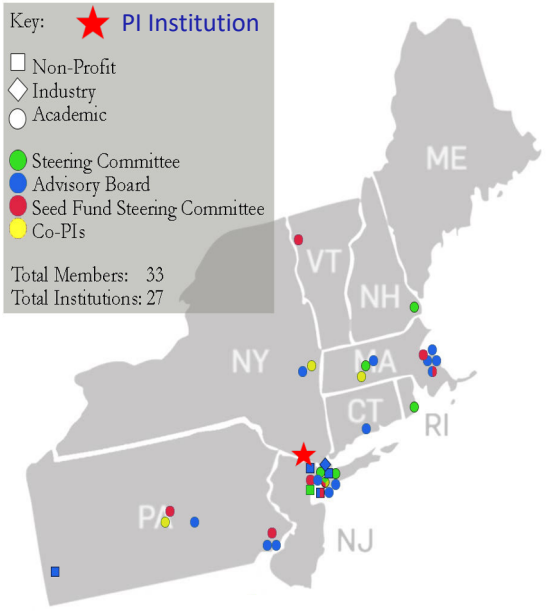
COVID Information Commons

COVID Information Commons

An Open Portal and Community for Research Collaboration

Florence D. Hudson
Executive Director and PI
COVID Information Commons
Northeast Big Data Innovation Hub at Columbia University

The Northeast Big Data Innovation Hub



- The Northeast Hub is a community convener, collaboration hub, and catalyst for data science innovation in the Northeast Region and around the world.
- Building a diverse, equitable, and inclusive community with accessible resources is central to our work.
- We are a community of over 8,300 individuals from 1,375 organizations, all 50 U.S. States and 61 countries. Join us!

Four Focus Areas guide community engagement and activities.

Education + Data Literacy

Health

Urban to Rural Communities

Responsible Data Science: Security + Privacy + Ethics

Funded by NSF grants #1550284, 1748395, 1916585, 2028999, 2139391, NIH AIM-AHEAD.

http://nebigdatahub.org/about/

BDHub Activities Leverage Data Science For Good

The CIC is the Northeast Hub's premier Health Focus Area Program

http://nebigdatahub.org/about/

The COVID Information Commons is supported by the National Science Foundation through awards #2028999 and #2139391.

NSF Convergence Accelerator

covidinfocommons.net | @CIC_COVID

COVID Information Commons (CIC) – Purpose + Plan

- An NSF-funded project designed to be a trusted resource for COVID research.
- NSF CIC RAPID Award funded in May 2020, Launched July 2020.
- Designed in collaboration with the 4 NSF Big Data Innovation Hubs (Northeast, Midwest, South, West), plus Columbia University Libraries and Columbia University IT.
- Designed to be Findable, Accessible, Interoperable, Reusable (FAIR).
- Includes NSF-funded COVID RAPID awards, leverages the NSF database and Machine Learning (ML) clustering tools.
- Uses Principal Investigator (PI) provided information to contextualize NSF abstracts, including ORCID ID, research websites, keywords.

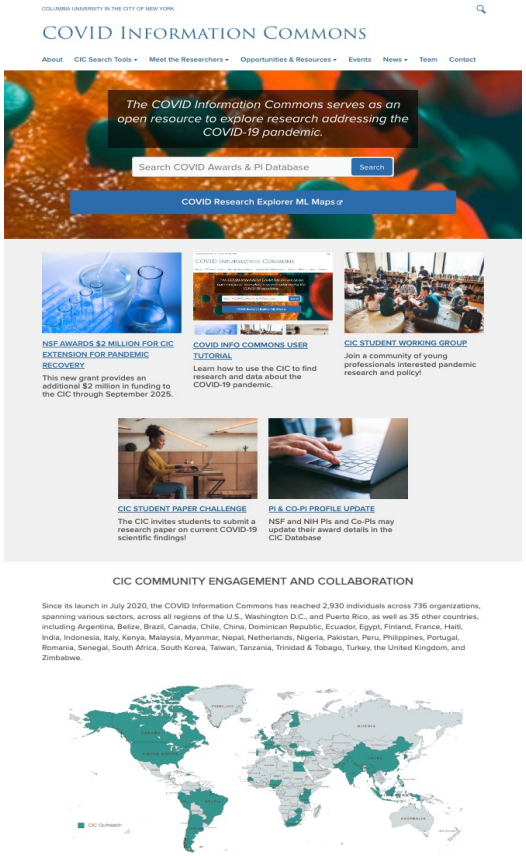
https://covidinfocommons.net

COVID Information Commons – Progress

- 2021 NSF $2M CIC Extension award funded significant expansion over 4 years.
  - ***10x corpus growth*** - now includes 10,000+ NSF ***and NIH*** COVID awards.
  - ***10x community growth*** - 2,930 individuals at 736 organizations in US + 35 countries.
  - Expanded CIC Student Paper Challenge
    - 3 undergraduate winners in 2021.
    - 5 winning teams in 2022 in both an undergraduate and a graduate student cohort.
    - Funded prizes for top 3 winning teams in each cohort (grad and undergrad).
  - English and Spanish transcripts/translations created with students and volunteers.
- Plan for advanced COVID data and metadata search and discovery tools.
- Plan to leverage DRYAD to archive the CIC corpus - https://datadryad.org/stash.

