



HEP STATISTICS IN LATIN AMERICA

MANUEL MORALES ALVARADO



UNIVERSITY OF
CAMBRIDGE



European Research Council
Established by the European Commission

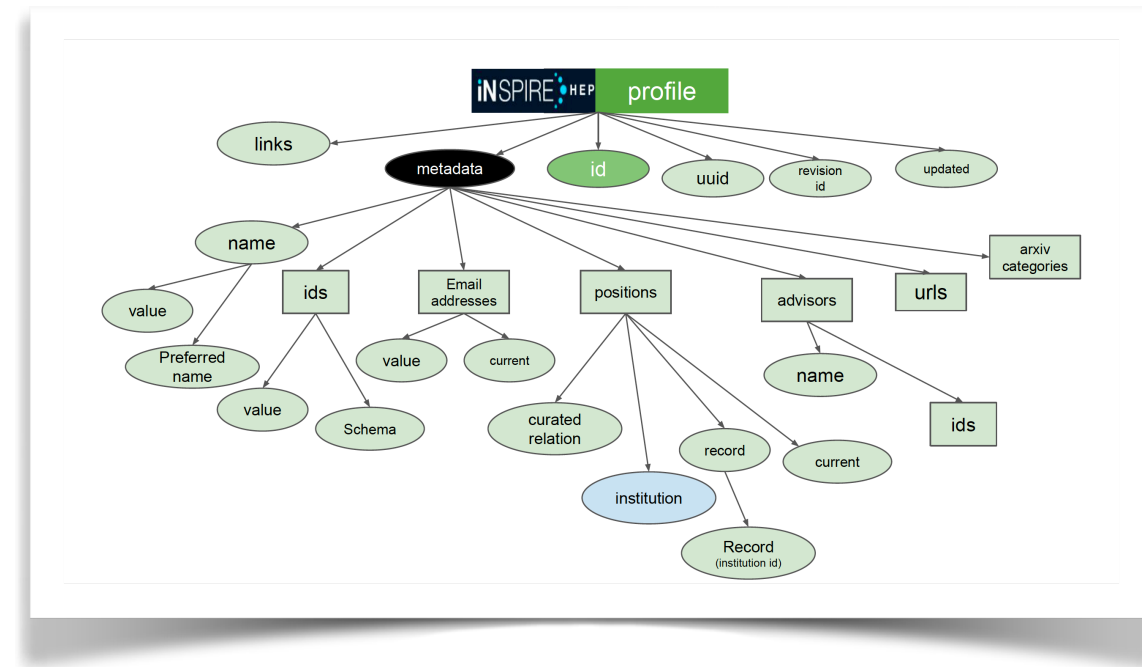


HEP STATISTICS IN LATIN AMERICA

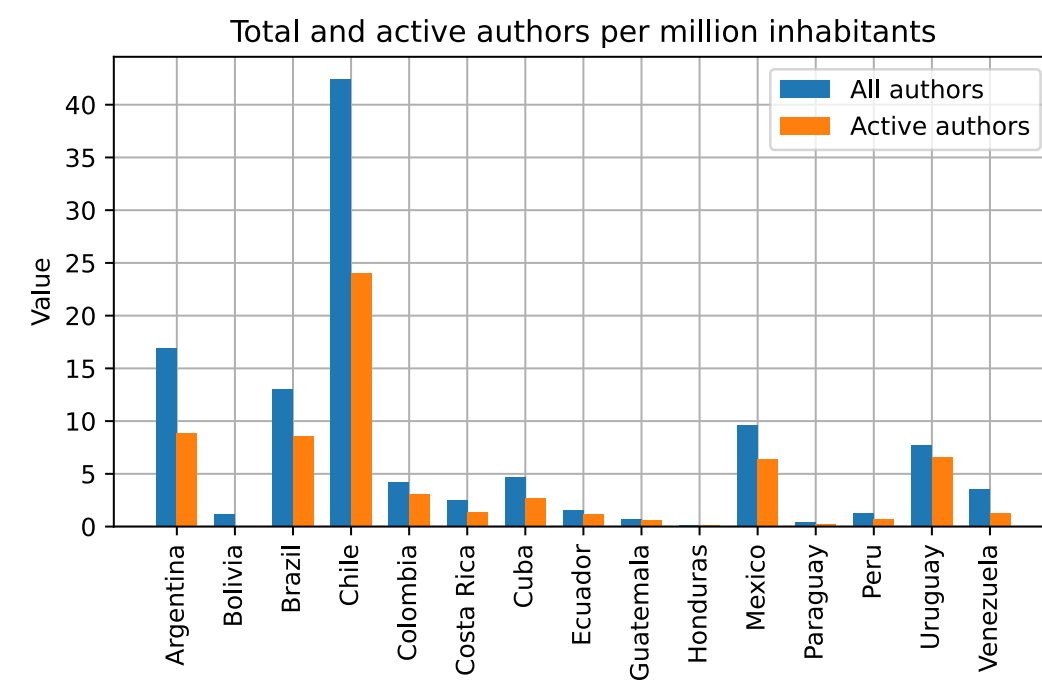
MANUEL MORALES ALVARADO

Thanks to D. Restrepo (U. de Antioquia)
for some of the slides!

