



COALITION FOR NETWORKED INFORMATION

Publication management & reporting with ReCiter

Paul Albert & Terrie Wheeler

Samuel J. Wood Library & Information Technologies and Services

November 15, 2023

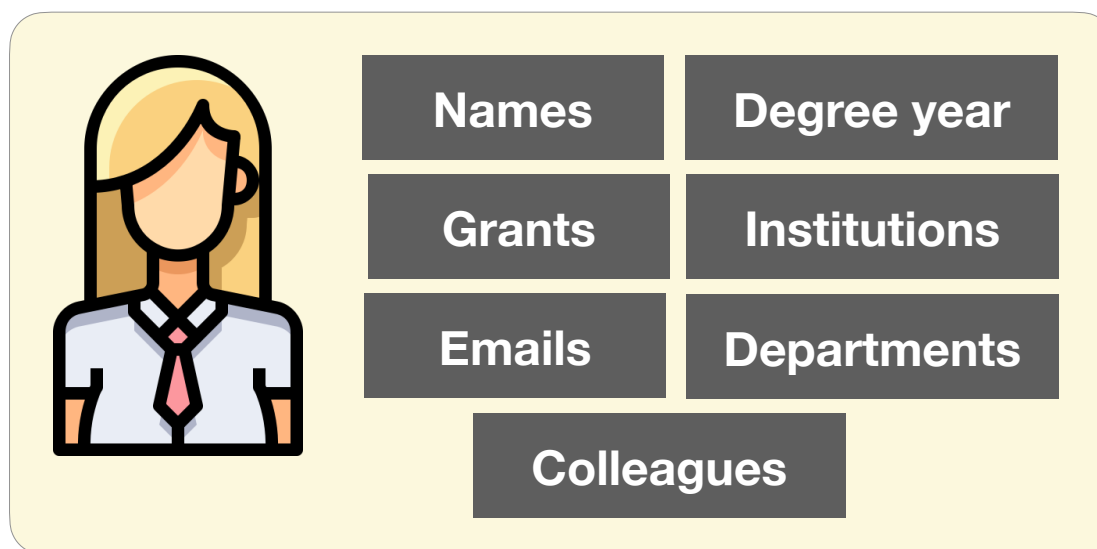


ReCiter is a powerful NIH-funded publications management system designed for academic medical institutions



How ReCiter works

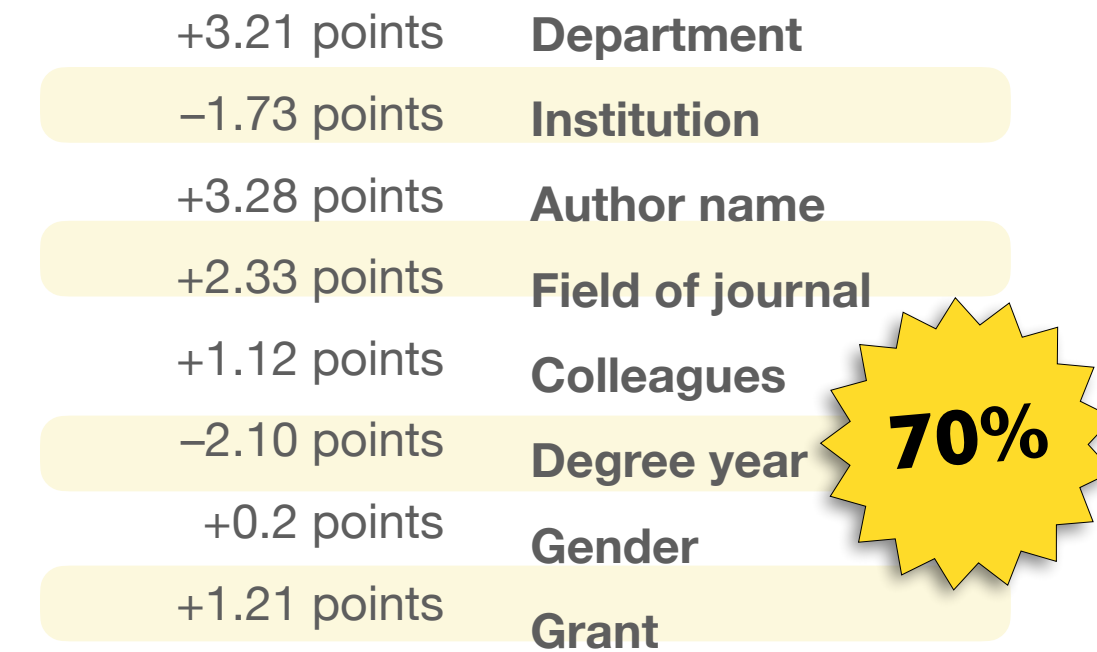
1. Use institutional systems to create a scholar's identity



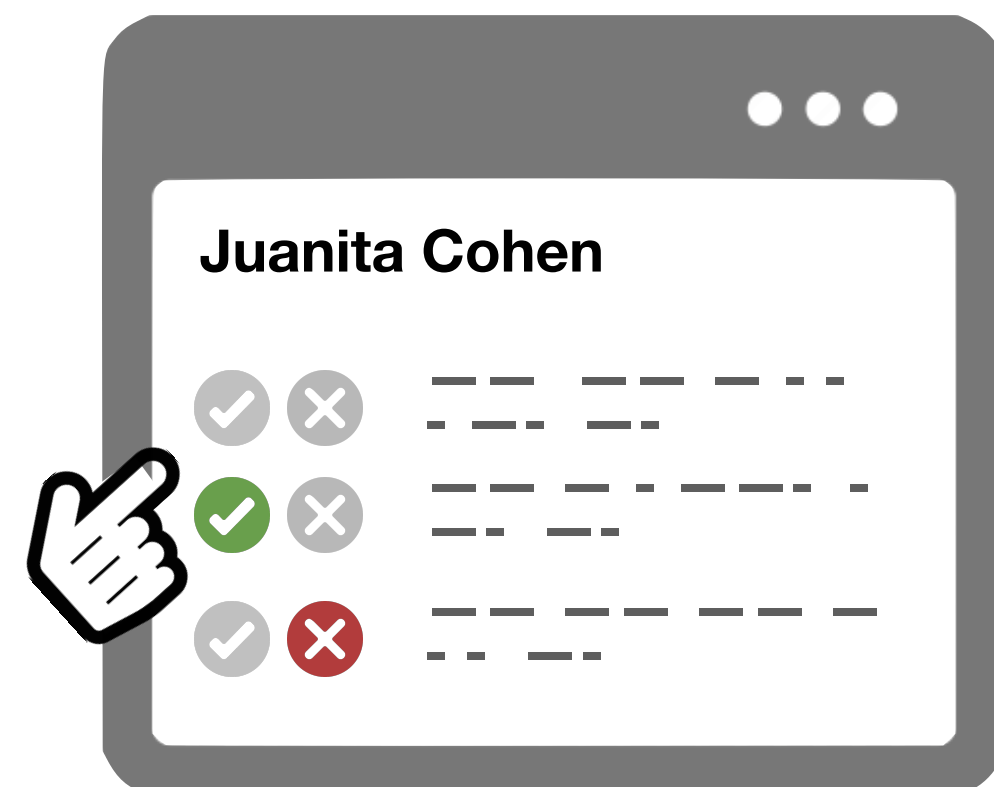
2. Use identity to retrieve candidate articles from PubMed



