

Author disambiguation and ORCID integration

Jan Åge Lavik & Javier Martin Montull
CERN

Invenio User Group Workshop
Jülich, Germany
17–20 November 2013

BibAuthorID

author disambiguation
paper attribution

WebAuthorProfile

author profiles

Sicking, Eva (60 papers relevant to High Energy Physics)

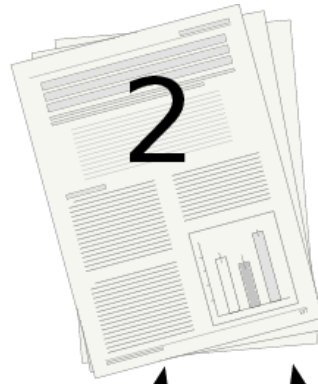
[This is me. Verify my publication list.](#)

Name variants	Papers	Citations (from papers in INSPIRE):		
Sicking, Eva (33)		Citations summary		
Sicking, E. (21)		Generated on 2013-04-05		
Sicking, E (6)		60 papers found, 60 of them citeable (published or arXiv)		
		Citation summary results	Citeable papers	Published only
		Total number of papers analyzed:	60	45
		Total number of citations:	2,514	2,427
		Average citations per paper:	41.9	53.9
		Breakdown of papers by citations:		
		Renowned papers (500+)	0	0
		Famous papers (250-499)	1	1
		Very well-known papers (100-249)	7	7
		Well-known papers (50-99)	8	8
Affiliations	Collaborations			
CERN (38)	ALICE Collaboration (58)			
Munster U. (11)	ALICE Collaboration (2)			
Munster U., ITP (1)				
UNDEFINED, Institut für Kernphysik, Westfälische Wilhelms-Universität Münster, Münster, Germany (1)				
Frequent co-authors (excluding collaborations)	Frequent keywords			
No Frequent Co-authors	ALICE (61)			
	CERN LHC Coll (53)			
	experimental results (48)			
	heavy ion: scattering (23)			
	lead (22)			
	2760 GeV-cms/nucleon (20)			
	n ⁰ : scattering (16)			

Short overview



Doe, J. Doe, J.



Doe, J. Doe, J.



Doe, J.



