La synthèse des connaissances sur la biodiversité : introduction aux méta-analyses et revues systématiques – 2024

Recherche de littérature : Bases de données & Equation de recherche

Sylvie Campagne
Sorbonne Université – Station Biologique de Roscoff

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 899546.

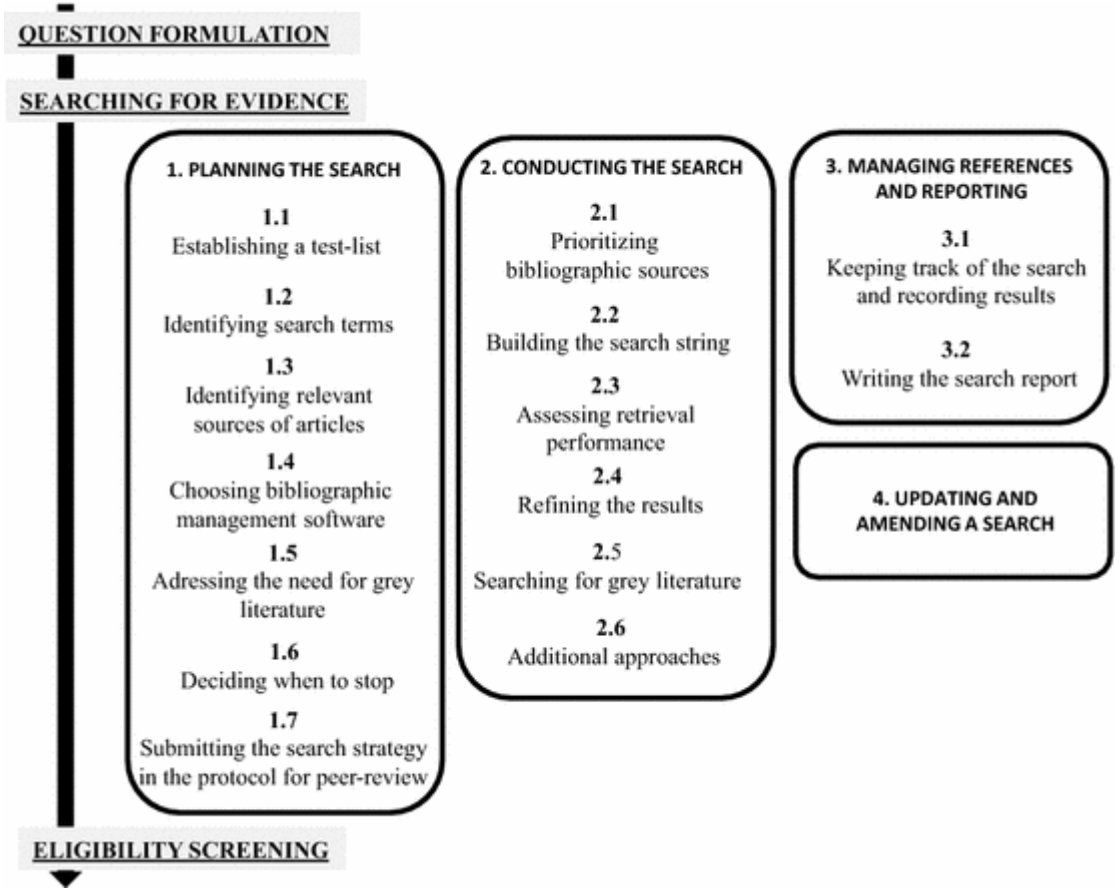
The search strategy

The goal:

Find the relevant bibliographic references !

The search strategy :

1. The search string
2. The bibliographic sources
3. The test-list



A guide to the planning, conduct, management and reporting of the searching phase of systematic reviews and systematic maps (after Livoreil et al. 2017).
https://environmentalevidence.org/information-for-authors/4-conducting-a-search/

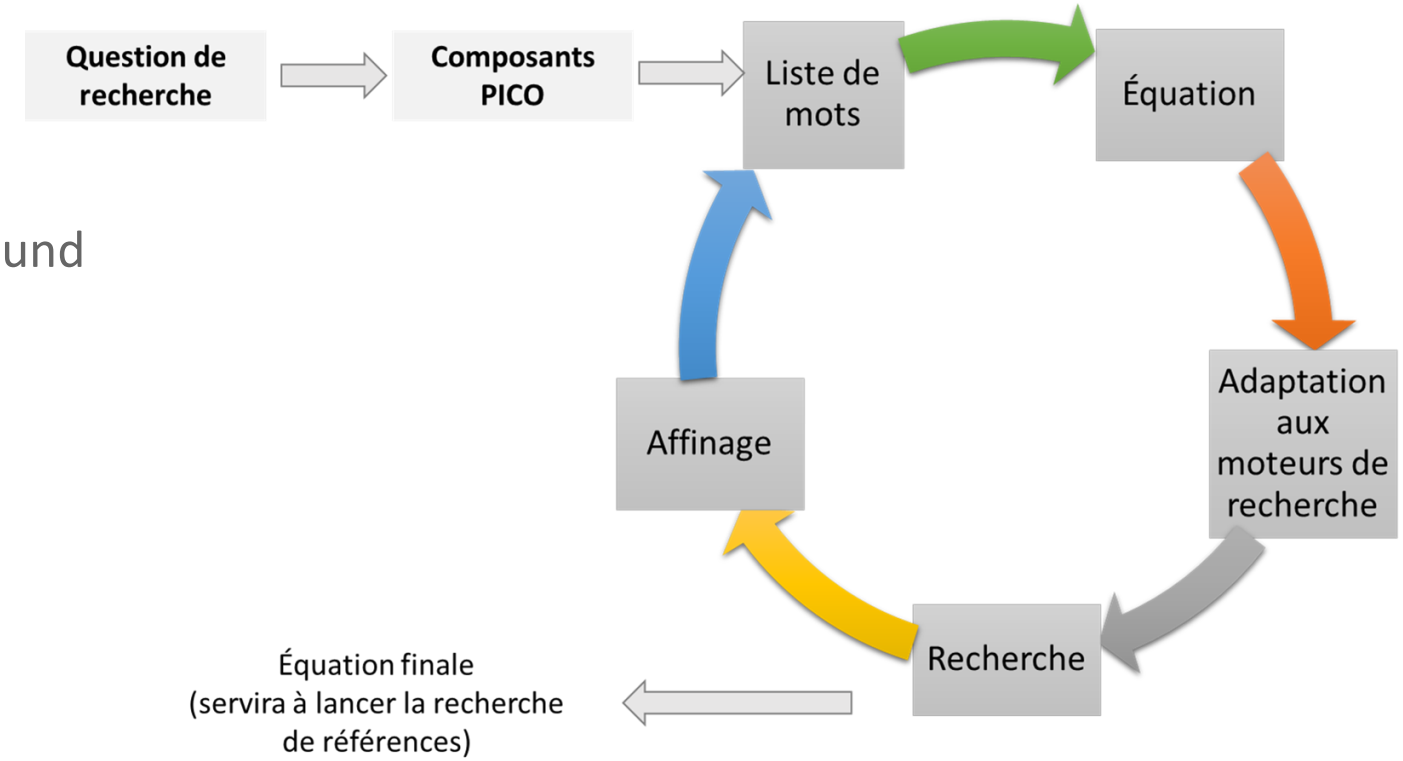
The search strategy

Starting with the research questions

Define the search string

search terms encompasses individual or compound words used in a search to find relevant articles

