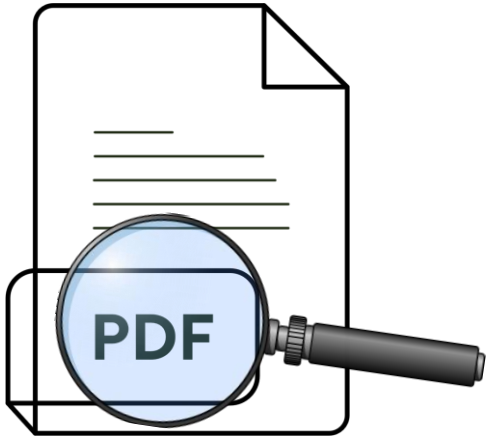
#semanticClimate
Transforming information into actionable knowledge

BRIC
a DBT Organization

Automated Literature Review



25 April, 2025

WII, Dehradun, Uttarakhand

Renu Kumari

Project Scientist III

NIPGR, New Delhi

Why do we need ALR?

- *Large no. of publications*
- *Time Efficiency*
- *covering a broader range of studies*
- *review stays up to date with the latest findings*
- *Consistency and Reproducibility*
- *Cost Efficiency*

Uses of Literature Review

- *Project grant writing*
- *Provides background information for a research problem*
- *Clinical Decision Support*
- *Identifies gaps in current knowledge.*
- *Defining Research Questions and Objectives*

TOOL DEVELOPERS & TESTERS!

- **pygetpapers (Ayush Garg)**
- **amilib (Peter Murray-Rust)**
- **Priti**
- **Moobashara Jawed**
- **Shabnam Bhabuiya**

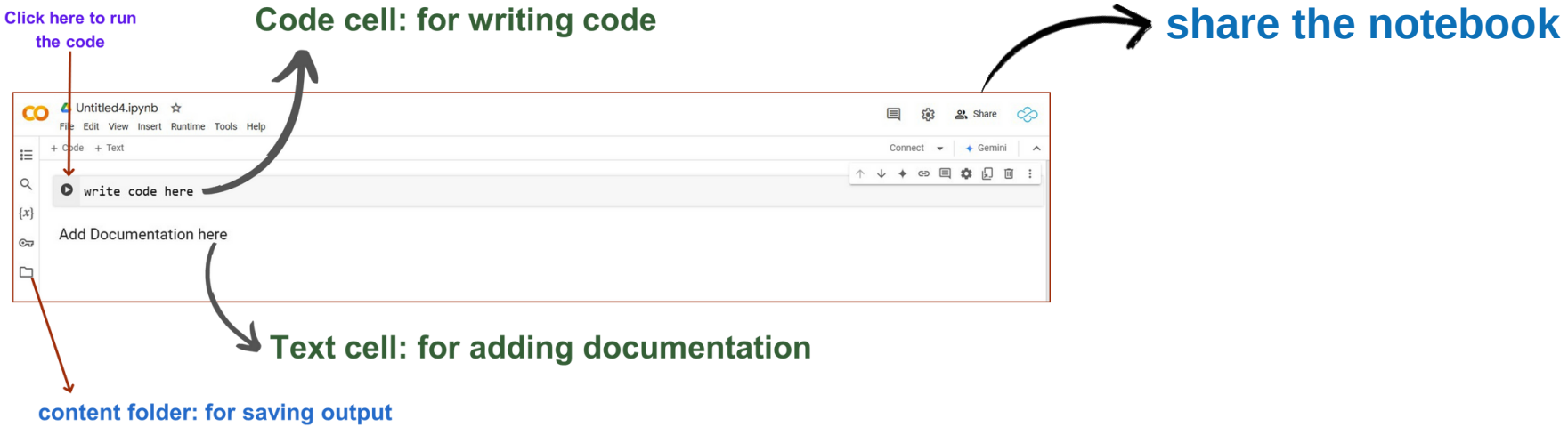
Introduction to Google Colab Notebook

- An open Jupyter E-Notebook environment
- No pain with setups, versions,
- Human Machine friendly
- Supports interactive programming
- Easy learn and explore new tools

semanticClimate.org

Google Colab (Collaboratory)

