

What I've learnt about Open Science by being a part of CEVOpen

Perspectives on Open Science

Shweata N. Hegde,
NIPGR Intern

Presentation structure

1. What we @ CEVOpen do?



2. What is Open Science?



3. My message and learnings!

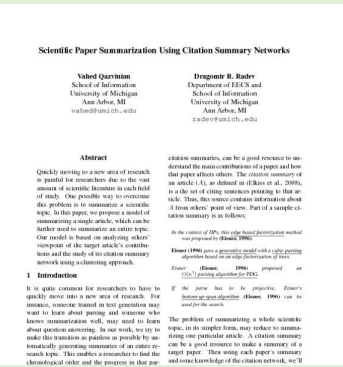


What we @CEVOpen do



Gain Insights

1000s plants
more!



PMCID	Chemicals	plants

Unstructured

Structured



Lantana Camara (An Invasive)



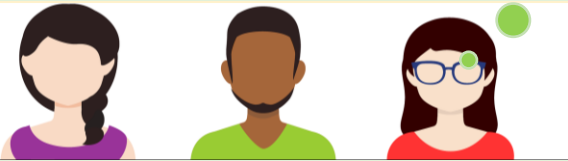
Tulsi

What we @CEVOpen do



Reduce time it takes to review literature

Months to review literature!



Rapidly download bulk of papers



Analyse and annotate



DONE
Minutes, if not hours!



Scientific Paper Summarization Using Citation Summary Networks

Vahed Qarvinian
School of Information
University of Michigan
Ann Arbor, MI
vahed@umich.edu

Dragomir R. Radev
Department of EECS and
School of Information
University of Michigan
Ann Arbor, MI
radev@umich.edu

Abstract

Quickly moving to a new area of research is painful for researchers due to the vast amount of scientific literature. One possible way to overcome this problem is to summarize a scientific topic. In this paper, we propose a model of summarizing a single article, which can be further used to summarize an entire topic. Our model is based on analyzing others' viewpoint of the target article's contributions and the study of citation summary network using a clustering approach.

Citation summaries, can be a good resource to understand the main contributions of a paper and how that paper affects others. The *citation summary* of an article (A), as defined in (Elkiss et al., 2008), is a set of citing sentences pointing to that article. Thus, this source contains information about A from others. A sample citation summary is as follows:

In the context of DP, this edge based factorization method was proposed by (Khan, 1996).

Elkiss (1996) gave a generative model with a cubic parsing algorithm based on an edge factorization of area.

Elkiss (Khan, 1996) proposed an $O(n^3)$ parsing algorithm for PEGs.

If the parse has to be projective, Elmer's bottom-up parse algorithm (Khan, 1996) can be used for the search.

1 Introduction

It is quite common for researchers to have to quickly move into a new area of research. For instant, text generation may want to learn about parsing, and someone who knows summarization well, may need to learn about question answering. In our work, we try to make this transition as painless as possible by automatically generating summaries of an entire research topic. This enables a researcher to find the chronological order and the progress in that par-

The problem of summarizing a whole scientific topic, in its singular form, may reduce to summarizing one particular article. A citation summary can be a good response to make a summary of a target paper's summary and some knowledge of the citation network, we'll

Visions for CEVOpen

De-jargonizing
science



Phytoremedi-
ation??

Look at ami's
annotation!

Examples here

Visions for CEVOpen

Make Scientific Literature
accessible to non-native
speakers



Ami with the help of
Wikidata can look up
terms for you, right
there!

I don't know
which
Ocimum
means

Annotating with images!

Nice example of using and doing open Science

WE make use of, and also contribute to...



Open literature



Open database - Wikidata



Open source tools

Impossible if ALL literature was closed and copyrighted

Open Science is
more than *just* open
publishing...

