UNIVERSITY AT ALBANY
State University of New York

Bringing Artificial Intelligence to Research Analytics

Research Highlighter-MatchMaker at SUNY UAlbany

Office of Strategic Initiatives, Data, Assessment, and Technology

Director: Tianning Huang

Data Analyst: Yuemin Li

© University at Albany, SUNY, 2024

Funding Opportunity Challenges

- **Lower success rate**
  - "In 2000, more than 30 percent of NIH grant applications got approved. Today, it's closer to 17 percent." (Belluz, Plumer, and Resnick 2016)
- **Greater pressure**
  - "In recent years, faculty have felt increasing pressure to obtain external funding to support their research." (Waaijer et al., 2018)
- **Time consuming**
  - "By some estimates, many top researchers spend 50 percent of their time writing grants." (Lofgren 2021)

UNIVERSITY AT ALBANY
State University of New York

Finding Funding Opportunities with Perfect Match

But How?

Current Solution: Funding Opportunity Databases

UNIVERSITY AT ALBANY
State University of New York

SP

SPI
spo
fou
staf
Univ

SPIN World's Largest Database of Sponsored Funding Opportunities

Login

Sign in using SPIN™ credentials

Username

Password

Remember me? What's this?

Login

Sign In Help

Need to create a new profile?

OR

Sign in using institutional credentials

InfoEd Global became an authorized service provider with the InCommon Federation to provide SPIN™ users with easier access. If your institution appears below or in the pick list, you may login using your institutional credentials. If your institution does not appear, contact your institutional SPIN™ administrator to inquire about InCommon federated access.

Select your organization from the list below:

University at Albany, State University of Ne...

Login

Help

UNIVERSITY AT ALBANY
State University of New York

PIVOT

Pivot is a database containing funding opportunities for all disciplines and project types. It includes information on federal, private, and international funding options all in one easy to search tool.

https://www.youtube.com/watch?v=Ko6hpM3L4TU&ab_channel=Pivot-RP

ExLibris | Pivot-RP

Issues with Funding Opportunity Databases

- Unawareness of funding opportunity databases
  - Miss the benefits provided by databases
- Keyword match recommendations
  - Possibility to miss the matched grants without matched keyword
  - Not taking personal research track records into consideration
- Faculty will need to monitor funding opportunities on their own, i.e. doing personal search
  - Time-consuming to do the search
  - Possibility to miss the matched grants

UNIVERSITY AT ALBANY
State University of New York

What is *Research Highlighter-MatchMaker*?

Match Research Strength with Federal Funding Priorities

- Identify critical mass with publication and proposal data
- Text mining of funding opportunities
- Make strategic decisions in overlapping area

© University at Albany, SUNY, 2024

UNIVERSITY AT ALBANY
State University of New York

What is *Research Highlighter-MatchMaker*?

- Identify research areas based on publication/proposal data
- Suggest potential research collaborations
- Recommend suitable funding opportunities by matching research track records

© University at Albany, SUNY, 2024

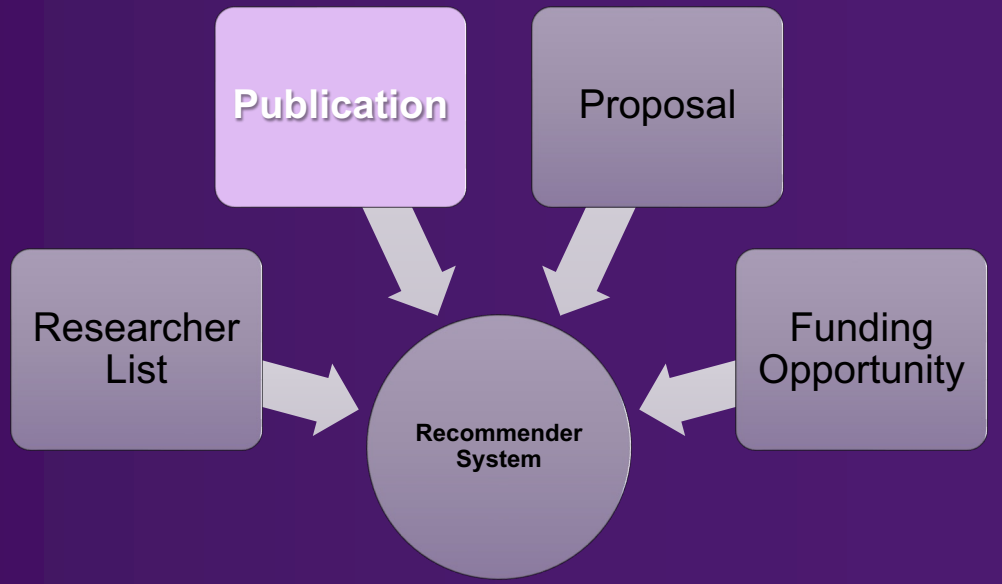
Collaborations

- The Office of Research, Innovation and Economic Development (ORIED) at the Research Foundation for SUNY
  - Temporary Pilot: Access to UAlbany's publication data by Elsevier's Scopus database
  - Submitted to NSF Growing Research Access for Nationally Transformative Equity and Diversity (GRANTED) Funding Opportunity
    - *"Growing Post-Award Research Support and Service Infrastructure for Nationally Transformative Equity and Diversity"*
  - Potential with expansion to SUNY system

© University at Albany, SUNY, 2024

Data

- Data Sources
  - Scopus publication abstracts (weekly update, temporary access)
    - *Manually identify Scopus ID of UAlbany researchers*
    - *Scopus API to extract publication abstracts*
    - *Cleaning: remove duplicated records*
    - *11,357 Scopus abstracts*