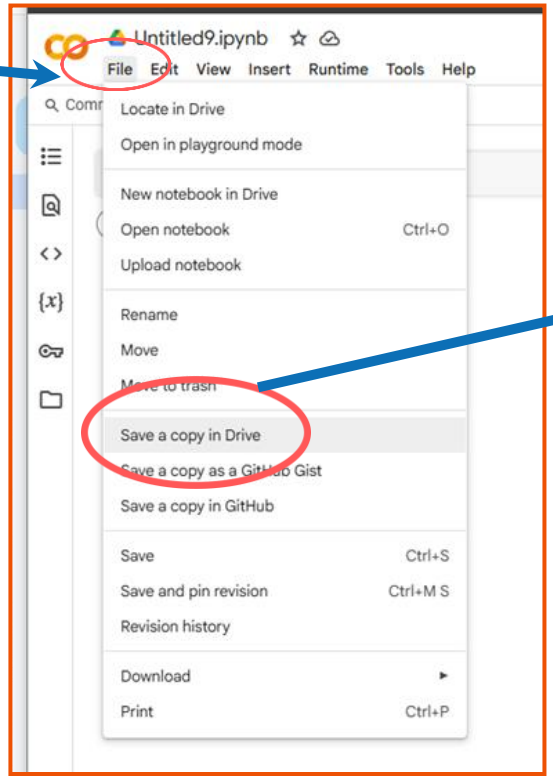
free, cloud-based platform to write and execute python-based codes



Need Google account to get started!

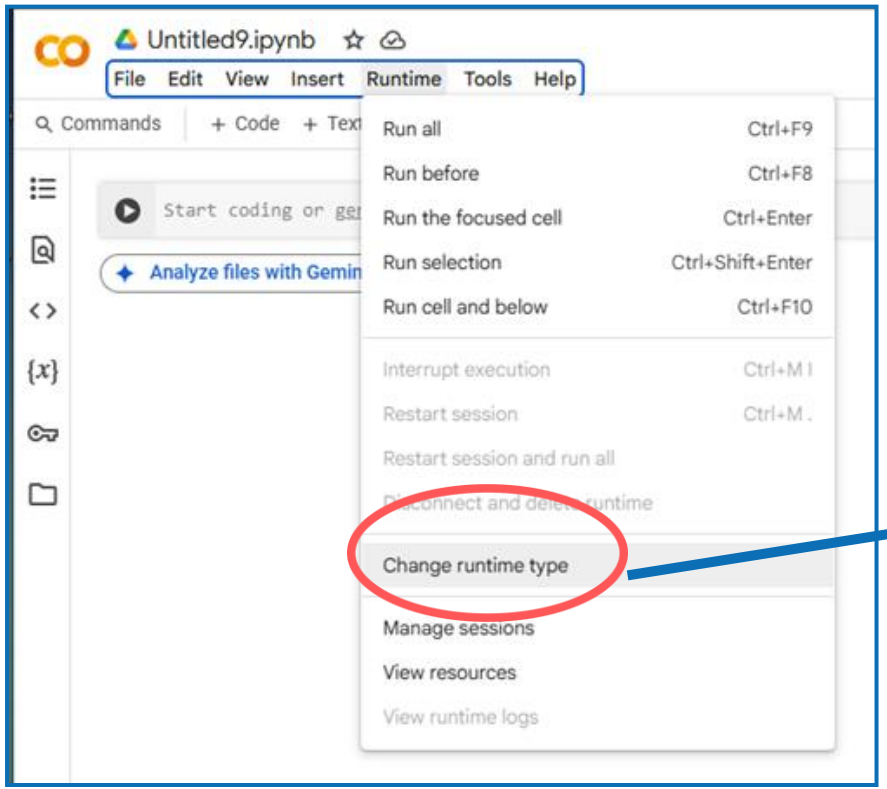
Google Colab (Collaboratory)

