

Apsis - Document Review Platform

Max Callaghan



August 13, 2018

The exponential growth in literature about climate change raises challenges for environmental assessments:

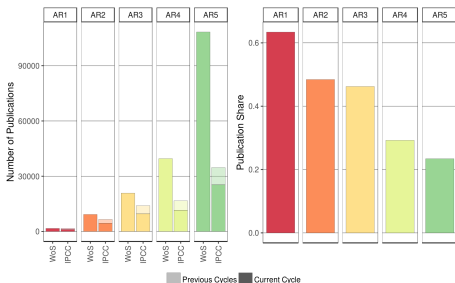
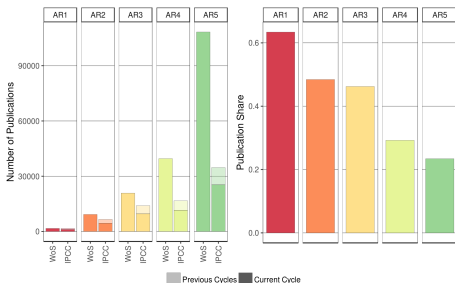


Figure: (?)

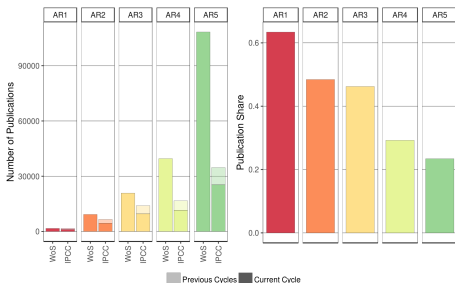
The exponential growth in literature about climate change raises challenges for environmental assessments:



- We need to develop ways of being more systematic in engaging with the literature

Figure: (?)

The exponential growth in literature about climate change raises challenges for environmental assessments:



- We need to develop ways of being more systematic in engaging with the literature
- We need more research on research results

Figure: (?)

The exponential growth in literature about climate change raises challenges for environmental assessments:

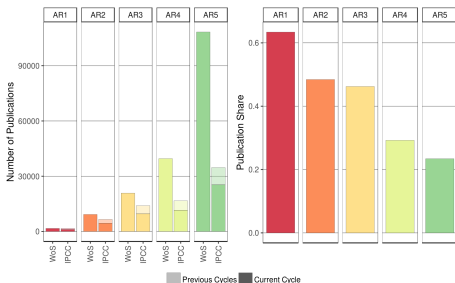
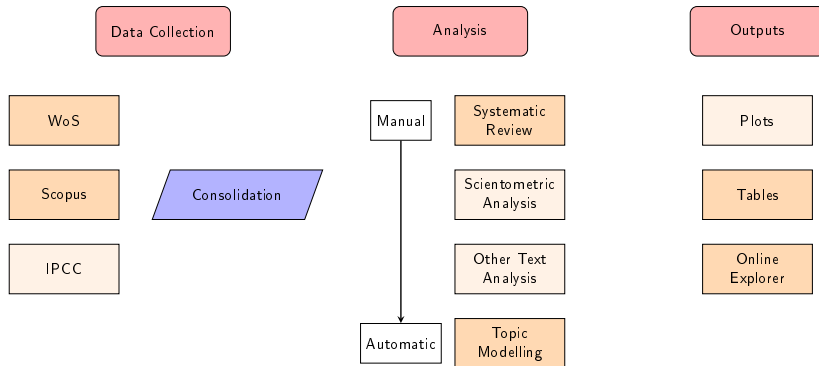


Figure: (?)

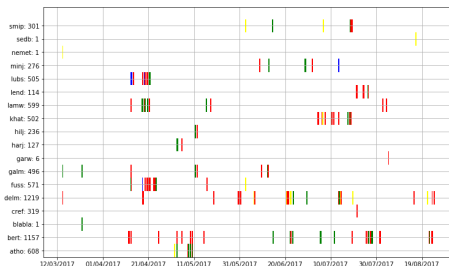
- We need to develop ways of being more systematic in engaging with the literature
- We need more research on research results
- We need ways of engaging with large amounts of text



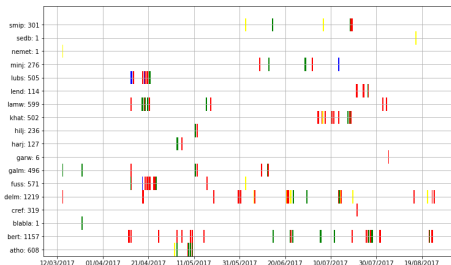
For our systematic review of NETs, we wanted to be as systematic in our search, selection and treatment of literature as possible. We developed a system to help us