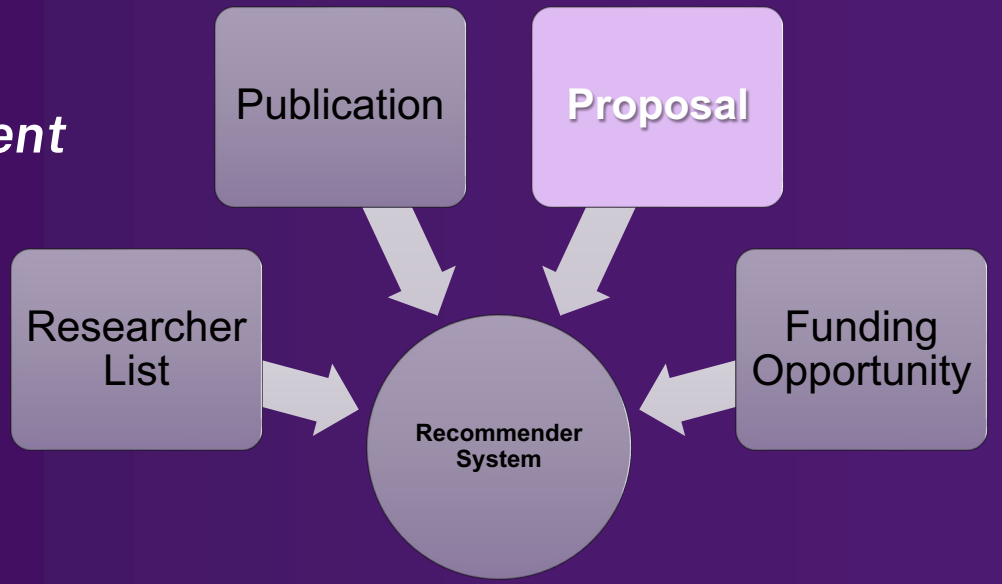


** Model run on 2/22/2024*

© University at Albany, SUNY, 2024

Data

- Data Sources
  - Scopus publication abstracts (weekly update, temporary access)
    - *Manually identify Scopus ID of UAlbany researchers*
    - *Scopus API to extract publication abstracts*
    - *Cleaning: remove duplicated records*
    - *11,357 Scopus abstracts*
  - Proposal abstracts (weekly update)
    - *Manually download from PACS, July 2021 – current*
    - *3,628 Proposal abstracts*

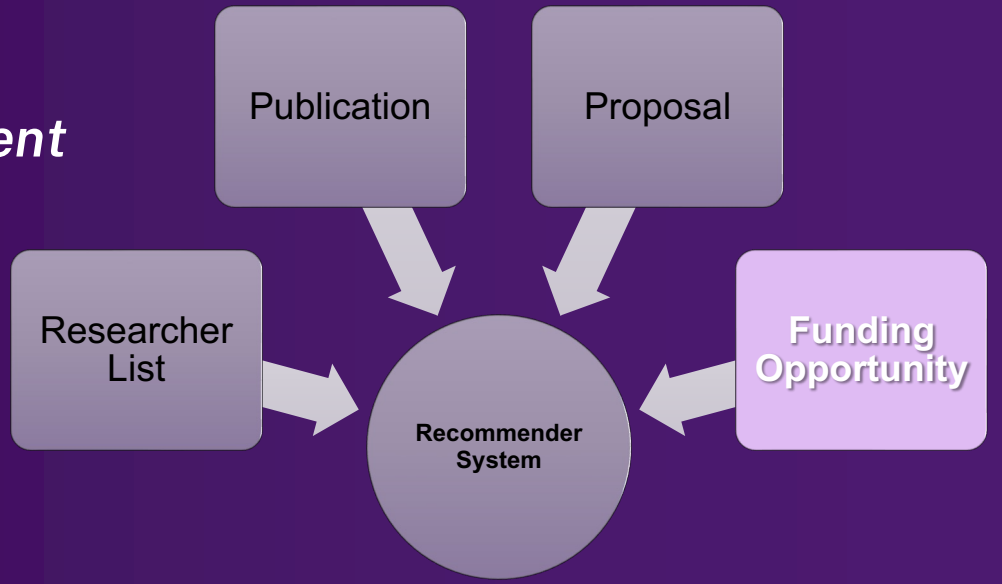


** Model run on 2/22/2024*

© University at Albany, SUNY, 2024

Data

- Data Sources
  - Scopus publication abstracts (weekly update, temporary access)
    - *Manually identify Scopus ID of UAlbany researchers*
    - *Scopus API to extract publication abstracts*
    - *Cleaning: remove duplicated records*
    - *11,357 Scopus abstracts*
  - Proposal abstracts (weekly update)
    - *Manually download from PACS, July 2021 – current*
    - *3,628 Proposal abstracts*
  - SPIN funding opportunities (daily update)
    - *Automatically update from SPIN API*
    - *1,901 records*
  - Total 16,828 records*

