

# Searching by Affiliation in ADS

*Carolyn Grant and the ADS Team*

ADS Users Group Meeting, 20-21 Nov. 2019



# Affiliations

## 1 year ago:

- Creation of canonical institution names, ids and abbreviations for facet (3800 inst + 2800 divisions)
- Assignment of raw affiliation strings to institutional ids (93% for astronomy, 70% for physics)
- Development of Python routine to match new and unmatched affiliations

# Affiliations

## Current status:

- Filter by facet on left panel

QUICK FIELD: [Author](#) [First Author](#) [Abstract](#) [Year](#) [Fulltext](#) [All Search Terms](#)

← Start New Search

(author:"Schmidt, Brian" |AND year:1980-2019)

Your search returned **1,178** results

Property  
+property:refereed

Sort Date

Export

Explore

- AUTHORS**
- Schmidt, B 1.1k
  - Schmidt, B 859
  - Schmidt, B P 80
  - Schmidt, Brian P 63
  - Schmidt, B E 48
  - Schmidt, B G 45
- [more](#)

- >  Cardini, A 480
  - >  Veltri, M 480
  - >  Graziani, G 479
  - >  Soler, F 478
- [more](#)

> COLLECTIONS

> REFEREED

- AFFILIATIONS**
- >  CNRS 288
  - >  U Heidelberg 269
  - >  U Oxford 268
  - >  UCamb 263
  - >  ITEP 254
- [more](#)

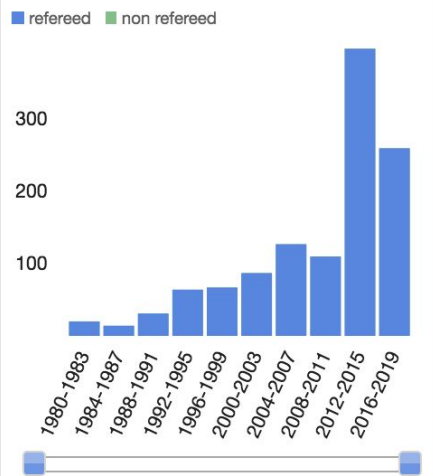
[Show highlights](#) [Show abstracts](#) [Hide Sidebars](#) [Go To Bottom](#)

1	2019PASP..131a4002H	2019/01	cited: 2			
<b>Carnegie Supernova Project-II: The Near-infrared Spectroscopy Program</b>						
Hsiao, E. Y.; Phillips, M. M.; Marion, G. H. <i>and 42 more</i>						
2	2019PASP..131a4001P	2019/01	cited: 3			
<b>Carnegie Supernova Project-II: Extending the Near-infrared Hubble Diagram for Type Ia Supernovae to z ~ 0.1</b>						
Phillips, M. M.; Contreras, Carlos; Hsiao, E. Y. <i>and 41 more</i>						
3	2018PhLB..787..124A	2018/12	cited: 2			
<b>Search for CP violation in <math>\Lambda_b^0 \rightarrow pK^-</math> and <math>\Lambda_b^0 \rightarrow \pi\pi^-</math> decays</b>						
Aaij, R.; Adeva, B.; Adinolfi, M. <i>and 836 more</i>						
4	2018JLTP..193..867D	2018/12				
<b>The CUORE Cryostat</b>						
D'Addabbo, A.; Alduino, C.; Bersani, A. <i>and 38 more</i>						
5	2018Icar..316...63H	2018/12	cited: 3			
<b>The Ac-5 (Fejokoo) quadrangle of Ceres: Geologic map and geomorphological evidence for ground ice mediated surface processes</b>						
Hughson, Kynan H. G.; Russell, C. T.; Williams, D. A. <i>and 19 more</i>						
6	2018Icar..316...46S	2018/12	cited: 2			
<b>Ceres' Ezinu quadrangle: a heavily cratered region with evidence for localized subsurface water ice and the context of Occator crater</b>						
Scully, Jennifer E. C.; Buczkowski, D. L.; Neesemann, A. <i>and 22 more</i>						
7	2018EPJC...78.1019A	2018/12				
<b>Evidence for an <math>\eta_c(1S) \pi^+ \pi^-</math> resonance in <math>B^0 \rightarrow \eta_c(1S) K^+ \pi^-</math> decays</b>						
Aaij, R.; Beteta, C.; Abellán, Adeva, B. <i>and 843 more</i>						

0 selected

Add papers to library

Years Citations Reads



Limit results to papers from

QUICK FIELD: Author First Author Abstract Year Fulltext All Search Terms

Start New Search

{author:"Schmidt, Brian" AND year:1980-2019}

Your search returned 143 results

Author Schmidt, B P
Author Schmidt, Brian P
Property +property:refereed

Date

Export

Explore

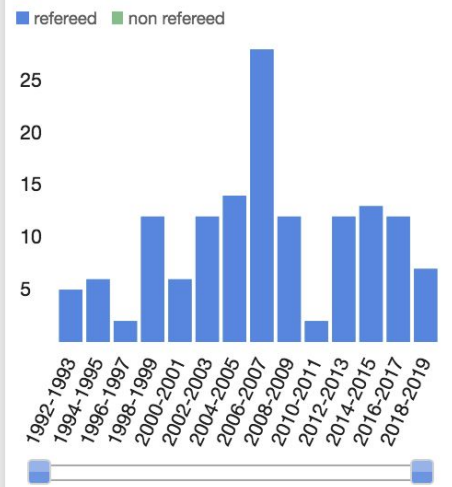
- AUTHORS
Schmidt, B 143
Schmidt, B P 80
Schmidt, Brian P 63
Filippenko, A 42
Kirshner, R 42
Challis, P 35
Suntzeff, N 35
COLLECTIONS
REFEREED
AFFILIATIONS
ANU 112
UCB 48
Manoa 46
ESO 43
Harvard U 42
KEYWORDS
PUBLICATIONS
BIB GROUPS