search string is a combination of search terms combined using Boolean operators



The search string

PICO/PECO elements (Richardson et al. 1995)

Population : effect *on what?*

Intervention / Exposure : *effect of what ?*

Comparator : *compared to what? to what reference?*

Outcome : effect *measured by what?*

(Context : *what type of study?*)

Question element	Definition
Population (of subjects)	Unit of study (e.g. ecosystem, species) that should be defined in terms of the statistical populations of subject(s) to which the intervention will be applied.
Intervention/exposure	Proposed management regime, policy, action or the environmental variable to which the subject populations are exposed.
Comparator	Either a control with no intervention/exposure or an alternative intervention or a counterfactual scenario.
Outcome	All relevant outcomes from the proposed intervention or environmental exposure that can be reliably measured

The search string

1 **Define the PICO based on the research question**

PICO/PECO elements (Richardson et al. 1995)

Population : effect *on what?*

Intervention / Exposure : *effect of what ?*

Comparator : *compared to what? to what reference?*

Outcome : effect *measured by what?*

(Context : *what type of study?*)

My PICO

Example with the Agri-TE project:
What is the effect of agricultural practices on biodiversity at the global level?

The search string

1 **Define the PICO based on the research question**

PICO/PECO elements (Richardson et al. 1995)

Population : effect *on what?*

Intervention / Exposure : *effect of what ?*

Comparator : *compared to what? to what reference?*

Outcome : effect *measured by what?*

(Context : *what type of study?*)

My PICO

Any unplanned/uncultivated taxon

Any agricultural practice

Agricultural witness or natural environment of ref.

Effect-size representing a biodiv metric.

Meta-analyses only

Example with the Agri-TE project:
What is the effect of agricultural practices on biodiversity at the global level?

The search string

2 **Establish the list of words that will be used to construct the search equation**

My search terms

biodiversity, soil fauna, birds, butterflies

tillage, fertilization, pesticides

croplands, forest

species richness, biomass, Shannon's entertainment

meta-analyses

My PICO

Any unplanned/uncultivated taxon

Any agricultural practice

Agricultural witness or natural environment of ref.

Effect-size representing a biodiv metric.

Meta-analyses only

Example with the Agri-TE project:
What is the effect of agricultural practices on biodiversity at the global level?

The search string

3 **Building the search string by adapting to search engines (eg: WoS)**

My search terms

biodiversity, soil fauna, birds, butterflies

tillage, fertilization, pesticides

croplands, forest

species richness, biomass, Shannon’s diversity

meta-analyses

My search string

TS= ((biodiversity OR soil fauna OR birds OR butterflies)

AND (tillage OR fertilizers **OR** pesticides **)**

AND (croplands OR forest)

AND (species richness OR biomass OR Shannon's diversity)

AND (meta-analyses))

Example with the Agri-TE project:
What is the effect of agricultural practices on biodiversity at the global level?

The search string

My search string

TS= ((biodiversity OR soil fauna OR birds OR butterflies)

AND (tillage OR fertilizers **OR** pesticides **)**

AND (croplands OR forest)

AND (species richness OR biomass OR Shannon's diversity)

AND (meta-analyses))

Topic (title, abstract, keywords, keywords plus)

Logical and Boolean Operators

Both depend on the bibliographic sources

Example with the Agri-TE project:
What is the effect of agricultural practices on biodiversity at the global level?

The search string

My search string

TS= ((biodiversity OR soil fauna OR birds OR butterflies)

AND (tillage OR fertilizers **OR** pesticides **)**

AND (croplands OR forest)

AND (species richness OR biomass OR Shannon's diversity)

AND (meta-analyses))

Topic (title, abstract, keywords, keywords plus)

Logical and Boolean Operators

Both depend on the bibliographic sources

Beware of database variations in the search equation!!!

- Some use a different language for searching
- For example, $ instead of *.
- Additional options (inside or nearby)
- Help files are useful!
- Check the options
- Seek specialist help if necessary
- SAVE EVERYTHING

The search string

4 **Test the search string**

My search string

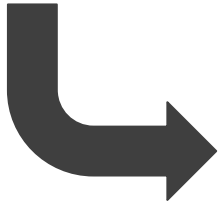
TS= ((biodiversity OR soil fauna OR birds OR butterflies)

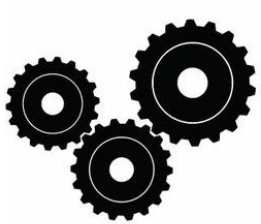
AND (tillage OR fertilizers **OR** pesticides **)**

AND (croplands OR forest)

AND (species richness OR biomass OR Shannon's diversity)

AND (meta-analyses))





200 results is not enough!
20,000 results is too much!
Refinement needed…

Example with the Agri-TE project:
What is the effect of agricultural practices on biodiversity at the global level?

The search string

5 **Refine the search string**

My search string

TS= ((biodiversity OR soil fauna OR birds OR butterflies)

AND (tillage OR fertilizers **OR** pesticides **)**

AND (croplands OR forest)

AND (species richness OR biomass OR Shannon's diversity)

AND (meta-analyses))

Exact Expression
"soil fauna"

Truncations
pesticide*, pesticide$

Exclusion
NOT (medical science OR economics)

Thematic
soil fauna OR (earthworms OR spiders OR collembola OR springtails)

Example with the Agri-TE project:
What is the effect of agricultural practices on biodiversity at the global level?