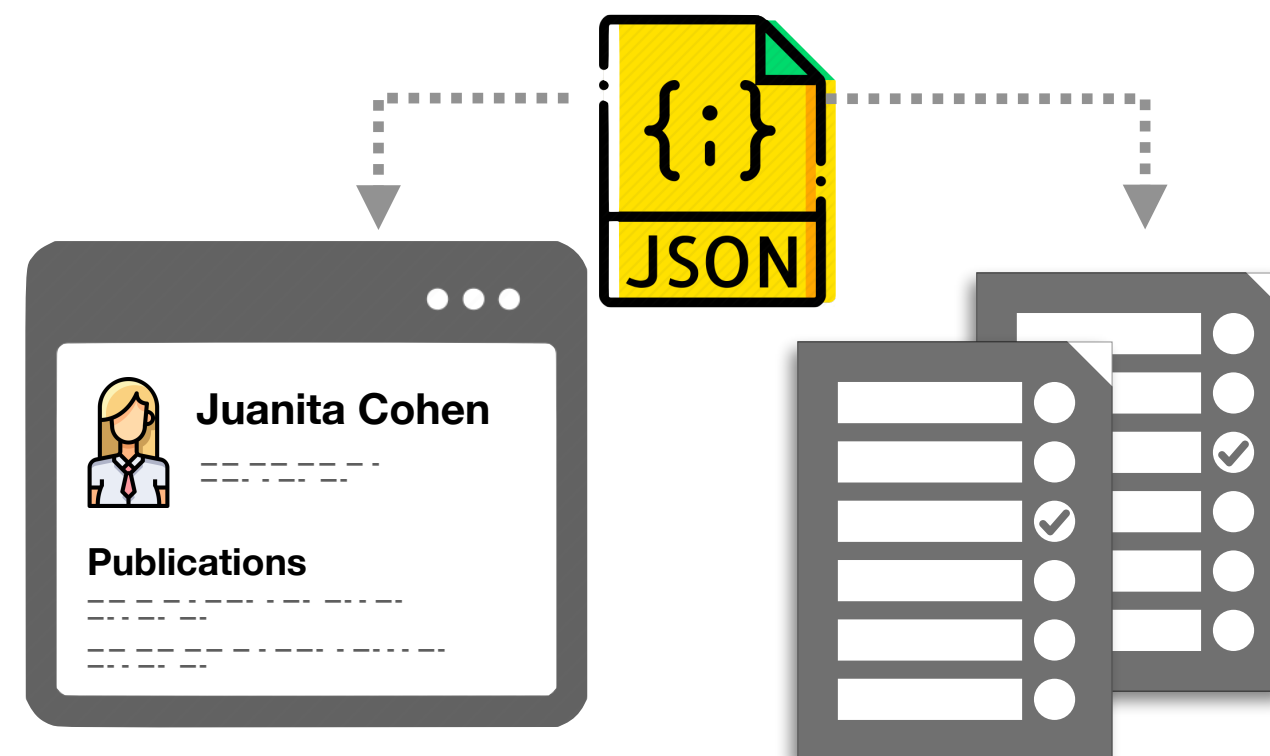
3. Use machine learning to estimate the likelihood scholar wrote each article



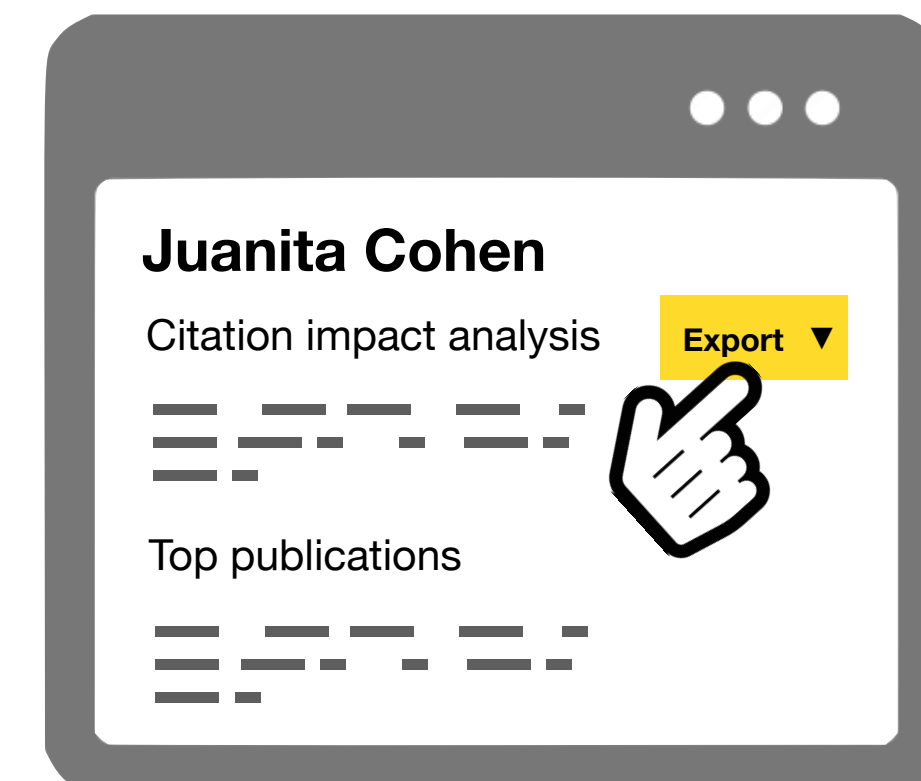
4. Collect feedback from faculty and librarians



