.4%)
Statistical plots	10 (27.8%)
legends	1 (2.8%)
Titiles and encapsulation of t…	1 (2.8%)
report pdf	1 (2.8%)
?	1 (2.8%)
Textual explanations about t…	1 (2.8%)
Make sure that the labels ar…	1 (2.8%)
Words written on arrows ma…	1 (2.8%)
Not on the graph but in the li…	1 (2.8%)
the graph's current problem i…	1 (2.8%)
detail acronym and larger gr…	1 (2.8%)

Axis values: 0, 5, 10, 15, 20

Questions about the dataset details:

Q6. How accessible are the details of a particular dataset?

36 responses

- 🔵 Not accessible at all
- 🔴 Slightly accessible
- 🟠 Moderately accessible
- 🟢 Easily accessible
- 🟣 Very easily accessible

Pie chart values: 33.3%, 19.4%, 8.3%, 36.1%

Q7. How informative are the details provided about the selected dataset?
36 responses

- Not informative at all
- Slightly informative
- Moderately Informative
- Informative
- Very informative

55.6%
8.3%
13.9%
19.4%
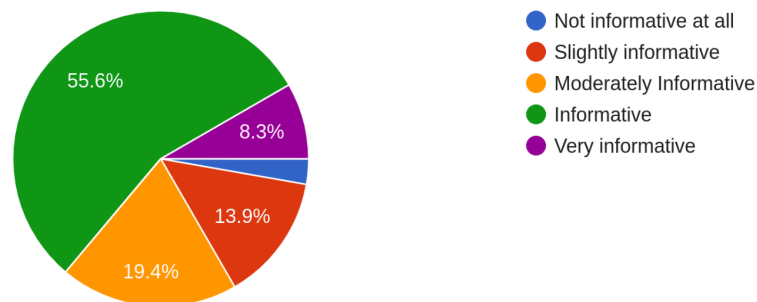
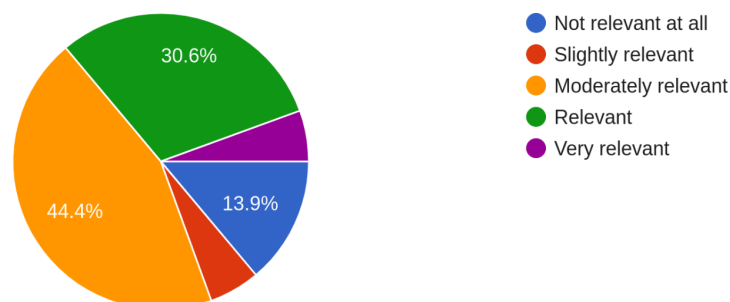## Step 3. 3.a) Extend AGROVOC external vocabulary,

3.b) Enable narrower and

3.c) Click on retrieve datasets.

→ The list of datasets matching linked concepts is displayed below

Q8. How relevant are the retrieved datasets compared to the initial search?
36 responses

- Not relevant at all
- Slightly relevant
- Moderately relevant
- Relevant
- Very relevant

30.6%
44.4%
13.9%

Scenario 2: User-defined Keyword search.

Please go to or refresh (Ctrl+r) the page:

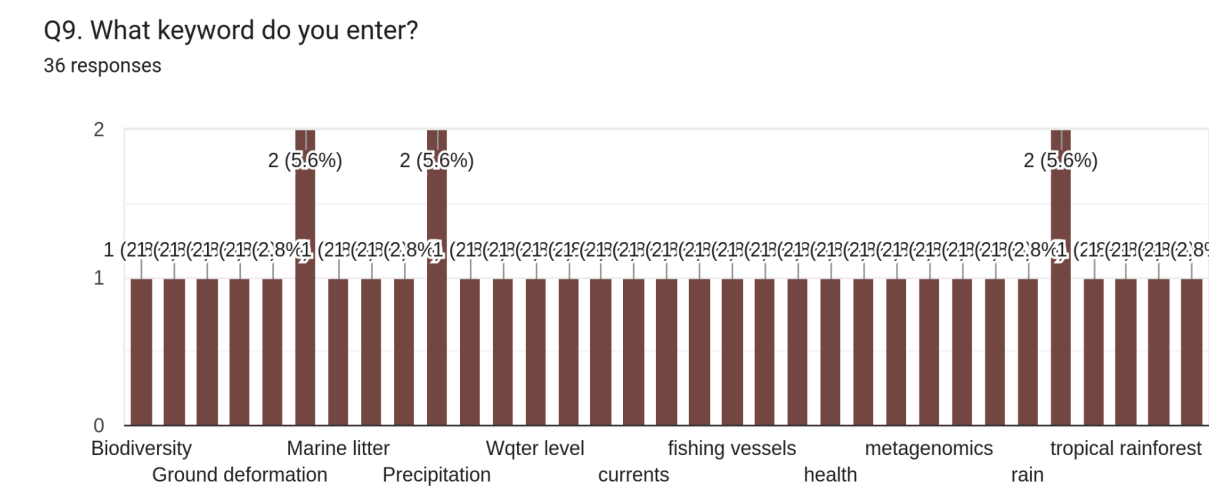## https://purl.org/earthsystemdatasetdiscovery

In the context of Earth System Dataset Search, follow this user experience:

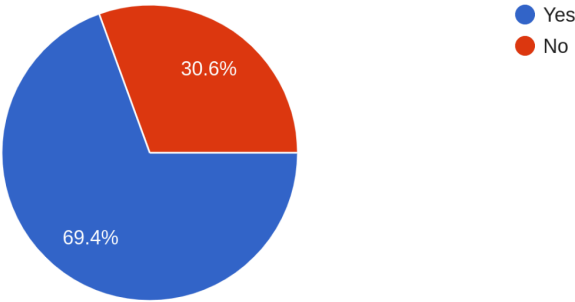**Enter a keyword that may be of interest in your domain or work.**

Explore the retrieved datasets

Enable (narrower/broader) linked concepts to the vocabularies of your choice

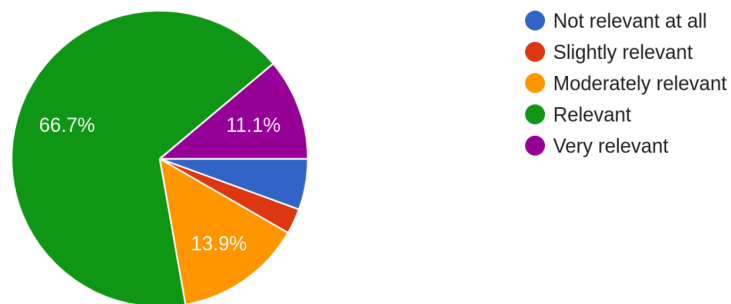Q9. What keyword do you enter?

36 responses

Q10. Did the retrieved datasets come from different data hubs?

36 responses

Q11. Are the retrieved datasets relevant to the search term?
36 responses

- Not relevant at all
- Slightly relevant
- Moderately relevant
- Relevant
- Very relevant

Q12. Did the exploration, enabling external vocabularies, help you to find additional relevant datasets?
36 responses

