

# Author items in Wikidata

Simon Cobb

WikiCite Virtual Conference, 26 October 2020



These slides are available under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License  
<https://creativecommons.org/licenses/by/4.0/>.

# Objectives

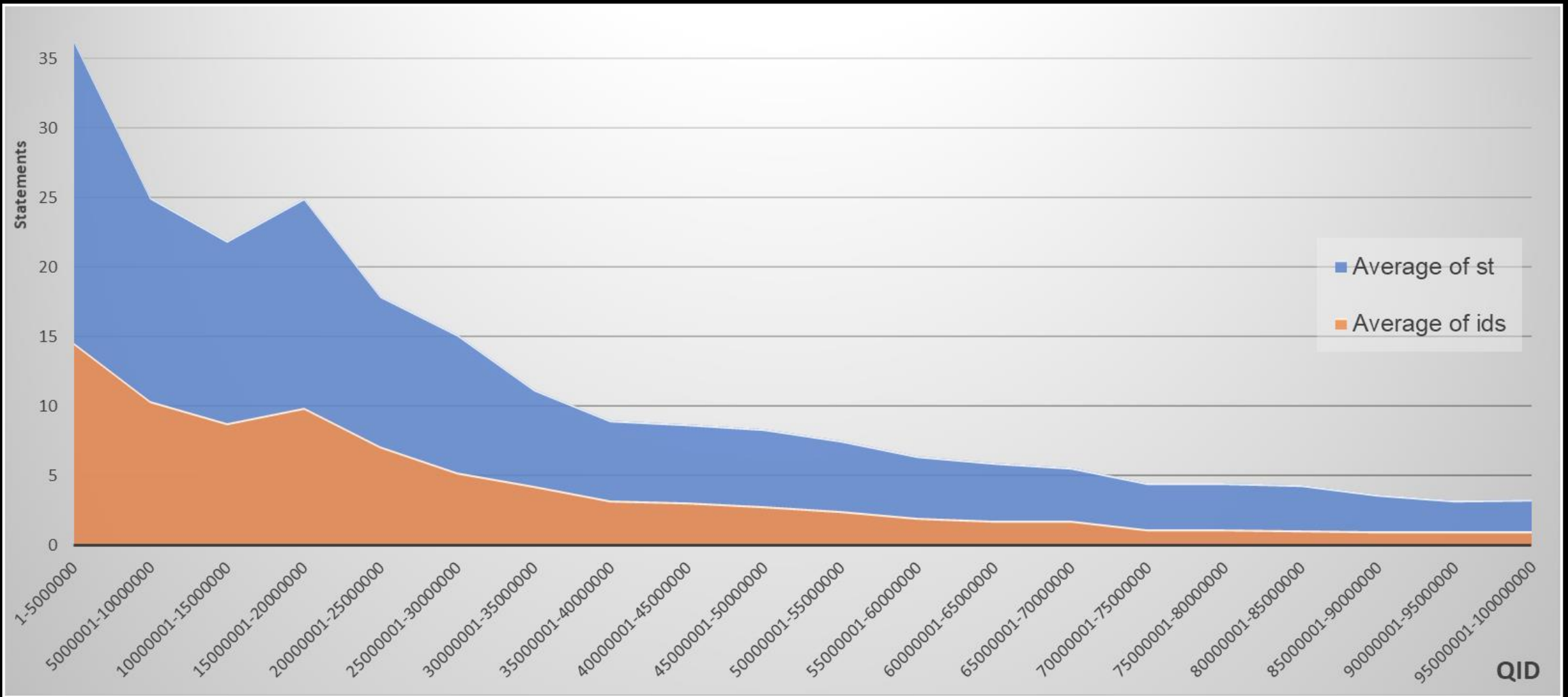
- Highlight issues in existing author items
- Share knowledge about data imports from ORCID
- Foster collaboration to improve and maintain author data

# Scope and overview

- Focus on items with ORCID ID (P496)
  - 1,625,666 entities in Wikidata with ORCID (in 28 Sept. 2020 dump)
  - Analysis of a ten percent sample (162,224 items)
    - Selected using 0xx checksum in ORCID IDs
  - Comparison of Wikidata items and ORCID records

# Property claims

	With claim	With claim %	No claim	No claim %
<b>P21 – sex or gender</b>	24841	15.31%	137383	84.69%
<b>P27 – country of citizenship</b>	2749	1.69%	159475	98.31%
<b>P31 – instance of</b>	162224	100.00%	0	
<b>P69 – educated at</b>	15686	9.67%	146538	90.33%
<b>P101 – field of work</b>	621	0.38%	161603	99.62%
<b>P106 – occupation</b>	129696	79.95%	32528	20.05%
<b>P108 – employer</b>	42248	26.04%	119976	73.96%
<b>P569 – date of birth</b>	3982	2.45%	158242	97.55%
<b>P570 – date of death</b>	40	0.02%	162184	99.98%
<b>P734 – family name</b>	15783	9.73%	146441	90.27%
<b>P735 – given name</b>	43135	26.59%	119089	73.41%
<b>P1412 – languages spoken, written or signed</b>	1006	0.62%	161218	99.38%
<b>P1416 – affiliation</b>	444	0.27%	161780	99.73%