Show highlights Show abstracts Hide Sidebars Go To Bottom

2019PASP.131a4002H 2019/01 cited: 2 Carnegie Supernova Project-II: The Near-infrared Spectroscopy Program
2019PASP.131a4001P 2019/01 cited: 3 Carnegie Supernova Project-II: Extending the Near-infrared Hubble Diagram for Type Ia Supernovae to z ~ 0.1
2018PASA...35...24W 2018/06 cited: 2 Discovery of the Most Ultra-Luminous QSO Using GAIA, SkyMapper, and WISE
2018NatSR...8.8561S 2018/06 cited: 5 Sensations from a single M-cone depend on the activity of surrounding S-cones
2018MNRAS.477..766D 2018/06 cited: 17 SMSS J130522.47-293113.0: a high-latitude stellar X-ray source with pc-scale outflow relics?
2018MNRAS.475.1046I 2018/03 cited: 46 On the nature of hydrogen-rich superluminous supernovae
2018PASA...35...10W 2018/02 cited: 41 SkyMapper Southern Survey: First Data Release (DR1)

0 selected
Add papers to library

Years Citations Reads



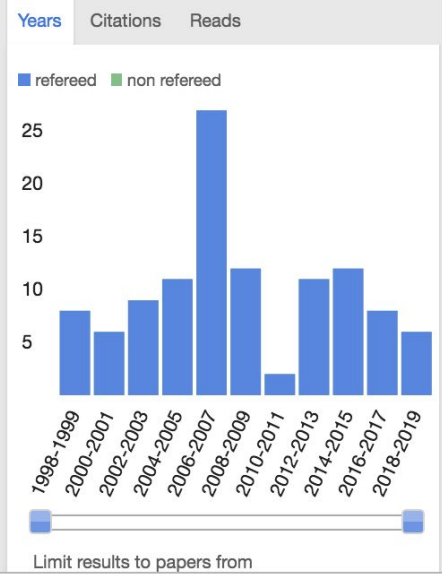
QUICK FIELD: [Author](#) [First Author](#) [Abstract](#) [Year](#) [Fulltext](#) [All Search Terms](#)[← Start New Search](#)(author:"Schmidt, Brian" AND year:1980-2019)  Your search returned **112** results[Author](#) Schmidt, B P [Author](#) Schmidt, Brian P [Property](#) +property:refereed [Affiliations](#) +ANU [Date](#) ▾[Export](#) ▾ [Explore](#) ▾

- ▼ **AUTHORS**
- ▼ Schmidt, B 112
  - Schmidt, B P 68
  - Schmidt, Brian P 44
- Filippenko, A 30
- Kulkarni, S 27
- Suntzeff, N 27
- Challis, P 26 [more](#)
- **COLLECTIONS**
- **REFEREED**
- ▼ **AFFILIATIONS**
- ▼ ANU 112
  - Res Sch Ast Astrop 97
  - CAASTRO 17
  - MSSSO 13
  - ANU 5
- UCB 47
- Manoa 46
- ESO 42
- CalTech 37 [more](#)

  [Go To Bottom](#)

Rank	ID	Year	Cited	Actions
1	2019PASP.131a4002H	2019/01	cited: 2	
<b>Carnegie Supernova Project-II: The Near-infrared Spectroscopy Program</b> Hsiao, E. Y.; Phillips, M. M.; Marion, G. H. <a href="#">and 42 more</a>				
2	2019PASP.131a4001P	2019/01	cited: 3	
<b>Carnegie Supernova Project-II: Extending the Near-infrared Hubble Diagram for Type Ia Supernovae to z ~ 0.1</b> Phillips, M. M.; Contreras, Carlos; Hsiao, E. Y. <a href="#">and 41 more</a>				
3	2018PASA...35...24W	2018/06	cited: 2	
<b>Discovery of the Most Ultra-Luminous QSO Using GAIA, SkyMapper, and WISE</b> Wolf, Christian; Bian, Fuyan; Onken, Christopher A. <a href="#">and 5 more</a>				
4	2018MNRAS.477..766D	2018/06	cited: 2	
<b>SMSS J130522.47-293113.0: a high-latitude stellar X-ray source with pc-scale outflow relics?</b> Da Costa, G. S.; Soria, R.; Farrell, S. A. <a href="#">and 17 more</a>				
5	2018MNRAS.475.1046I	2018/03	cited: 46	
<b>On the nature of hydrogen-rich superluminous supernovae</b> Inserra, C.; Smartt, S. J.; Gall, E. E. E. <a href="#">and 34 more</a>				
6	2018PASA...35...10W	2018/02	cited: 41	
<b>SkyMapper Southern Survey: First Data Release (DR1)</b> Wolf, Christian; Onken, Christopher A.; Luvaul, Lance C. <a href="#">and 14 more</a>				
7	2017PASA...34...69A	2017/12	cited: 28	
<b>Follow Up of GW170817 and Its Electromagnetic Counterpart by Australian-Led Observing Programmes</b> Andreoni, L.; Ackley, K.; Cooke, J. <a href="#">and 122 more</a>				