5. Use web services to expose to downstream systems (e.g., VIVO)

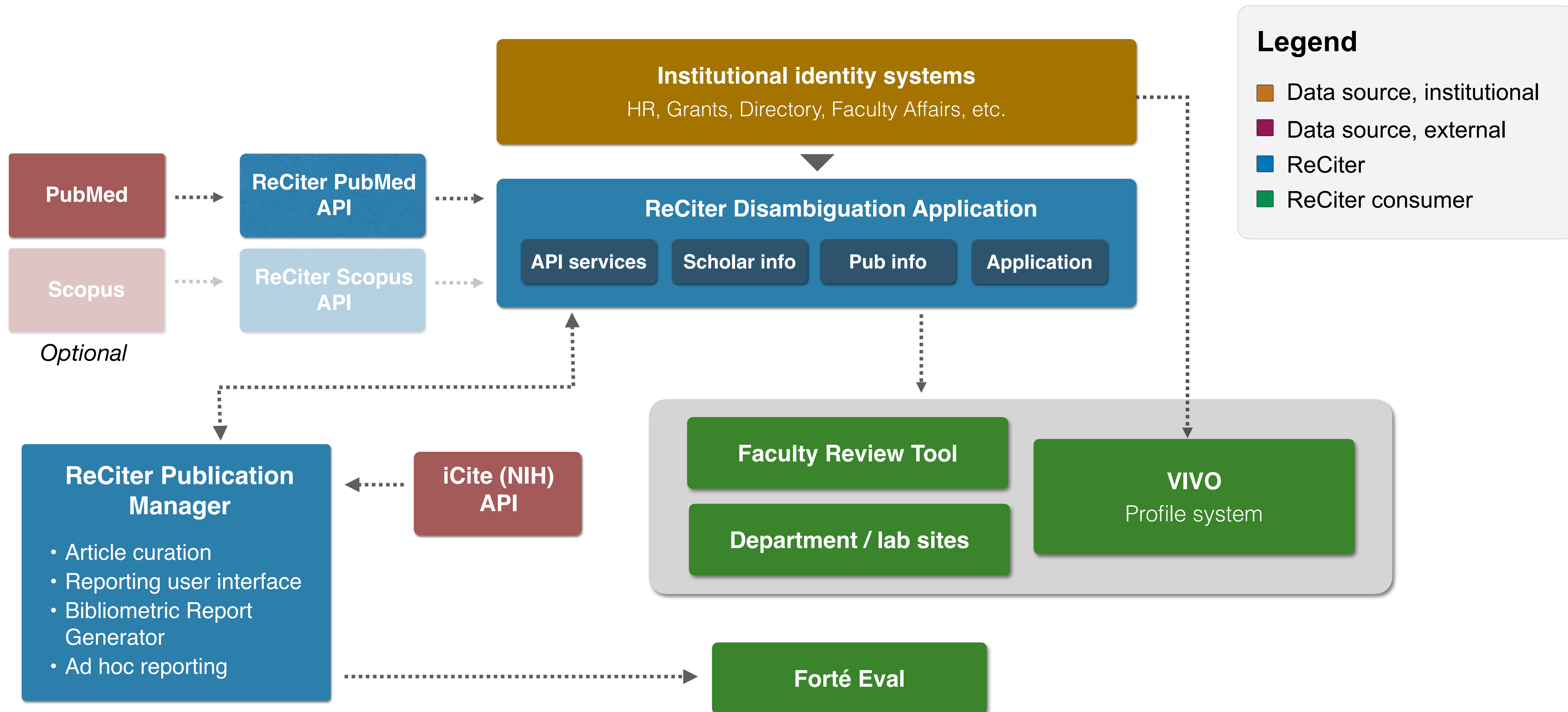


6. Output reports & bibliometric analyses





Key components in ReCiter ecosystem





WCM has a lot useful identity data about its people

Name variants

- Jochen Buck

Years of Degree

- MD - 1983
- PhD - 1984

NIH Grant #s

- 038722
- 055255
- 059913
- 062328
- 064842
- 025810
- 062948
- 073546

Department(s)

- Pharmacology
- Biochemistry

Current/prior affiliations

- Weill Cornell Medical College
- Weill Graduate School of Medical Sciences
- Memorial Sloan-Kettering Cancer Center
- Cornell University

Email

- jobuck@med.cornell.edu
- [personal email] (not shown)

Relationships

- Alexander M Stessin (co-investigator/mentor)
- Ana Diaz (co-investigator)
- Carl Nathan (co-investigator)
- Eric ***** (mentee)
- Erin M ***** (mentee)
- Jacob Bitterman (co-investigator)
- Jacob Louis Bitterman (mentee)
- James A. ***** (co-investigator/mentee)
- Jeanne Marie ***** (mentee)

- Jonathan J ***** (mentee)
- Jonathan ***** (mentee)
- Kenneth ***** (mentee)
- Kenneth Hess (co-investigator)
- Kirk W. Deitsch (co-investigator)
- Laureen Colis (employee in same unit)
- Lavoisier S. Ramos (co-investigator)
- Lonny R. Levin (co-investigator/employee in same unit)
- Margarita ***** (mentee)
- Martin Tresguerres (co-investigator)
- Mime Kobayashi (co-investigator)
- Nawreen Rahman (co-investigator/mentee)
- Nicole B. ***** (mentee)
- Samie R. Jaffrey (co-investigator)
- Tatiana Litvin (co-investigator)



Whose academic output do we monitor at WCM?

ReCiter tracks publications by ~20,000 people with certain roles, active or historical, at Weill Cornell Medicine including:

- **Academic appointees** – Anyone who has or had an academic Weill Cornell Medical College appointment including those who were active before 2015. This includes full-time faculty, voluntary faculty, postdocs, etc.
- **Doctoral students** – Anyone who was or is in a terminal degree educational program in Weill Cornell Medical College or the Graduate School. This includes MD candidates, MD-PhD candidates, and PhD candidates.
- **NYP residents**
- **Members of the Clinical & Translational Science Center (CTSC)**

All publications are from PubMed, which has over 35 million publication records.




Curation

ReCiter Publication Management System
Signed in as paa2013 [Logout](#)

compact mode <

- Find People
- Curate Publications
- Create Reports
- Perform Analysis
- Manage Module v

Curate Publications



Jochen Buck

Professor of Pharmacology
Pharmacology

Inferred keywords: Adenylyl Cyclases Cyclic AMP Bicarbonates Vitamin A Signal Transduction Spermatozoa Carbon Dioxide Retinoids

Adenylyl Cyclase Inhibitors Hydro-Lyases

[View Profile](#)

Refresh Suggestions

You have provided feedback on **1** record(s).

Suggested 1
Accepted 121
Rejected 22

Accept All
Reject All

Filter...

Sort by v

Accept
 Reject

Matching
Score
10

Soluble adenylyl cyclase mediates nerve growth factor-induced activation of Rap1.

Alexander M Stessin, Jonathan H Zippin, Margarita Kamenetsky, Kenneth C Hess, Jochen Buck, Lonny R Levin

The Journal of biological chemistry • 2006 Apr 20 • Academic Article

PMID: [16627466](#) • [DOI](#) • [Show History](#)

[+ Show evidence behind this suggestion](#)



Curation (cont'd)

Undo ✖ Reject

Matching Score
10

Alexander M Stessin, Jonathan H Zippin, Margarita Kamenetsky, Kenneth C Hess, **Jochen Buck**, Lonny R Levin
 The Journal of biological chemistry • 2006 Apr 20 • Academic Article
 PMID: [16627466](#) • [DOI](#) • [Show History](#)
 - [Hide evidence behind this suggestion](#)

Evidence	Institutional Data	Article Data
Name 3.31 points	Jochen Buck	Jochen Buck
Clustering 15.87 points	-	Score of article without clustering: 32.86 Average score of cluster: 15.87
Grants 9.012 points	NS-55255 GM-62328 HD-38722 AI-64842	R01 NS055255 GM62328 HD38722 R01 AI064842
Email 8 points	jobuck@med.cornell.edu	jobuck@med.cornell.edu
Relationships 6.50 points	Alexander M Stessin Co-investigator Kenneth Hess Co-investigator Lonny R. Levin Co-investigator CTSC Grant Associate Shared organizational unit CTSC Protocol Associate	Alexander M Stessin Kenneth Hess Lonny R. Levin
Journal category 1.88 points	Pharmacology Organizational Unit	Biochemistry & Molecular Biology Journal Category
Target author's institutional affiliation 1.8 points	-	Weill Cornell Medicine Scopus Department of Pharmacology, New York, New York 10021. Electronic address: jobuck@med.cornell.edu. Pubmed
Degree year discrepancy 0.9 points	1983 - Doctoral	2006
Co-authors's institutional affiliation	Weill Cornell Medicine Individual Affiliation	5 author(s) from Weill Cornell Medicine 60007997 Scopus



Scoring articles

Did Curtis Cole write, “A scalable method for supporting multiple patient cohort discovery projects using i2b2”?