OUTLINE



Organisation, INSPIRE-HEP API, analysis code



Results

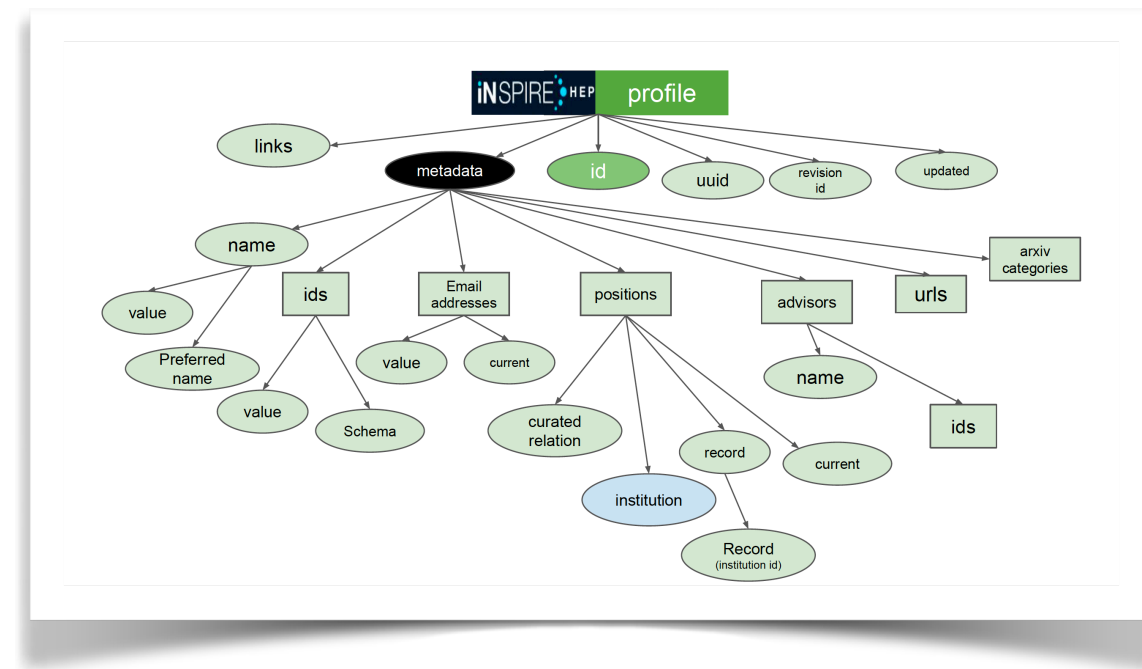


Conclusions and outlook

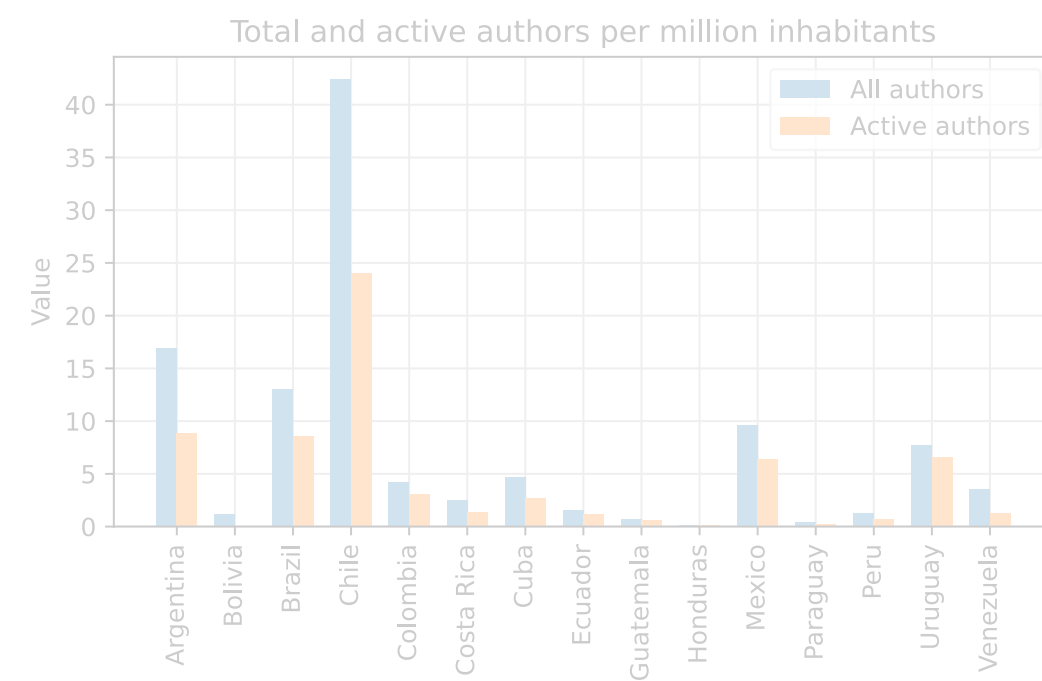
Generate a public database, analysis, and report to assess the status of HEP research and researchers in Latin America