0 selected

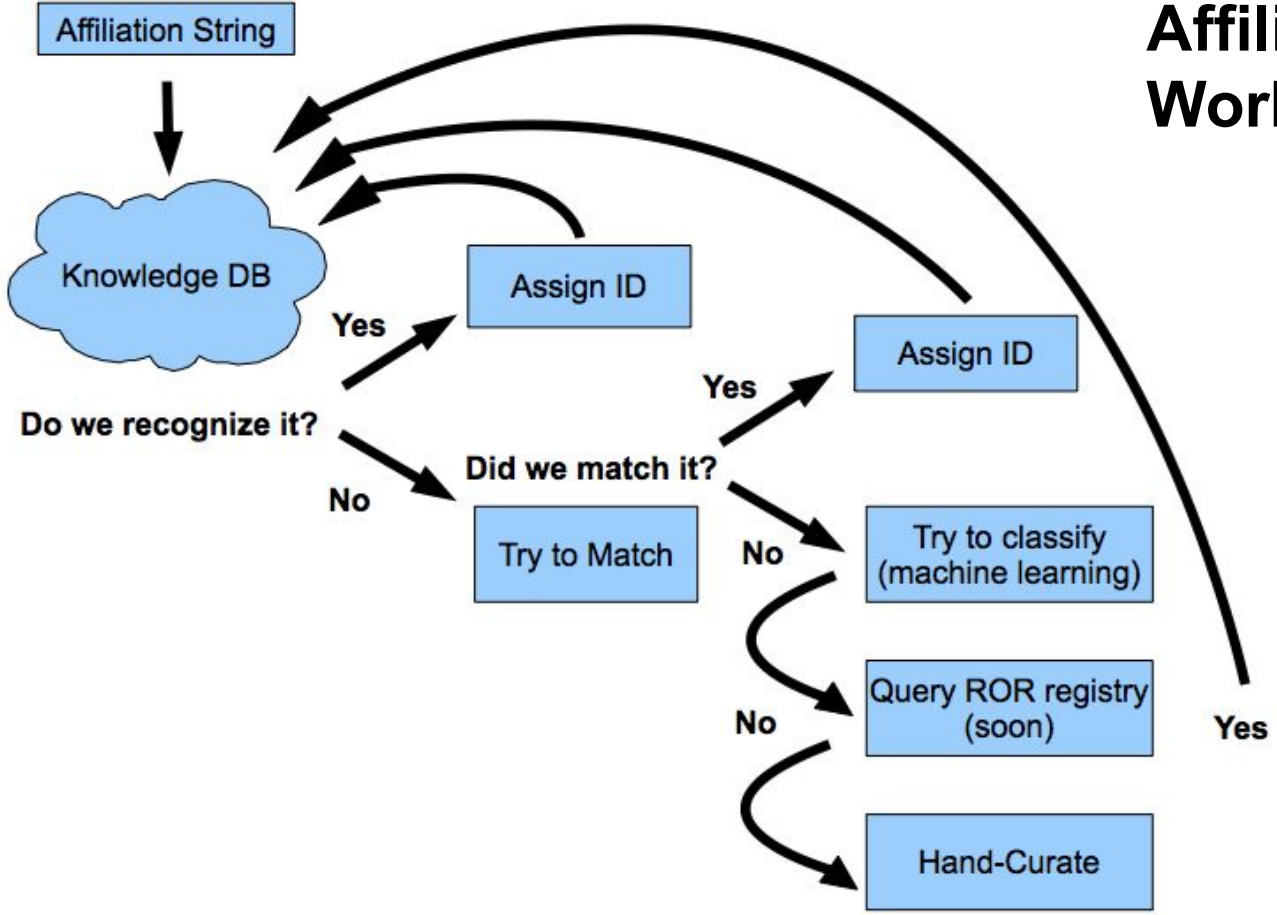
[Add papers to library](#) ▾

# Affiliations

## Current status:

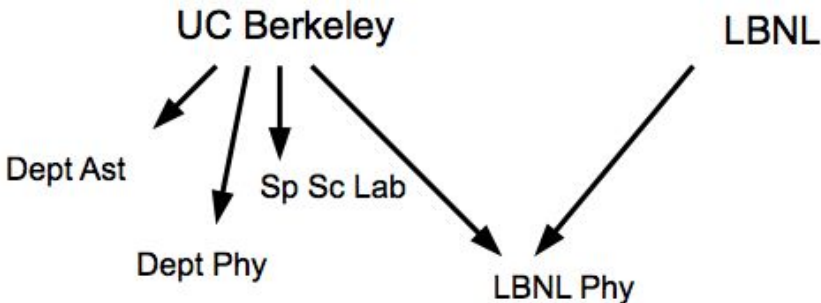
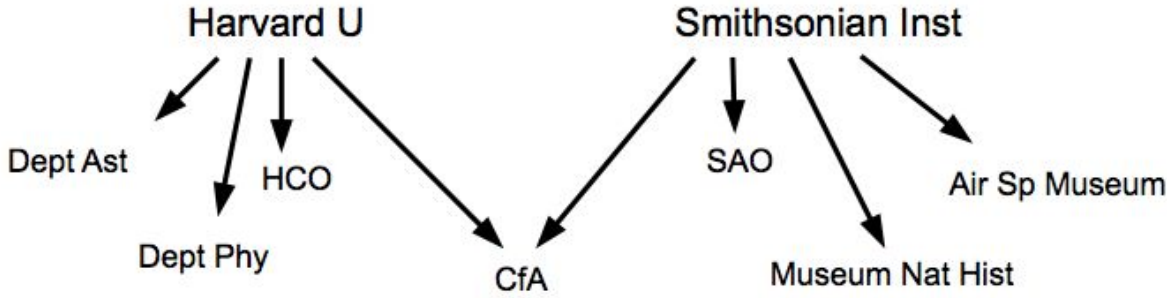
- Filter by facet on left panel
- Two-tiered affiliations, automatically matched on input