Type of evidence	Data from institution	Data from article	Points
Name	Curtis Leland Cole	Curtis L Cole	3.31
Relationships	Marcos Davila, Thomas Campion <i>co-investigators</i>	Marcos Davila, Thomas Campion <i>co-authors</i>	4.21
Grants	TR-457	UL1 TR000457	2.25
Journal category	Population Health Sciences <i>department</i>	Medical Informatics <i>journal category</i>	1.88
Target author's institutional affiliation	Weill Cornell Medicine	Weill Cornell Medicine	1.80
Departmental affiliation	Information Technologies and Services	Information Technologies & Services Department, Weill Cornell Medicine, New York, NY, USA	1.80
Co-author's institutional affiliation	Weill Cornell Medicine	6 author(s) from Weill Cornell Medicine	0.90
Degree year discrepancy	1994 - Doctoral	2018	0.90
Clustering	Score of article without clustering: 16.62 Average score of cluster: 15.47		-0.67
Candidate article count		1192	0.24
Inferred gender of name	Male - 99% probability	Male - 99% probability	0
Total		16.62 – Raw score, 10 – Standardized score	



Reporting

reciter.weill.cornell.edu/report

ReCiter Publication Manager Signed in as paa2013 [Logout](#)

compact mode <

- Find People
- Curate Publications
- Create Reports**
- Manage Users
- Configuration

Create Reports

Author Filters

Author Organization Institution Person Type Author Position

Type letters to search

- (friedman) marcus, Julie (Cardiology (Medicine))
- (Kinney) Nelson, Joanna (Pediatric Cardiology)
- Aaroe, Ashley (Neurology)
- Aaron, Tzvi
- Aaronson, Jaime (Anesthesiology)
- Aaronson, Paul (Urology)
- Abadeer, Andrew
- Abadi, Danielle (General Pediatrics (Pediatrics))

[See More](#)

[Search](#) [Reset](#)

[Sort by](#) [Export to CSV](#) [Export to RTF](#)

Page 1 of 11450

Author Correction: Evolution of structural rearrangements in prostate cancer intracranial metastases.
 Francesca Khani, William F Hooper, Xiaofei Wang, Timothy R Chu, Minita Shah, Lara Winterkorn [...] Himisha Beltran
 NPJ precision oncology • 2023 Nov 02 • Erratum
 PMID: [37919447](#) • [DOI](#)

Myocardial infarction and physical function: the REasons for Geographic And Racial Differences in Stroke prospective cohort study.
 Emily B Levitan, Parag Goyal, Joanna Bryan Ringel, Orysya Soroka, Madeline R Sterling, Raegan W Durant [...] [Monika M Safford](#)
 BMJ public health • 2023 Oct 12 • Academic Article
 PMID: [37920711](#) • [DOI](#)

Acute Appendicitis in the Epicenter of the COVID-19 Pandemic: A New York City Single-Center Experience.



Bulk reporting

The screenshot shows a Microsoft Excel spreadsheet with the following data in column A:

	A	B	C	D	E	F	G	H	I	J	K	L
1	aab9028											
2	aad2004											
3	aag9014											
4	aah2003											
5	aam2011											
6	aam9008											
7	aas9008											
8	aas9027											
9	aba9034											
10	aba9036											
11	abd4001											
12	abg9017											
13	abh9006											
14	abj2006											
15	abj9004											
16	abm2013											
17	abs2001											
18	abs9026											
19	abt4002											
20	acg2005											
21	achadbur											
22	ack2003											
23	acl9007											
24	acp9008											
25	acr2213											
26	acr9004											
27	acy2001											
28	adc9001											
29	adk9016											

The interface includes a ribbon with tabs for Paste, Font, Alignment, Numbers, General, Conditional Formatting, Insert, Delete, Editing, Analyze Data, and Sensitivity. The status bar at the bottom shows 'Count: 200' and a zoom level of '150%'.



Export to CSV



Create Reports

Author Filters

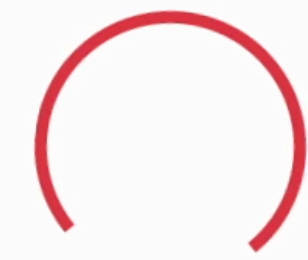
Author Organization Institution Person Type Author Position

Article Filters

Date Type Journal Journal Rank

Search

Reset





Export to RTF

>

Create Reports

Author Filters

Author Organization Institution Person Type Author Position

Type 3 letters to search

- Anwar, Md rafi (Cardiothoracic Surgery)
- Ashrafi, Ghazaleh ()
- Ashrafi, Soha (Neuroscience)
- Guirguis, Nayera (Doctor of Medicine)
- Maarafiya, Muna (Pediatrics)
- Rafiaa, Amer (General Internal Medicine (Medicine))
- Rafie, Christopher (Emergency Medicine)
- Rafii. Daniel (General Internal Medicine (Medicine))

[See More](#)

Search Reset

by Export to CSV Export to RTF

© 2023 by Weill Cornell Medical College. All rights reserved.



Bibliometrics

Create Reports

Author Filters

Author

Organization

Institution

Person Type

Author Position

lyden



Lyden, John (Orthopaedic Surgery)

[See More](#)

Lyden, David (Hematology/Oncology (Pediatrics))

Search

Reset

20 articles

Sort by

Export to CSV

Export to RTF

Show records 20

Page 1 of 1

Extracellular vesicle and particle isolation from human and murine cell lines, tissues, and bodily fluids.

Linda Bojmar, Han Sang Kim, Gabriel C Tobias, Fanny A Pelissier Vatter, Serena Lucotti, Kofi Ennu Gyan [...], David Lyden

STAR protocols • 2020 Dec 22 • Academic Article

PMID: [33786456](#) • DOI • Citation count (NIH): 1 • Percentile Rank: 5.1 • Relative Citation Ratio (NIH): 0.11

Tumour-regulated anorexia preceding cachexia.

Gang Wang, Haiying Zhang, David Lyden

Nature cell biology • 2021 Feb 01 • Academic Article

PMID: [33558727](#) • DOI • Citation count (NIH): 2 • Journal Rank: 11.38



Custom reporting

Mutual admiration analysis: The pairs of active WCM full-time faculty, in which Person A cites Person B the most, and vice versa.