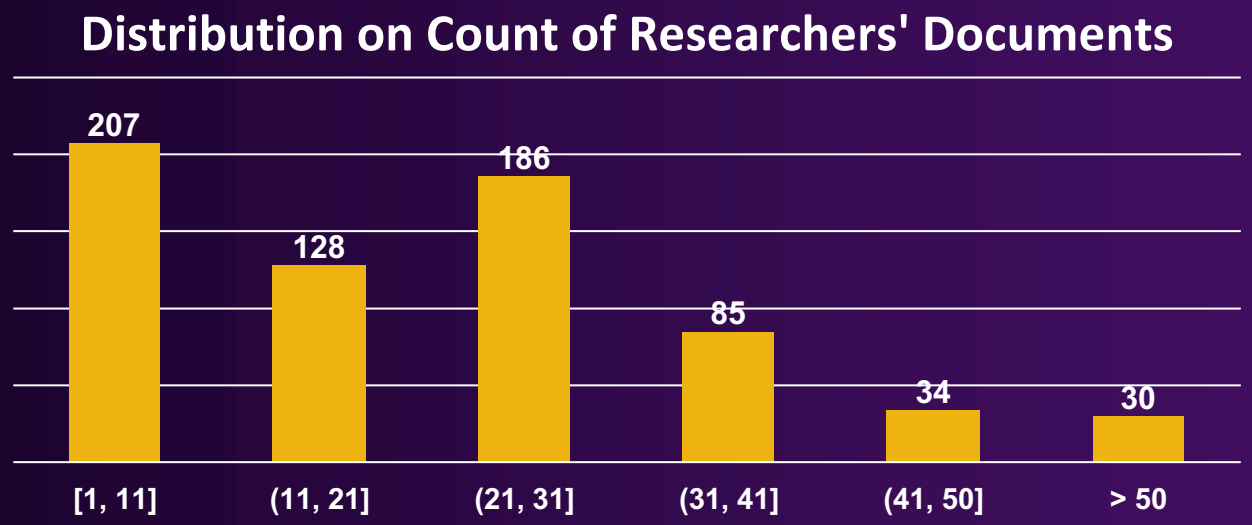


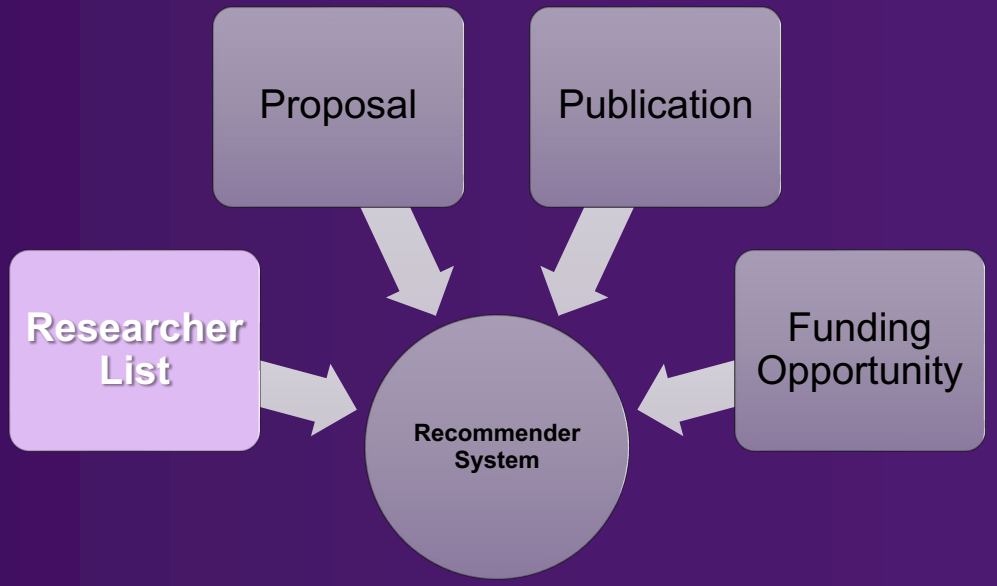
** Model run on 2/22/2024*

© University at Albany, SUNY, 2024

Data

- Data Sources
  - UAlbany Researchers: 711*
  - Titles: TT-track faculty (Assistant/Associate/Full Professor), Librarian, Research Scientist/Associate, Dean/Director/Chair
  - Quarterly update



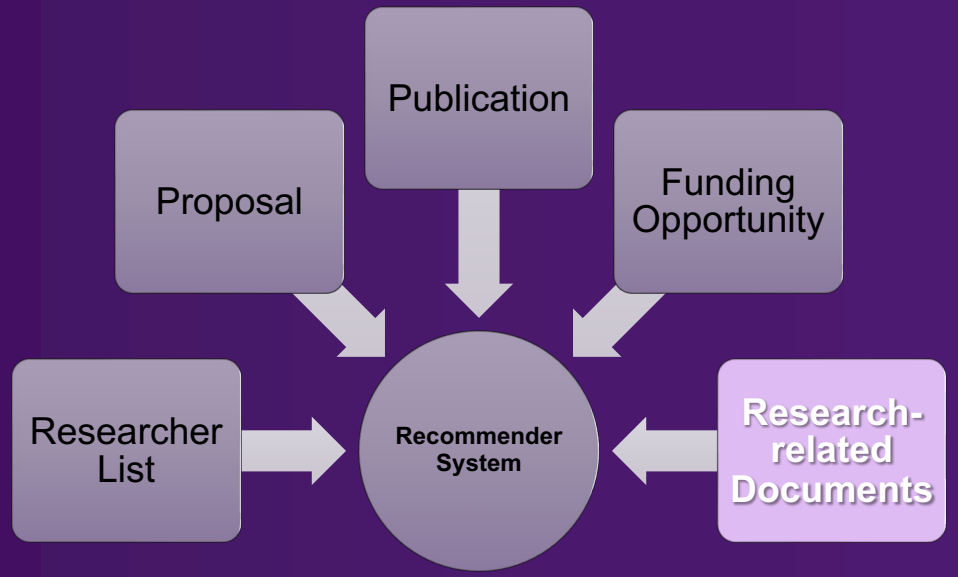


** Model run on 2/22/2024*

© University at Albany, SUNY, 2024

Data

- Potential Data Sources
- Other research-related documents
  - IRB
  - Technology Transfer/IP
  - Syllabus
- ORCID
  - Affiliate to auto-populate data
  - Expand scope of research data
  - Validity, Reliability, Consistency