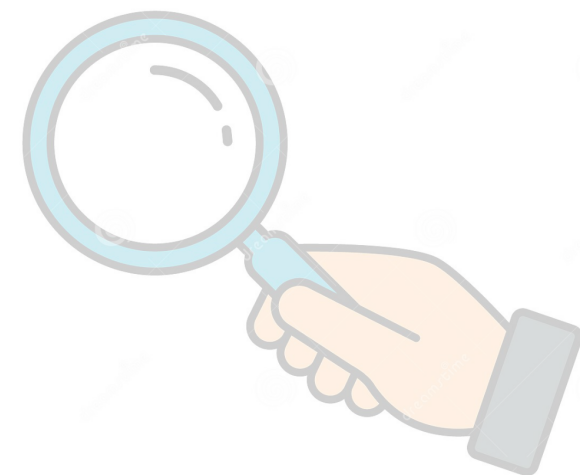
OUTLINE



Organisation, INSPIRE-HEP API, analysis code

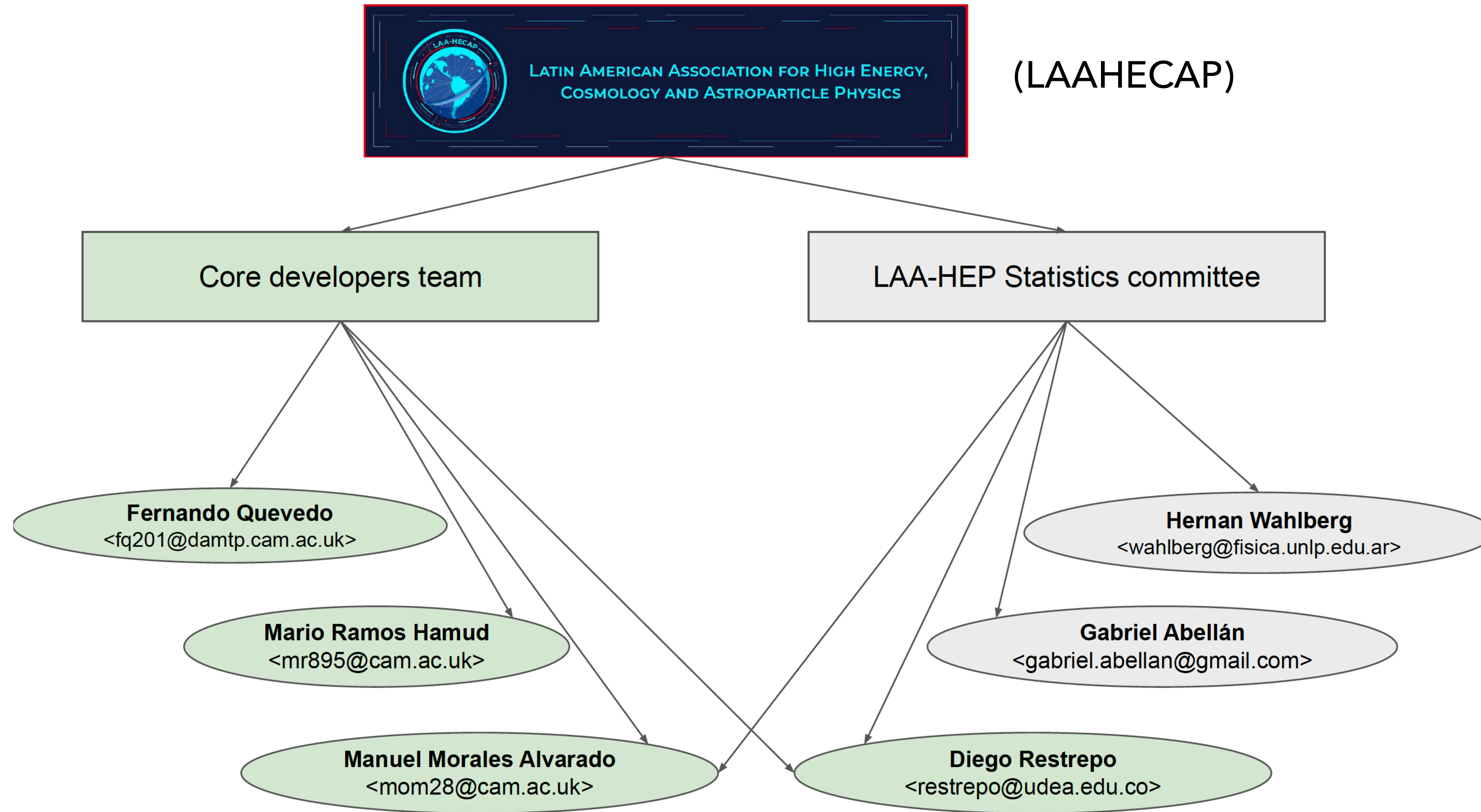


Results



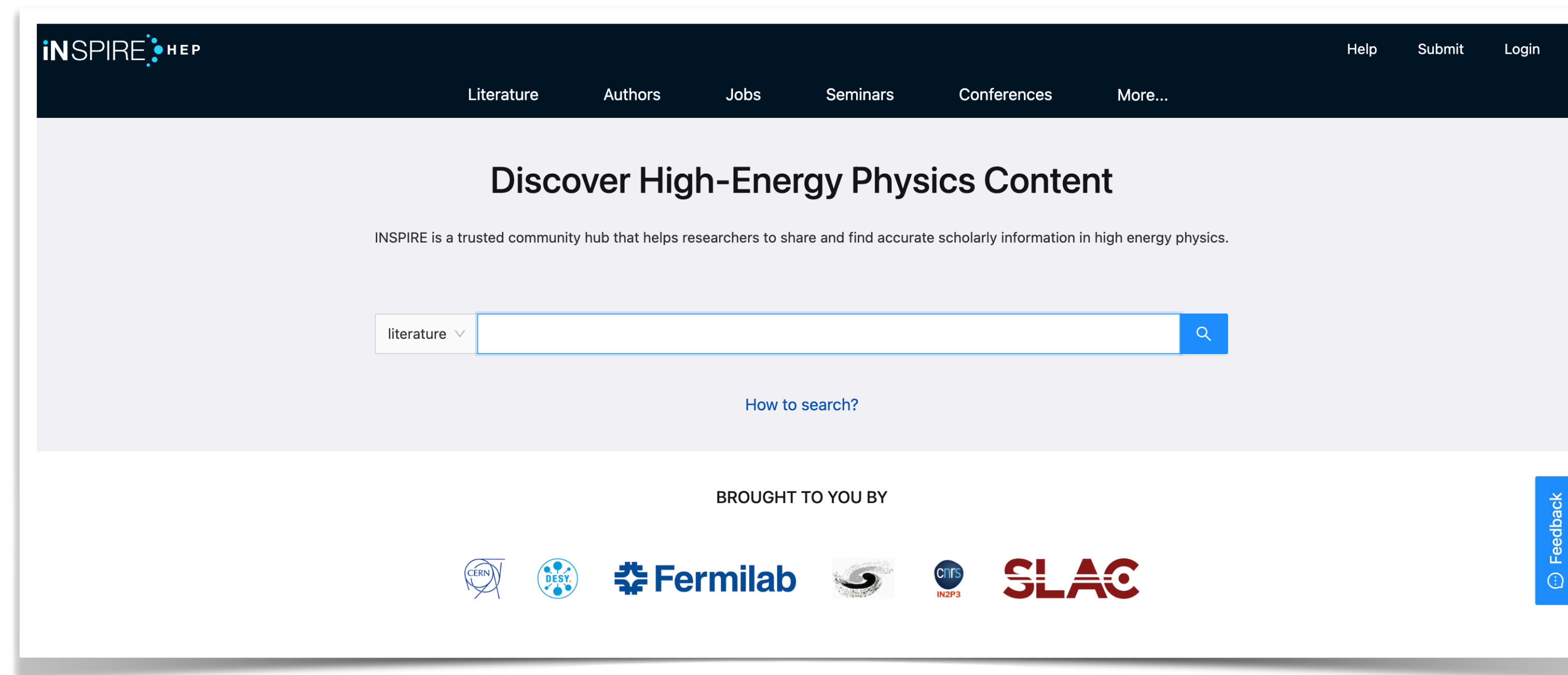
Conclusions and outlook

ORGANISATION



INSPIRE-HEP

- We use the <https://inspirehep.net/> data for this analysis