Click to File



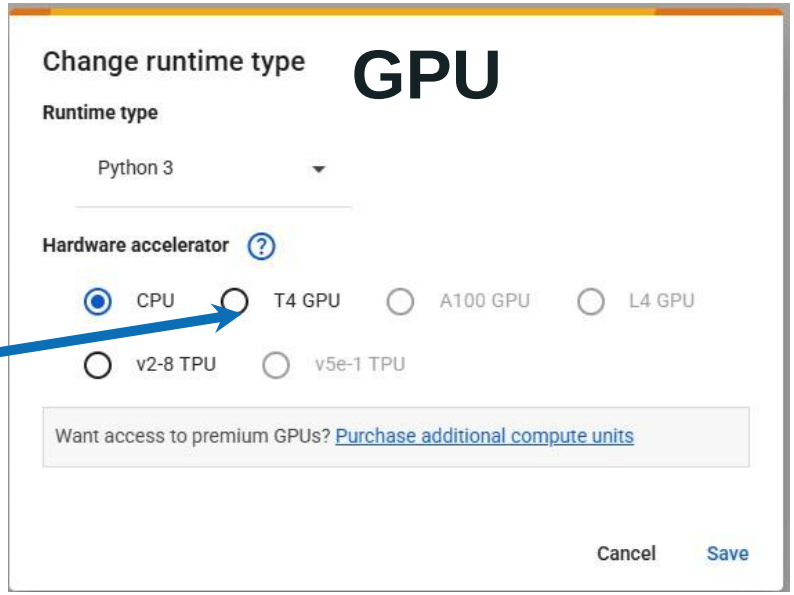
Save a copy in Drive

Google Colab (Collaboratory)