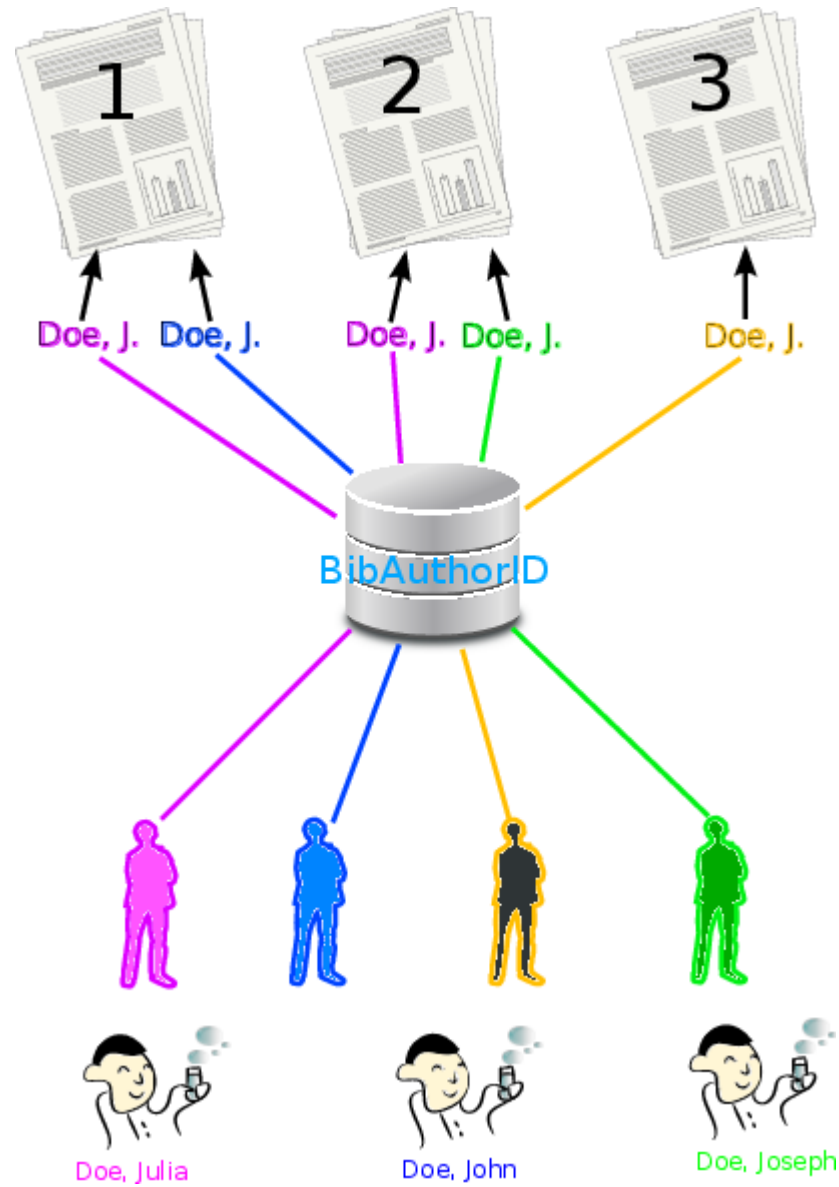
Doe, Julia

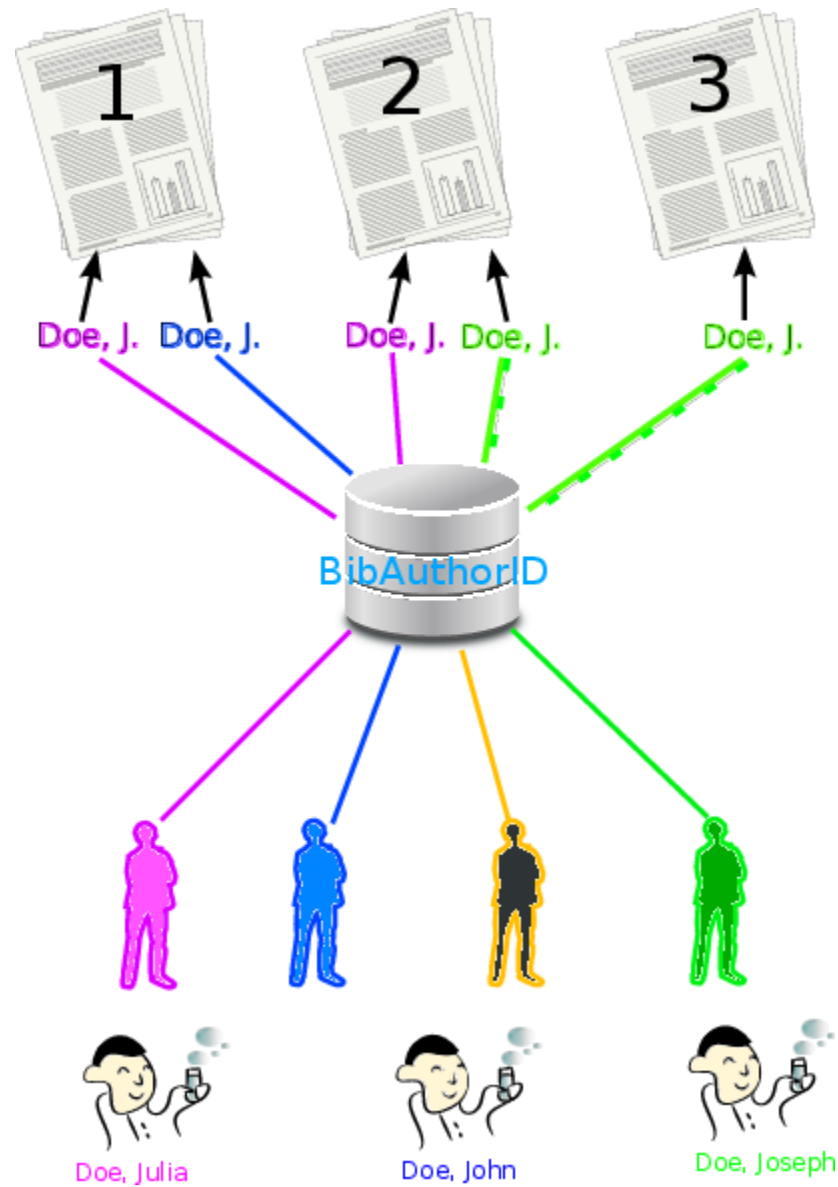


Doe, John



Doe, Joseph





[This is me. Verify my publication list.](#)

Ellis, Jonathan Richard (956 papers relevant to High Energy Physics)

[This is me. Verify my publication list.](#)

Name variants	Papers	
Ellis, Jonathan Richard (1)	All papers	Single authored
Ellis, Jonathan R. (2)	All papers 956	234
Ellis, John.R. (1)	Book 2	0
Ellis, John R. (842)	ConferencePaper 233	158
Ellis, John (72)	Introductory 21	19
Ellis, J.R. (12)	Lectures 46	36
Ellis, J. (25)	Published 626	54
Ellis, J (1)	Review 105	72
	Thesis 1	1
	Proceedings 16	0

Affiliations	Collaborations
CERN (849)	CPLEAR Collaboration (3)
SLAC (32)	ILC Collaboration (2)
King's Coll. London	CHEEP Study Groups (1)

Citations (from papers in INSPIRE):

Citations summary

Generated on 2013-06-25

956 papers found, 781 of them citeable (published or arXiv)

Citation summary results	Citeable papers	Published only
Total number of papers analyzed:	781	626
Total number of citations:	57,070	54,295
Average citations per paper:	73.1	86.7
Breakdown of papers by citations:		
Renowned papers (500+)	12	12
Famous papers (250-499)	40	39

Attribute papers for: S.Carli.1

[View Profile](#)
[Manage Profile](#)
[Attribute Papers](#)
[Help](#)

[View name variants](#)

Carli, W. (2)
 Carli, Samuele (1)

[Papers \(3\)](#)
[Papers removed from this profile \(0\)](#)