- Not relevant at all
- Slightly relevant
- Moderately relevant
- Relevant
- Very relevant
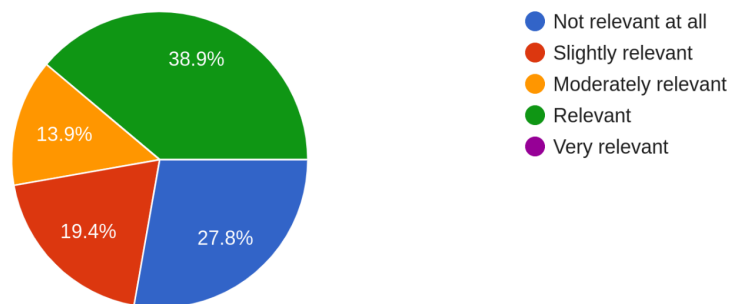
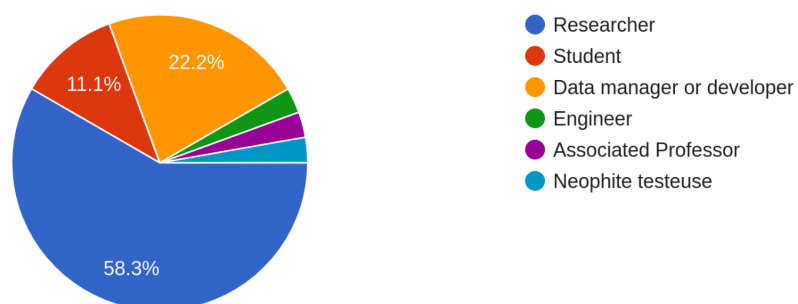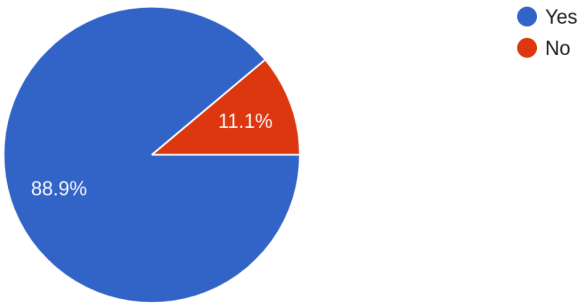## User Profile Information

Q13. What is your current status?
36 responses

- Researcher
- Student
- Data manager or developer
- Engineer
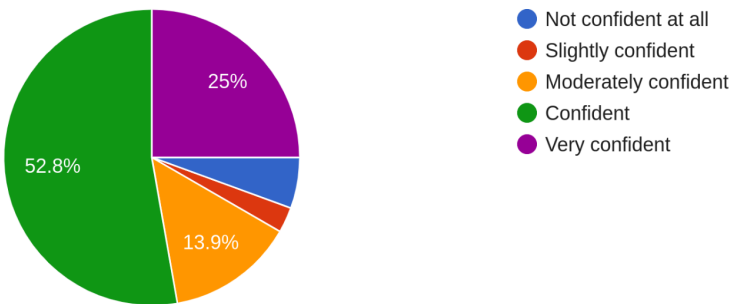- Associated Professor
- Neophite testeuse

Q14. Do you have a scientific domain of expertise?

36 responses

- Yes
- No

11.1%

88.9%

Q15. Could you please enter your domains of expertise separated by a comma?

36 responses

2 (5.6%)

1 (2.8%)

1 (2.8%)

| , | Biostatistics, epi… | GISc; Urban ana… | Macroecology, B… | Remote Sensing | Software engine… | artificial intellige… | data science, m… | energy modelling | lung physiology | remote sensing | zoonosi… |
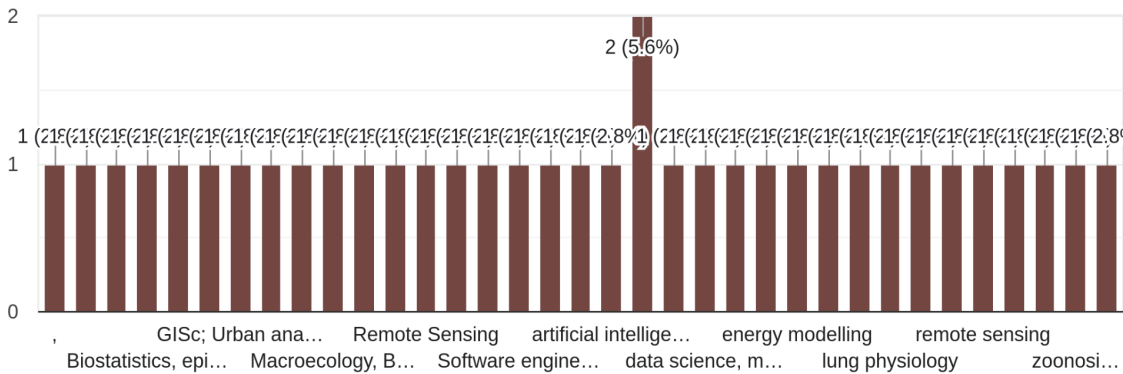
Q16. How confident do you feel about adopting new numerical tools?