Person 1	Person 2	Publication count
Teresa Milner (Brain and Mind Research Institute)	Virginia Pickel (Brain and Mind Research Institute)	130
Timothy McGraw (Biochemistry)	Frederick Maxfield (Biochemistry)	67
Sabine Ehrt (Microbiology and Immunology)	Carl Nathan (Microbiology and Immunology)	61
Theodore Schwartz (Neurological Surgery)	Antonio Bernardo (Neurological Surgery)	51
Neil Bander (Urology)	Cora Sternberg (Medicine)	40
Olivier Elemento (Physiology and Biophysics)	Ari Melnick (Medicine)	38
Barbara Hempstead (Medicine)	Francis Lee (Psychiatry)	32
Khaled Machaca (Physiology and Biophysics)	Tobias Meyer (Cell and Developmental Biology)	31
Mony de Leon (Radiology)	Costantino Iadecola (Brain and Mind Research Institute)	24
Christopher Triggle (Pharmacology)	Dietrich Busselberg (Physiology and Biophysics)	23
Antonio Bernardo (Neurological Surgery)	Theodore Schwartz (Neurological Surgery)	22
Harel Weinstein (Physiology and Biophysics)	Olaf Andersen (Physiology and Biophysics)	20
Louis Aronne (Medicine)	Alpana Shukla (Medicine)	20



Available at <https://reciter.weill.cornell.edu>

Access

- *Anyone* with an active account at WCMC, NYP, WCMC-Q, or HSS should be able to log in to and use Publication Manager here: <https://reciter.weill.cornell.edu>
- Must be on campus or VPN

Roles

- **Reporter** – anyone who can successfully authenticate at one of the above institutions is automatically assigned this role
- **Self-curator** – anyone who has a matching CWID or email on record
- **Curator all** – several WCM librarians can curate for everyone

Weill Cornell Medicine

ReCiter Publication Manager

Please select your institution from the list below.

- Weill Cornell Medicine
- NewYork-Presbyterian Hospital
- Weill Cornell Medicine - Qatar
- Hospital for Special Surgery

Remember my choice for:



Credit

Work

- Development: Sarbajit Dutta, Mahender Jangari, Jie Lin, Rajeev Ranjan, Michael Bales, Judy Gabeskiria
- Management: Terrie Wheeler, Curtis Cole, Mohammad Mansour, Phillip Lamont, Steve Johnson

Funding

ReCiter Publication Manager has been funded by:

- National Institutes of Health, National Library of Medicine, under Cooperative Agreement UG4LM012347 with the University of Massachusetts, Worcester
- Lyris through its Catalyst fund

The ReCiter suite of applications has been funded by :

- The National Institutes of Health National Center for Advancing Translational Sciences through grant number UL1TR002384 (CTSC)



Follow up

- Contact publications@med.cornell.edu or Paul Albert at paa2013@med.cornell.edu about:
 - Features
 - Bugs
 - API access
 - Custom reports
 - Requests for presentations
- Look at the open source code: <https://github.com/wcmc-its/reciter>
- Read the paper in PLoS ONE

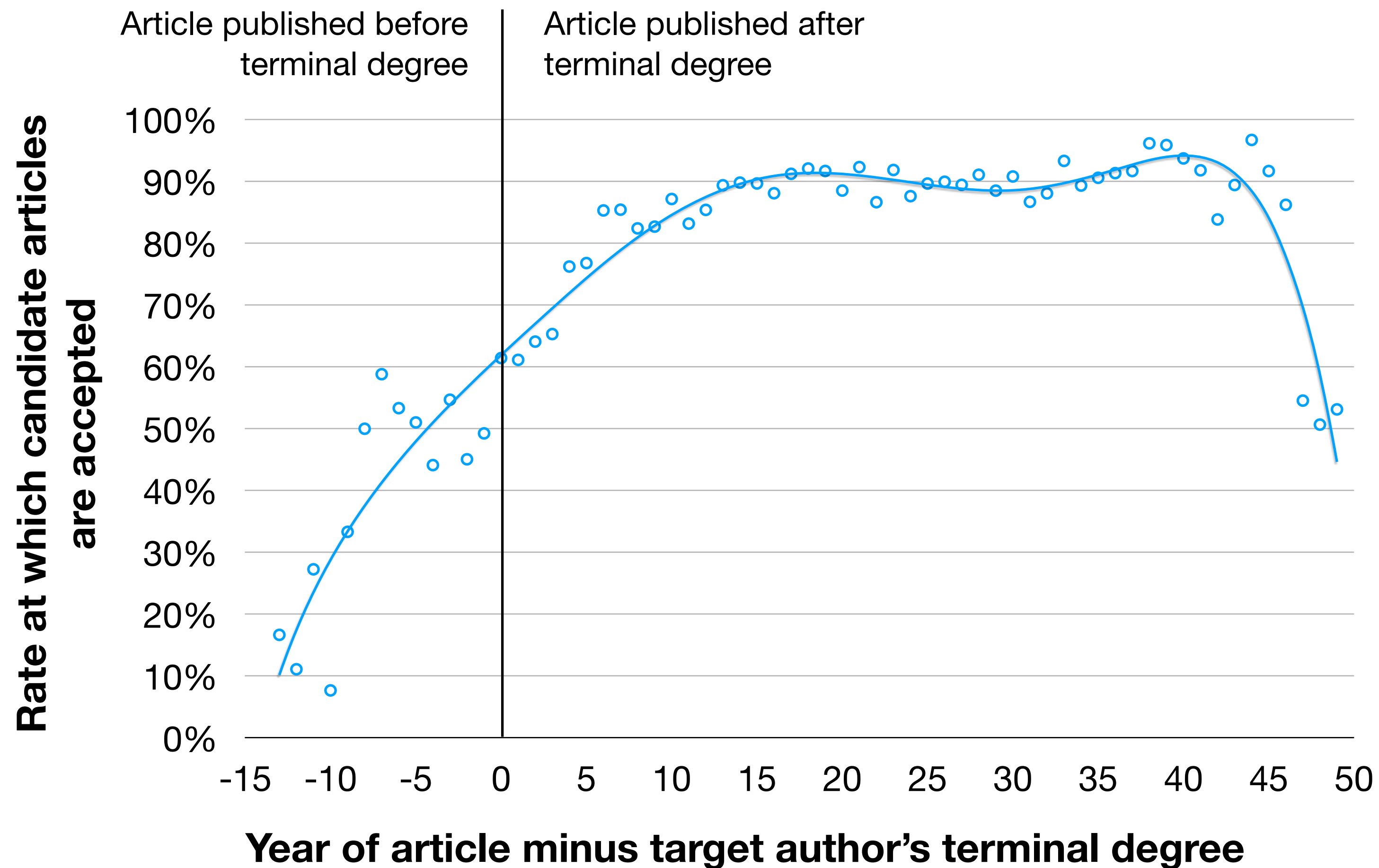
The screenshot shows the PLOS ONE article page for "ReCiter: An open source, identity-driven, authorship prediction algorithm optimized for academic institutions". The page includes a navigation bar with "PUBLISH", "ABOUT", "BROWSE", and "SEARCH" options. The article is marked as "OPEN ACCESS" and "PEER-REVIEWED". It is a "RESEARCH ARTICLE" with 0 saves, 0 citations, 73 views, and 0 shares. The authors listed are Paul J. Albert, Sarbajit Dutta, Jie Lin, Zimeng Zhu, Michael Bales, Stephen B. Johnson, Mohammad Mansour, Drew Wright, Terrie R. Wheeler, and Curtis L. Cole. The article was published on April 1, 2021, with a DOI of 10.1371/journal.pone.0244641. A navigation menu at the bottom of the article highlights "Article" and includes options for "Authors", "Metrics", "Comments", and "Media Coverage". The abstract text is visible, starting with "Academic institutions need to maintain publication lists for thousands of faculty and other scholars. Automated tools are essential to minimize the need for direct feedback from the scholars themselves who are practically unable to commit necessary effort to keep the data accurate. In relying exclusively on clustering techniques, author disambiguation applications fail to satisfy key use cases of academic institutions. Algorithms can perfectly group together a set of publications authored by a common individual, but, for them to be useful to an academic institution, they need to programmatically and recurrently map articles to thousands of scholars of interest *en masse*. Consistent with a savvy librarian's approach for generating a scholar's list of publications, identity-driven authorship prediction is the process of using information about a scholar to quantify the likelihood that person wrote certain articles. ReCiter is an application that attempts to do exactly that. ReCiter uses institutionally-maintained identity data such as name". On the right side, there are buttons for "Download PDF", "Print", "Share", and "Check for updates". Below these are "Subject Areas" including "Librarians", "Metadata", and "Support vector machines".