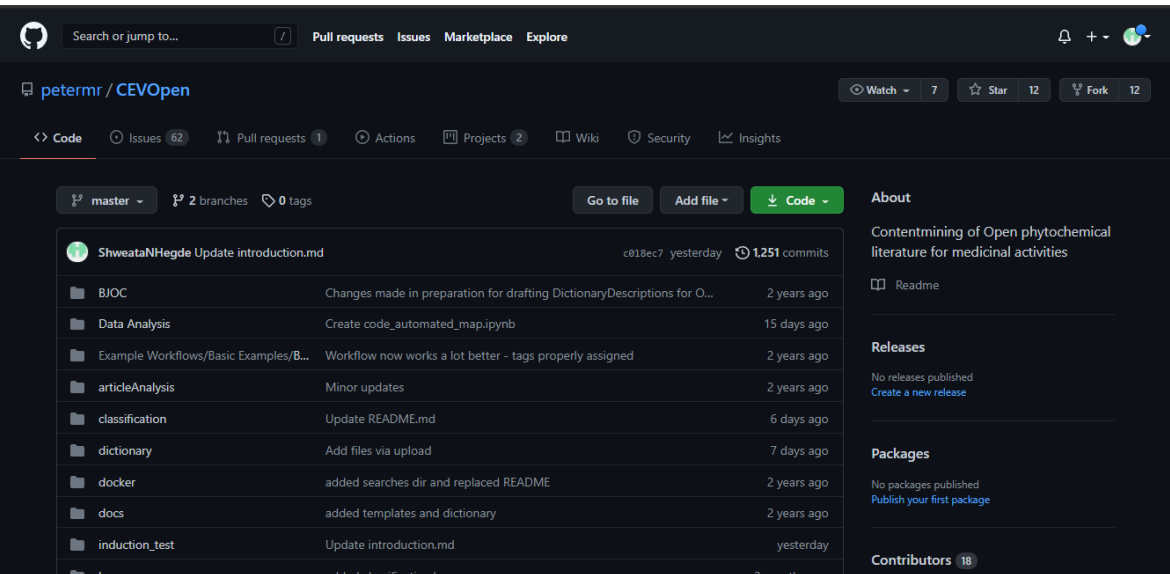
Open Science is...

definition

PMR would have or will talk more

Challenges

It's also about...



Open Notebook Philosophy

1.

How we work

Work updated in real-time!
Go back to any version!

You can try out, modify or access any material from our project!

<https://github.com/petermr/CEVOpen>
<https://github.com/petermr/pygetpapers>
<https://github.com/petermr/pyami>