# Affiliation Workflow





# Two tiers of affiliations



# MIT

Lincoln Lab

4.7%

Space Science

5.3%

Engineering

6.1%

Earth Science

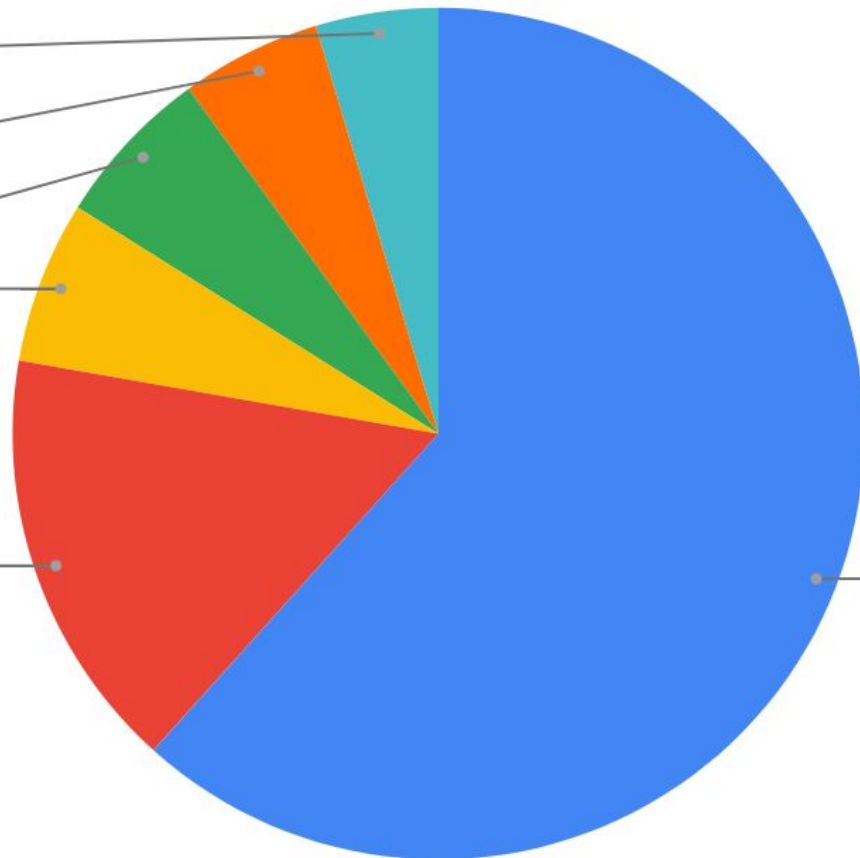
6.2%

Physics

16.1%

MIT parent

61.7%

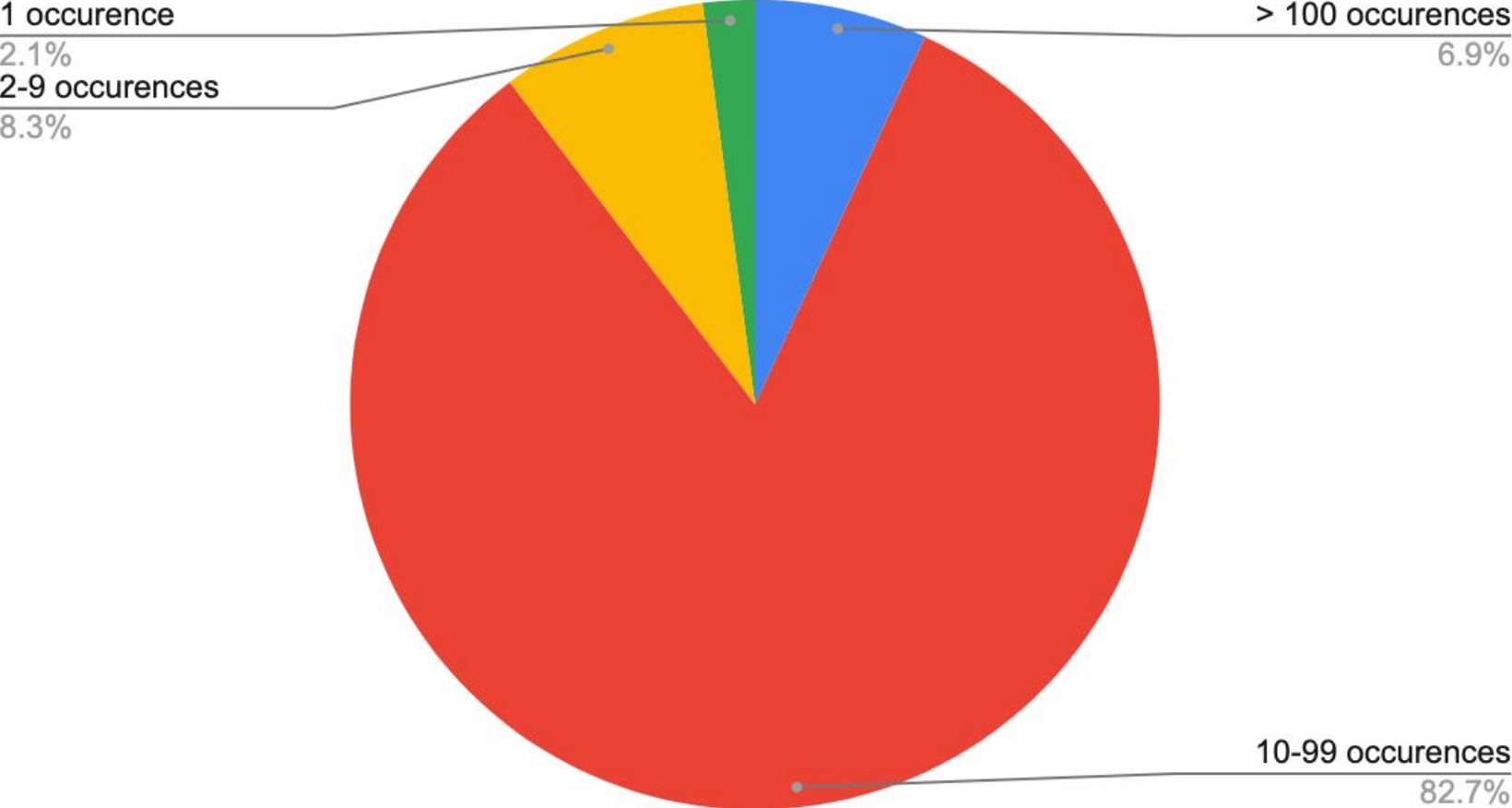


# Affiliations

## Current status:

- Filter by facet on left panel
- Two-tiered affiliations, automatically matched on input
- 19.3 million matched affiliation strings (of 36.6 million)

# 19.3 million matched affiliation strings



# Matched affiliations:

## Top matches:

- 348K CERN
- 139K Argonne National Lab
- 130K Fermilab
- 126K JINR (Dubna)
- 116K DESY
- 115K Brookhaven National Lab
- 106K JPL
- 101K MIT
- 100K IHEP, Beijing (CAS)

## Top matches (parent institutions):

- 662K INFN
- 356K CERN
- 291K Max Planck Society
- 269K Helmholtz Association
- 260K Caltech
- 256K Chinese Acad of Sci
- 218K Russian Acad of Sci
- 171K UC Berkeley
- 168K Stanford
- 168K MIT

(847K DOE, also NASA, CNRS)