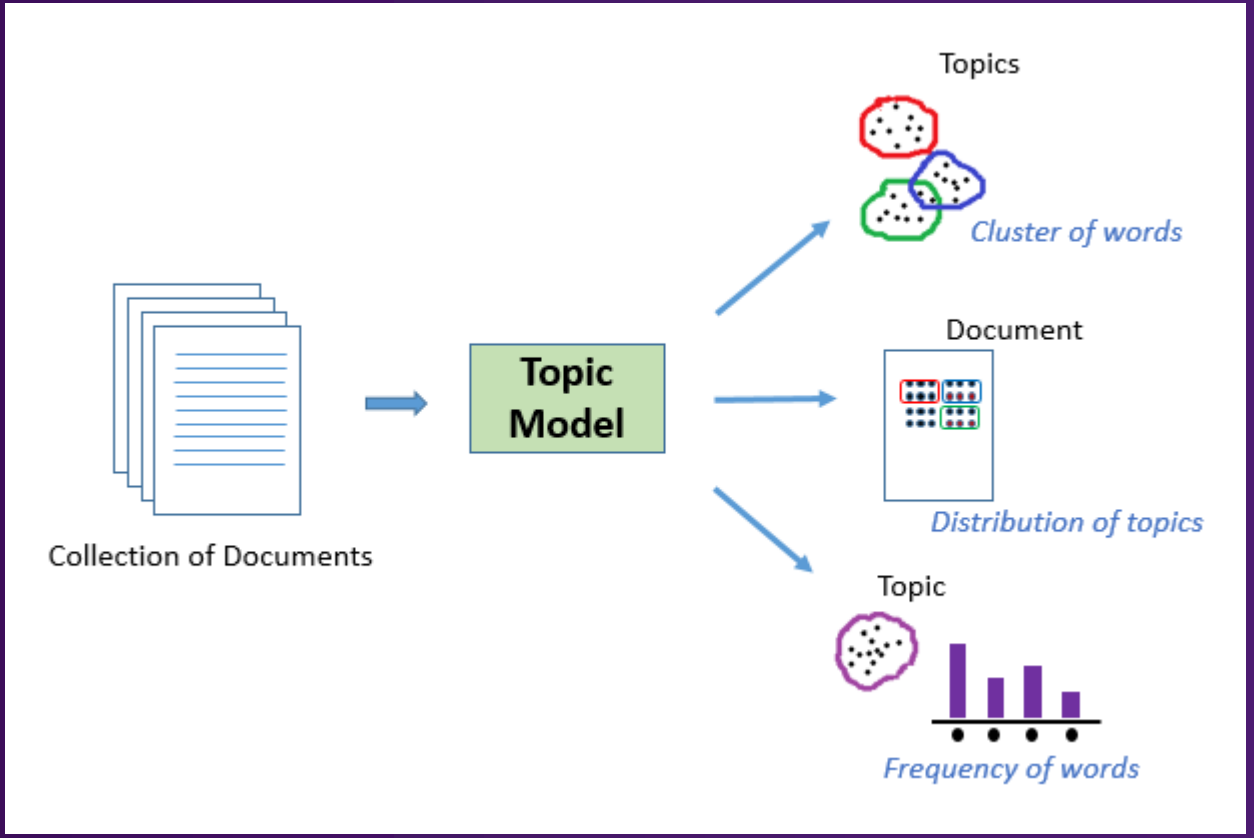


© University at Albany, SUNY, 2024

Method – Topic Modeling

- Unsupervised Learning: machine learning algorithms to analyze and cluster unlabeled datasets

- uses pre-trained large language model to classify documents into easily interpretable topics based on context of words in each document

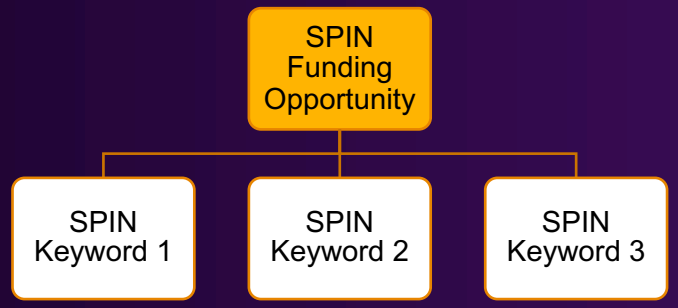


© University at Albany, SUNY, 2024

Method – Keyword Matrix

- **Keyword Algorithm**
  - SPIN Keyword System: Hierarchy Categorization

  - Each funding opportunity from SPIN is tagged with a set of keywords.



Exact String Match:
Search through publication/proposals data to find the matched keywords

© University at Albany, SUNY, 2024

Method – Keyword Matrix

- **Keyword Algorithm**
  - Calculate the matching score
    - #times this match occurs X
    - whether the author is first author X
    - #matched document/total document count
    - *i.e. AI appears 3 times in 1 researcher's 1 out of 40 records as the first author: (100+50+25) X 3 X 1/40*
  - Recommend funding opportunities based on this score could provide more personalized recommendations.

© University at Albany, SUNY, 2024

UNIVERSITY AT ALBANY
State University of New York

What is *Research Highlighter-MatchMaker*?