https://covidinfocommons.net

COVID Information Commons – The Portal



CIC portal includes

- 10,000+ NSF and NIH COVID-related Awards
- COVID Awards & PI database search tool
- COVID Research Explorer ML Maps tool
- PI-provided information on research websites, links to datasets, ORCID ID hotlinks
- User tutorial and demo videos
- Links to events, challenges and working groups
- Links to global resources
    - 80+ datasets from 6 continents
    - 60+ groups and guides

https://covidinfocommons.net

COVID Information Commons - The Community

" Your site and the ability to come together is marvelous. I thank you especially for thinking about this and bringing us together. People will be able to use your [CIC] site as a proper, safe, true information source." – Nora Garza, Laredo College

September COVID-19 Research Lightning Talks: Webinar and Q&A

July 2020 launch webinar

- 178 attendees + 2 PI lightning talks
- 40 more PIs offered to do lightning talks

Created CIC Community

- Now 2,930 CIC Community Members
- Members across the US + 35 countries

Ongoing CIC Community Webinars

- 26 CIC hosted webinars to date
- 128 researcher lightning talks
- 10,800+ total views (live + YouTube)

https://covidinfocommons.net

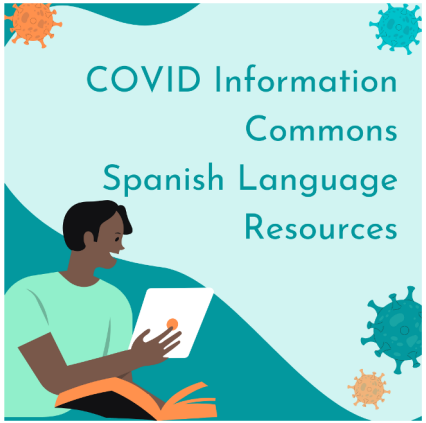
CIC Accessibility & Language Resources

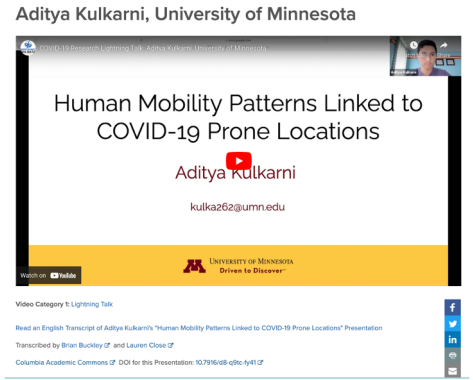
The lightning talk videos are posted on the NEBDHub **YouTube** website!

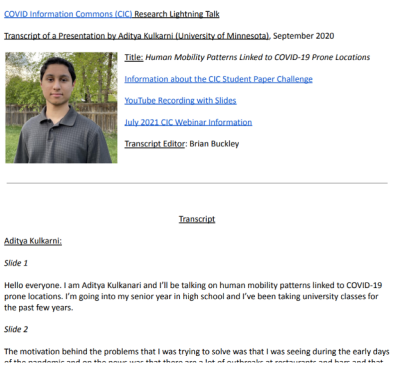
Review past talks and transcripts: bit.ly/cic-meet-the-researchers

With help from our students, staff, and volunteers, all talks are also available in **written English and Spanish**.

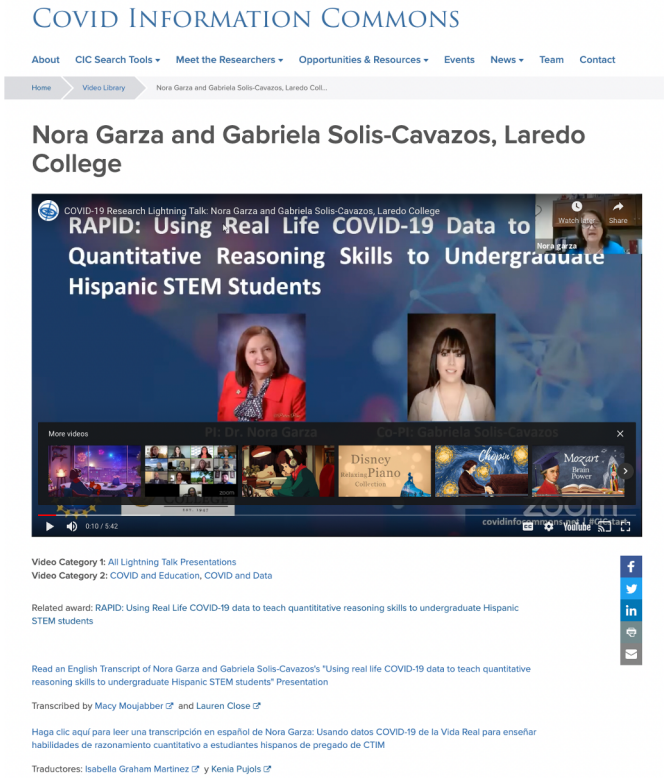
Volunteer to join the effort: info@covidinfocommons.net