- To search and download (bulk) metadata from Web of Science (WoS) and Scopus
- To combine, compare and manage these queries and the documents associated with them
- To manage (centrally) the screening of documents by internal and external collaborators
- To run analysis based on user-entered tagging of documents and metadata from the WoS/Scopus

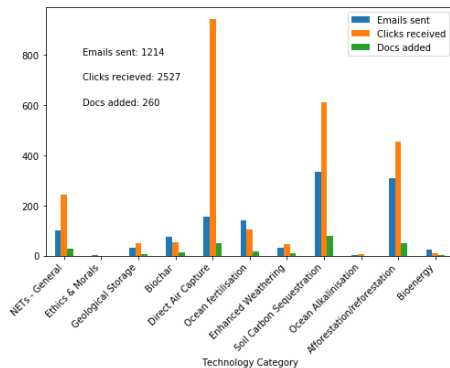
We downloaded over 400 queries, and a team of 18 users reviewed hundreds of documents each.



We downloaded over 400 queries, and a team of 18 users reviewed hundreds of documents each.



We used the results to automatically email all authors of relevant documents



Based on the labels, we could efficiently characterise this bibliographic coupling network

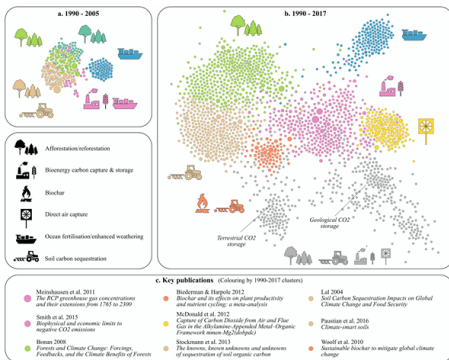


Figure: (?)

And after collecting further information, we could characterise the documents

Based on the labels, we could efficiently characterise this bibliographic coupling network

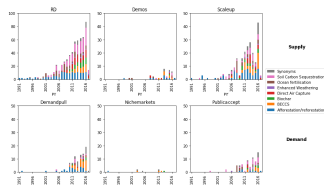
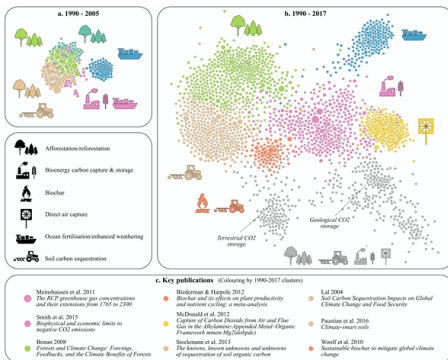


Figure: (?)

Figure: (?)

And after collecting further information, we could characterise the documents

Based on the labels, we could efficiently characterise this bibliographic coupling network

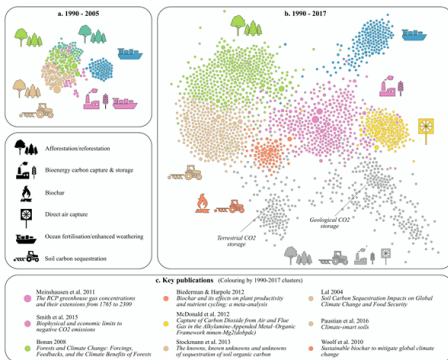


Figure: (?)

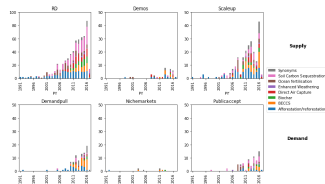


Figure: (?)

And summarise information from them

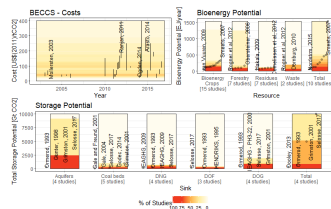
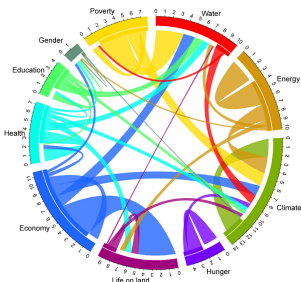


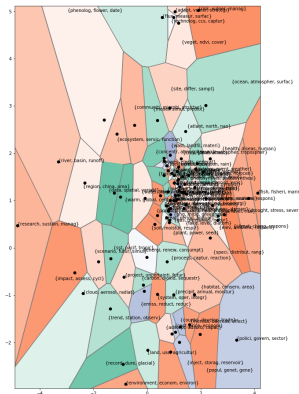
Figure: (?)