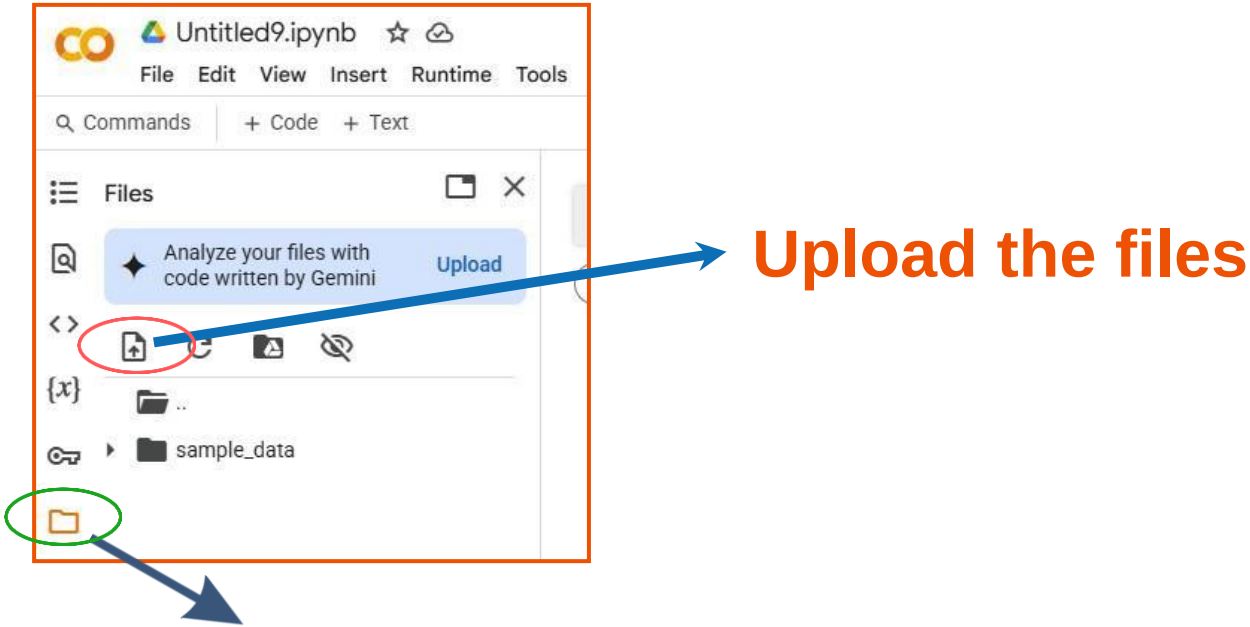


T4

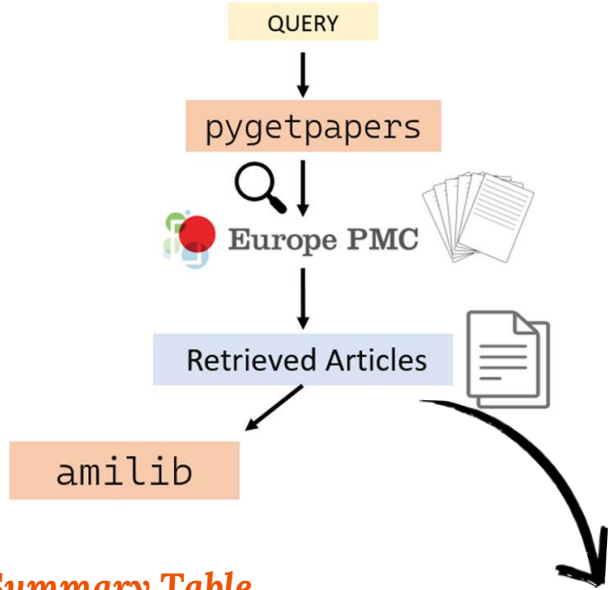
GPU



Google Colab (Collaboratory)



Workflow



Summary Table

title	authorString	journalInfo.journal.title	pubYear	abstractTe
mprovement in the Phytochemical Content and Biological Properties of Stevia rebaudiana (Bertoni) Bertoni Plant Using Endophytic Fungi Fusarium fujikuroi.	• Devi R • Abdulhaq A • Verma R • ... • Yadav R • Mohan S.	Plants (Basel, Switzerland)	2023	This study a medicinal pl endophytes twenty fung
Plants and Plant-Derived Molecules as Natural Immunomodulators.	• Zebeaman M • Tadesse MG • Bachheti RK • Bachheti A • Gebeyhu R • Chaubey KK	BioMed research international	2023	Backgro of plants ha to increasin immunomo
In vitro production of steroidal saponin, total phenols and antioxidant activity in callus suspension culture of Paris polyphylla Smith: an important Himalayan medicinal plant.	• Rawat JM • Pandey S • Rawat B • ... • El-Gazzar AM • El-Saber Batiha G.	Frontiers in plant science	2023	Paris pol native to the recently dw medicine. T

Text Summarization

ot cause of many environmental problems . Schola
to how paying attention to co Conservationists can
mmunities on weaving knowledge systems, lived exp
thropology, geography, and Indigenous studies . W
ropocentric ways of seeing and studying animals h
dge systems, and ontologies that closely reflect mu
traditional ecological knowledge, and non-Western
al and intelligent In traditional communities, people
nd mediums of ancestral spirits . When wild animals
as a societal moral compass, guid-ging people to a
als When Indigenous people recognize wildlife as c
ls . Sustaining coexistence implies supporting more
wildlife as entangled and interconnected (Parathia
es, goods, and To address the complexity of humar
nd coex-existence will depend on prioritizing and lis
Society, 17(3), 283–296. Carter, N. H., & Linnell

Demo Link to Notebook 1

Shabnam Barbhuiya

1 **Summarization of Single PDF**

CLICK HERE:

https://colab.research.google.com/drive/1el5Zjogk7DXqqeuBzGMqFDBGTvyWg1Pm?usp=sharing

Different Steps of the notebook 1

Step 1: Install Required Libraries

Step 2:Download research papers using pygetpapers

Step 3: Create Datatables for Your retrieved articles using amilib

Step 4: Display/Visualise Datatable (HTML output)

Step 5: Import Libraries,Initialize the Tokenizer & Model

Step 6: Clean and Summarize PDF Content
(Remove References + Extract DOI)

Step 7: Loading the PDF & Display the Final Summary.