The search string

5 **Refine the search string**

Iterative process that can (must?) be long

Ex: Foo et al. (2021)

Final search equation

Search Engines

Initial search string 159 results	1	*TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation") AND ("immune challeng*" OR "immunochalleng*" OR "infect*")) NOT (load OR human OR people))*
Add inclusion terms 4,360 results	2	TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation" *OR "reproductive compensation" OR "fitness"*) AND ("immune challeng*" OR "immunochalleng*" OR "infect*" *OR lipopolysaccharide OR lps OR phytohemagglutinin OR pha OR "sheep red blood cells" OR srbc OR implant OR vaccin**)) NOT (load OR human OR people))
Edit inclusion term 493 results	3	TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation" OR "reproductive compensation" *OR "reproductive fitness"*) AND ("immune challeng*" OR "immunochalleng*" OR "infect*" OR lipopolysaccharide OR lps OR phytohemagglutinin OR pha OR "sheep red blood cells" OR srbc OR implant OR vaccin*)) NOT (load OR human OR people))
Add inclusion terms 2,489 results	4	TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation" OR "reproductive compensation" OR "reproductive fitness" *OR "reproductive investment" OR "Life History Trade-Off*" OR "life history"*) AND ("immune challeng*" OR "immunochalleng*" OR "infect*" OR lipopolysaccharide OR lps OR phytohemagglutinin OR pha OR "sheep red blood cells" OR srbc OR implant OR vaccin*)) NOT (load OR human OR people))
Change inclusion term 1,819 results	5	TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation" OR "reproductive compensation" OR "reproductive fitness" OR "reproductive investment" OR "Life History Trade-Off*" *~~OR "life history"~~ OR "trade off"*) AND ("immune challeng*" OR "immunochalleng*" OR "infect*" OR lipopolysaccharide OR lps OR phytohemagglutinin OR pha OR "sheep red blood cells" OR srbc OR implant OR vaccin*)) NOT (load OR human OR people))
Delete inclusion term 1,155 results	6	TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation" OR "reproductive compensation" OR "reproductive fitness" OR "reproductive investment" OR "reproductive success" OR "Life History Trade-Off*" *~~OR "trade off"~~*) AND ("immune challeng*" OR "immunochalleng*" OR "infect*" OR lipopolysaccharide OR lps OR phytohemagglutinin OR pha OR "sheep red blood cells" OR srbc OR implant OR vaccin*)) NOT (load OR human OR people))
Add inclusion term 1,429 results	7	TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation" OR "reproductive compensation" OR "reproductive fitness" OR "reproductive investment" OR "reproductive success" OR "Life History Trade-Off*" *OR "Phenotypic Plasticity"*) AND ("immune challeng*" OR "immunochalleng*" OR "infect*" OR lipopolysaccharide OR lps OR phytohemagglutinin OR pha OR "sheep red blood cells" OR srbc OR implant OR vaccin*)) NOT (load OR human OR people))
Add exclusion terms 1,141 results	8	TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation" OR "reproductive compensation" OR "reproductive fitness" OR "reproductive investment" OR "reproductive success" OR "Life History Trade-Off*" OR "Phenotypic Plasticity") AND ("immune challeng*" OR "immunochalleng*" OR "infect*" OR lipopolysaccharide OR lps OR phytohemagglutinin OR pha OR "sheep red blood cells" OR srbc OR implant OR vaccin*)) NOT (load OR human OR people *OR men OR women OR infant* OR rat OR rats OR mouse OR mice OR pig* OR pork OR beef OR cattle OR sheep OR lamb* OR chicken* OR calf* OR horse**))

Pilot 100 papers to check hit rate. 6% hit rate. Continue refining.

Final search string 1,567 results (~10% hit rate)	TS = ((("terminal investment" OR "reproductive effort" OR "fecundity compensation" OR "reproductive compensation" OR "reproductive fitness" OR "reproductive investment" OR "reproductive success" OR "Life History Trade-Off*" OR "Phenotypic* Plastic*" OR "pre-copulatory NEAR/5 trait*" OR "sexual NEAR/5 weapon*" OR "sexual NEAR/5 ornament*" OR "post-copulatory NEAR/5 trait*" OR "ejaculate quality" OR "sperm quality" OR "mating effort" OR "parental care") AND ("immune challeng*" OR "immunochalleng*" OR "infect*" OR lipopolysaccharide OR lps OR phytohemagglutinin OR pha OR "sheep red blood cells" OR srbc OR implant* OR vaccin* OR nylon OR sephadex)) NOT (load OR human OR people OR men OR women OR infant* OR rat OR rats OR mouse OR mice OR pig* OR pork OR beef OR cattle OR sheep OR lamb* OR chicken* OR calf* OR horse* OR infective))

The bibliographic sources

Bibliographic sources capture any source of references, including electronic bibliographic databases, those sources which would not be classified as databases (e.g. the Internet via search engines), hand searched journals, and personal contacts.

- Bibliographic
  - eg WoS, Scopus, Pubmed
- Web search tools
  - eg Google, Google Scholar
- Grey literature sources
  - Organizational websites
  - Thesis repositories

The bibliographic sources

Bibliographic databases

- Web of Science
- Scopus
- Agricola
- AGRIS (FAO)
- Academic Search Premier
- Biological Abstracts
- CAB Abstracts
- etc.

Web search tools

!!! Depend of connection parameters!!!!

The bibliographic sources

Grey literature sources

- 'File drawer' research / unpublished research results

– Unfinished/published/accepted articles

– The theses

– The “uninteresting” results

- Non-academic studies

– Technical reports

– Government documents

– Internal reports

all results not intended for academic publication

The bibliographic sources

Grey literature sources

- 'File drawer' research / unpublished research results

– Unfinished/published/accepted articles

– The theses

– The “uninteresting” results

- Non-academic studies

– Technical reports

– Government documents

– Internal reports

all results not intended for academic publication

How to find them?