- INSPIRE-HEP is an open access database
- First design conceived in 1967: SPIRES (Stanford Physics Information Retrieval System)
- SPIRES-HEP: the first website in America, and the first WWW database



DATA STRUCTURES

Lists = []

```
[2]: l=['cat', 'dog', 'cow']  
l[1]
```

```
[2]: 'dog'
```

Dictionaries = {key, value}

```
[3]: d={'animal':'cat', 'fruit':'mango'}  
d['fruit']
```

```
[3]: 'mango'
```

They can be nested!

```
[4]: ds={'animals':l, 'fruits':['mango', 'banana']}  
ds['animals'][1]
```

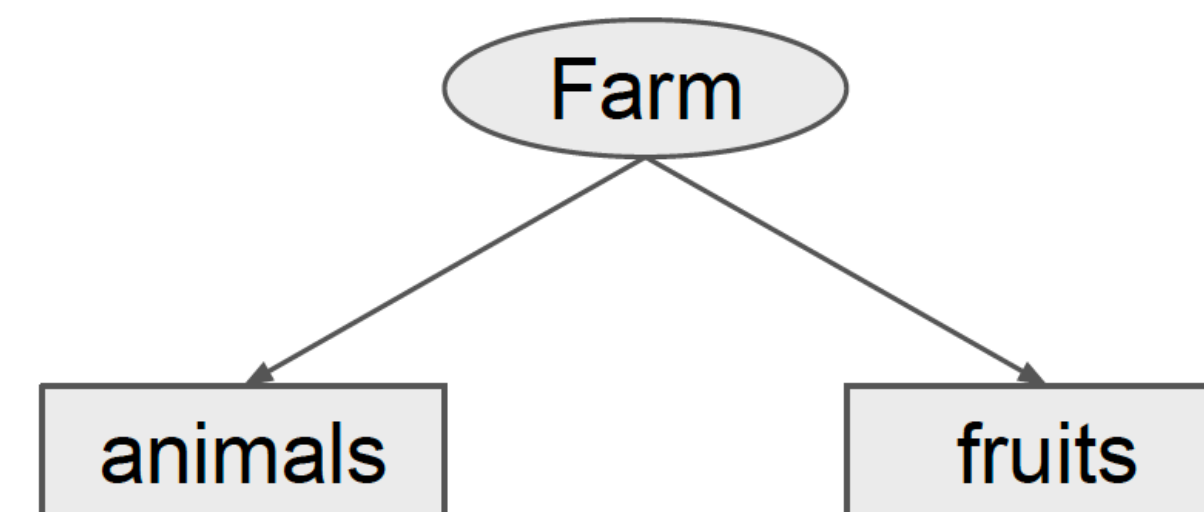
```
[4]: 'dog'
```

```
[5]: from IPython import display
```

```
[6]: display.JSON(ds)
```