36 responses

- Not confident at all
- Slightly confident
- Moderately confident
- Confident
- Very confident

25%

52.8%

13.9%

Q17. How familiar are you about the terms of open linked data and FAIR principles?
36 responses

- ● Not familiar at all
- ● Slightly familiar
- ● Moderately familiar
- ● Familiar
- ● Very familiar

30.6%
22.2%
13.9%
8.3%
25%
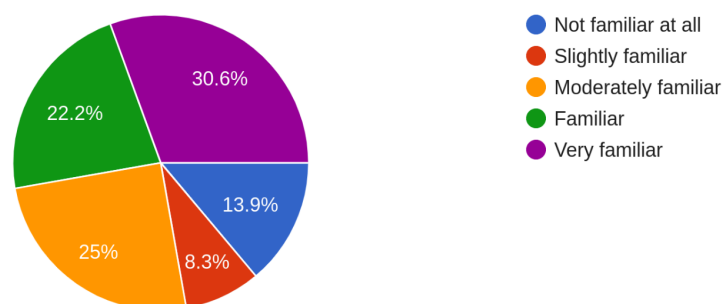
Q18. Which of the Data-Terra data-hubs have you used?: (multiple choice)
36 responses

ODATIS	6 (16.7%)
AERIS	3 (8.3%)
FormaTerre	1 (2.8%)
THEIA	15 (41.7%)
None of them	14 (38.9%)
OMERE	1 (2.8%)
AMMA-CATCH	1 (2.8%)
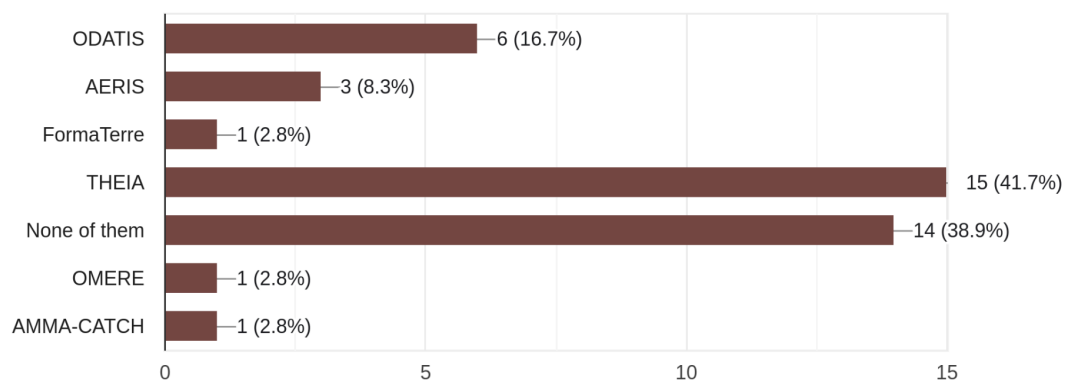
Please feel free to add any comments about your user experience or this survey:

need a formation !

Dans la section "Dataset metadata title " tous les titres ne sont pas coupés au même nombre de caractères. Certains titres font une 15aine de caractères de long quand d'autres en font une 50aine.

Merci pour votre beau travail !

I'm involved in THEIA Land. I don't understand the difference between THEIA-OZCAR and THEIA-HYDRO. Just this comment to share some information to display the best view in the 'front end'. THEIA SPATIAL ?. I think , only THEIA Name will be the best entry to share the results

La partie résultat sur les métadonnées devrait être plus centrale, la zone actuelle dédiée à ces résultats est trop petite. Cela rend notamment la lecture des résultats plus difficile avec des titres tronqués, on doit cliquer sur le résultat pour avoir le titre entier, si plusieurs titres commencent de la même manière on est obligé de cliquer sur chacun d'eux pour voir la différence.