Author items:  
average number of statements and external identifiers

occupation  
(P106)  
=  
researcher  
(Q1650915)  
X  
1.66 million

<https://w.wiki/imW>

The screenshot displays the Wikidata page for the property 'researcher' (Q1650915). The page is organized into several sections, each with a list of values and associated actions:

- occupation**: researcher (0 references). Actions: edit, add reference, add value.
- field of work**: computer data processing (0 references), open knowledge (0 references), library science (0 references). Actions: edit, add reference, add value.
- Commons category**: Simon Cobb (0 references). Actions: edit, add reference, add value.
- Identifiers**: ORCID ID (0000-0002-2874-8657, 0 references). Actions: edit, add reference.

Two red circles are drawn around the 'occupation' property and the 'Identifiers' section.

Item [Discussion](#) [View](#) [History](#) [Recent changes](#)

# French Society of Pediatric Hematology and Immunology (Q79331041)

Revision as of 17:50, 29 August 2020 by Luckyz (talk | contribs) (Created claim: Property:P106: Q1650913, batch #40731) (diff) — Older revision | Latest revision (diff) | Newer revision — (diff)

researcher

[↗ In more languages](#)  
[Configure](#)

Language	Label	Description	Also known as
English	French Society of Pediatric Hematology and Immunology	researcher	
British English	No label defined	No description defined	
Welsh	No label defined	No description defined	
German	No label defined	No description defined	

[All entered languages](#)

## Statements

**instance of** 👤 human + 0 references

**occupation** 👤 researcher + 0 references

## Identifiers

**ORCID ID** 👤 0000-0003-2289-8096 + 0 references

Wikipedia (0 entries)

Wikibooks (0 entries)

Wikinews (0 entries)

Wikiquote (0 entries)

Wikisource (0 entries)

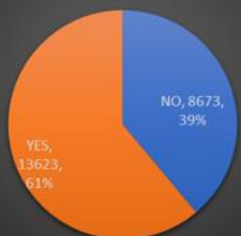
Wikiversity (0 entries)

Wikivoyage (0 entries)

Wiktionary (0 entries)

Other sites (0 entries)