[Select All](#) |
 [Select None](#) |
 [Invert Selection](#) |
 [Hide successful claims](#)

↪
 [Yes, those papers are by this person.](#)
[No, those papers are not by this person](#)
[Assign to other person](#)

Search: <input type="text"/>						
	Paper Short Info	Author Name	Affiliation	Date	Experiment	Actions
<input type="checkbox"/>	1. An Energy Upgrade of the REX ISOLDE Linac O. Kester, H. Bongers, W. Carli, S. Emhofer, D. Habs, K. Rudolph, T. Sieber (Munich U.).	Carli, W.	Munich U.	2003-05	N.A.	✓ Yes, this paper is by this person. ✗ No, this paper is <i>not</i> by this person ➔ Assign to another person
<input type="checkbox"/>	2. OBSERVATIONS ON THE MUNICH MP ARC DISCHARGE CONDITIONING TEST N.F. Ziegler, P.H. Stelson (Oak Ridge), W. Carli, H. Munzer, G. Korschinek (Munich, Tech. U.).	Carli, W.	Munich, Tech. U.	1984	N.A.	✓ Yes, this paper is by this person. ✗ No, this paper is <i>not</i> by this person ➔ Assign to another person
<input type="checkbox"/>	3. Authormagic in INSPIRE Author Disambiguation in Scholarly Communication Travis C. Brooks (SLAC), Samuele Carli (CERN & Florence U.), Sünje Dallmeier-Tiessen (CERN & Humboldt U., Berlin), Salvatore Mele (CERN), Henning Weiler (CERN & Erlangen - Nuremberg U.). 08 June 2011	Carli, Samuele	CERN, Florence U.	Not available	N.A.	✓ Yes, this paper is by this person. ✗ No, this paper is <i>not</i> by this person ➔ Assign to another person