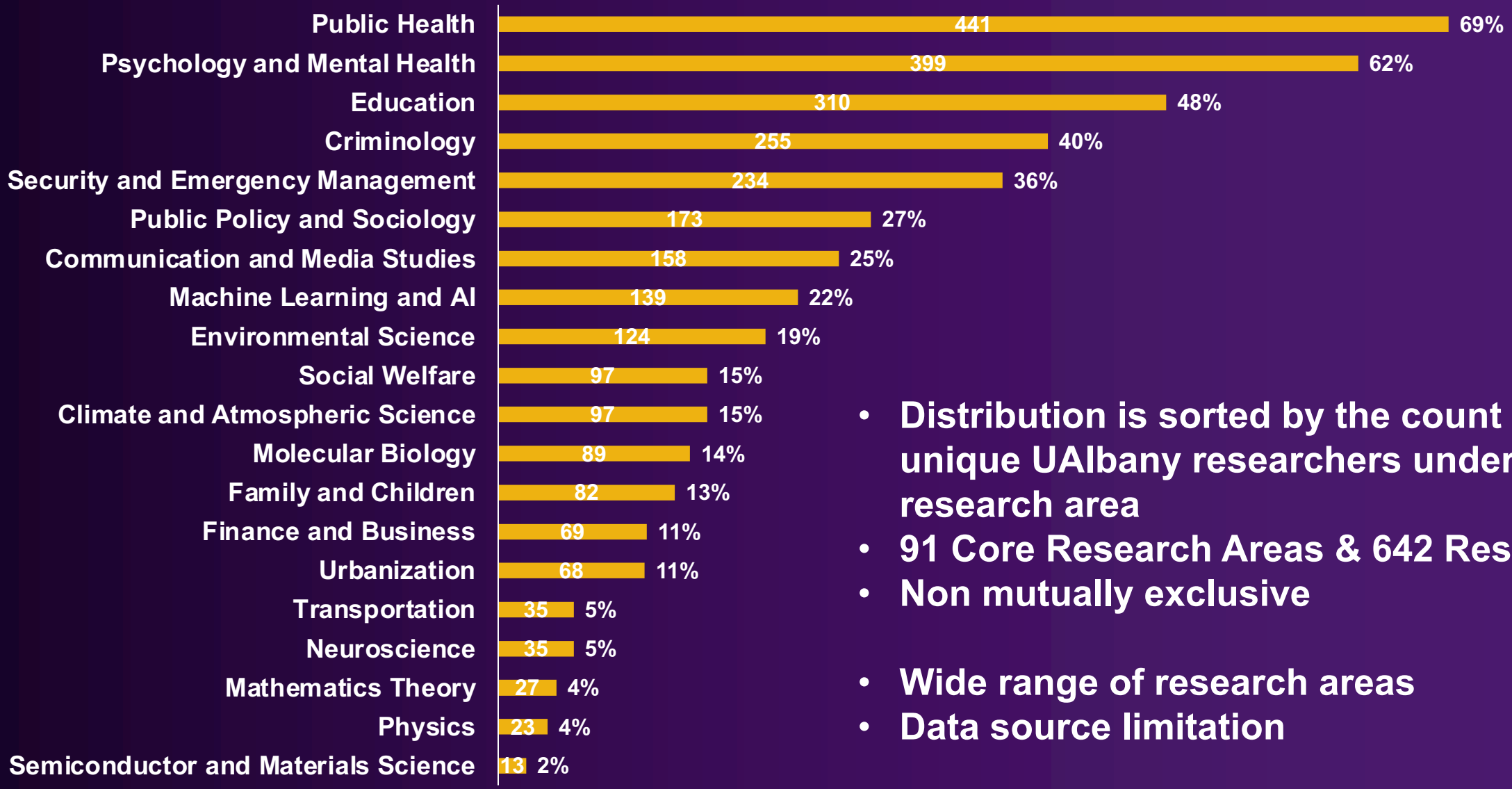
- **Identify research areas based on publication/proposal data**
- **Suggest potential research collaborations**
- **Recommend suitable funding opportunities by matching research track records**

© University at Albany, SUNY, 2024

UNIVERSITY AT ALBANY
State University of New York

Result

Topic Distribution By UAlbany Researchers



- Distribution is sorted by the count of unique UAlbany researchers under each research area
- 91 Core Research Areas & 642 Researchers
- Non mutually exclusive