Employments: PID in Wikidata



Employments: Ringgold in Wikidata



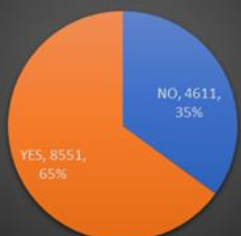
Employments: GRID in Wikidata



Employments: Fundref in Wikidata



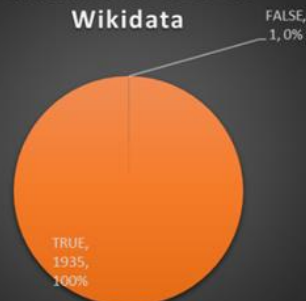
Educations: PID in Wikidata



Educations: Ringgold in Wikidata



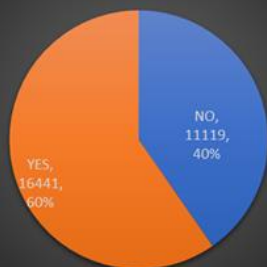
Educations: GRID in Wikidata



Educations: Fundref in Wikidata



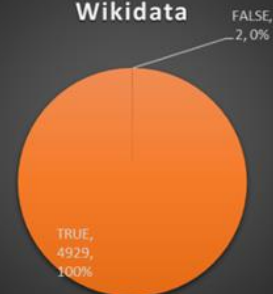
Total: PID in Wikidata



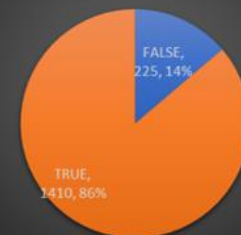
Total: Ringgold in Wikidata



Total: GRID in Wikidata

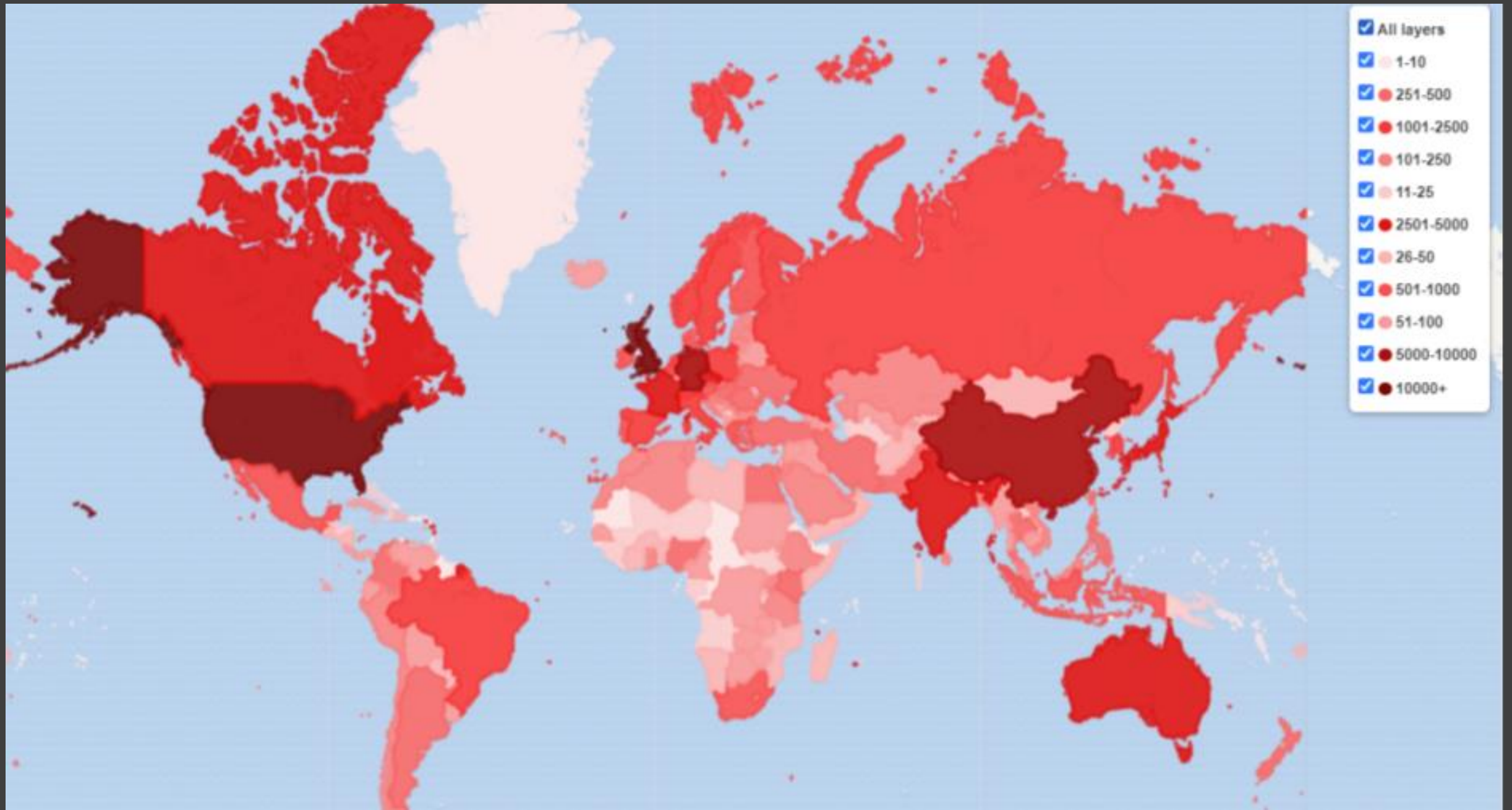


Total: Fundref in Wikidata



Persistent identifiers (PIDs) for organisations in ORCID and Wikidata





# Problems in Wikidata

- Incomplete dataset
- Errors
- Deprecated identifiers

## Magda Matias

### ORCID iD

<https://orcid.org/0000-0003-0875-8011>

 [Print view](#) 

### Also known as

Paraíso Matias

### Country

Portugal

### Other IDs

Scopus Author ID: 57188720540

### Employment (10)

Sort

Universidade de Lisboa Instituto Superior Técnico: Lisboa, Lisboa, PT

2018-09-03 to 2019-02-28 | Bolseira de Investigação no Estudo da Paisagem da Chamusca (CERIS)

Employment

Source: Magda Matias

★ Preferred source

Universidade de Lisboa Instituto Superior Técnico: Lisboa, Lisboa, PT

2017-11-14 to 2018-07-30 | Bolseira de Investigação no projecto AQUIREMOTE (CERIS)

Employment

Source: Magda Matias

★ Preferred source

Universidade de Lisboa Instituto Superior Técnico: Lisboa, Lisboa, PT

2017-05-01 to 2017-10-31 | Bolseira de Investigação no projecto GeoNerC (CERIS)