We are using **sshleifer/distilbart-cnn-12-6,**

you can try other models such as:

- facebook/bart-large-cnn
- nsi319/legal-pegasus
- google/pegasus-pubmed

See more from Hugging Face Summarization Models

Different parameters for query in pygetpapers

User input (title or keywords) required to fetch the paper.

Number of papers to download

```
[ ] !pygetpapers --query '"wildlife" AND "biodiversity"'  --pdf --limit 5 --output downloaded_file --save_query
```

query

Output folder

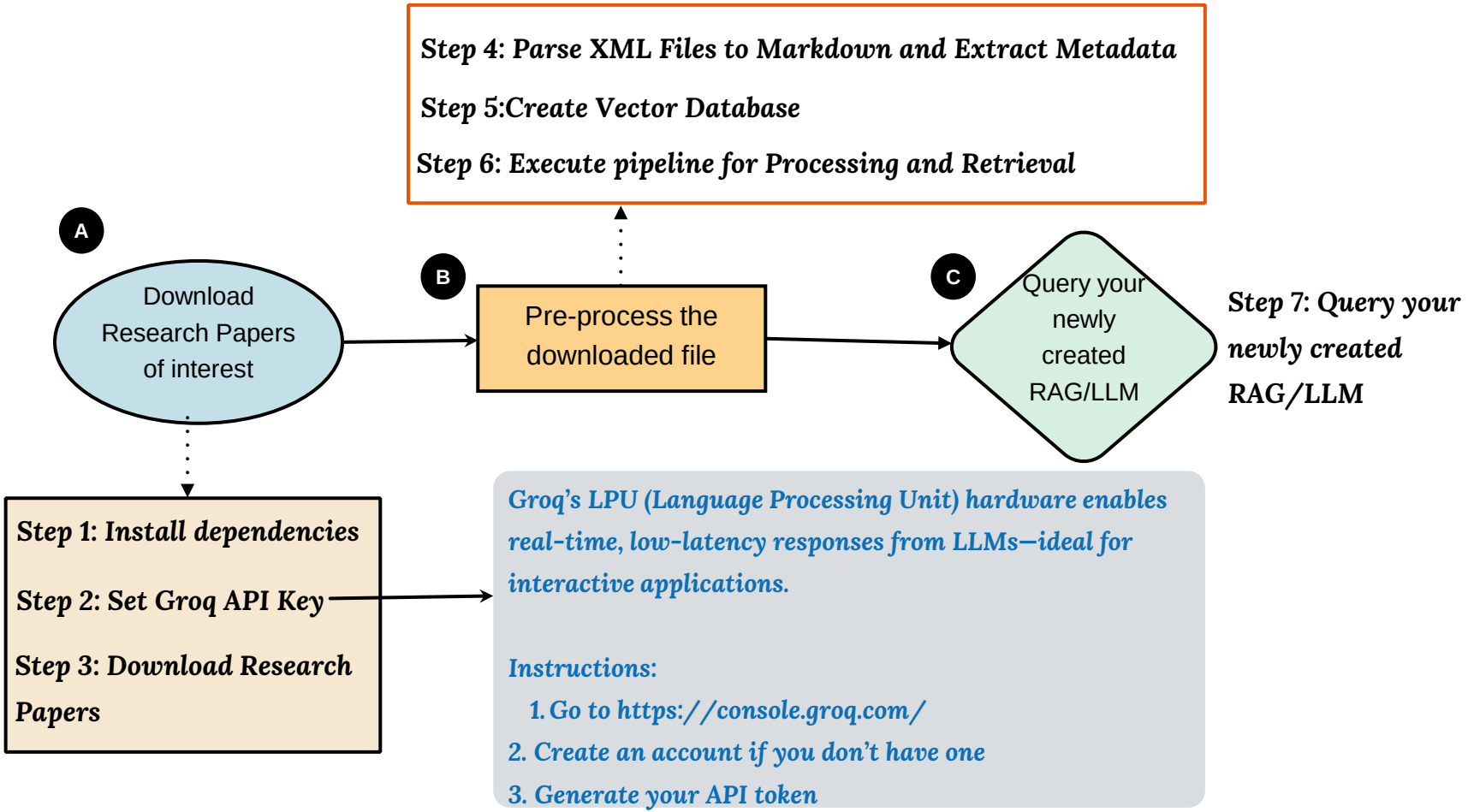
Demo Link to Notebook 2

2 **Querying the corpus with RAG/LLM**

CLICK HERE:

https://colab.research.google.com/drive/1RteHNh-ZROSSxja7tYRaKVCwT5wWOeVP?usp=sharing