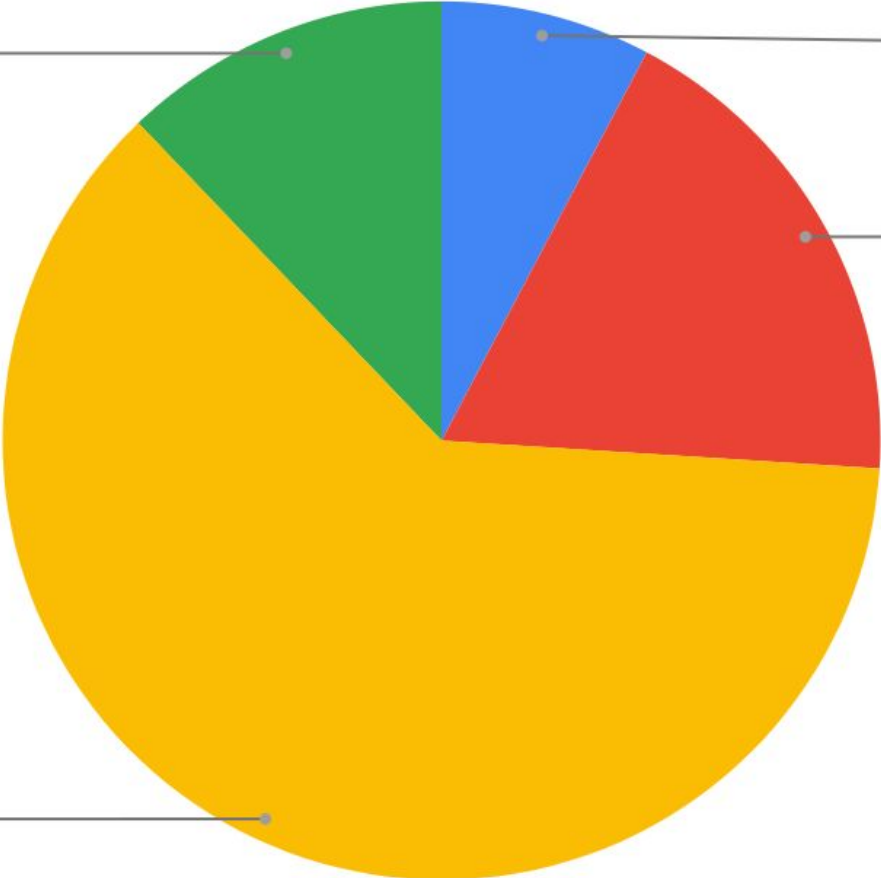
# 17.3 million unmatched affiliation strings

1 occurrence  
12.1%

> 100 occurrences  
7.7%

10-99 occurrences  
18.3%

2-9 occurrences  
61.8%



# Unmatched affiliations:

## Top unmatched include:

- Institute of High Energy Physics
- Institute for High Energy Physics
- These authors contributed equally to this work
- To whom any correspondence should be addressed
- Corresponding author
- .
- ,
- Presenting author

Plus others we update and reprocess.

# Affiliations

## Current status:

- Filter by facet on left panel
- Two-tiered affiliations, automatically matched on input
- 19.3 million matched affiliation strings
- Search by institution (canonical affiliation)



# Institutions vs. Affiliations

Institution is the “canonical” version of an affiliation, part of a controlled dictionary. Affiliation is simply a string search.

Use institution for broader scope

- `aff:"CfA"` = 3900
- `aff:"CfA" OR aff:"Harvard Smithsonian"` = 32K
- `inst:"CfA"` = 32K

There's a learning curve!

- `aff:"Harvard"` = 82K
- `inst:"Harvard"` = 0
- `inst:"Harvard U"` = 60K
- `inst:"Harvard U/Dep Ast"` = 750
- `inst:"Harvard U/Dep Phy"` = 8300

Autocomplete will help.

QUICK FIELD: [Author](#) [First Author](#) [Abstract](#) [Year](#) [Fulltext](#) [All Search Terms](#)

← Start New Search

inst:"UCLA"

Your search returned **35,712** results

- ▼ AUTHORS
  - >  Russell, C 2.9k
  - >  Lee, J 1.1k
  - >  Kim, J 1.1k
  - >  Yang, Y 1.1k
  - >  Lyons, L 1k
  -
- ▼ COLLECTIONS
- ▼ REFEREED
- ▼ AFFILIATIONS
  - >  UCLA 35.7k
    - UCLA 19.9k
    - IGPP 7.8k
    - Dep Phy Ast 7.3k
    - Dep Atm Oc Sci 1.6k

[Go To Bottom](#)

- 1  2019GSL.....6....3M 2019/12   

**A survey of submesoscale currents**  
McWilliams, James C.
- 2  2019APh...111...35A 2019/09 cited: 2   

**Monte Carlo studies for the optimisation of the Cherenkov Telescope Array layout**  
Acharya, A.; Agudo, I.; Angüner, E. O. *and 467 more*
- 3  2019MNRAS.486.2780Y 2019/06   

**Atmospheric mass-loss from high-velocity giant impacts**  
Yalinewich, Almog; Schlichting, Hilke
- 4  2019MNRAS.486.2254D 2019/06   

**Hunting for brown dwarfs in the globular cluster M4: second epoch HST NIR observations**  
Dieball, A.; Bedin, L. R.; Knigge, C. *and 5 more*
- 5  2019JAE0..78....1Y 2019/06   

**Supraglacial rivers on the northwest Greenland Ice Sheet, Devon Ice Cap, and Barnes Ice Cap mapped using Sentinel-2 imagery**  
Yoon, Kyoung; Smith, Lawrence C.; Cole, Andrew *and 4 more*

0 selected

Years Citations Reads

■ refereed ■ non refereed

Year	Referenced (k)	Non-Referenced (k)	Total (k)
2019	~0.5	~0.5	~1.0
2020	~1.5	~1.5	~3.0
2021	~2.5	~2.5	~5.0
2022	~4.0	~4.0	~8.0
2023	~7.0	~5.0	~12.0

QUICK FIELD: [Author](#) [First Author](#) [Abstract](#) [Year](#) [Fulltext](#) [All Search Terms](#)

inst:"UCLA/IGPP"

Your search returned **7,309** results

- AUTHORS**
- Russell, C 1.9k
  - Angelopoulos, V 874
  - Strangeway, R 451
  - Raymond, C 342
  - Luhmann, J 336
- 

- COLLECTIONS**
- IGPP 7.3k
  - UCLA 332
  - Dep Atm Oc Sci 330
  - Dep Phy Ast 223

[Go To Bottom](#)

- 1  2019JVGR..376...27S 2019/05   

**Age of the dacite of Sunset Amphitheater, a voluminous Pleistocene tephra from Mount Rainier (USA), and implications for Cascade glacial stratigraphy**  
 Sisson, Thomas W.; Schmitt, Axel K.; Danišik, Martin *and 4 more*
- 2  2019Icar..323....1Y 2019/05   

**Near-equilibrium isotope fractionation during planetesimal evaporation**  
 Young, E. D.; Shahr, A.; Nimmo, F. *and 4 more*
- 3  2019RAA....19...50S 2019/04   

**Reminiscences of my life in science**  
 Schubert, Gerald
- 4  2019PhRvE..99d3204C 2019/04   

**In situ spacecraft observations of a structured electron diffusion region during magnetopause reconnection**  
 Cozzani, Giulia; Retinò, A.; Califano, F. *and 18 more*
- 5  2019M&PS...54..699J 2019/04   