Access CIC researcher Lightning Talk videos, English and Spanish Transcripts



Video Category 1: All Lightning Talk Presentations
Video Category 2: COVID and Education COVID and Data
Related award: RAPID: Using Real Life COVID-19 data to teach quantitative reasoning skills to undergraduate Hispanic STEM students

Read an English Transcript of Nora Garza and Gabriela Solis-Cavazos's "Using real life COVID-19 data to teach quantitative reasoning skills to undergraduate Hispanic STEM students" Presentation
Transcribed by Macy Moujabber and Lauren Close

Haga clic aquí para leer una transcripción en español de Nora Garza: Usando datos COVID-19 de la Vida Real para enseñar habilidades de razonamiento cuantitativo a estudiantes hispanos de pregado de CTIM
Traductores: Isabella Graham Martinez y Kenia Pujols

https://covidinfocommons.net

Researcher Lightning Talk videos, English and Spanish Transcripts Example

COVID Information Commons (CIC) Research Lightning Talk

Transcript of a Presentation by Nora Garza and Gabriela Solis-Cavazos (Laredo College), September 16, 2020

Title: RAPID: Using Real Life COVID-19 data to teach quantitative reasoning skills to undergraduate Hispanic STEM students

Nora R Garza CIC Profile

NSF Award #: 2032954

YouTube Recording with Slides

September 2020 CIC Webinar Information

Transcript Editor: Macy Moujabber

Transcript

Katie Naum:

Next Nora Garza and Gabriela Solis-Cavazos of Laredo College. So, I'll let you guys take it away.

Nora Garza:

Slide 1

Thank you very much and thank you all for hosting us with the COVID Information Commons. It's nice to be able to share. I'd also like to thank Laredo College and my RAPID team of faculty members that are part of this group and the RAPID research assistant students that are coming on board as we speak. Gabriela Solis-Cavazos is the co-PI for this grant. Also thank NSF and the HSI program which is funding this RAPID grant. Gabriela Solis is our undergraduate research coordinator for a grant that we have where students are already participating in undergraduate research experiences with faculty partners. So, Gabby for the next slide-

Centro de Información de COVID (CIC): Charlas científicas relámpago

Transcripción de una presentación de Nora Garza y Gabriela Solis-Cavazos (Laredo College), 16 de septiembre de 2020

Título: RAPID: Usando datos de la Vida Real COVID-19 para enseñar habilidades de razonamiento cuantitativo a estudiantes hispanos de pregrado de CTIM

Perfil de Nora R Garza en la base de datos del CIC

Subvención de La Fundación Nacional de Ciencias (NSF, por sus siglas en inglés) #: 2032954