- Wide range of research areas
- Data source limitation

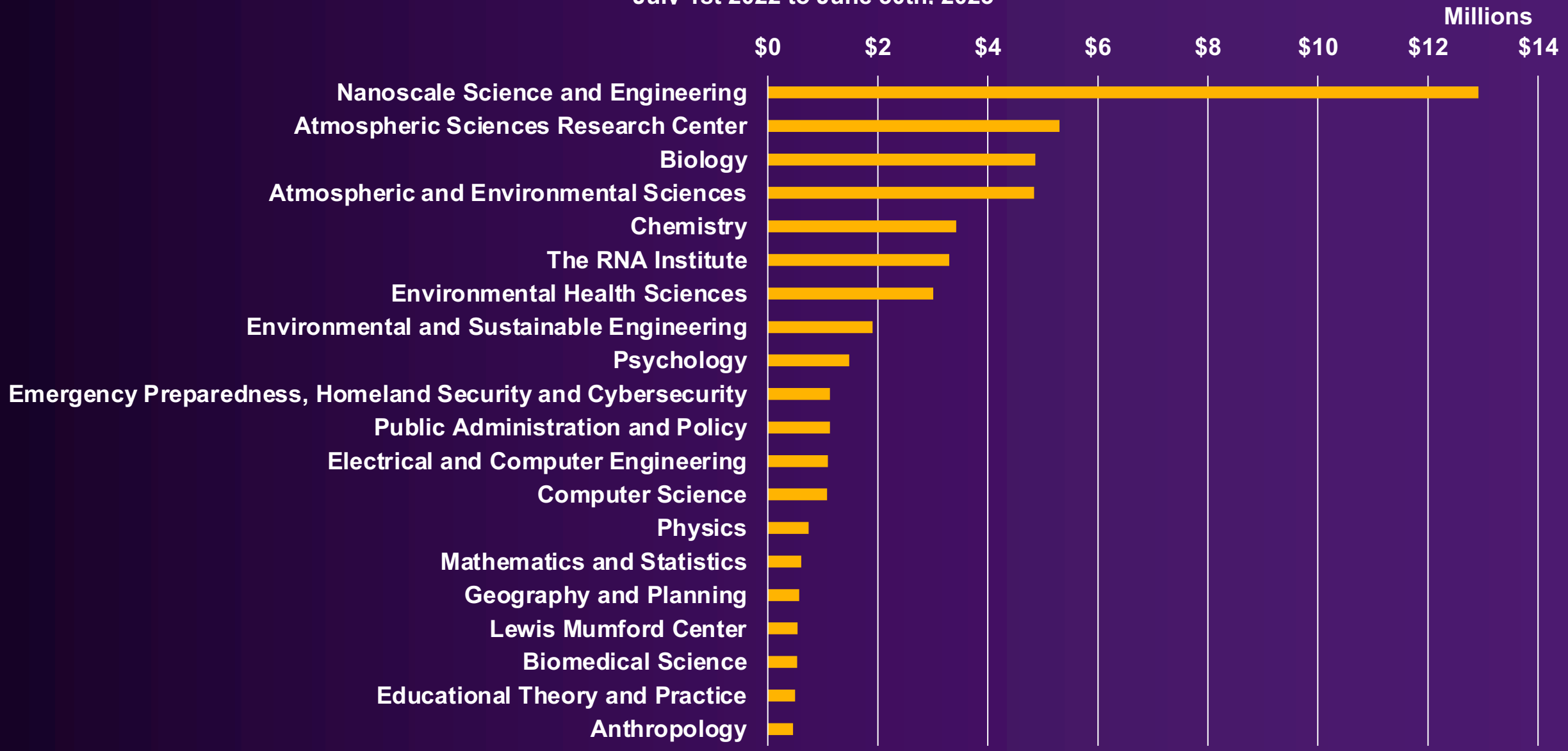
** Model run on 2/22/2024*

© University at Albany, SUNY, 2024

UNIVERSITY AT ALBANY
State University of New York

Sponsored Program Expenditures by Top 20 Unit in FY2022-23

* Julv 1st 2022 to June 30th. 2023



UNIVERSITY AT ALBANY
State University of New York

NSF HERD R&D Expenditures by Fields in FY2022-23 (Preliminary)*

** Department of Nanoscale Science and Engineering is not included.*

UNIVERSITY AT ALBANY
State University of New York

What is *Research Highlighter-MatchMaker*?

- Identify research areas based on publication/proposal data
- Suggest potential research collaborations
- Recommend suitable funding opportunities by matching research track records

© University at Albany, SUNY, 2024

UNIVERSITY AT ALBANY
State University of New York

Pilot Testing

Weekly Email Recommendation to UAlbany Researchers

The Office of Strategic Initiatives, Data, Assessment, and Technology in the Division for Research and Economic Development is launching a pilot testing of our Research Highlighter-Matchmaker project that aims to provide tailored funding opportunities for UAlbany researchers like you based on previous publication and grant proposal track records using AI/NLP techniques.

The following funding opportunities are recommended to you based on your research track record and the latest funding opportunities from SPIN in **Immunology**, e.g., rage, hiv, hcv.

- RFA-AI-23-039 -- Advancing Development of Diagnostics for Congenital and Adult Acquired Syphilis (R21 Clinical Trial Not Allowed), National Institute of Allergy and Infectious Diseases/NIH/DHHS
- Notice of Special Interest (NOSI): Enhancing Research on Deciphering Mechanisms of COVID-19-Associated Coagulopathy, National Institutes of Health/DHHS
- Notice of Special Interest (NOSI): Advancing Research in Lipoprotein(a) and Cardiovascular Disease (R01 Clinical Trials Allowed) and (R01 No Clinical Trials Allowed), National Heart, Lung, and Blood Institute/NIH/DHHS
- Notice of Special Interest (NOSI): Enhancing Research on Deciphering Mechanisms of COVID-19-Associated Coagulopathy, National Institutes of Health/DHHS
- Notice of Special Interest (NOSI): HIV Drug Resistance Assays and Actionable Data Dissemination Strategies, National Institutes of Health/DHHS

If you are interested in pursuing any of the funding opportunities above, please submit a Proposal Request Form at your earliest convenience.

Should you have any questions, please do not hesitate to contact sdat@albany.edu. We will soon launch a web-based application that will allow you to see additional matching results and your colleagues on campus who have shared research interests with you in the near future.

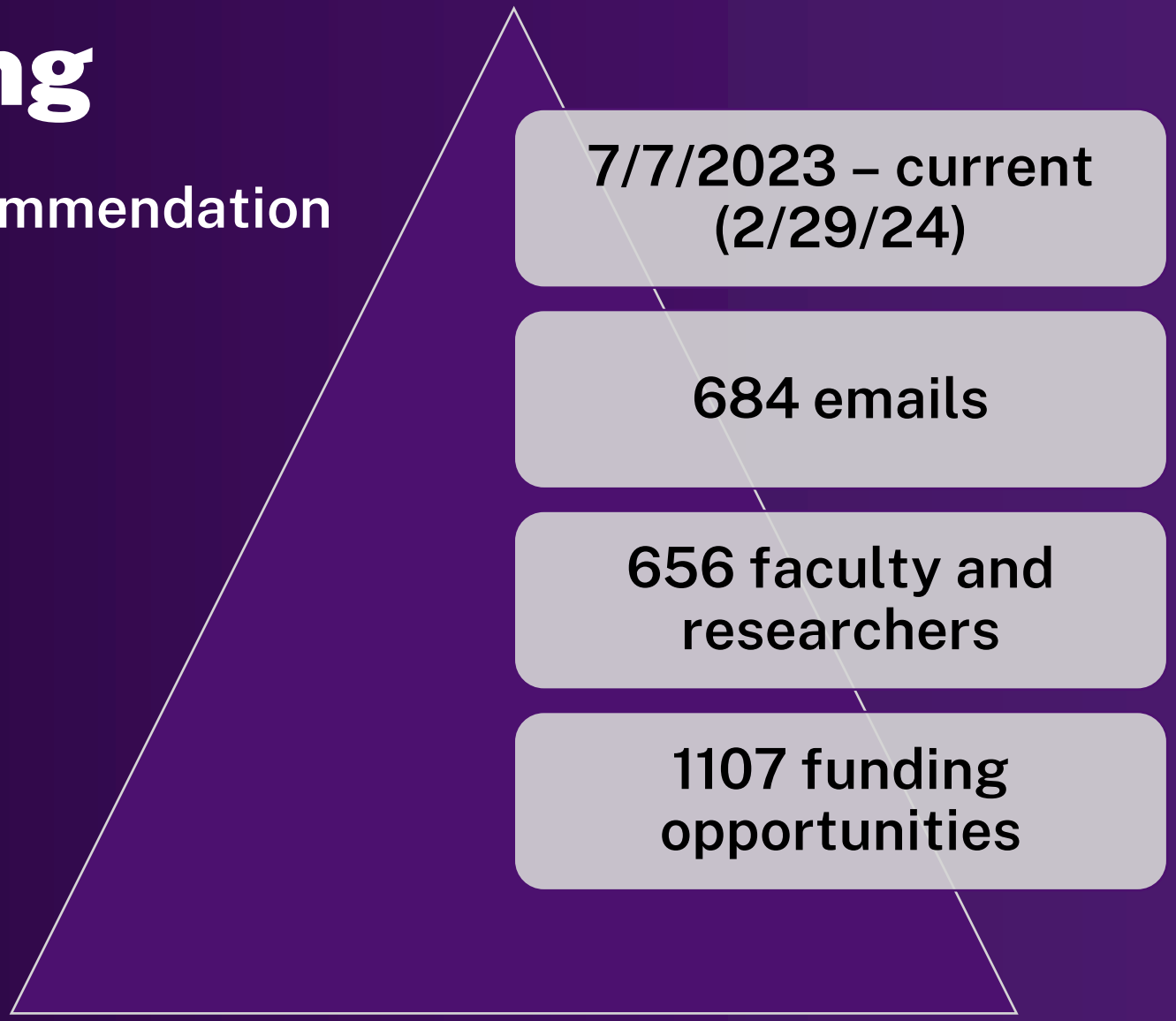
Please kindly let us know how you think about our recommendations by participating in a one-minute survey. Your input will help us improve the model and provide you with better recommendations in the future.