**The Greater California subarcuate foreland the origin of the...**

0 selected



QUICK FIELD: [Author](#) [First Author](#) [Abstract](#) [Year](#) [Fulltext](#) [All Search Terms](#)

← Start New Search

((inst:"UCLA/IGPP") AND year:2008) ✕ 🔍

Your search returned **406** results

☰ Date ▾

📄 Export ▾

📊 Explore ▾

- ▼ AUTHORS
  - ☐ Russell, C 142
  - ☐ Angelopoulos, V 82
  - ☐ Glassmeier, K 55
  - ☐ McFadden, J 52
  - ☐ Larson, D 42
  - [more](#)
- > COLLECTIONS
- > REFEREED
- ▼ AFFILIATIONS
  - ☐ UCLA 406
  - ☐ UCB 121
  - ☐ TU Braunschweig 69
  - ☐ GSFC 65
  - ☐ JHU 49

☐ [Show highlights](#) [Show abstracts](#) [Hide Sidebars](#) [Go To Bottom](#)

1	☐	2008SSRv..141..557P	2008/12	cited: 12	📄 ☰ 📄
<b>The Time History of Events and Macroscale Interactions during Substorms (THEMIS) Education and Outreach (E/PO) Program</b>					
Peticolas, L. M.; Craig, N.; Odenwald, S. F. <i>and 8 more</i>					
2	☐	2008SSRv..141..509L	2008/12	cited: 76	📄 ☰ 📄
<b>First Results of the THEMIS Search Coil Magnetometers</b>					
Le Contel, O.; Roux, A.; Robert, P. <i>and 15 more</i>					
3	☐	2008SSRv..141..453A	2008/12	cited: 104	📄 ☰ 📄
<b>First Results from the THEMIS Mission</b>					
Angelopoulos, V.; Sibeck, D.; Carlson, C. W. <i>and 30 more</i>					
4	☐	2008SSRv..141..413M	2008/12	cited: 142	📄 ☰ 📄
<b>The Upgraded CARISMA Magnetometer Array in the THEMIS Era</b>					
Mann, I. R.; Milling, D. K.; Rae, I. J. <i>and 15 more</i>					
5	☐	2008SSRv..141..389R	2008/12	cited: 84	📄 ☰ 📄
<b>THEMIS Ground-Based Magnetometers</b>					
Russell, C. T.; Chi, P. L.; Deaerborn, D. L. <i>and 6 more</i>					

0 selected

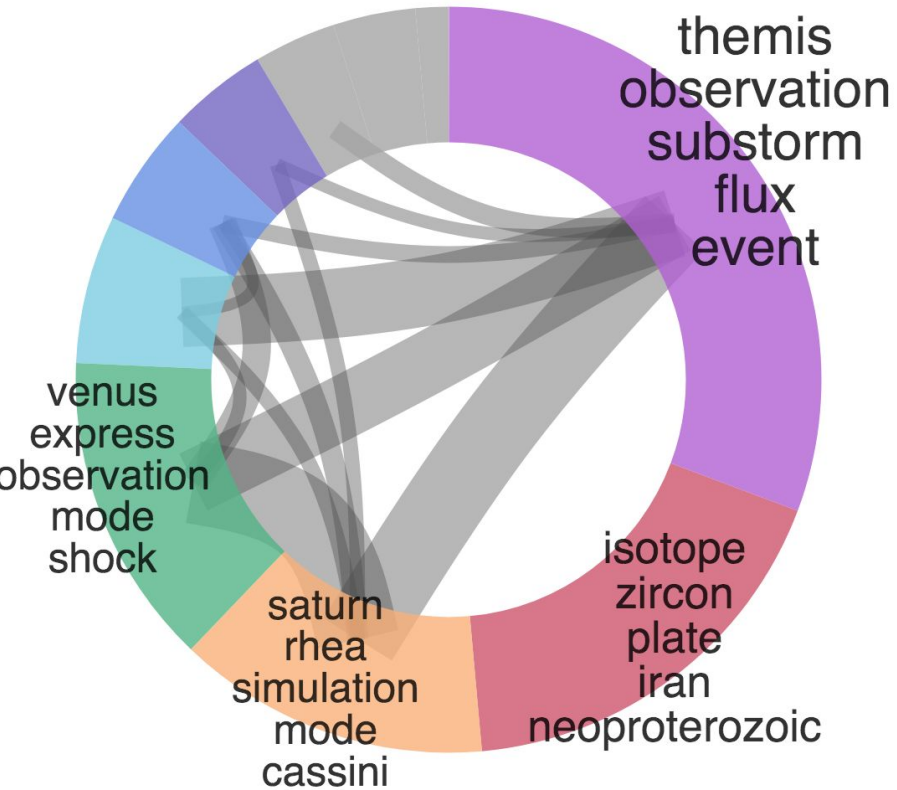
[📄 Add papers to library](#) ▾

Years Citations Reads

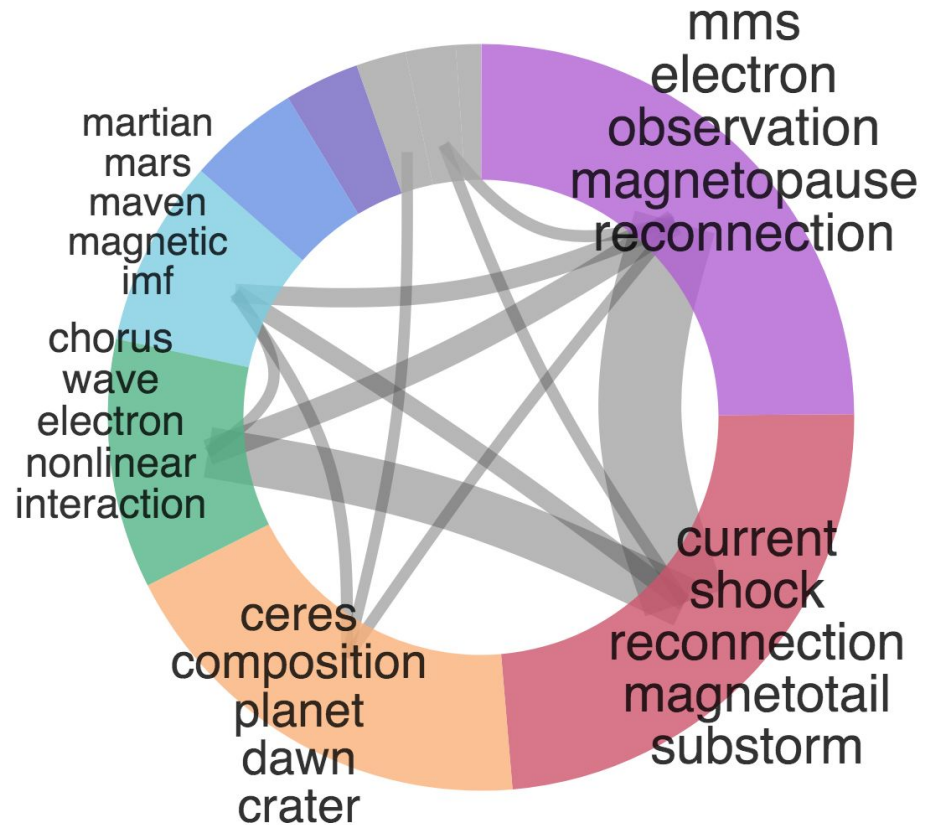
Too little data to make a useful graph.