Grabación de YouTube con diapositivas

Seminario web del CIC de Septiembre 2020

Editora de la transcripción: Macy Moujabber

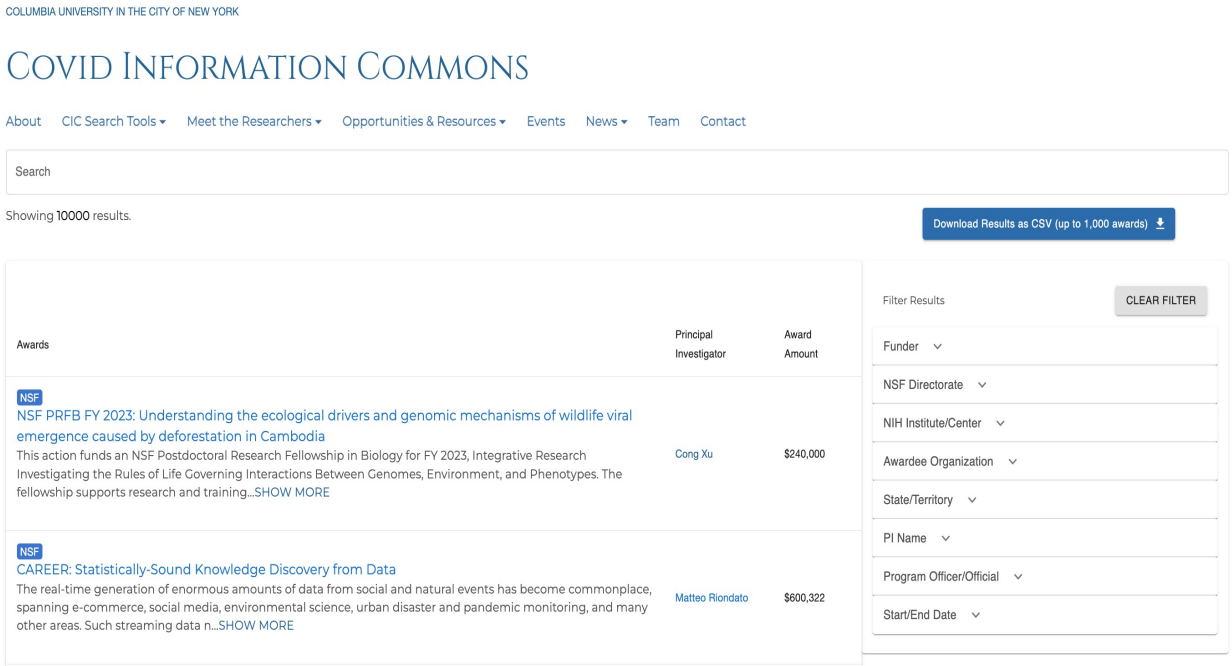
Traductora: Isabella Graham Martínez

Diapositiva 1

Muchas gracias y gracias a todos por recibirnos en el COVID Information Commons. Es muy bueno poder compartir. También me gustaría darles las gracias a Laredo College y a mi equipo RAPID de miembros de la facultad que forman parte de este grupo y a los estudiantes asistentes de investigaciones RAPID que se están uniendo al equipo. Gabriela Solis-Cavazos es la co-PI de esta subvención. También agradezco a NSF y al programa HSI que está financiando esta subvención RAPID. Gabriela Solís es nuestra coordinadora de investigación de pregrado para una subvención que tenemos donde los estudiantes ya están participando en experiencias de investigación de pregrado con socios de la facultad. Así que Gabby para la próxima diapositiva -

https://covidinfocommons.net

COVID Awards & Researcher (PI) Database



CIC Database includes 10,000+ NSF & NIH COVID-related awards.

Browse the Database: https://bit.ly/cic-award-search

Update your CIC PI Profile: https://cic-apps.datascience.columbia.edu/survey/

NSF + NIH COVID PI Profiles

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

COVID INFORMATION COMMONS

About | CIC Search Tools | Meet the Researchers | Opportunities & Resources | Events | News | Team | Contact

Florence D Hudson

Institution:
Columbia University

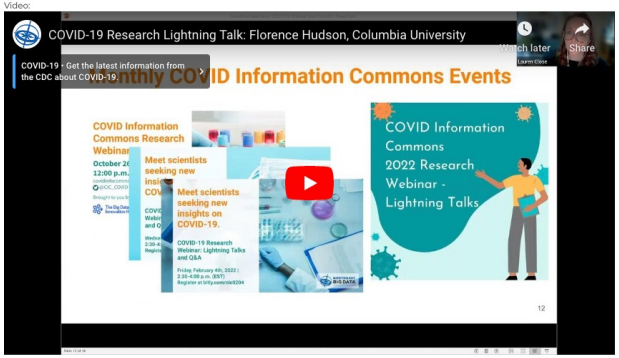
Email:
fh2417@columbia.edu

ORCID ID:
https://orcid.org/0000-0003-0896-2127

Websites:
https://covidinfocommons.net

https://nebigdatahub.org/about/

Video:

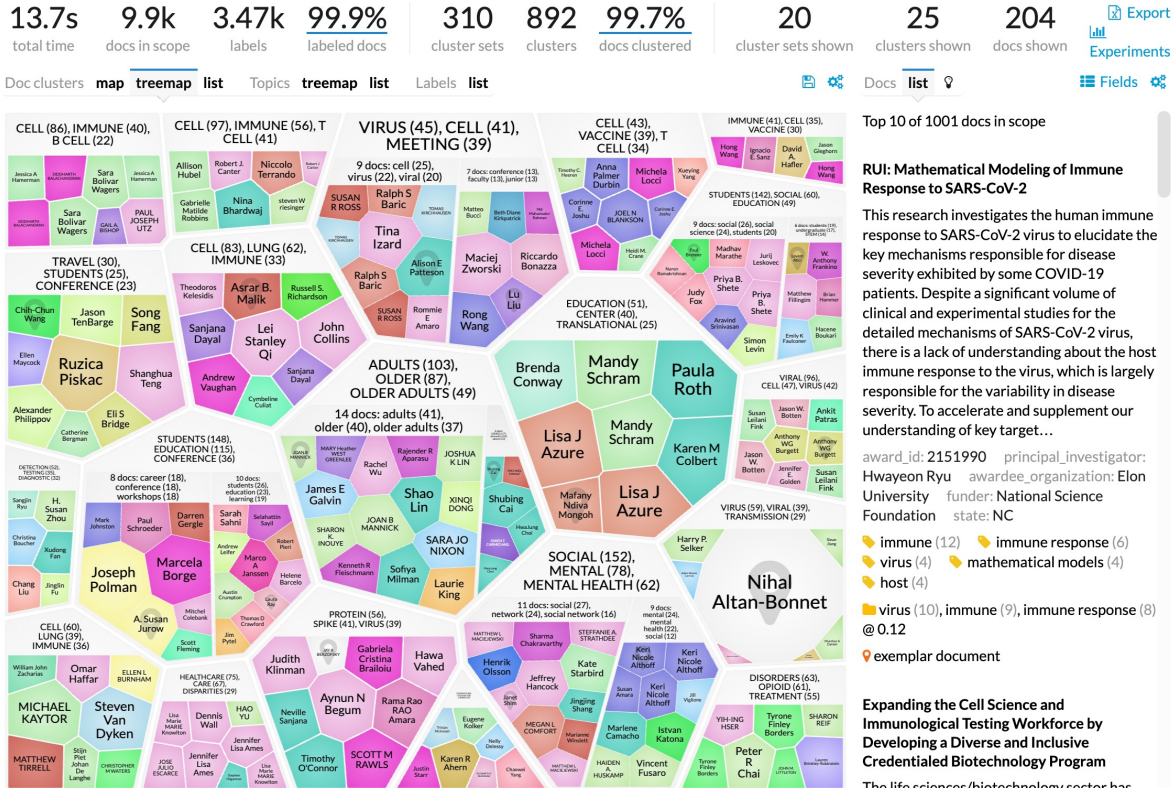


Awarded COVID Grants:
CIC-E: COVID Information Commons Extension for Pandemic Recovery