Appendix



Likelihood a candidate article was authored by a given person relative to their year of doctoral degree





Using a critical mass of identity info, authorship identity is rarely ambiguous

 Jochen Buck (WCM) almost certainly wrote

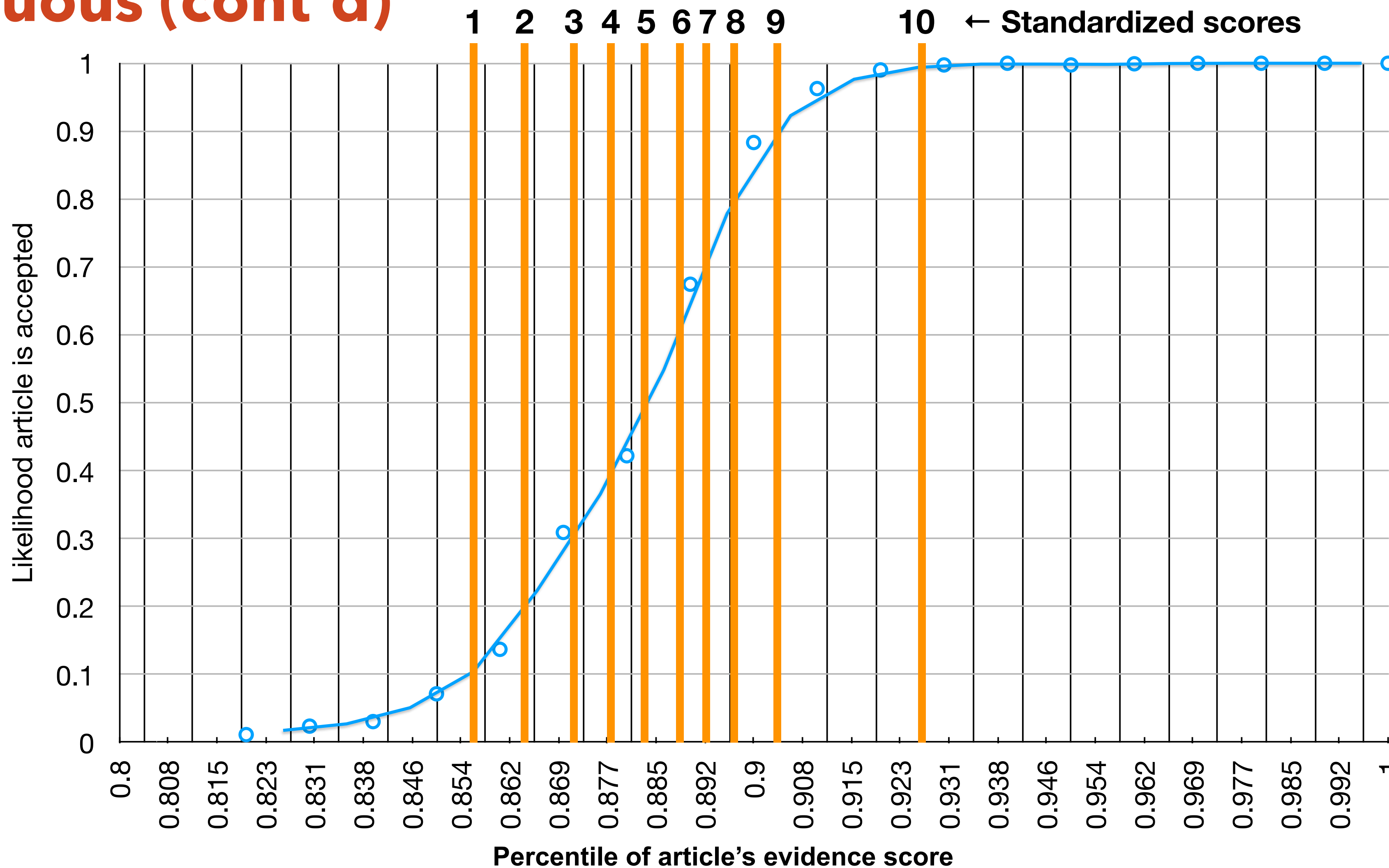
 Ambiguous

 Almost certainly did not write

1378327	15322879	24254994	30401788	30029576	21576236	26517124	15061498	16363787	17826347	1378327	15322879	24254994	30401788	30029576	21576236	26517124	15061498	16363787	17826347	1378327	15322879	24254994	30401788	30029576
1535365	15608121	24282227	29590062	29544531	8976941	28001054	26342983	24751434	19392950	1535365	15608121	24282227	29590062	29544531	8976941	28001054	26342983	24751434	19392950	1535365	15608121	24282227	29590062	29544531
1655814	15619637	24324443	26793798	20683624	27323809	26961873	27065983	22670108	19144814	1655814	15619637	24324443	26793798	20683624	27323809	26961873	27065983	22670108	19144814	1655814	15619637	24324443	26793798	20683624
1698557	15758154	24567411	30292553	29503642	25969580	26663923	25190508	25294045	12609998	1698557	15758154	24567411	30292553	29503642	25969580	26663923	25190508	25294045	12609998	1698557	15758154	24567411	30292553	29503642
1698611	16002394	24814023	29807838	30072966	28246421	16250004	19381559	20412585	23060493	1698611	16002394	24814023	29807838	30072966	28246421	16250004	19381559	20412585	23060493	1698611	16002394	24814023	29807838	30072966
1749937	16043520	24874272	31468926	20732380	33016598	30688560	25224905	20200045	23726029	1749937	16043520	24874272	31468926	20732380	33016598	30688560	25224905	20200045	23726029	1749937	16043520	24874272	31468926	20732380
2187837	16054031	25002117	31333191	12556599	28864766	26704418	26778309	17160486	22844459	2187837	16054031	25002117	31333191	12556599	28864766	26704418	26778309	17160486	22844459	2187837	16054031	25002117	31333191	12556599
2332732	16250004	25009261	31905741	22827224	29390769	22095623	28956554	24841487	21890900	2332732	16250004	25009261	31905741	22827224	29390769	22095623	28956554	24841487	21890900	2332732	16250004	25009261	31905741	22827224
2484676	16303561	25308880	28736043	31274453	31046326	26851860	27497900	32017718	15758154	2484676	16303561	25308880	28736043	31274453	31046326	26851860	27497900	32017718	15758154	2484676	16303561	25308880	28736043	31274453
2992460	16400172	25340964	28659159	30214721	29195430	27595695	27091322	12680472	21944716	2992460	16400172	25340964	28659159	30214721	29195430	27595695	27091322	12680472	21944716	2992460	16400172	25340964	28659159	30214721
3033079	16627466	26037077	31961918	28552203	27251416	28041734	12744875	23273337	17895388	3033079	16627466	26037077	31961918	28552203	27251416	28041734	12744875	23273337	17895388	3033079	16627466	26037077	31961918	28552203
3365688	16934836	26101266	25150148	31722126	28541183	16318692	22998876	12475901	32765563	3365688	16934836	26101266	25150148	31722126	28541183	16318692	22998876	12475901	32765563	3365688	16934836	26101266	25150148	31722126
3749500	16964251	26961873	31659025	30405105	29271744	28859004	26995127	25151881	16340675	3749500	16964251	26961873	31659025	30405105	29271744	28859004	26995127	25151881	16340675	3749500	16964251	26961873	31659025	30405105
3864530	17524362	27323809	23636265	26422337	27853987	27621408	25002117	25993056	24583003	3864530	17524362	27323809	23636265	26422337	27853987	27621408	25002117	25993056	24583003	3864530	17524362	27323809	23636265	26422337
3866799	17591988	27547922	31478135	17226886	20121413	26091858	20443056	24573549	21666074	3866799	17591988	27547922	31478135	17226886	20121413	26091858	20443056	24573549	21666074	3866799	17591988	27547922	31478135	17226886
7642543	17680672	27627854	30639013	29960454	26871723	20689673	17524362	23845647	22308401	7642543	17680672	27627854	30639013	29960454	26871723	20689673	17524362	23845647	22308401	7642543	17680672	27627854	30639013	29960454
8243889	18630896	27670898	30814064	28353475	29857731	25795452	23521048	20018667	16043520	8243889	18630896	27670898	30814064	28353475	29857731	25795452	23521048	20018667	16043520	8243889	18630896	27670898	30814064	28353475
8245778	18695009	28686538	31404529	30126598	8966294	27872617	18630896	23450375	16124216	8245778	18695009	28686538	31404529	30126598	8966294	27872617	18630896	23450375	16124216	8245778	18695009	28686538	31404529	30126598