Showing 1 to 3 of 3 entries

[First](#)
[Previous](#)
[1](#)
[Next](#)
[Last](#)

INSPIRE use case

~9 million author names (signatures)

on 1 million papers

for ~50 000 actual persons

INSPIRE use case

Disambiguation takes ~10 days on Intel
X5650@2.67GHz with 24GB Ram

Periodical update of clusters may be needed

Uses paper claims to refine results

**500 000* signatures
manually claimed**

to 7000* distinct author profiles

**Numbers as of Summer 2013*

Disambiguation & clustering



Doe, J.
 Somewhere Univ.
 ORCID-0000
 2012
 Detectors



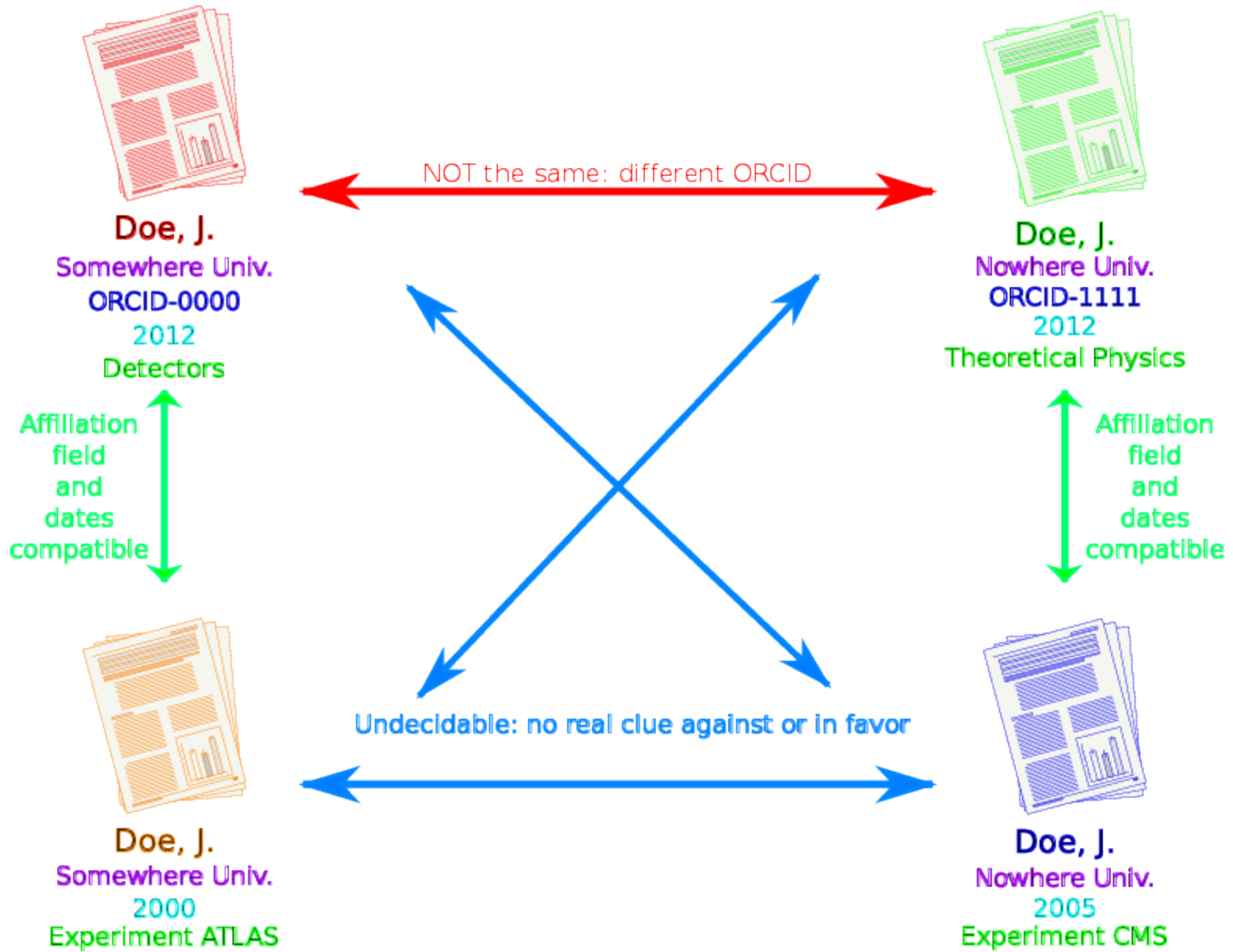
Doe, J.
 Nowhere Univ.
 ORCID-1111
 2012
 Theoretical Physics