RAPID: COVID Information Commons (CIC)

Keywords:
Data science | analytics | health | COVID | coronavirus | cybersecurity | connected healthcare | Medical Internet of Things | IoT | TIPPSS | Trust | Identity

COVID Info Commons
Research Explorer Machine Learning Maps



- Find COVID research and researchers by Principal Investigator (PI) name, institution, award number, state
- Clustered by topic using Machine Learning
- Input domain areas or keywords to limit the view to target areas

Query the CIC by Keyword

Lingo4G Explorer

Analyze

Filters

find settings, e.g. 'overlap'

ANALYSIS SCOPE

Defines the set of documents to process.

Query

epidemiology

Limit scope size

Max number of documents to analyze

10 k

Query parser

enhanced

DOCUMENT MAP

Configuration of the documents 2d map.

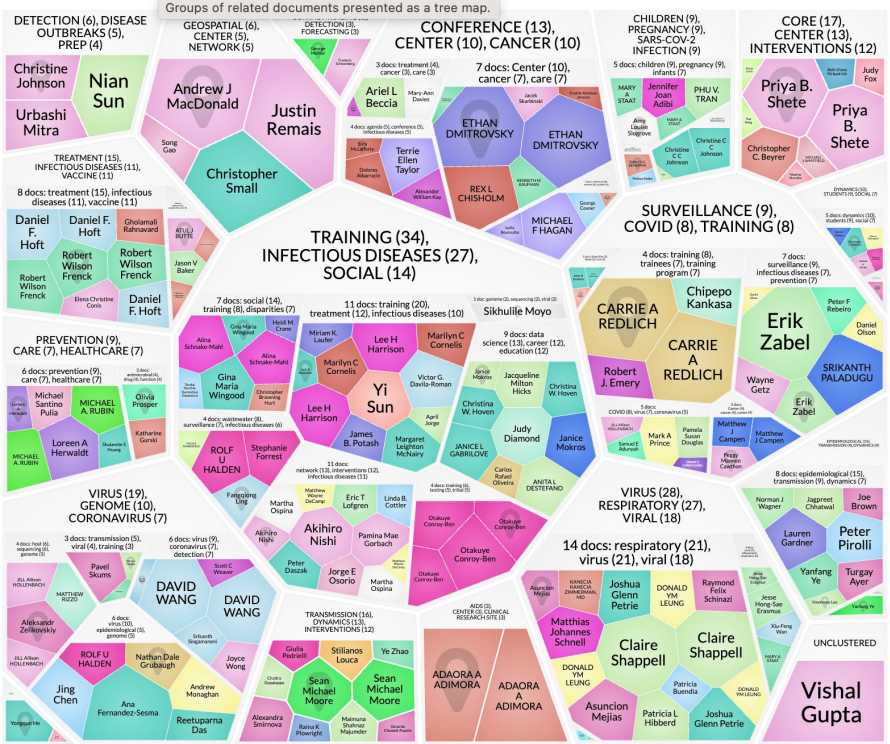
291ms total time | 387 docs in scope | 305 labels | 100% labeled docs | 18 cluster sets | 42 clusters | 99.5% docs clustered | 204 docs shown

Export

Experiments

Doc clusters map treemap list | Topics treemap list | Labels list

Groups of related documents presented as a tree map.



Docs list | Fields

Top 10 of 387 docs in scope

COVID-Inspired Data Science Education through Epidemiology

The ongoing COVID-19 pandemic provides a starting point for empowering young people to understand uses of data science through epidemiology. Through this program, 400 underserved youth nationwide will engage in a 15-hour out-of-school multimedia program centered on a project-developed text, The Case of the COVID Crisis, which is integrated with data activities, modeling, animations, and career exploration. Participants will: 1) Learn to use data tools and models to track the spread of infectious…