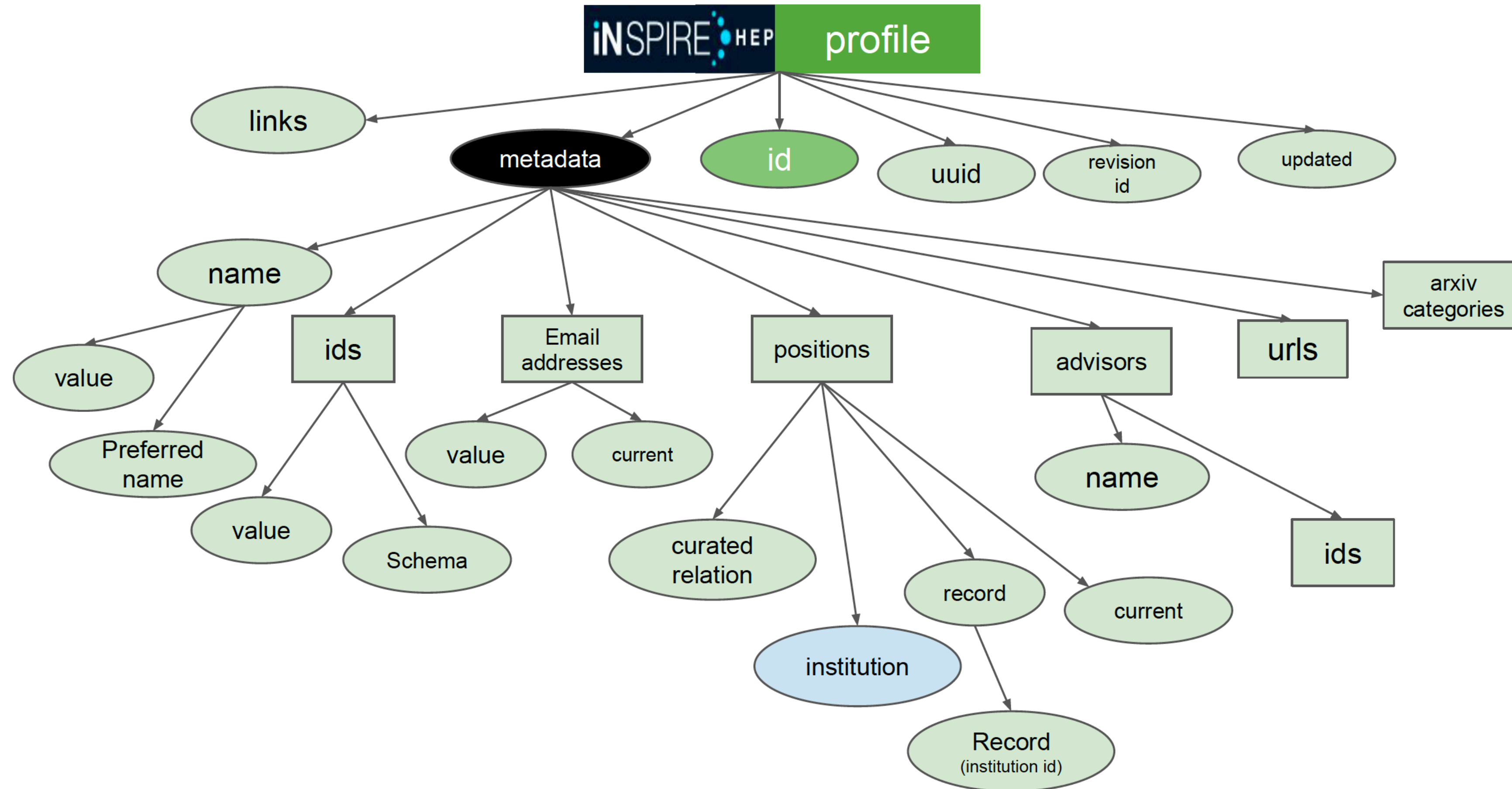
```
[6]: ▼ root:  
    ▼ animals: [] 3 items  
        0: "cat"  
        1: "dog"  
        2: "cow"  
    ► fruits: [] 2 items
```

```
[ ]:
```