Le graphique avec les relations sémantiques n'est pas forcément facile à comprendre au premier usage, d'autant plus que là aussi certaines expressions sont tronquées dans certains cas (ex : quand on active narrower). Une fois que l'on a compris le fonctionnement, il devient pertinent lorsque la recherche du mot clé dans la métadonnée est plus complexe qu'une recherche simple (pas uniquement indexé via un keyword).
Présence de doubles ascenseurs dans l'interface ce qui rend la navigation plus compliquée.
Cette interface sera certainement plus adaptée pour des scientifiques que du grand public.

- keywords are one way to search data, other criteria are also of primary importance (geographical: place, geographical extent, ...; temporal: time period, date; etc.);
- a search tool in the results of the query would be useful in order to avoid to scan all the results, but on the hand, a "ctrl-f" on the page works and significantly helps ...

I think that more information about the retrieved datasets should be displayed (in a formated way) in the "Retrieved dataset metadata" section: geographic location, time periods, format? To help users see quickly which datasets are relevant or not.
I don't know if the separation between the different data hubs (theia, odatis ..) are easy to understand for scientists that don't know them.
More classic filters could be preferred to reduce the number or retrieved datasets (location, time etc), but maybe that's not the point in this study.
The "Theia description" link in the "Selected dataset details" section does not display anything in the Theia-Ozcar portal (eg:
https://in-situ.theia-land.fr/description/dataset/TheiaOZCAR.ERUN_DAT_Hydrogeochemistry-PF CI1300).
It took me some time to see that the results of the open link discovery were displayed at the bottom of the page in the "Open linked dataset metadata retrieval", because I had to scroll to see it. But that's just a question of ergonomy, maybe again that's not the point in this study :)
It seems that the open link discovery only works with Agrovoc vocabulary (maybe it's normal).
Finally I think that the Explanatory Graph may be really difficult to understand for scientists, or with clear explanations about how we can read it and use it.

The explanatory graph is difficult to read, probably because at the moment it does not give back the details from nodes or links.
The design of the interface is not very appealing and can be consufing (I was trying to type the search term on the title of that box.
It would also be good that the menu from the hub does not rolls up (and disappear) when rolling down the retrieved datasets.
And is not evident that when narrowing the search we will get a new box with retrieved datasets, as I was expecting the first box with retreieved datasets just to update and give a subset with the narrower search.

Nice tool ! But I would need a quick starter guide to explore it correctly and assess its full potential.

Maybe a 3 entry point approach for data seeking : location (x,y,z) and time.

Having a counter for how many "narrower" and/or "broader" terms there are in the different vocabularies would make it easier to navigate (or even have clickable links to some of those terms).

The interface is unclear, for example I tried to put my search term in the title of the search box because it's the same typography. And all the graphs are hard to read because the labels and arrows overlap.

I prefer the names of data-hub titles to always appear at the top when I scroll to the bottom of the list.

A good tool for open data research

This project is very interesting and will certainly simplify data retrieval. Are there any plans to integrate biodiversity data?
The best thing would be for this type of tool to exist tomorrow, at least at European level.
The display of dataset details could be better enhanced (e.g. with a thumbnail display).

Ma réponse au test n'est pas forcement pertinente car trop neophite

good, but I need start from the basic

My observations are expressed in the questions concerned. They are as follows;

Step 1. Enter the keyword "Temperature" as the search term and click "retrieve datasets".

→ the retrieved datasets appear in the list at the bottom.

My Response: The background around the search bar should have a separate colour. It is a bit difficult to find something when everything looks white.

Step2. Click on a specific dataset to have more details.

→ Relationships linking the search term to the retrieved dataset are shown as a explanatory graph.

→ Details of the selected dataset are displayed on the right.

My Response: The graph is clumsy with overlapping text.

Q8. How relevant are the retrieved datasets compared to the initial search?*

My Response: I did not see any link, and I did not observe any change (I don't know if it's my browser).

need either sort facility of key word to choose among data base lists

I searched for "MORPHOLOGY" looking for some connection between species morphological features or patterns and other databases such as occurrence area or epidemiological contexts. However, this type of data may not have been included in this database and even in the project at all. Therefore my user experience may be biased.