– Calls for evidence (social media, networks)

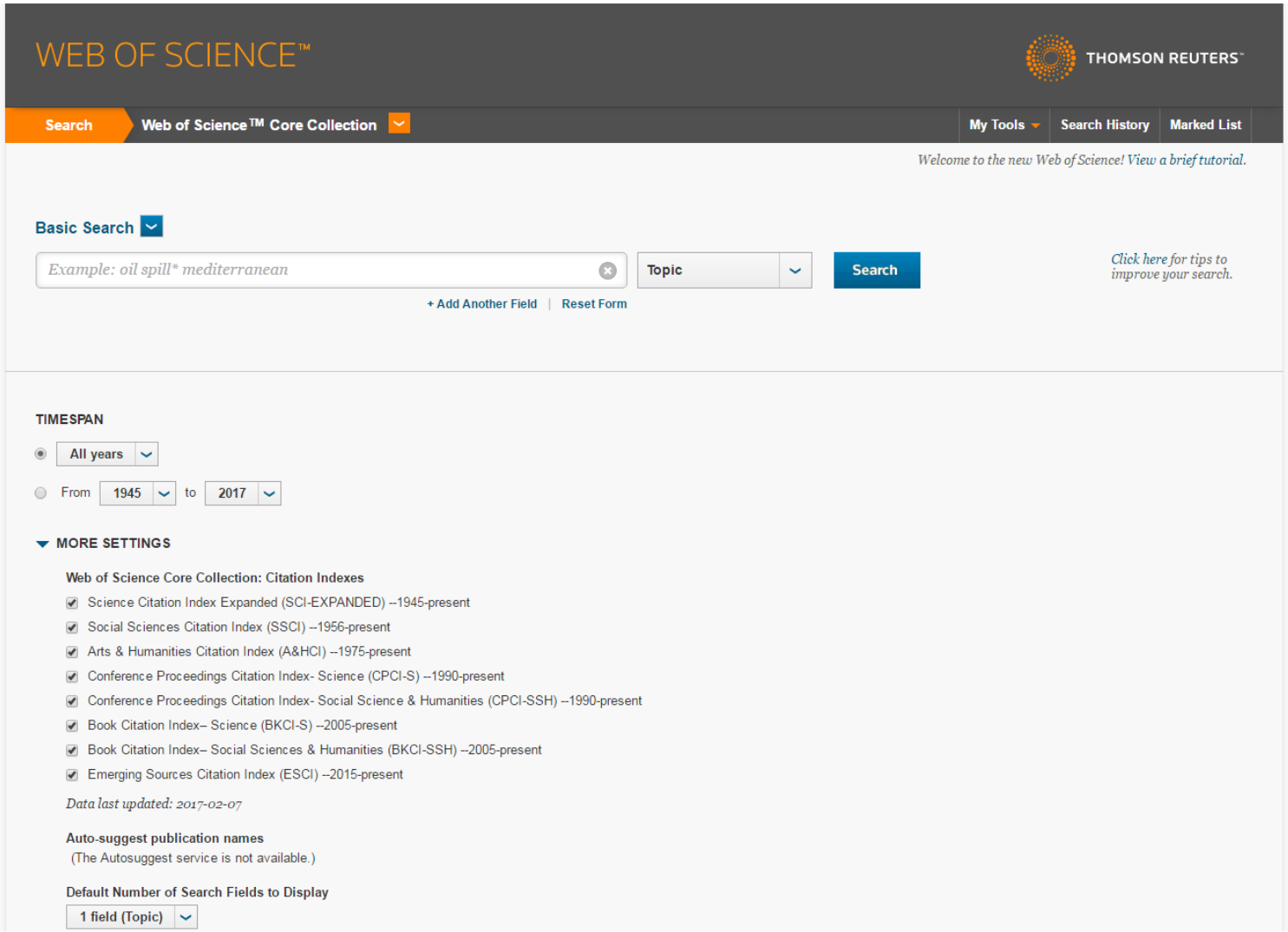
– Thesis databases (eg eThOS)

– Google Scholar, Google

– Pre-print servers (eg ArchivX)