IAMs & SDGs with IIASA



- colleagues from IIASA now trialling tool

My PhD work + additional APSIS projects (??)



Other MCC projects

- Nils' Meta-analysis
- Barbora and colleagues' meta-analyses
- Discourse analysis with *Governance* group

`https://apsis.mcc-berlin.net/scoping`

Data Collection

Each time we download a query, we go through a tunnel to PIK (where we have access as Guest Researchers to WoS and Scopus) and instruct the computer to perform a search, and download the results in the maximum chunk size you are allowed (500 or 2000). Both companies prefer that only humans use their website.

Please therefore do search online, and only download what you need

Longer-term workarounds:

- Negotiate machine-access with WoS or Scopus (Elsevier !)
- Remove scraping feature from website and ask users to upload WoS/Scopus files directly

Data Collection

Each time we download a query, we go through a tunnel to PIK (where we have access as Guest Researchers to WoS and Scopus) and instruct the computer to perform a search, and download the results in the maximum chunk size you are allowed (500 or 2000). Both companies prefer that only humans use their website.

Please therefore do search online, and only download what you need

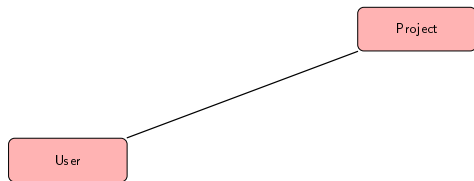
Longer-term workarounds:

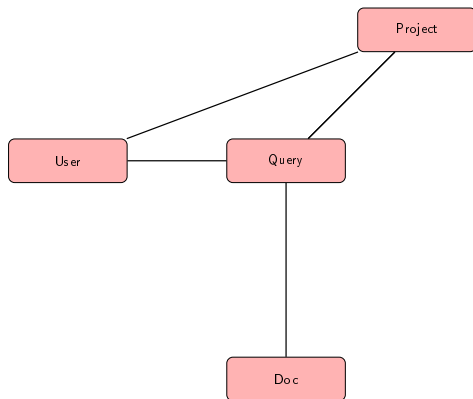
- Negotiate machine-access with WoS or Scopus (Elsevier !)
- Remove scraping feature from website and ask users to upload WoS/Scopus files directly

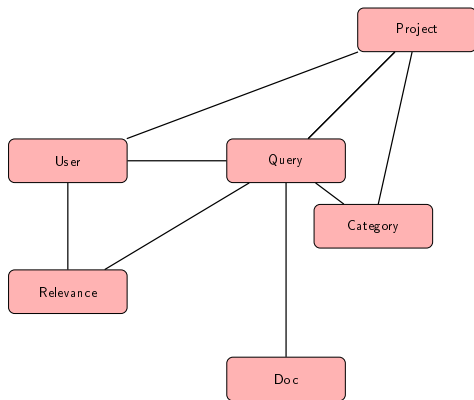
In development

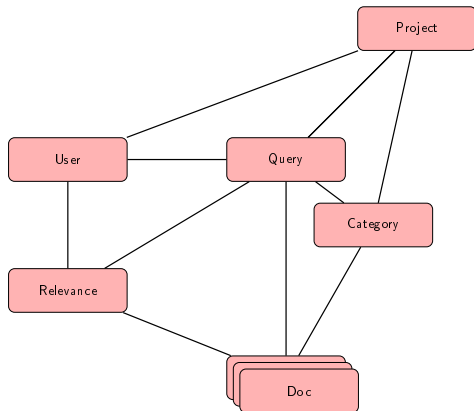
This is lots of pieces of work tied together, written as a side-project to my PhD. I like improving it, and adding more features, but I sometimes break it by accident, and I don't always have time or ability to fix it.