Good luck!

© University at Albany, SUNY, 2024

Pilot Testing

- Weekly Email Recommendation



© University at Albany, SUNY, 2024

Pilot Testing Feedback

Weekly Email Recommendation - Comments

Dead on!

It is really helpful to receive these emails. RFAs and related notices are generally in my area and help me to learn about opportunities I would have missed.

Thanks for these emails and they are very helpful in learning about new opportunities!

Thank you very much. This was very helpful.

** Survey result as of 2/29/2024*

© University at Albany, SUNY, 2024

Pilot Testing Feedback

Weekly Email Recommendation - Comments

- The recommendations had nothing to do with my research.
- Suggestions have not been aligned with my research interests and areas of expertise at all. It might be helpful to ask for keywords aligned with our research interests to better tailor suggestions. I do appreciate the attempts to match me with possible funding sources and look forward to future improvements with the system.
- Maybe limit to related industries - philosophy is very different from marketing.

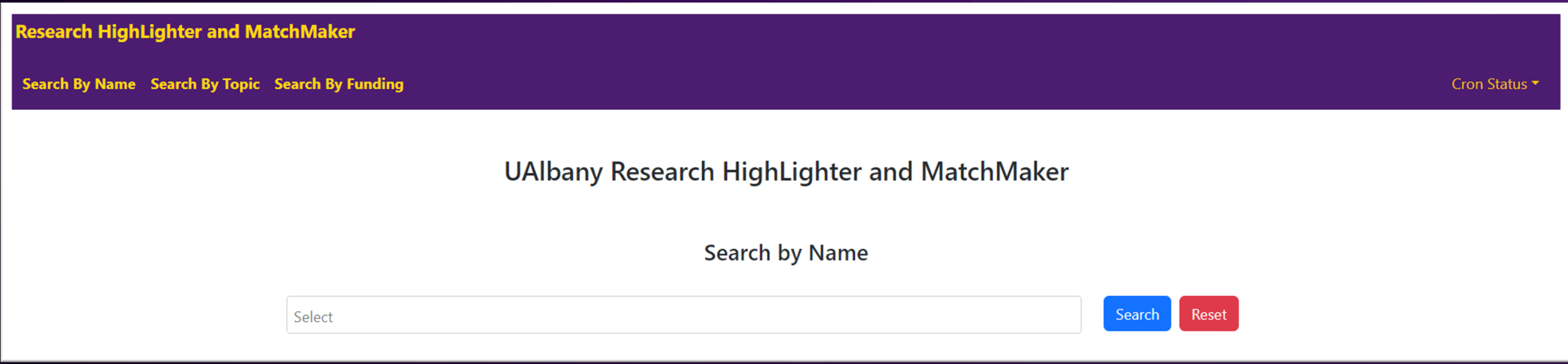
** Survey result as of 2/29/2024*

© University at Albany, SUNY, 2024

UNIVERSITY AT ALBANY
State University of New York

Front-End Portal

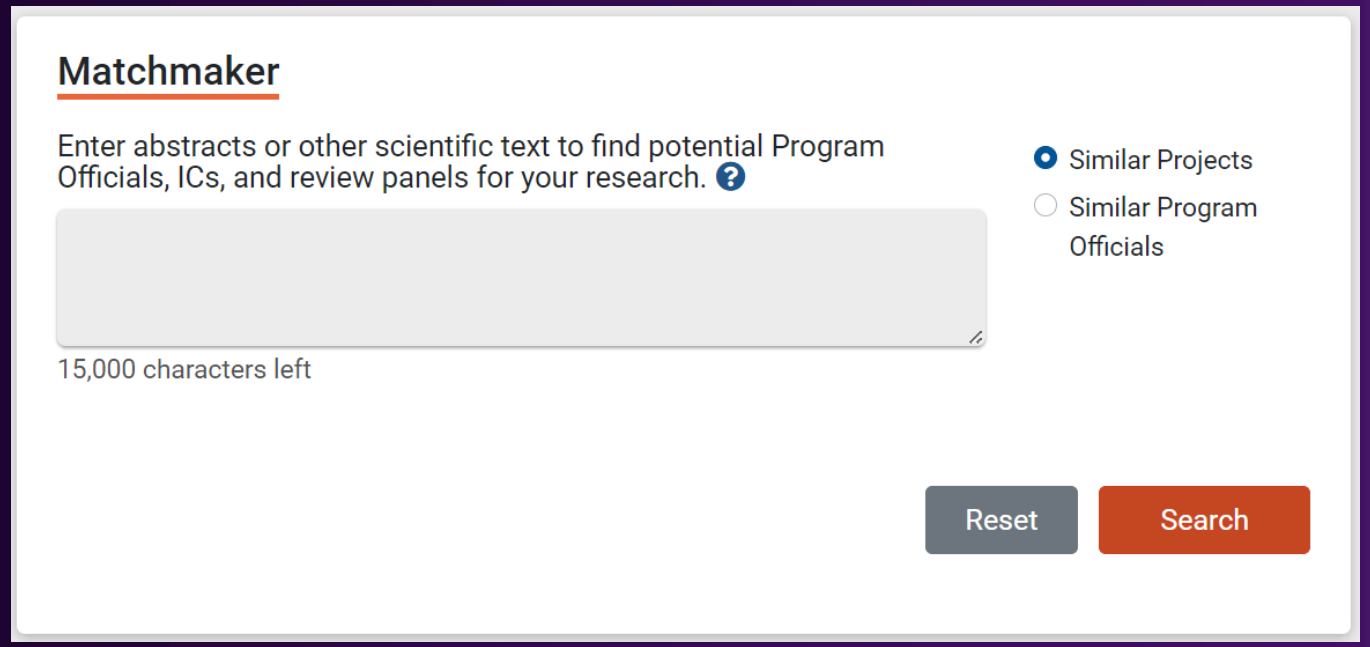
- Coming Soon
  - Search by Name: Obtain funding opportunity recommendation by direct Name input
  - Search by Topic: Obtain researchers info and funding opportunity by direct Topic input
  - Search by Funding: Obtain matched researchers info by funding opportunity