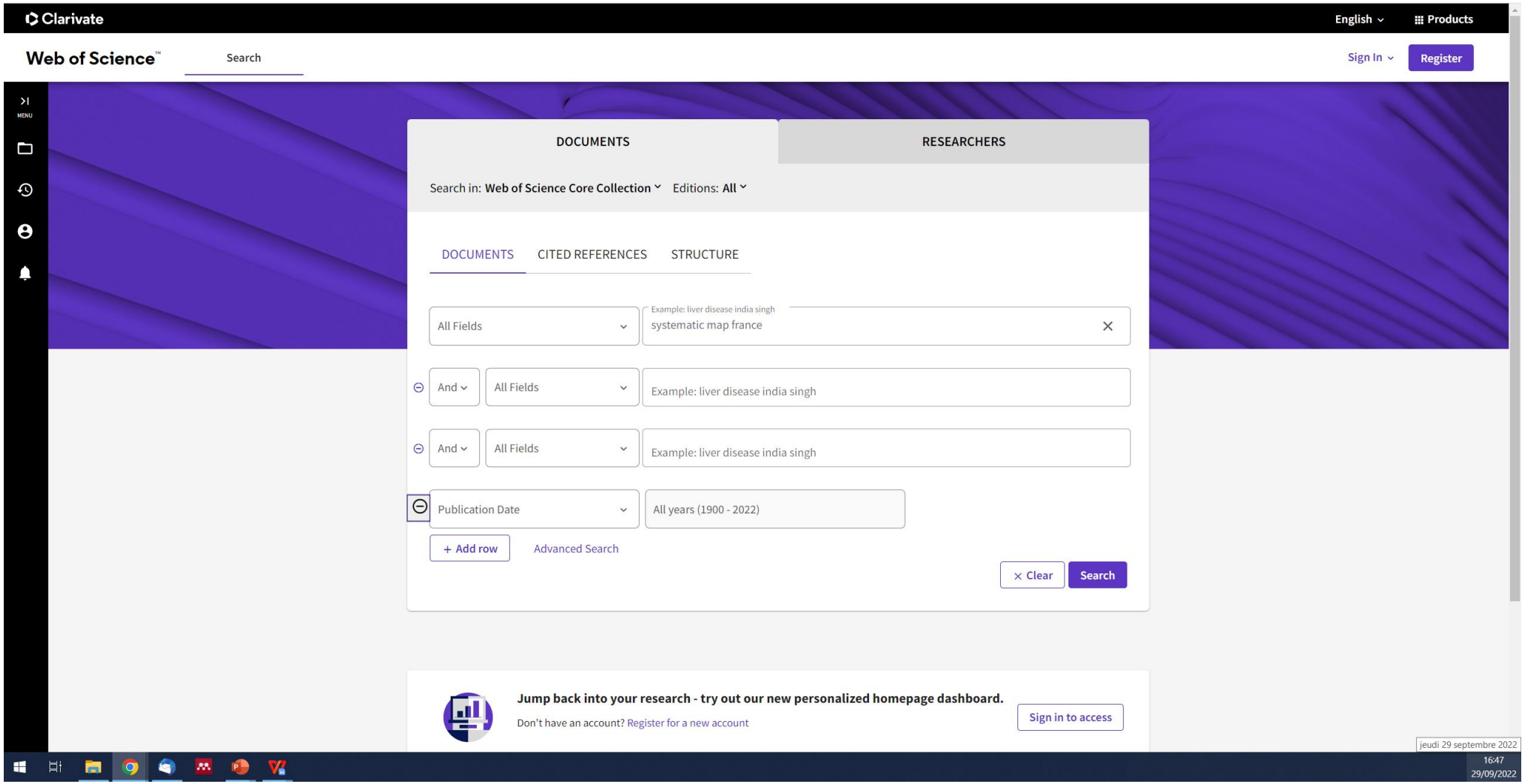
– Organizational websites

Example with web of science

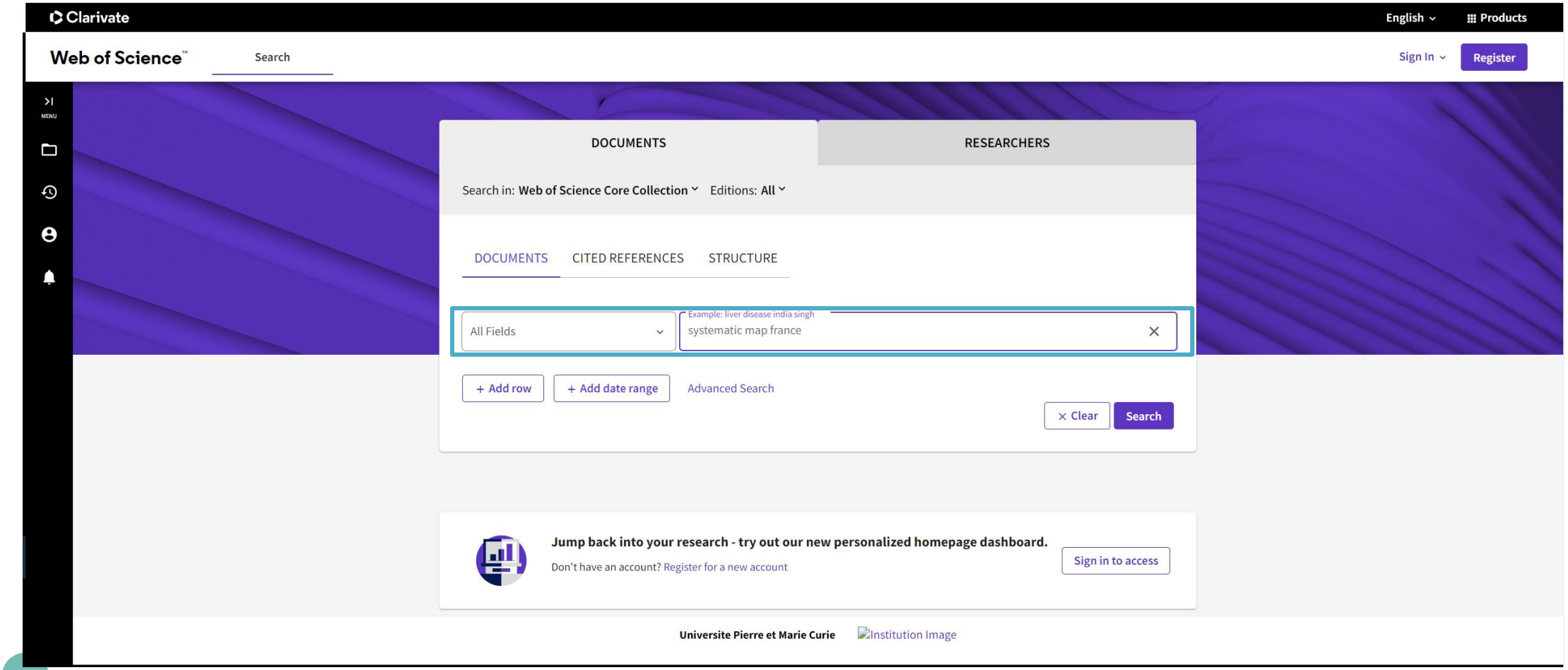


Example with web of science

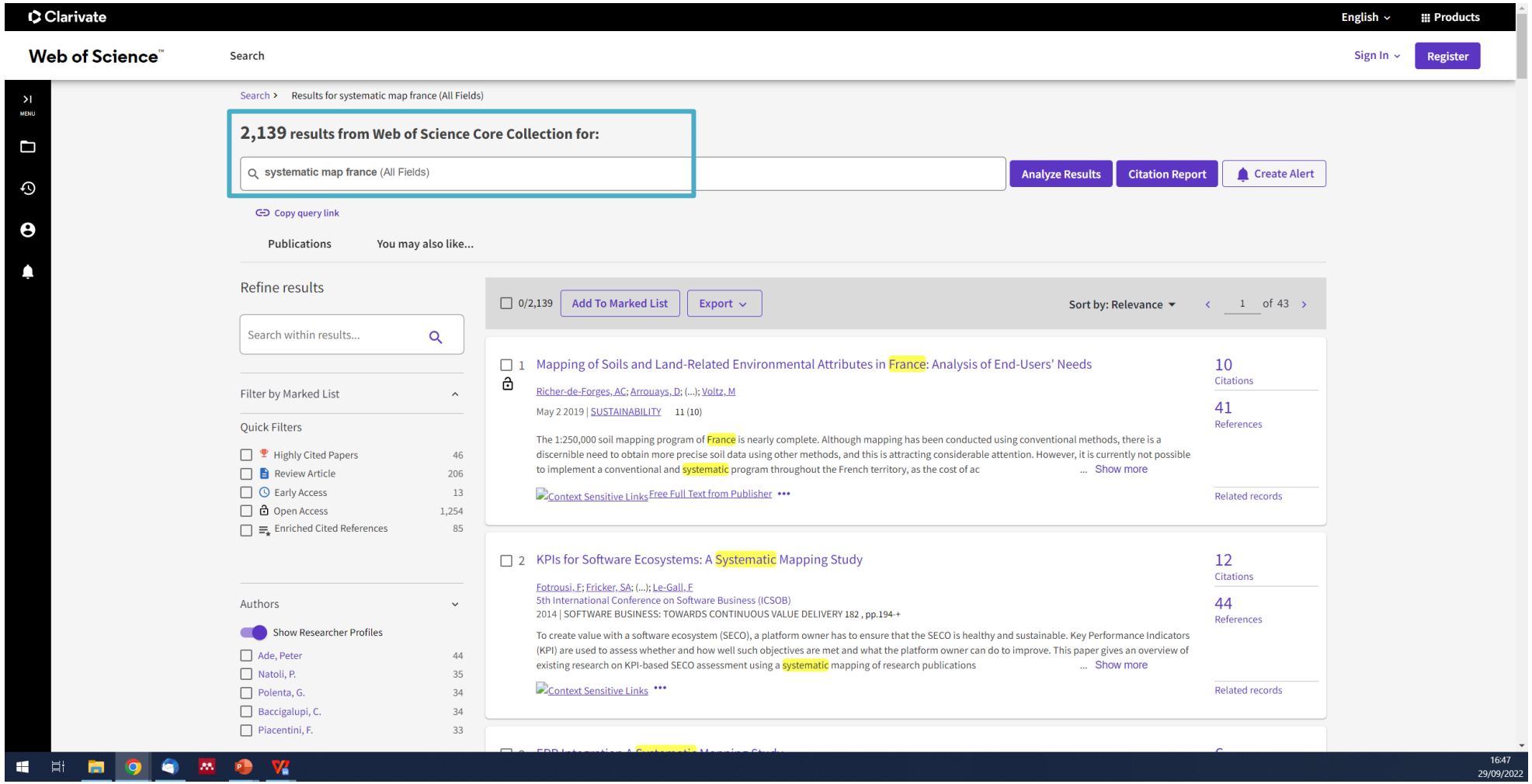
https://www.webofscience.com/wos/woscc/basic-search