# UCLA/IGPP

2008



2018

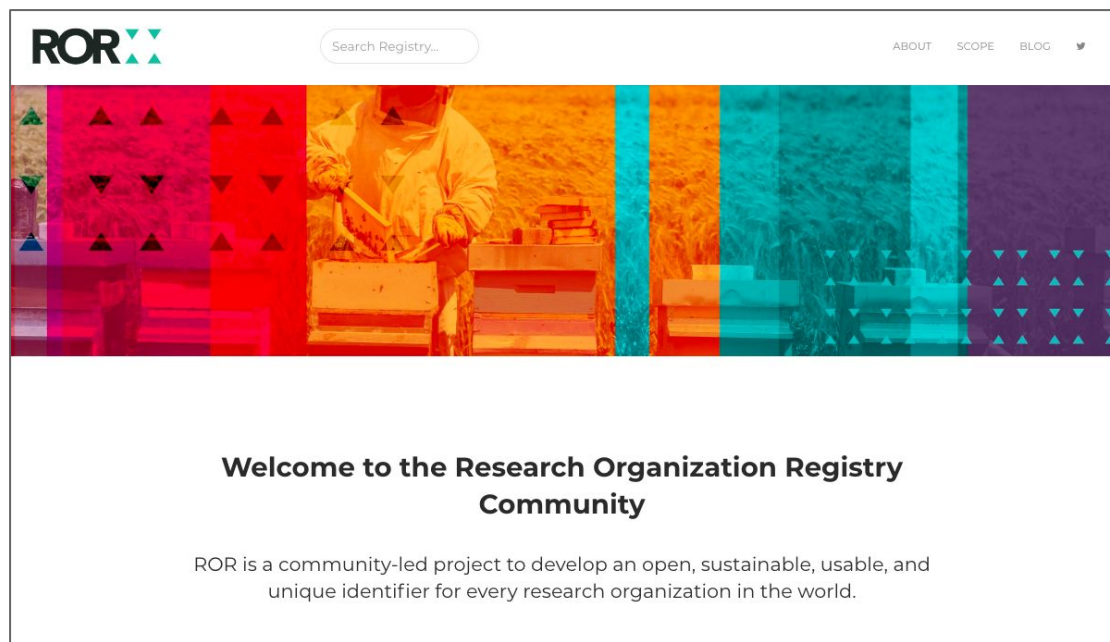


# Affiliations

## Current status:

- Filter by facet on left panel
- Two-tiered affiliations, automatically matched on input
- 19.3 million matched affiliation strings
- Search by institution (canonical affiliation)
- Working on integration with ROR

# ROR MVR

A screenshot of the ROR MVR website homepage. The page features a colorful header with a central image of a person in a white lab coat working in a laboratory. The header includes the ROR logo on the left, a search bar with the text "Search Registry..." in the center, and navigation links for "ABOUT", "SCOPE", "BLOG", and a Twitter icon on the right. Below the header, the main content area has a white background with the following text:

**Welcome to the Research Organization Registry  
Community**

ROR is a community-led project to develop an open, sustainable, usable, and unique identifier for every research organization in the world.

# ROR project team

Outreach,  
communications,  
strategic  
direction

Community  
advisors



University of California  
**CDL**  
California Digital Library



**DataCite**  
FIND, ACCESS, AND REUSE DATA

**DIGITAL**  
science



# ROR metadata

ROR ID

(URL resolves to the ROR record)

 <https://ror.org/03yrm5c26>

Name + alternate names/abbreviations

California Digital Library  
CDL

WEBSITE  
<http://www.cdlib.org/>

OTHER IDENTIFIERS  
GRID [grid.463323.3](#)  
ISNI [0000000119575136](#)  
Wikidata [Q5020447](#)

Other identifiers  
for the  
organization  
(if available)

UNITED STATES ARCHIVE

Other metadata  
(more to come!)

# Still to do

- Integrate publisher-provided identifiers
- Search by other identifiers (ROR, ISNI, etc)
- Integrate with ROR on a department level
- Improvement of user experience (e.g. hover over, auto-complete)
- Couple with authors
- Improve splitting affiliations with multiple institutions

# Great, I can search by Institutions. What about Collaborations?

*Carolyn Grant and the ADS Team*

ADS Users Group Meeting, 20-21 Nov. 2019



# Collaboration challenges

- Normalization
  - 134 different variations of “ATLAS Collaboration”
  - Use of “the”, “for the”, “on behalf of”, “and the”, etc.
- Expansion
  - Individual authors want credit
  - Collaboration members change over time
- Extraction
  - Not always in the metadata (appendix, affiliation)
  - Sometimes twice in the metadata
  - Journals inconsistent with placement
- Not Authors
  - Search implications
  - Citation implications