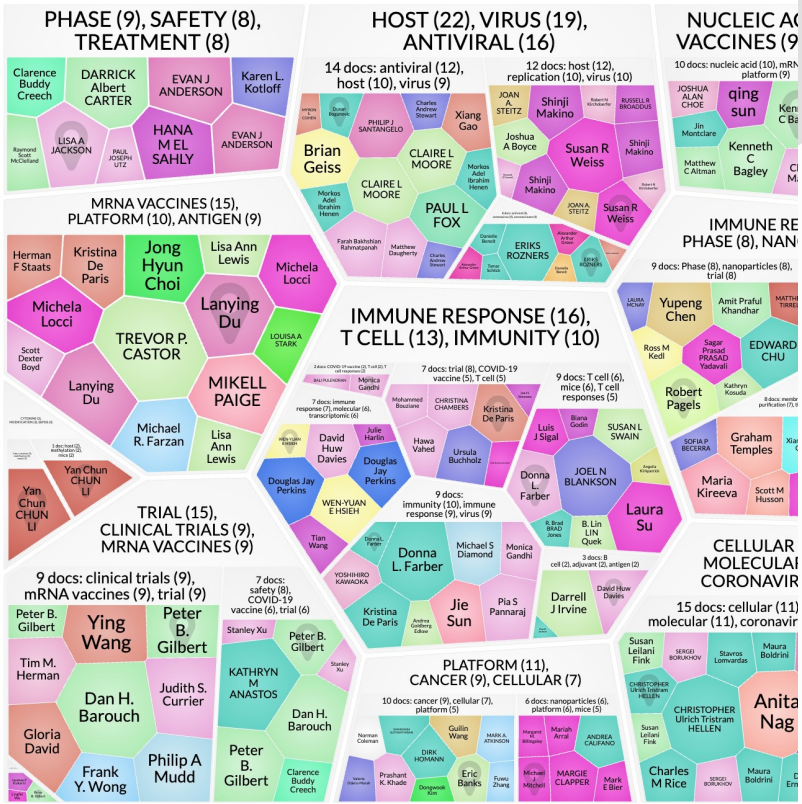
…; and 3) Gain confidence in their ability to use data to study and communicate to local audiences about epidemiological challenges. The program is also aimed at encouraging youth's interest in the myriad careers of the data-rich discipline of epidemiology. The project will achieve these goals through a multifaceted partnership involving Science Education Solutions, Tumblehome Books, Imagine Science/STEM Next, Concord Consortium, Jackson Laboratory (JAX), Strategic Learning Partners for…

…by: 1) Elucidating the ways in which youth use datasets and data tools to ask epidemiological questions, examine patterns, and make predictions; 2) Studying how youth become motivated to engage in work in the

CIC Research ML Explorer - View by PI or Institution

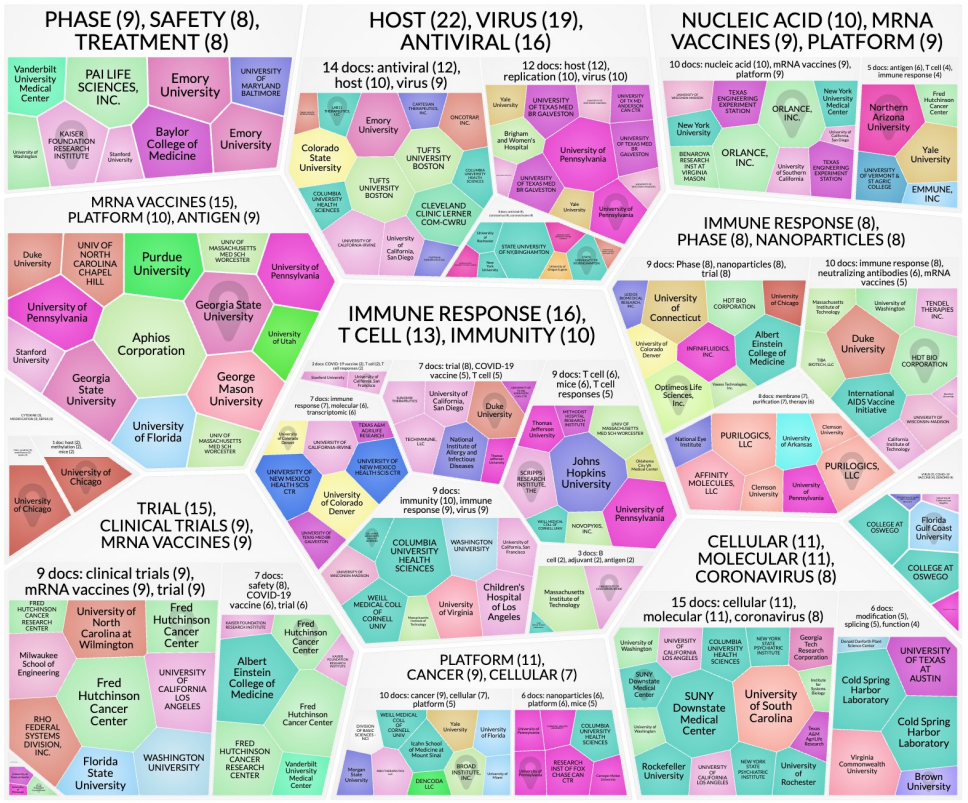
0ms	214	195	100%	11	26	100%
total time	docs in scope	labels	labeled docs	cluster sets	clusters	docs clustered

Doc clusters **map** **treemap** **list** Topics **treemap** **list** Labels **list**



0ms	214	195	100%	11	26	100%	200
total time	docs in scope	labels	labeled docs	cluster sets	clusters	docs clustered	docs shown

Doc clusters **map** **treemap** **list** Topics **treemap** **list** Labels **list**