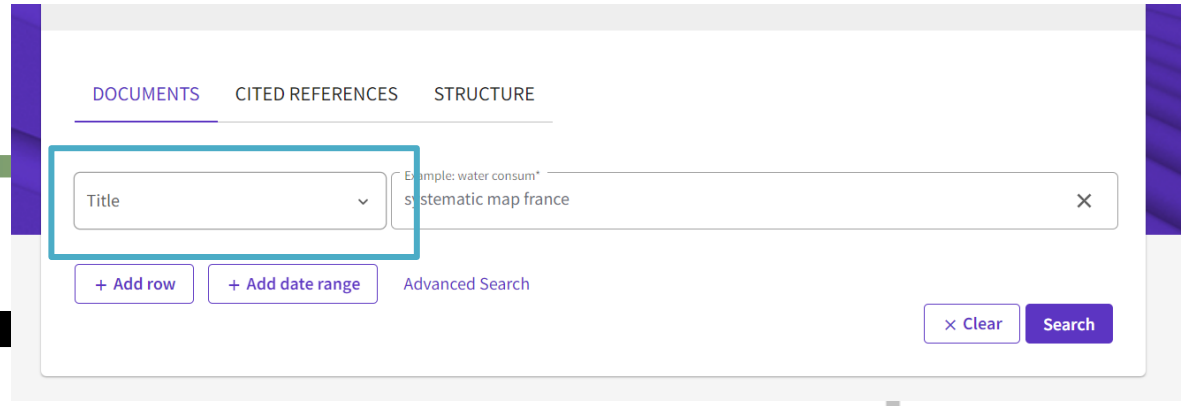
Example with web of science

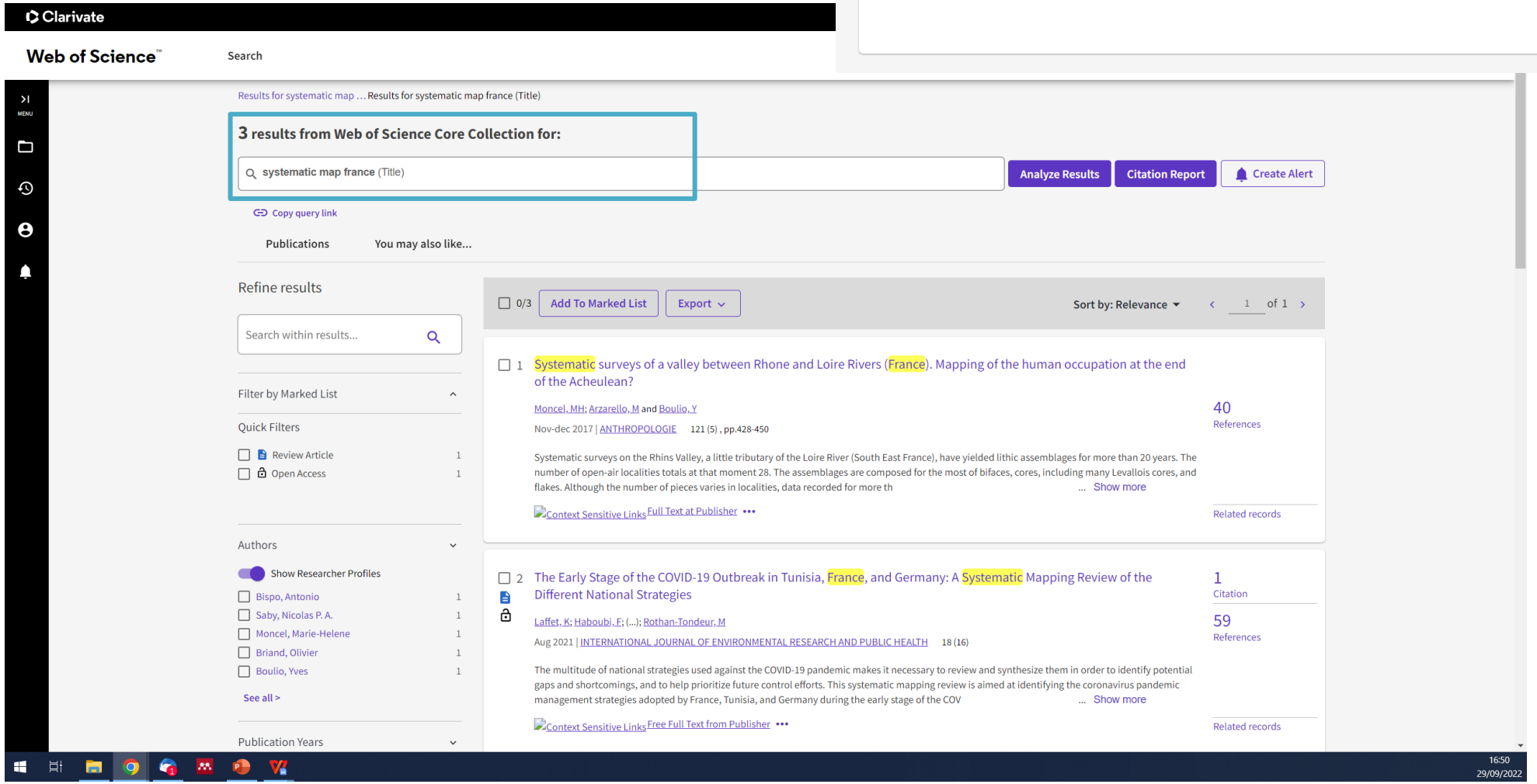


Example with web of science

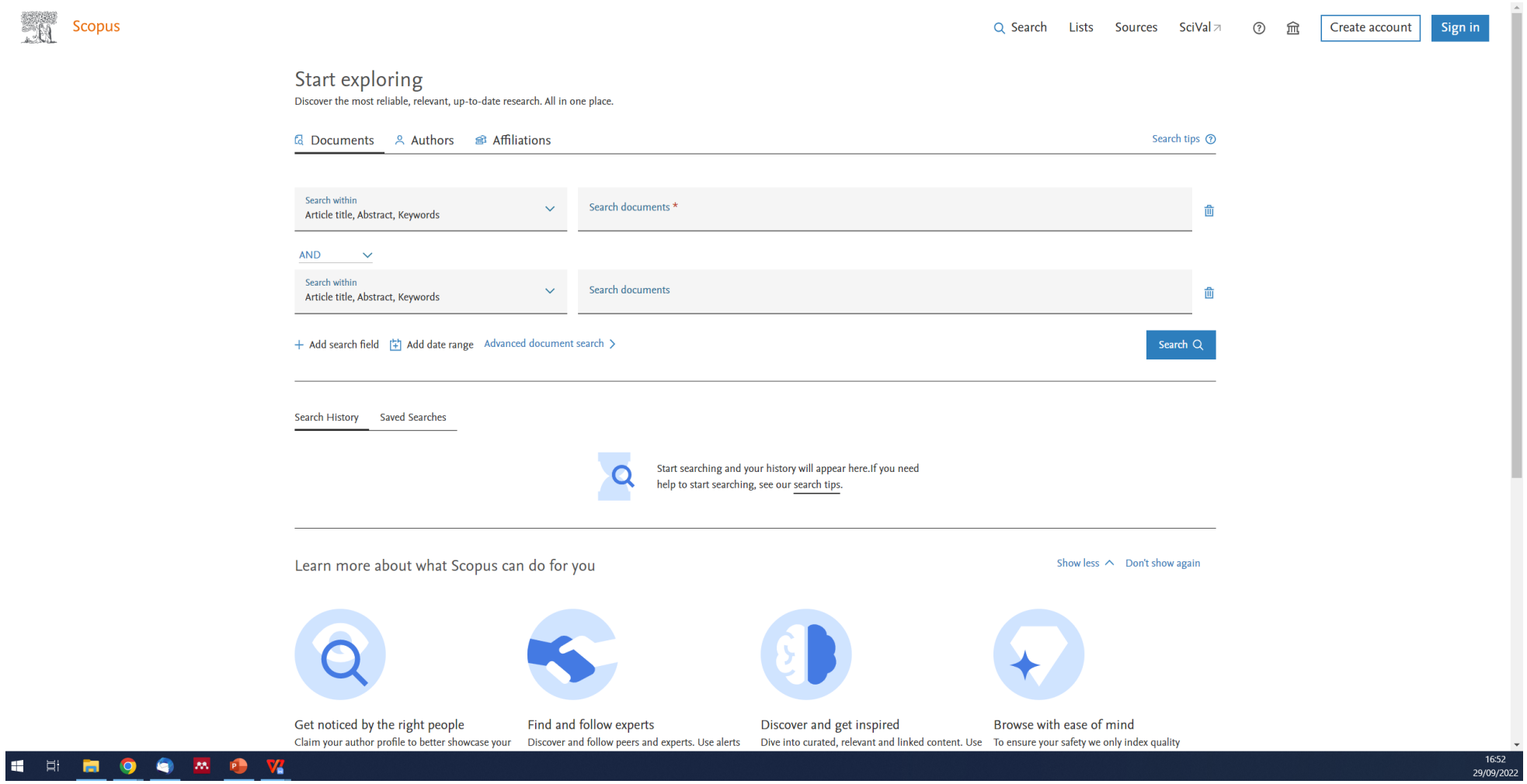


Example with web of science

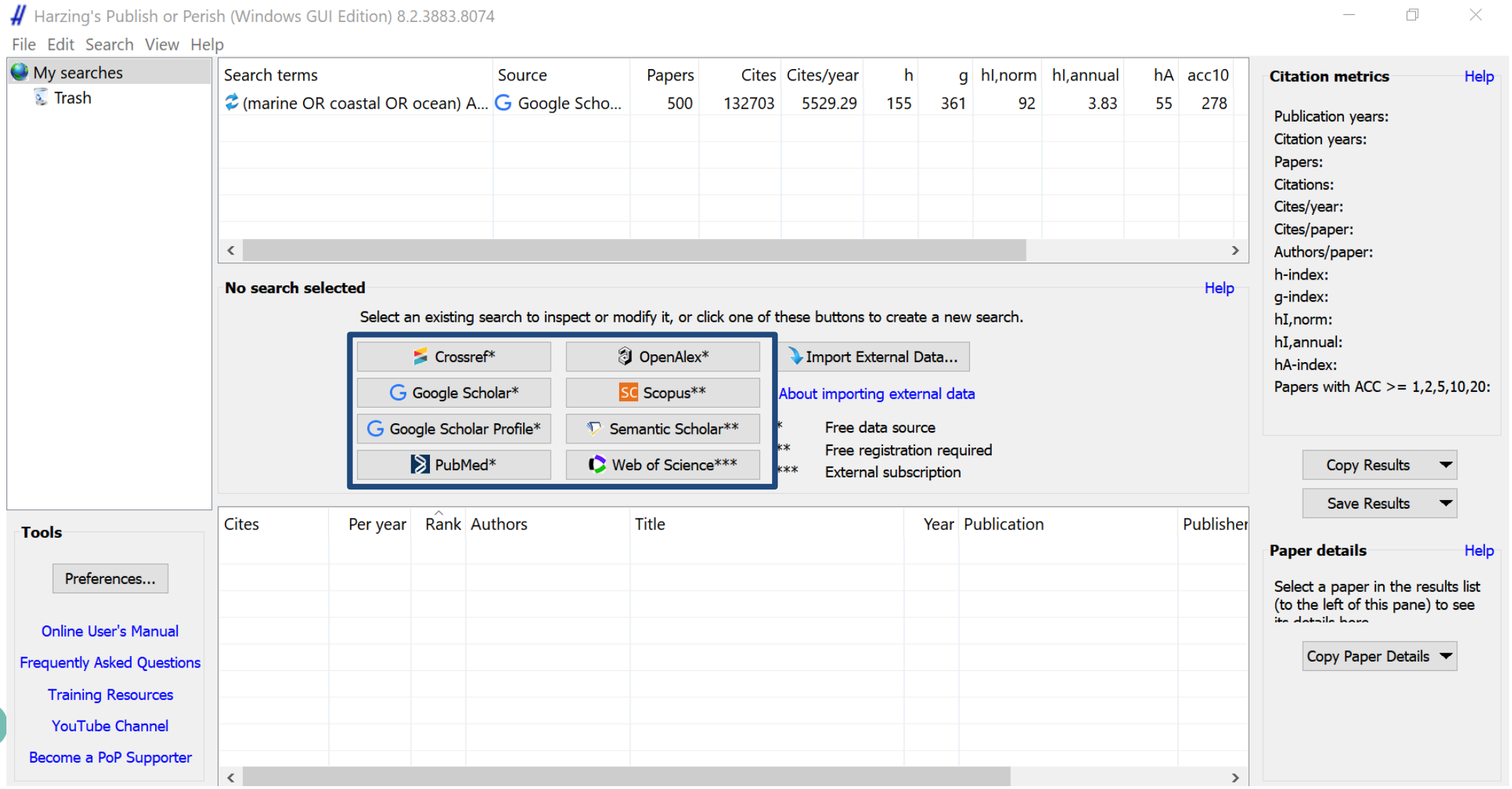




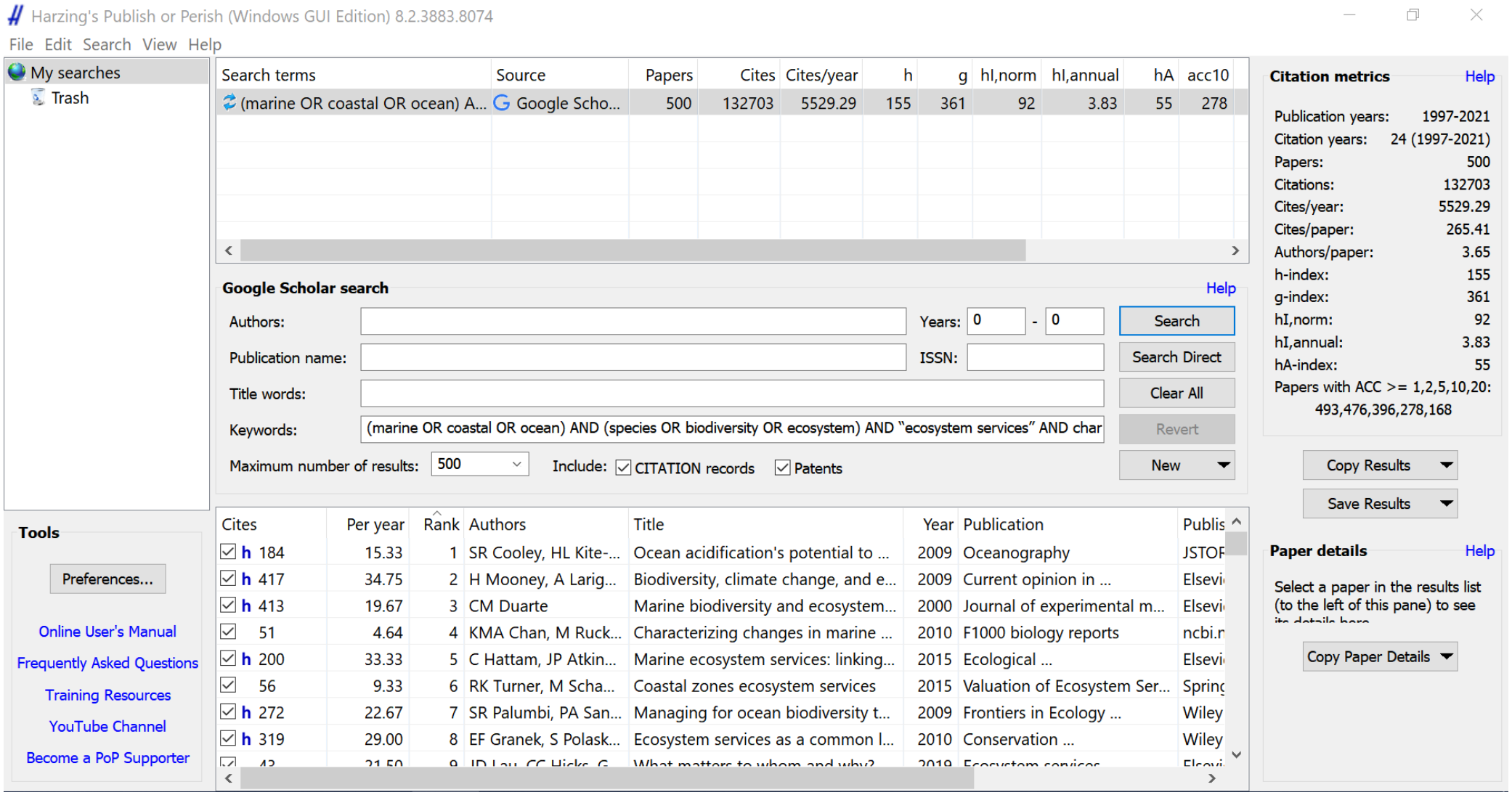
Example with Scopus



Example with Publish or Perish



Example with Publish or Perish



The test list

<u>Test-list</u> : studies that you wish to include in your systematic review and which you know meet the inclusion criteria.

→ Discuss the list (involve partners/co-authors/colleagues) to construct it and then consolidate it

→ Extract metadata

→ Order of magnitude, ca. 30 items

<u>Interest</u> : verify the capacity of a research equation to capture studies corresponding to the aim of our systematic review.

→ Calculate the miss rate = the % of items belonging to the test list not captured by the equation

It must be minimized, ie the equation must approach 100% of the captured test-list… Refinement possible.

The search strategy

Complementary measures of the efficiency of the equation

- ***Miss-rate*** : thanks to the test-list, must be minimized

- ***Hit-rate*** : Percentage of relevant articles, calculated on a sample
  (for example, on 100 randomly selected results)
  → aim **for at least 10%**

- ***Number of results*** : Aim for between 1000 and 3000.

  Adapt depending on the search engine used and/or the strategy employed.

Thank you for your attention !!!

Sylvie Campagne
Sorbonne Université – Station Biologique de Roscoff
sylviecampagne@gmail.com

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 899546.