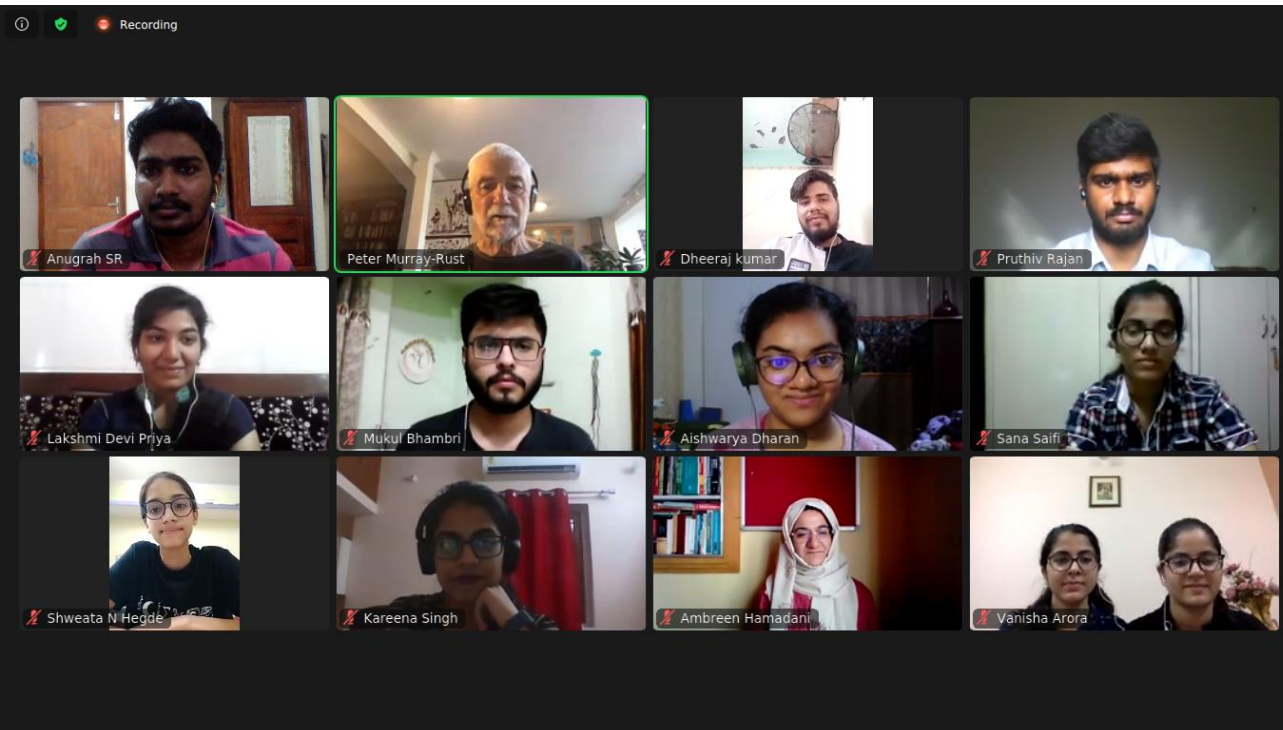
Community and Accountability



ISSUES	Ambreen	Vaishali	Priya	Rajan	Vanisha	Sana	Kareena	Charles
corpus (between 200-950)	Ready	Ready	Ready	Ready	Ready	Ready	Ready	Ready
dictionary	Ready	Ready	Ready	Ready	Ready	Ready	Ready	In progress
annotated (Pos/Neg) viral epidemics	8.90%	40/10	In progress	111/116	In progress	36/14	121/29	5/35
specialist subsets (e.g. methods) e.g. sections	Ready	Ready	Ready	Ready	Ready	Ready	Ready	Haven't Started
notebooks	Ready	In progress	In progress	In progress	In progress	Haven't Started	In progress	Haven't Started
machine learning / NLP	In progress	In progress	Haven't Started	In progress	In progress	In progress	In progress	In progress
display	In progress	Haven't Started	Haven't Started	Haven't Started	Haven't Started	Haven't Started	In progress	In progress
language variants	Ready	Haven't Started	Ready	Ready	Haven't Started	In progress	Haven't Started	Haven't Started
ami search	Ready	In progress	Ready	Ready	Ready	Ready	Ready	Haven't Started

Work as a community
Individual accountability

It's also about...



cambiohack

Putting every work you do
openly on the web increases
discoverability and
serendipitous encounters!

3.

Whom we include and collaborate with

Transition to Open Science
practice in *ANY* ways
possible or **promote!**

Context

Life Science Undergrad

Joined openVirus -> 11 months ago (Sep. 2020)

No coding experience!



You don't have to be an expert
to start out!

No restrictions!
No gatekeeping!

Reflecting on learnings!

Documenting your work **openly** helped me...

 Forgo *perfectionism*: Everything's work-in-progress

 Experiment: Document failures!

Reflecting on learnings!

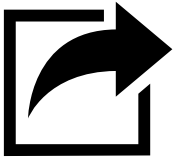
The **open science community** has helped me...



Learn: Coding, best-practices from others!



Repurpose: No going down the reinvention well again and again!



Share: Without worries!



Communicate: Reading and writing better documentation



Long way to go!
Grateful that OpenVirus and CEVOpen are a part
of my journey!

Extremely grateful to mentors and the Team!

Dr Peter Murray-Rust

Dr Gitanjali Yadav

OpenVirus Team

CEVOpen Team