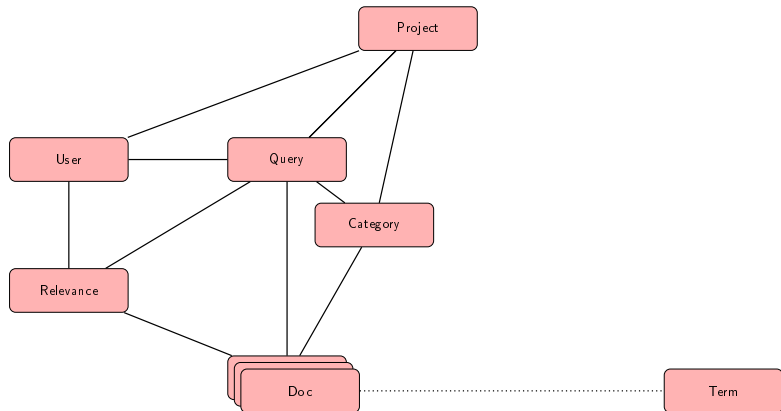
Project

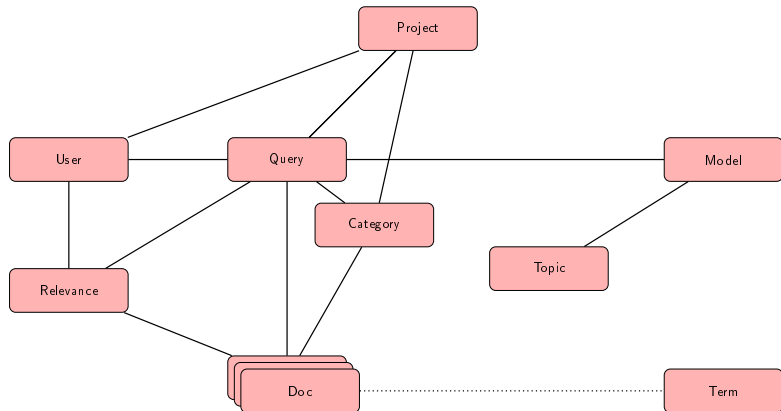


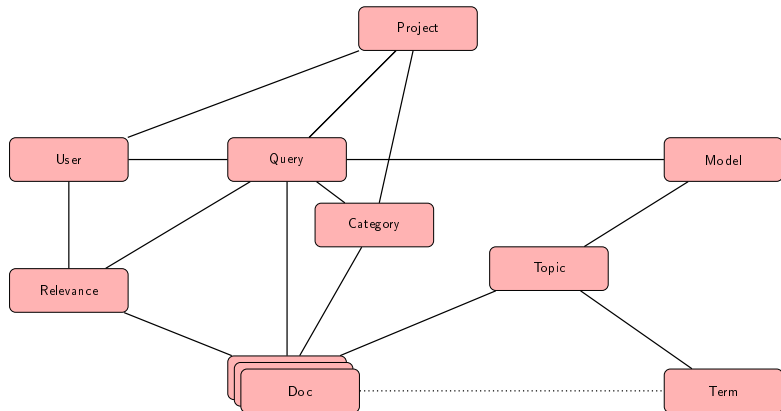


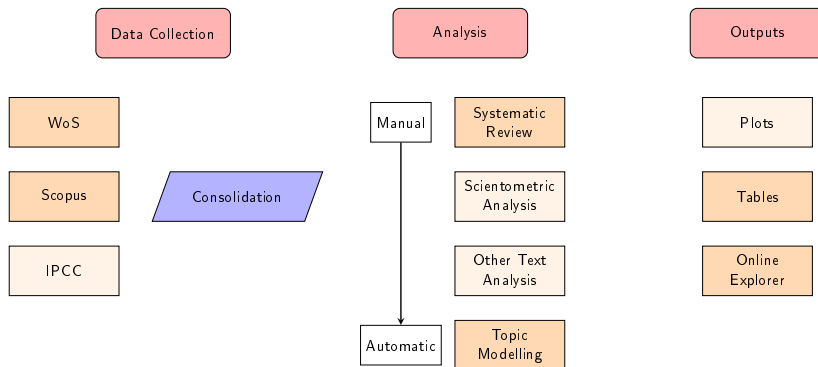












- Does it make sense to bring these bits of analysis online into one platform?
- Can more features be added?
- Are there other uses for the information we've collected and store in the database?
- Are there better ways to show the information we have / organise the platform

Code: <https://github.com/mcallaghan/tmv> (You can raise “issues” there)

Documentation: <https://github.com/mcallaghan/tmv/wiki/Scoping-Documentation> (fairly comprehensive but out of date as of August 13, 2018)