© University at Albany, SUNY, 2024

Next Step

- Search by Text
  - Recommendations based on Input Text
  - NIH Matchmaker tool in RePORTER
  - Beyond NIH, sponsors of all kinds



© University at Albany, SUNY, 2024

Next Step

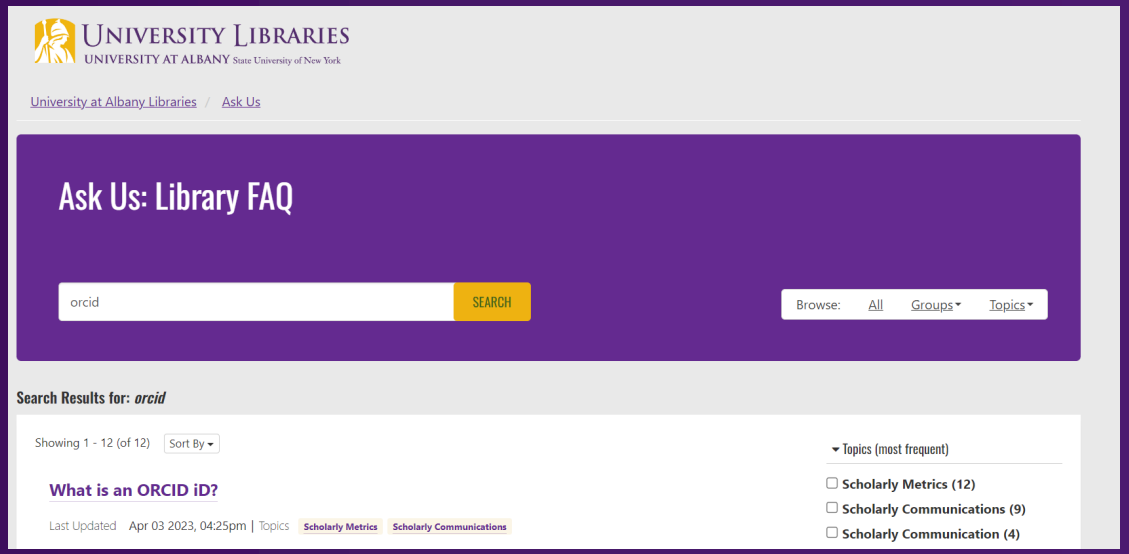
Expansion in Data Source and Use Cases

- Technology Transfer/IP
  - Venture Capital
  - Corporate Collaborators
- IRB, Syllabus
  - Students
  - Researchers

© University at Albany, SUNY, 2024

Next Step

- Expansion in Data Source and Use Cases
- ORCID
  - SUNY ORCID Task Force
  - Affiliate to auto-populate data
  - Expand scope of research data
  - Validity, Reliability, Consistency



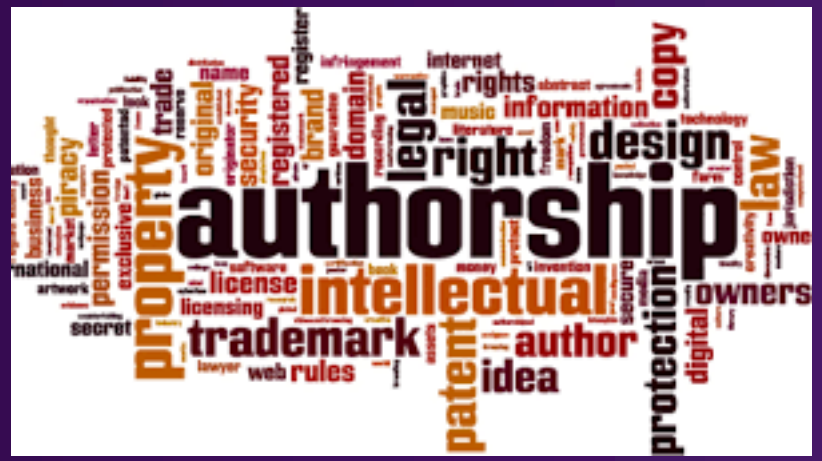
ORCID
Connecting research and researchers

© University at Albany, SUNY, 2024

Next Step

Expansion in Data Source and Use Cases

Author Contribution



© University at Albany, SUNY, 2024

UNIVERSITY AT ALBANY
State University of New York

Research Highlighter-MatchMaker Connects Researchers with Funding Opportunities & Potential Collaborators

Welcome for more feedbacks!

© University at Albany, SUNY, 2024

UNIVERSITY AT ALBANY
State University of New York

Thank You!

Office of Strategic Initiative, Data, Assessment, and Technology

Tianning Huang, Director

thuang3@albany.edu

Yuemin Li, Data Analyst

yli53@albany.edu

MSC, 1400 Washington Ave, Albany NY 12222