Different Steps of the notebook 2



Examples : questions format

What species are mentioned in this paper?

What are the key findings of the study {PAPER TITLE}?

What species are mentioned in this paper?

What compounds were isolated from the plant?

What experimental techniques were used in this paper {PAPER TITLE}?

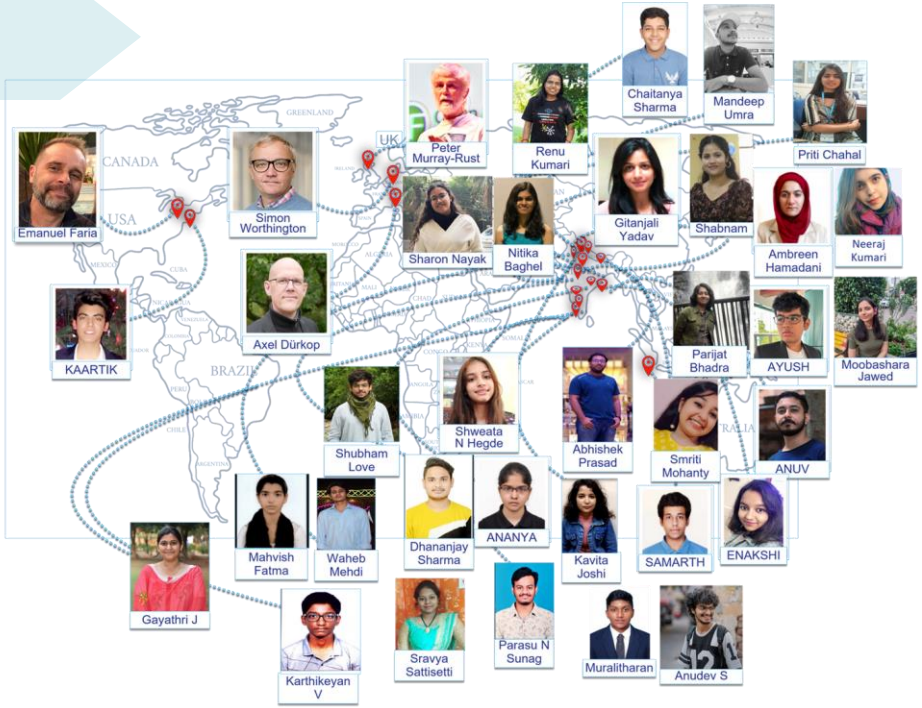
Thank you

BRIC
a DBT Organization

NIPGR

भारतीय वन्यजीव संस्थान
Wildlife Institute of India

#semanticClimate
Transforming information into actionable knowledge



Website : [https://semanticclimate.github.io/p/en/]

email : semanticclimate@gmail.com

X : [@semanticClimate]

LinkedIn : [@semantic Climate]

Git hub : [https://github.com/semanticclimate]