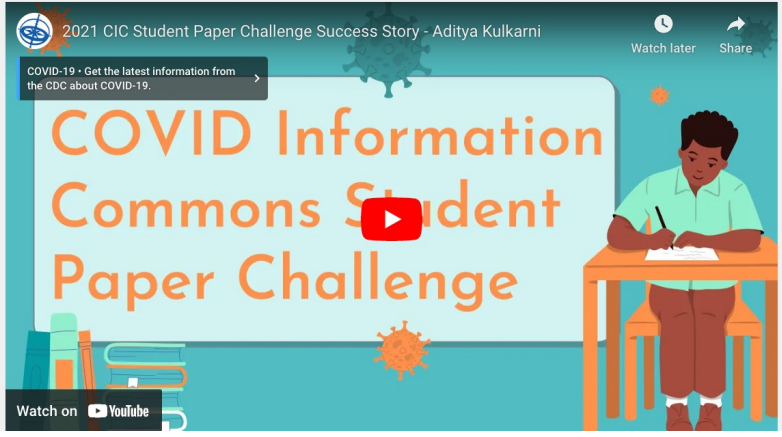
CIC Student Paper Challenge

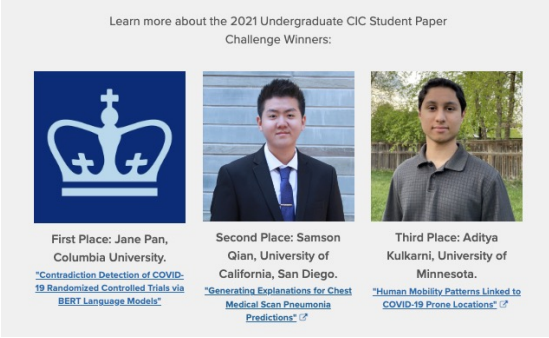
2021 & 2022 CIC STUDENT PAPER CHALLENGE WINNERS

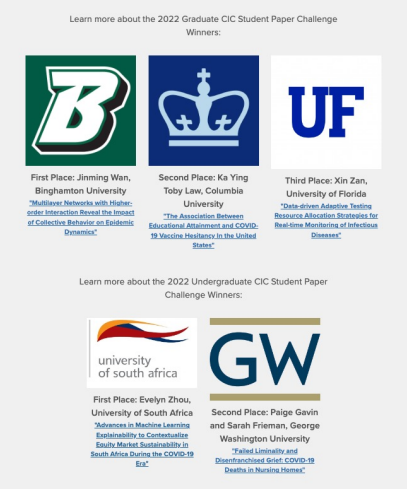
The inaugural CIC Undergraduate Student Paper Challenge was launched in December 2020 to bring students into the CIC community and encourage their professional development and training for their research careers.

In 2022, the CIC hosted its second annual Student Paper Challenge, broadening its reach to undergraduate, graduate, and recently-graduated students. With students participating from around the globe, the winning submissions represented higher education institutions from across the U.S. and South Africa.

The student winners were offered the opportunity to present their research in a virtual setting to the Northeast Big Data Innovation Hub and COVID Information Commons communities. The feedback was overwhelmingly positive: one of the student presenters was offered research funding by one of the PIs attending the webinar, and a potential internship in 2022 at another university.







https://covidinfocommons.datascience.columbia.edu/content/2023-cic-student-paper-challenge

2023 CIC Student Paper Challenge

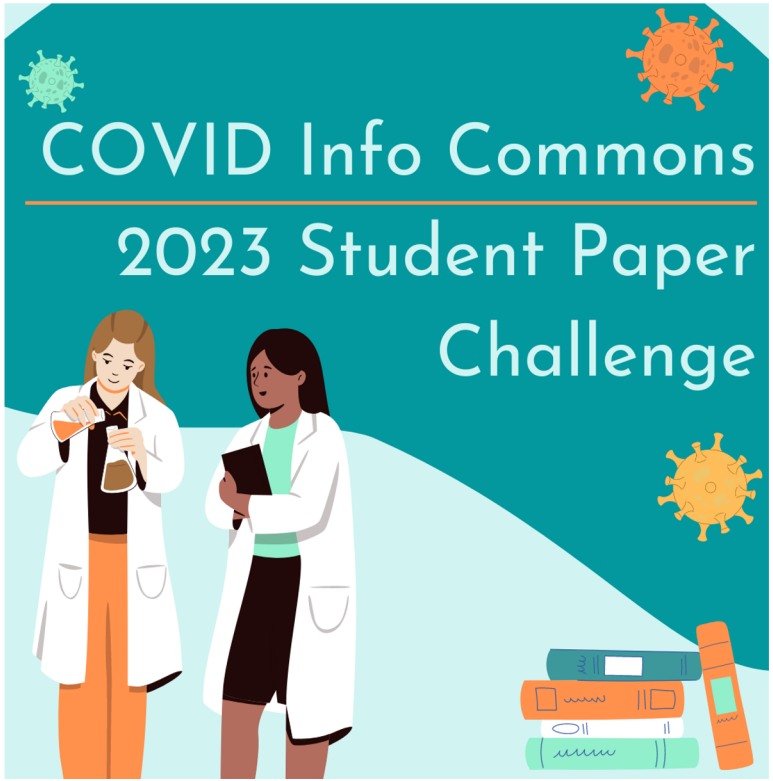
Undergraduate, community college, graduate students and recent graduates are invited to investigate and synthesize COVID-19 research into a short 4 to 5 page paper.