8292589	18806876	28864766	30742847	29729922	27587106	28089854	14769862	20130594	24212568	8292589	18806876	28864766	30742847	29729922	27587106	28089854	14769862	20130594	24212568	8292589	18806876	28864766	30742847	29729922
8340762	19126549	29314906	29444475	30532663	28518406	27550735	26883688	22136728	16303561	8340762	19126549	29314906	29444475	30532663	28518406	27550735	26883688	22136728	16303561	8340762	19126549	29314906	29444475	30532663
8643496	19254571	29807057	31444059	15239689	22184500	26181492	16381572	12037129	22226839	8643496	19254571	29807057	31444059	15239689	22184500	26181492	16381572	12037129	22226839	8643496	19254571	29807057	31444059	15239689
8643497	20018667	30401788	28353478	26321657	27627854	26101266	18628535	24130358	18806876	8643497	20018667	30401788	28353478	26321657	27627854	26101266	18628535	24130358	18806876	8643497	20018667	30401788	28353478	26321657
8663216	20130594	31750304	29124910	32274424	25971659	28456402	22775536	21821557	16002394	8663216	20130594	31750304	29124910	32274424	25971659	28456402	22775536	21821557	16002394	8663216	20130594	31750304	29124910	32274424
8760808	20383722	32065141	25972781	31429807	29081589	26342989	27301771	21665942	15608121	8760808	20383722	32065141	25972781	31429807	29081589	26342989	27301771	21665942	15608121	8760808	20383722	32065141	25972781	31429807
8866697	20410468	32259482	25340964	31594556	27671553	25153530	19062876	25465568	24147415	8866697	20410468	32259482	25340964	31594556	27671553	25153530	19062876	25465568	24147415	8866697	20410468	32259482	25340964	31594556
9067641	20683624	32307523	29769359	31660970	27547922	21754966	25734415	19824671	19584166	9067641	20683624	32307523	29769359	31660970	27547922	21754966	25734415	19824671	19584166	9067641	20683624	32307523	29769359	31660970
9294188	21124988	32614044	31117690	31219151	24091307	32960928	28222197	18695009	19126549	9294188	21124988	32614044	31117690	31219151	24091307	32960928	28222197	18695009	19126549	9294188	21124988	32614044	31117690	31219151
9857081	21490586	32914502	28941207	30419937	28980046	22315070	21894962	18840617	30818607	9857081	21490586	32914502	28941207	30419937	28980046	22315070	21894962	18840617	30818607	9857081	21490586	32914502	28941207	30419937
9874775	21544217	2992460	16214650	31255136	28162819	32614044	20544375	21149488	19268584	9874775	21544217	2992460	16214650	31255136	28162819	32614044	20544375	21149488	19268584	9874775	21544217	2992460	16214650	31255136
10258484	21665942	3365688	28282774	31799067	28987586	17630455	24567411	23686265	24814023	10258484	21665942	3365688	28282774	31799067	28987586	17630455	24567411	23686265	24814023	10258484	21665942	3365688	28282774	31799067
10463596	21666074	24814023	29913083	10875147	28527793	16688388	25009261	25382805	23302439	10463596	21666074	24814023	29913083	10875147	28527793	16688388	25009261	25382805	23302439	10463596	21666074	24814023	29913083	10875147
10737962	21724846	3864530	21490586	31519883	28686538	27994984	26426713	12858237	16054031	10737962	21724846	3864530	21490586	31519883	28686538	27994984	26426713	12858237	16054031	10737962	21724846	3864530	21490586	31519883
10869379	21994938	26037077	29397171	28821805	18049252	22647526	24996279	20167007	14512417	10869379	21994938	26037077	29397171	28821805	18049252	22647526	24996279	20167007	14512417	10869379	21994938	26037077	29397171	28821805
10915626	22027830	29314906	29345312	31169584	24129574	23611875	19880469	17487523	15223560	10915626	22027830	29314906	29345312	31169584	24129574	23611875	19880469	17487523	15223560	10915626	22027830	29314906	29345312	31169584
11092933	22058150	20383722	24324443	30827744	28066335	24845529	22761895	14650645	17108528	11092933	22058150	20383722	24324443	30827744	28066335	24845529	22761895	14650645	17108528	11092933	22058150	20383722	24324443	30827744
11323722	22226839	2484676	31491235	21449968	28454560	30599657	15627520	16627466	20807825	11323722	22226839	2484676	31491235	21449968	28454560	30599657	15627520	16627466	20807825	11323722	22226839	2484676	31491235	21449968
11551811	22761895	28678784	29411859	29356454	19723729	30681924	19744021	21289350	16964251	11551811	22761895	28678784	29411859	29356454	19723729	30681924	19744021	21289350						