Employment

Source: Magda Matias

★ Preferred source

Universidade de Lisboa Instituto Superior Técnico: Lisboa, Lisboa, PT

2015-08-23 to 2016-11-30 | Bolseira de investigação no projecto RESULTADOS (CERENA)

Employment

Source: Magda Matias

★ Preferred source

Universidade de Lisboa Instituto Superior Técnico: Lisboa, Lisboa, PT

2013-09-23 to 2015-08-22 | Bolseira de investigação no projecto de I&D AULIS (CERENA)

Employment

Source: Magda Matias

★ Preferred source

Universidade de Lisboa Instituto Superior Técnico: Lisboa, Lisboa, PT

2011-03-15 to 2013-09-22 | Bolseira de investigação no projecto RIVERSAR (ICIST)

Employment

Source: Magda Matias

★ Preferred source

Universidade de Lisboa Instituto Superior Técnico: Lisboa, Lisboa, PT

2009-03-15 to 2011-03-14 | Bolseira de Investigação no projecto MIAVITA (Departamento de Física)

Employment

## Ricardo Cesar Giorgetti Landim

### ORCID iD

<https://orcid.org/0000-0001-6801-3519>

[Print view](#)

### Also known as

Ricardo Landim

### Websites & Social Links

[Mendeley profile](#)

[LinkedIn](#)

### Country

Brazil

### Keywords

Dark energy, Dark matter, Cosmology

### Other IDs

ResearcherID: C-8399-2014

Scopus Author ID: 56429071100

Ciência ID: 6A11-52DF-ACA2

### Email

[ricardo.landim@tum.de](mailto:ricardo.landim@tum.de)

[rlandim@if.usp.br](mailto:rlandim@if.usp.br)

### ▼ Employment (8)

Sort

#### Technical University Munich: Munich, DE

2019-05 to present | Postdoctoral researcher

Employment

Source: Ricardo Cesar Giorgetti Landim

★ Preferred source

#### SLAC National Accelerator Laboratory: CA, CA, US

2018-06 to 2019-01 | Postdoc visitor

Employment

Source: Ricardo Cesar Giorgetti Landim

★ Preferred source

#### Universidade de São Paulo Instituto de Física: Sao Paulo, SP, BR

2017-03-01 to 2018-05-31 | Postdoctoral researcher

Employment

Source: Ricardo Cesar Giorgetti Landim

★ Preferred source

#### Universidade do Porto Faculdade de Ciências: Porto, Porto, PT

2017 to 2018 | Visitor

Employment

Source: Ricardo Cesar Giorgetti Landim

★ Preferred source

#### Universidade de São Paulo: Sao Paulo, SP, BR

2016-07 to 2016-11 | Teaching assistant - Quantum Mechanics (Institute of Physics)

Employment

### Organization identifiers

Ringgold: 67887

Universidade de Santo Amaro: Sao Paulo, SP, BR

Other organization identifiers provided by Ringgold

ISNI: 0000000101066835

### Added

2016-09-19

### Last modified

2016-09-19

## Wendy Garrett

### ORCID iD

<https://orcid.org/0000-0002-5092-0150>

[Print view](#)

### Websites & Social Links

[Garrett Lab Website](#)

### Country

United States

### Keywords

microbiome; microbiota; immunology; cancer; inflammation

### Other IDs

Scopus Author ID: 7102162233

### ▼ Employment (3)

Sort

Harvard T.H. Chan School of Public Health: MA, MA, US

| (Immunology & Infectious Diseases)

Employment

Source: Wendy Garrett

★ Preferred source



Harvard Medical School: Boston, MA, US

Employment

Source: Wendy Garrett

★ Preferred source



Dana Farber Cancer Institute: Boston, MA, US

Employment

Source: Wendy Garrett

★ Preferred source



### ▼ Education and qualifications (2)

Sort

Yale University School of Medicine: New Haven, CT, US

| MD PhD (Cell Biology)

Education

Source: Wendy Garrett

★ Preferred source



Yale College: New Haven, CT, US

| BS MS (Molecular Biophysics & Biochemistry)

Education

Source: Wendy Garrett

★ Preferred source



# Suggested next steps

- Seek community consensus on minimum acceptable standard for author items.
- Define author data requirements for a variety of use cases.
- Review and update the FundRef identifiers in Wikidata to ensure complete coverage.
- Identify and remove incorrect Ringgold IDs from Wikidata.
- Continue importing Ringgold identifiers to existing Wikidata items.
- Review and validate data in existing author items.
- Evaluate options for storing put codes from ORCID in Wikidata reference statements.
- Organise an author items online workshop to facilitate discussion and collaboration.
- Establish a author items WikiProject and/or special interest group (SIG).

Questions or comments?