Winners will receive a prize and may be offered the chance to present their work at a future symposium.

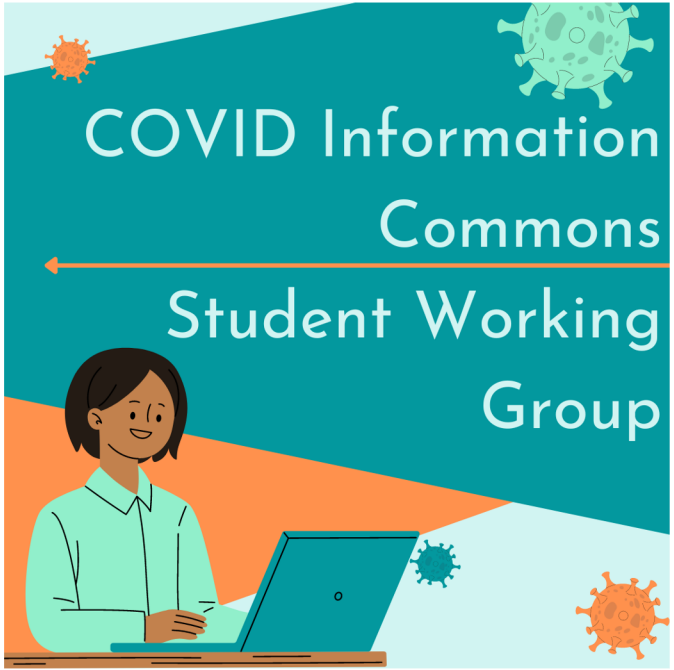
Monthly virtual student mentoring sessions.

Professionals and grad students invited to be mentors and judges!

Final submissions due **August 14th, 2023**!
Learn more: https://bit.ly/2023-cic-spc



CIC Student Working Group



A **community of young professionals and researchers** interested in the future of COVID-19 research begun in late 2022 to further student research and collaboration.

Biweekly Meetings

Website: bit.ly/CICStudentWG
Mailing List: bit.ly/wg-email-signup
Slack: bit.ly/slack-cic

CIC Research Working Group

Join a collaborative **Working Group of scientists and researchers** interested in the future of COVID-19 research.

Monthly Meetings

Website: bit.ly/CICResearchWG
Mailing List: bit.ly/wg-email-signup
Slack: bit.ly/slack-cic

COVID-19 Research Webinars: Lightning Talks and Q&A

Meet the scientists seeking new insights on COVID-19

April 2023 CIC Lightning Talks

Ellen Foxman, Yale University. Host response-based screening for unexpected or emerging respiratory viruses. Funded by NIH National Institute of Allergy and Infectious Diseases.

Ioannis Paschalidis, Boston University. Predictive Models of COVID-19 Severity and Patient Outcomes. Funded by NSF Computer and Information Science and Engineering (CISE) PIPP Phase 1.

Jeffrey Townsend, Yale University. Infection by SARS-CoV-2 with alternate frequencies of mRNA vaccine boosting. Funded by NSF Biological Sciences.

Chang-Yu Wu, University of Miami. Environmental Surveillance to Assess Aerosol Transmission Pathways of COVID-19 Enabled by On-The-Spot Sampling and Detection. Funded by NSF Engineering.

Niema Moshiri, University of California, San Diego. Massively scalable reference-guided Multiple Sequence Alignment of viral genomes. Funded by NSF Biological Sciences.

July 2023 CIC Lightning Talks

- **Carlos Badenes-Olmedo**, Universidad Politécnica de Madrid. Drugs4Covid: Knowledge Graph about Drugs used in the Clinical Control of the Coronavirus.
- **Niu Gao**, Public Policy Institute of California. Impact of COVID-19 on Science Education: Early Evidence from California. Funded by NSF Education and Human Resources.
- **Hong Qin**, University of Tennessee, Chattanooga. PIPP Phase I: Develop and Evaluate Computational Frameworks to Predict and Prevent Future Coronavirus Pandemics. Funded by NSF Computer and Information Science and Engineering (CISE).
- **Evelyn Yemurai Zhou**, University of South Africa. Advances in Machine Learning Explainability to Contextualize Equity Market Sustainability in South Africa During the COVID-19 Era. 2022 COVID Information Commons Undergraduate Student Paper Challenge 1st Place Winner.
- **Xin Zan**, University of Florida. Data-driven Adaptive Testing Resource Allocation Strategies for Real-time Monitoring of Infectious Diseases. 2022 COVID Information Commons Graduate Student Paper Challenge 3rd Place Winner.

Thank you!

Florence D. Hudson, Executive Director and PI
NSF Northeast Big Data Innovation Hub and COVID Information Commons
Florence.Hudson@columbia.edu

Stay in the loop about upcoming events: covidinfocommons.net/events

Sign up for our mailing list: bit.ly/HubNewsSignup

Join the CIC community on Slack: bit.ly/joincicslack

Follow us on Twitter: @CIC_COVID, #cicwebinar

Email us if you are interested in presenting
your NSF or NIH-funded COVID research at a future CIC webinar:
info@covidinfocommons.net