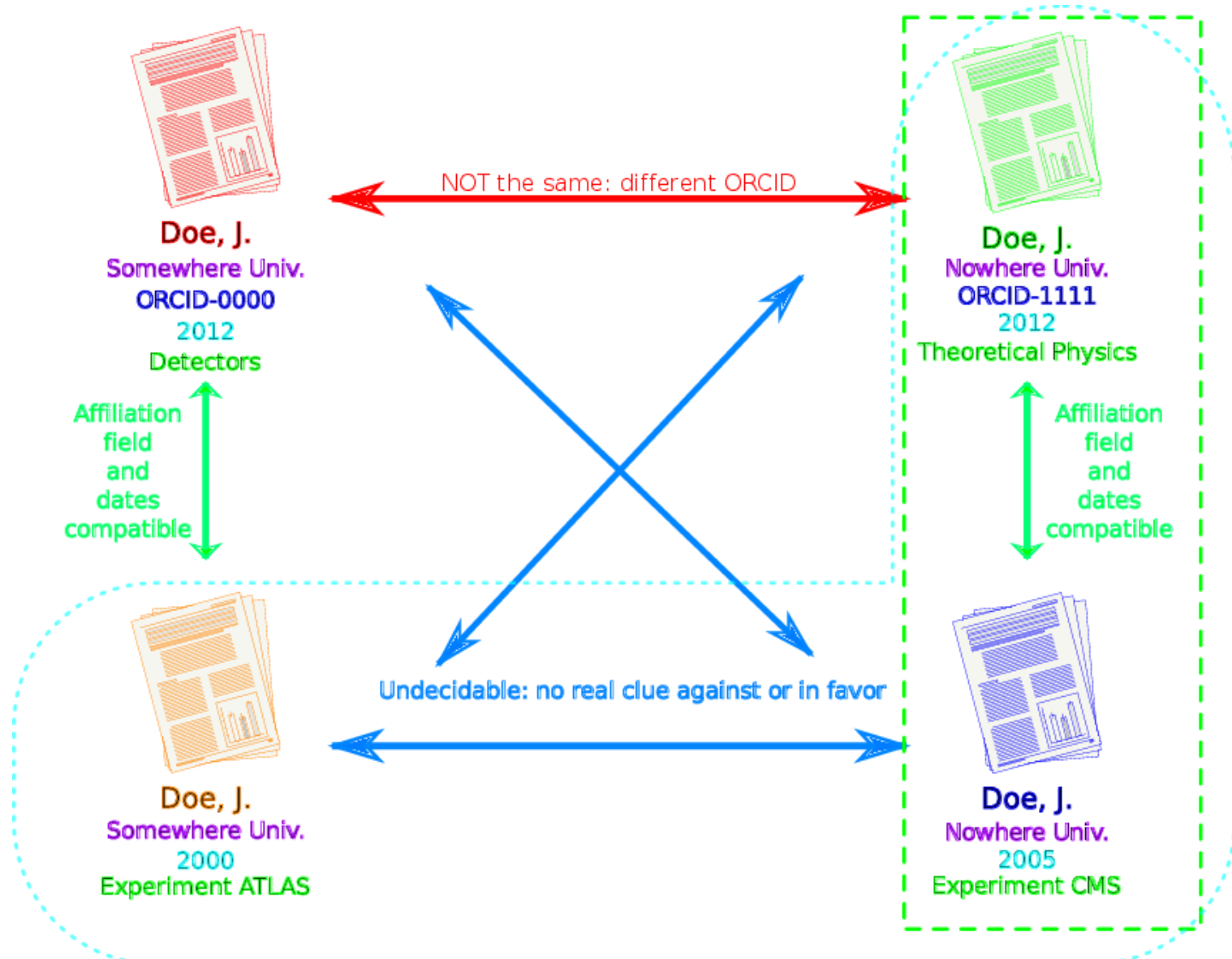


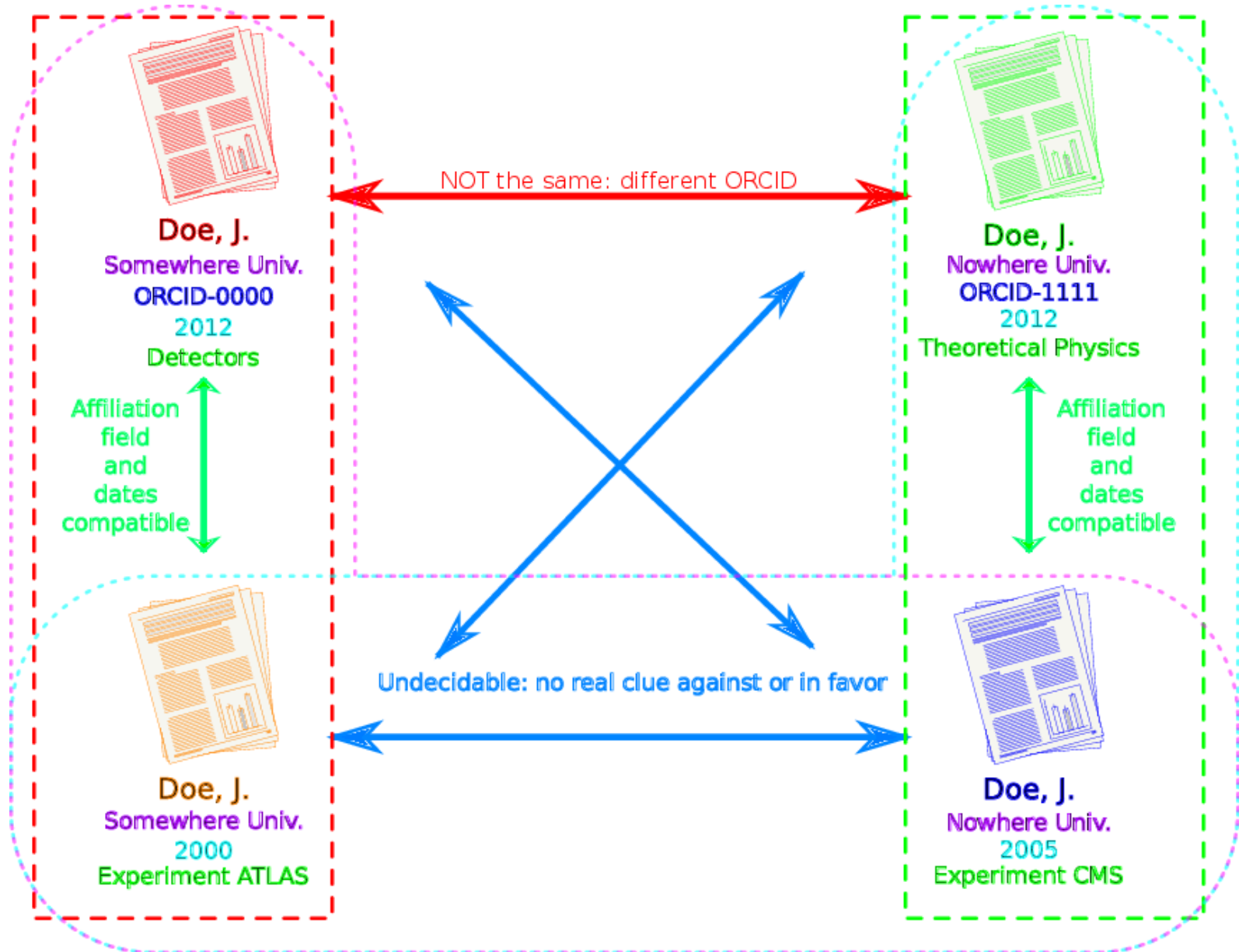
Doe, J.
 Somewhere Univ.
 2000
 Experiment ATLAS