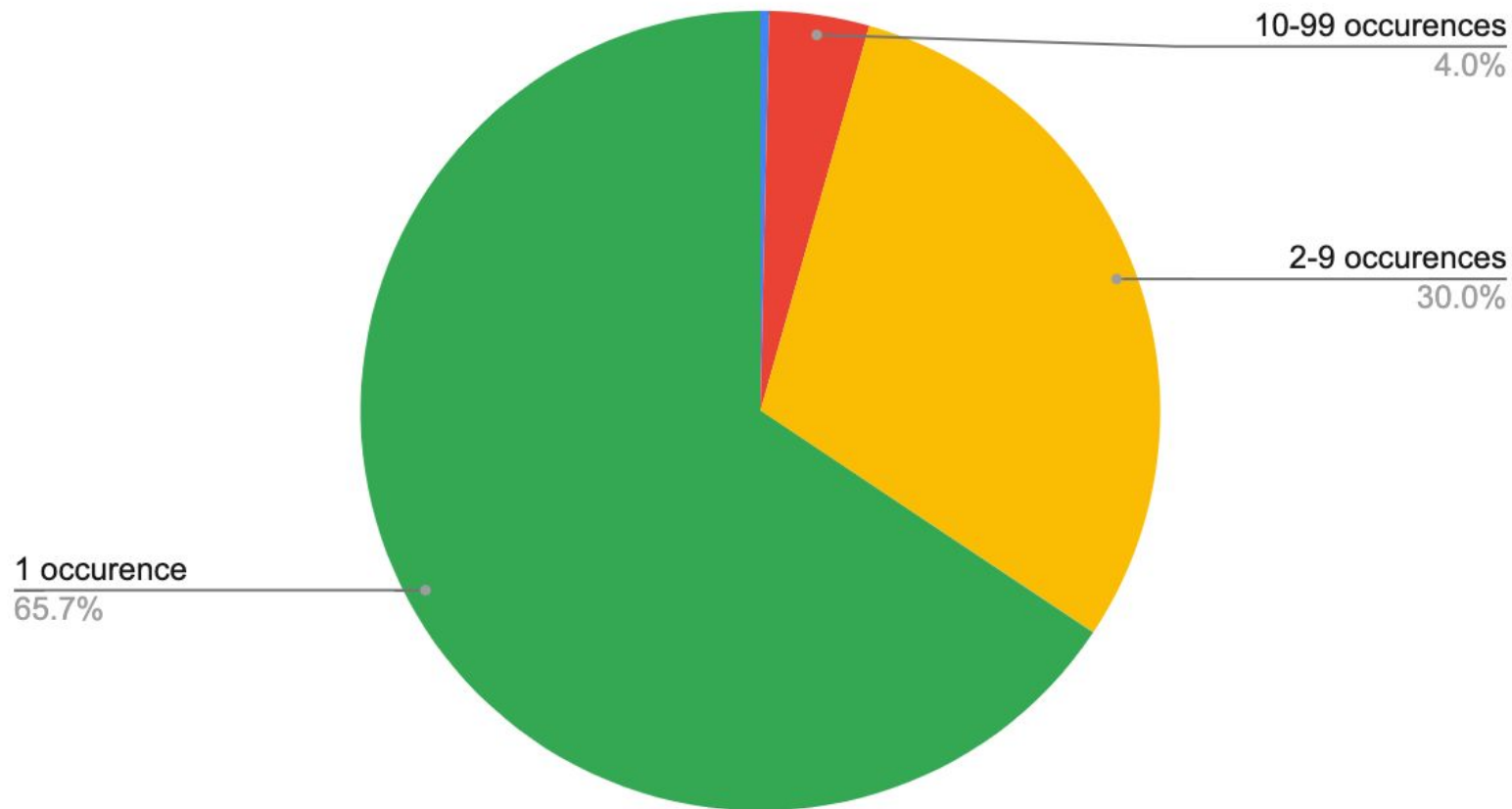
1582 Atlas Collaboration  
355 ATLAS Collaboration  
21 Atlas Muon Collaboration  
15 ATLAS Liquid Argon Calorimeter Group  
14 Atlas Sct Collaboration  
14 ATLAST Concept Study Team  
12 ATLAS Muon Collaboration  
9 Atlas Pixel Collaboration  
8 Atlas Liquid Argon Calorimeter Group  
8 ATLAS Tile Calorimeter Group  
8 ATLAS TDAQ Collaboration  
5 H-ATLAS Team  
5 ATLAS Pixel Collaboration  
4 The ATLAS Collaboration  
4 Herschel-ATLAS Consortium  
4 Atlas3d Team  
4 Atlas Tdaq Collaboration  
4 Atlas Experiment Collaboration  
4 ATLAST Team

etc. (134)

291 Fermi LAT Collaboration  
229 Fermi-LAT Collaboration  
82 Fermi Large Area Telescope Collaboration  
26 Fermi Collaboration  
23 Fermi GBM Collaboration  
19 Fermi-Lat Collaboration  
18 Fermi/LAT Collaboration  
15 Fermi Team  
11 Fermilab E653 Collaboration  
10 Fermi GBM Team  
7 Fermi/GBM Collaboration  
6 Fermilab E791 Collaboration  
5 Fermi/Lat Collaboration  
5 Fermi Pulsar Search Consortium  
5 Fermi Mission Team  
5 Fermi Gamma-ray Burst Monitor Team  
4 Fermilab E989 Collaboration  
4 Fermilab E665 Collaboration  
4 Fermi Large Area Space Telescope Collaboration

etc. (117)

## 20K unique collaboration strings



## **Collaboration low-hanging fruit**

- Data cleanup (capitalization, "The")
- Basic synonyms

## **Collaboration middle branches?**

- Incorporating our partner metadata (INSPIRE)
- Handling differently from authors (index and search)

## **Collaboration top of the tree?**

- Expand into individuals

# And what about Holdings?

*Matthew Templeton and the ADS Team*

ADS Users Group Meeting, 20-21 Nov. 2019





# Thank you!

[cgrant@cfa.harvard.edu](mailto:cgrant@cfa.harvard.edu)



## ISNI Relationships:

- isMemberOf
- hasMember
- isUnitOf
- hasUnit
- isSupersededBy
- supersedes
- isAffiliatedWith
- isRelatedTo
- formerName
- laterName
- see also from
- no relation

### Recommended additions\*:

- acquired
- acquired by
- isGovernedBy
- Governs
- isHostedBy
- Hosts
- isPartneredWith



\*Smith-Yoshimura, Karen et al. "Addressing the Challenges with Organizational Identifiers and ISNI", OCLC Research Report, 2016, <https://www.oclc.org/research/publications/2016/oclcresearch-organizational-identifiers-and-isni-2016.html>