Using a critical mass of identity info, authorship identity is rarely ambiguous (cont'd)





Librarians' time spent curating publications at WCM

- From Oct. 2021 to Jun. 2023, two WCM librarians spent a total of ~71 minutes / week curating publication lists
- Includes daily updates of 6,000+ people of high interest
 - full-time faculty (n = 2,200)
 - doctoral students (n = 1,550)
 - postdocs (n = 500)
 - residents (n = 1,850)
- Also includes ad hoc curation for:
 - alumni
 - external mentors
 - members of our local CTSA

Person	Date	Start	End	Duration
Drew	4/24/2023	8:31	8:38	0:07
Drew	4/25/2023	8:21	8:23	0:02
Drew	4/26/2023	8:00	8:09	0:09
Drew	4/27/2023	8:26	8:34	0:08
Drew	4/28/2023	9:20	9:27	0:07
Paul	4/29/2023	9:00	9:30	0:30
Drew	5/1/2023	8:11	8:23	0:12
Drew	5/2/2023	8:45	8:54	0:09
Drew	5/3/2023	8:07	8:14	0:07
Drew	5/4/2023	8:22	8:29	0:07
Paul	5/4/2023	9:00	9:09	0:09
Drew	5/5/2023	9:01	9:04	0:03
Drew	5/9/2023	7:18	7:29	0:11
Drew	5/10/2023	8:12	8:21	0:09
Drew	5/11/2023	9:08	9:17	0:09
Drew	5/12/2023	8:06	8:17	0:11
Drew	5/15/2023	8:01	8:07	0:06
Drew	5/16/2023	7:52	8:06	0:14
Drew	5/17/2023	8:24	8:36	0:12



Why not use ORCID for disambiguation and reporting?

- ReCiter doesn't use ORCID for scoring candidate articles
- Problems with ORCID:
 - Even “compliant” publication may only include ORCID for a handful of authors
 - Duplicate profiles
 - Missing a lot of older publications
 - Delay between publication and availability via ORCID
 - Acting as a proxy requires securing explicit permission
 - No concept of rejected
 - No concept of author position
 - Faculty can hide their profiles (as 13% do)

The screenshot shows the 'Account settings' page for an ORCID account. The 'Visibility preferences' section is highlighted with a red box. It contains three radio button options: 'Everyone' (87% of users choose this), 'Trusted parties' (5% of users choose this), and 'Only me' (8% of users choose this). A dotted line connects the text 'Faculty can hide their profiles (as 13% do)' from the list on the left to the 'Only me' option in the screenshot.

Setting	Action
Email and notification preferences	Edit
Language display preferences	Edit
Password	Edit
Visibility preferences	Hide

By default, who should be able to see information added to your ORCID iD?

- Everyone** (87% of users choose this)
- Trusted parties** (5% of users choose this)
- Only me** (8% of users choose this)



Data in ReCiter is shared via web service

re-citer-controller Operations on ReCiter API. ⌵

GET `/reciter/article-retrieval/by/uid` Article retrieval by UID.

This api returns all the publication for a supplied uid.

Parameters Try it out

Name	Description
uid <small>required</small> string (query)	uid
totalStandardizedArticleScore number (query)	totalStandardizedArticleScore
filterByFeedback string (query)	filterByFeedback

Responses Response content type: `application/json` ⌵

Code	Description
200	

Access available by request



Bibliometric Report Generator

Analysis of Juanita Cohen's research articles

This is a bibliometric analysis of research articles from PubMed authored by Juanita Cohen (person identifier: jcohen) in Pharmacology. It was prepared by the Weill Cornell Medical Library on March 28, 2022. The terms used in this report are described in the "Explanation" section at the end.

Summary

For the purpose of this analysis, Cohen has authored **105** research articles appearing in PubMed, **12** of which were written as a first author and **23** as a senior author. Any articles written in the last two years or prior to 1980 are excluded from this analysis but are included in either the "Select recent articles" or "Select older articles" sections.

Articles' NIH percentile by author position:

First or senior - The five and ten most influential first- or senior-authored research articles have an average NIH percentile of **94.7** and **89.8**, respectively (higher is better). There are 343 WCM-employed full-time faculty at the full professor level with five or more first- or senior-authored research articles, and 329 who have written ten or more. Within these groups, Cohen's percentiles rank at **#134** and **#140**, respectively (lower is better).

Senior - The five and ten most influential senior-authored research articles have an



Export RTF file from web interface

1. Williams M, Bray BE, Greenes RA, McCusker J, Middleton B, Perry G ... **Wheeler T**. Summary of fourth annual MCBK public meeting: Mobilizing computable biomedical knowledge-metadata and trust. *Learning health systems*. 2021;6:e10301. PMID: 35036558. [Full text](#).
2. Champion TR, Sholle ET, Pathak J, Johnson SB, Leonard JP, **Cole CL**. An architecture for research computing in health to support clinical and translational investigators with electronic patient data. *Journal of the American Medical Informatics Association : JAMIA*. 2022;29:677-685. PMID: 34850911. [Full text](#).
3. Yin AL, Guo WL, Sholle ET, Rajan M, Alshak MN, Choi JJ ... **Cole CL**. Comparing automated vs. manual data collection for COVID-specific medications from electronic health records. *International journal of medical informatics*. 2021;157:104622. PMID: 34741892. [Full text](#).
4. **Albert PJ**, Dutta S, Lin J, Zhu Z, Bales M, Johnson SB ... **Cole CL**. ReCiter: An open source, identity-driven, authorship prediction algorithm optimized for academic institutions. *PloS one*. 2021;16:e0244641. PMID: 33793563. [Full text](#).
5. Cusick MM, Sholle ET, Davila MA, Kabariti J, **Cole CL**, Champion TR. A Method to Improve Availability and Quality of Patient Race Data in an Electronic Health Record System. *Applied clinical informatics*. 2020;11:785-791. PMID: 33241548. [Full text](#).
6. **Cole CL**, Sengupta S, Rossetti N, Collins S, Vawdrey DK, Halaas M, Maddox TM ... Estrin D. Ten principles for data sharing and commercialization. *Journal of the American*



Use identity metadata to construct lookup queries and retrieve candidate articles from PubMed

Case	Example query
One name on record, fewer than 2,000 candidate records (lenient query)	Albert P[au] OR paa2013@med.cornell.edu OR palbert1@gmail.com
One name on record, more than 2,000 candidate records (strict query)	<ul style="list-style-type: none">• Wang Y[au] AND [grants on record]• Wang Y[au] AND [departments on record]• Wang Y[au] AND [institutions on record]• Wang Y[au] AND [relationships on record]• Wang YA[au]• yiwang331@med.cornell.edu OR yiwang@hotmail.com
Compound names (e.g., Carol Storey-Johnson or Selin Somersan Karakaya)	Storey-Johnson C[au] OR Storey C{au] OR Johnson C[au]