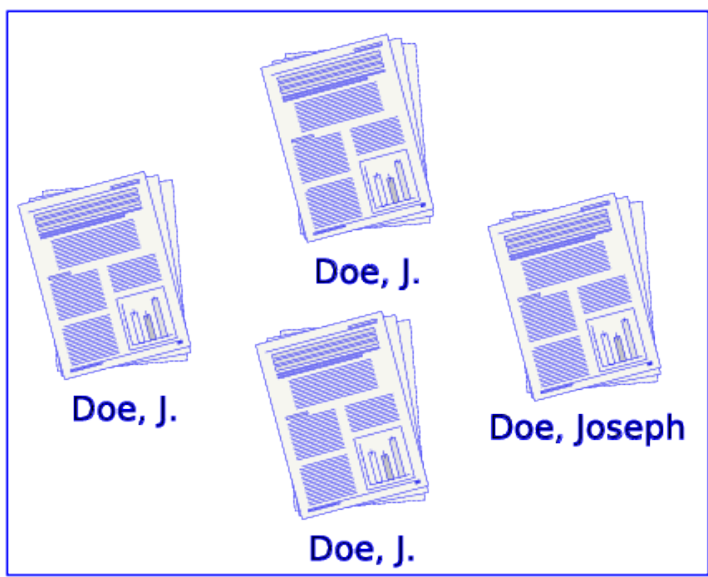


Doe, J.
 Nowhere Univ.
 2005
 Experiment CMS









Doe, J.

Doe, J.

Doe, Joseph

Doe, J.

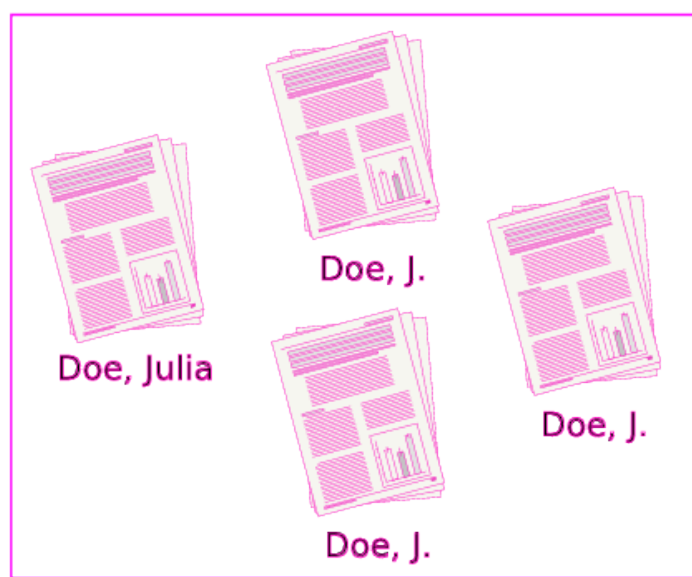


Doe, J.

Doe, J.

Doe, J.

Doe, J.



Doe, Julia

Doe, J.

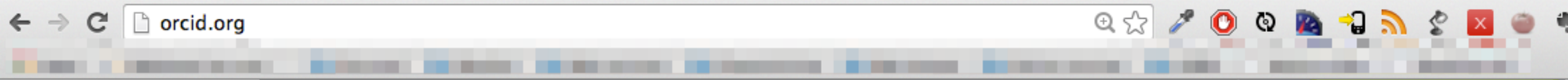
Doe, J.

Doe, J.

What is ORCID?

?





ORCID

Connecting Research and Researchers

FOR RESEARCHERS

FOR ORGANIZATIONS

ABOUT

HELP

SIGN IN

DISTINGUISH YOURSELF IN THREE EASY STEPS

ORCID provides a persistent digital identifier that distinguishes you from every other researcher and, through integration in key research workflows such as manuscript and grant submission, supports automated linkages between you and your professional activities ensuring that your work is recognized. [Find out more.](#)



REGISTER

Get your unique ORCID identifier [Register now!](#)
Registration takes 30 seconds.



ADD YOUR INFO

Enhance your ORCID record with your professional information and link to your other identifiers (such as Scopus or ResearcherID or LinkedIn).



USE YOUR ORCID ID

Include your ORCID identifier on your Webpage, when you submit publications, apply for grants, and in any research workflow to ensure you get



LATEST NEWS

Wed 11/13/2013
Universities Now
Creating ORCID iDs
for their Researchers
and Scholars

Fri 11/01/2013
ORCID Celebrates Its
First Year

Mon 10/28/2013
ORCID Public Data
File Now Available

- Open Researcher and Contributor ID
- Unique and persistent
- Each researcher:
 - Owns the ORCID record
 - Can add name variants, affiliations, publications, datasets...
 - Manages data sharing (private, restricted, open)
- Can be linked to other IDs



ORCiD-INVENIO integration

- Right now:
 - Has been developed under the ODIN-project (ORCiD and DataCite Interoperability Network)
 - Tailored to the author profiles in INSPIRE

Cranmer, Kyle S.

View Profile **Manage Profile** Manage Publications Search Profiles Help

Login with your arXiv.org account

Please log in through arXiv to manage your profile.

Login into Inspire through arXiv.org

Manage publications

Assign publications to your INSPIRE profile to keep it up to date.

Manage publication list

Connect this profile to an ORCID

This profile has not been connected to an ORCID account yet.

Suggest an ORCID for this profile:

Submit Suggestion

HepNames data

Kyle S. Cranmer (New York U.)

Both users and curators can add ORCID to profiles

The publication list retrieved from ORCID is shown in the author profile

Cranmer, Kyle S.

View Profile **Manage Profile** Manage Publications Search Profiles Help

Publications list

Publications in INSPIRE:

1. Search for Quantum Black-Hole Production in High-Invariant-Mass Lepton+Jet Final States Using Proton-Proton Collisions at $\sqrt{s} = 8$ TeV and the ATLAS Detector
2. Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using 4.6 fb⁻¹
3. Higgs Working Group Report of the Snowmass 2013 Community Planning Study
4. Search for long-lived stopped R-hadrons decaying out-of-time with pp collisions using the ATLAS detector
5. Measurement of the mass difference between top and anti-top quarks in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector
6. Search for charmed nearly mass-degenerate with the lightest neutralino based on

Citations (from papers in INSPIRE)

Citations summary

Generated on 2013-11-15

419 papers found, 409 of them citeable (public)

Citation summary results

Total number of papers analyzed:
 Total number of citations:
 Average citations per paper:
 Breakdown of papers by citations:
 Renowned papers (500+)

On the technical level

- Updates on the OAuth library
 - It was using an old version of the standard
 - Moved to Rauth
- ORCID API Library is available for other integrations
 - Supports different levels:
 - Members/non-members

Latest changes in INSPIRE branch:

<http://invenio-software.org/repo/invenio-inspire-ops/log/?h=prod>

webauthorprofile_orcidutils.py

```
def get_dois_from_orcid(orcid_id):
```

```
def _get_access_token_from_orcid
```

```
def _get_public_data_from_orcid_public
```

Next steps...

- Many publishers are starting to require ORCID^s on submission
 - It will start appearing on metadata
- Support on INVENIO without INSPIRE's author pages
- Go get an ORCID @ orcid.org

QUESTIONS ?

Samuele Carli
samuele.carli@cern.ch