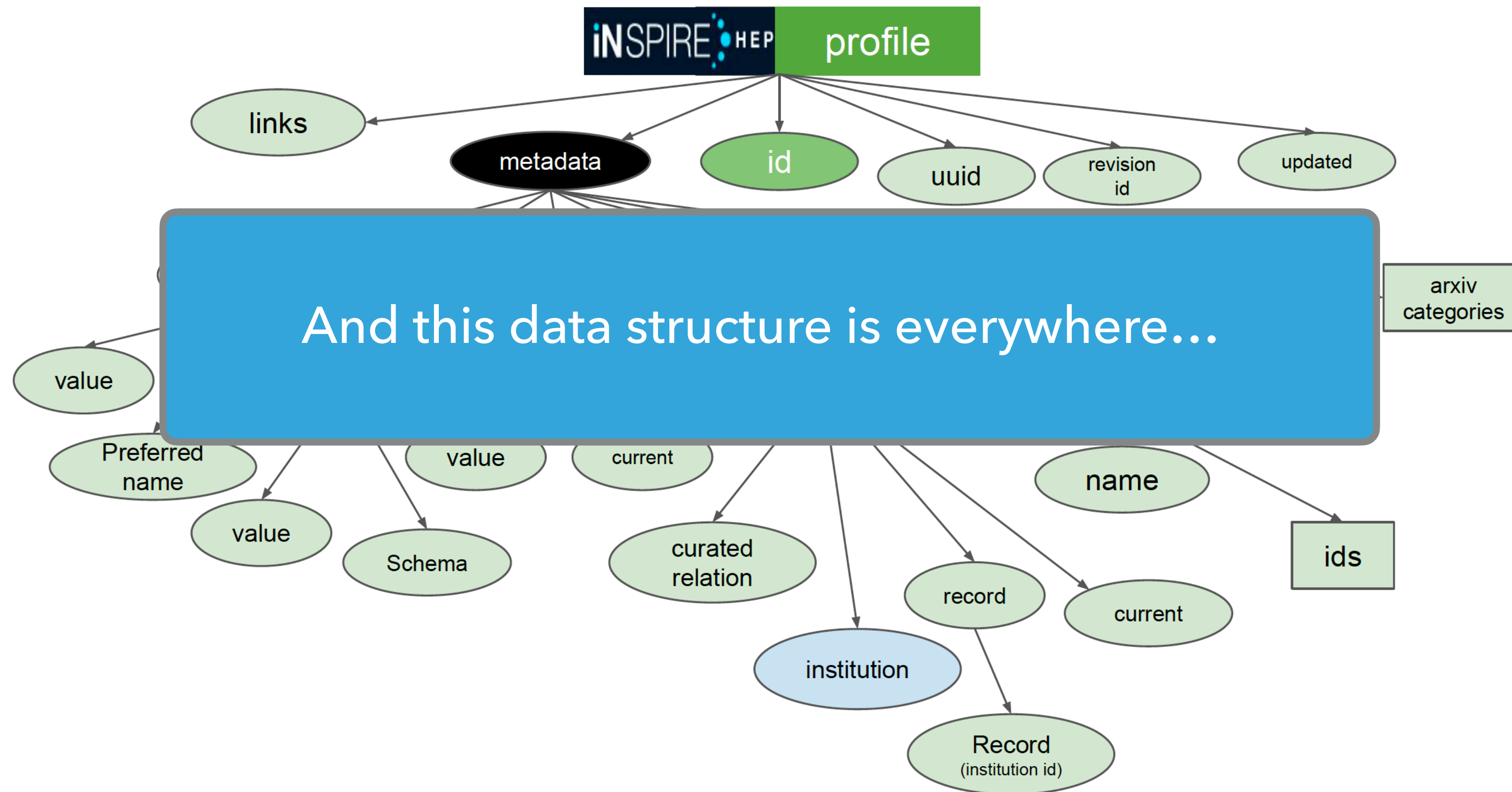
INSPIRE-HEP APPLICATION PROGRAMMING INTERFACE (API)

A profile on INSPIRE-HEP gives us access to different information



INSPIRE-HEP API

A profile on INSPIRE-HEP gives us access to different information



INSPIRE-HEP API FOR THE PHYSICS COMMUNITY

This is us



INSPIRE-HEP API FOR THE PHYSICS COMMUNITY

This is us



With this we have
generated ...

DATABASE AND CODE

Completely open source!

<https://github.com/restrepo/inspireauth/tree/main>

The screenshot shows the GitHub interface for the repository 'restrepo/inspireauth'. The repository is public and has 1 watch and 0 forks. The current branch is 'exploration', which is 35 commits ahead and 38 commits behind the 'main' branch. The file explorer shows the following files and folders:

File/Folder	Description	Commit Date
__pycache__	Added routine to calculate the total citations for a given country	4 months ago
articles_countries_per_year	Added emails on excel and routines to obtain them	2 weeks ago
combined_plots	Added functions to parse the countries as pd.Series and combined hi...	5 months ago
data	Added database for experimental collaborations	2 months ago
email_lists	Added emails on excel and routines to obtain them	2 weeks ago
plots_for_tables	Added emails on excel and routines to obtain them	2 weeks ago
.gitignore	Updated .gitignore	5 months ago
INSPIREHEP.ipynb	Added statistic per inhabitants	4 months ago

The repository statistics on the right side of the page are:

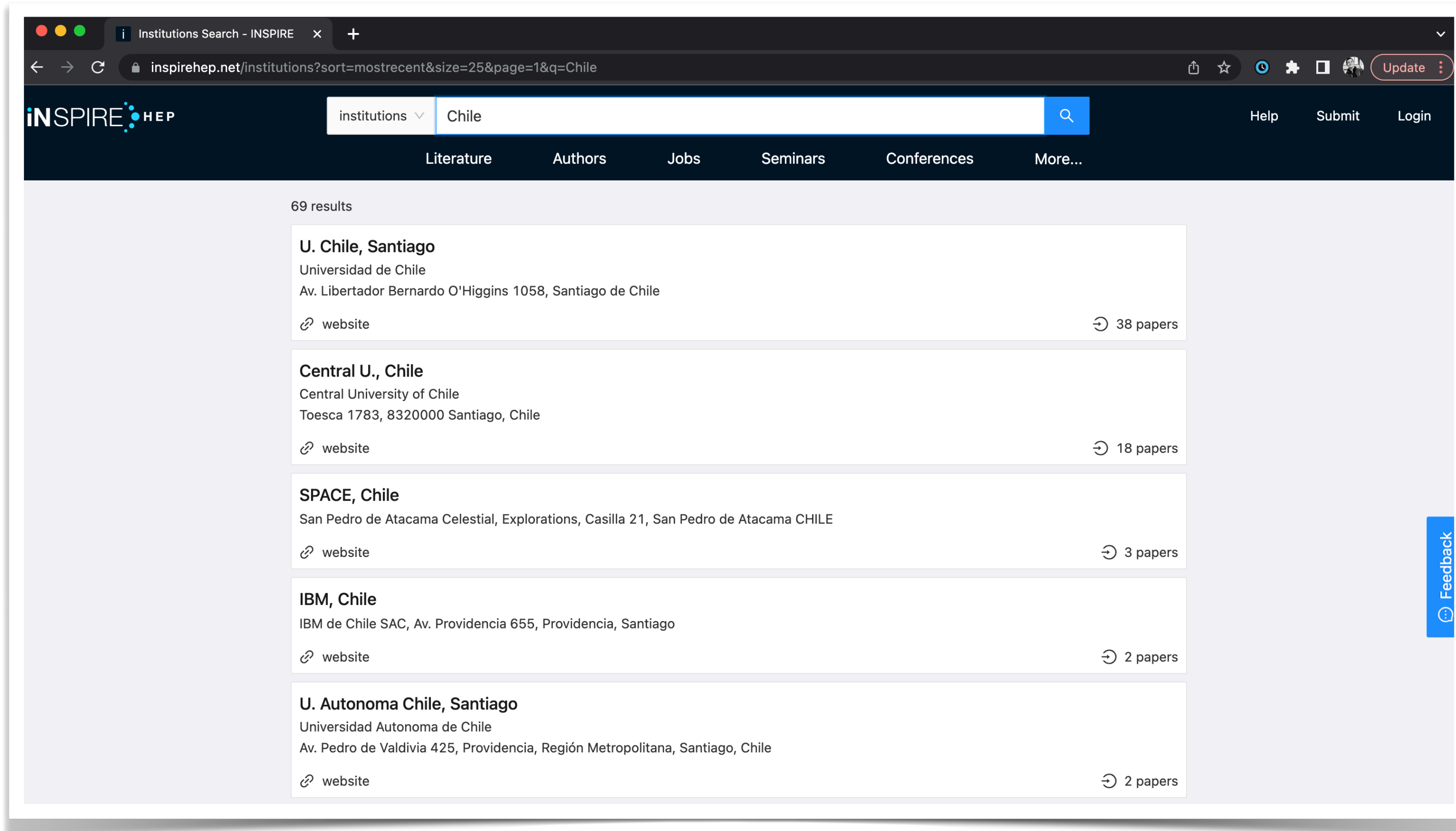
- About: Get authors info from INSPIRE-HEP
- Readme
- 0 stars
- 1 watching
- 0 forks
- Releases: No releases published. [Create a new release](#)
- Packages: No packages published. [Publish your first package](#)

A short overview of the logic of the database...

A SHORT EXERCISE

“Extract the list of all the authors from Chile from INSPIRE-HEP”

1 - Access the institutions from Chile



The screenshot shows the INSPIRE-HEP website's search results for institutions in Chile. The browser address bar shows the URL: `inspirehep.net/institutions?sort=mostrecent&size=25&page=1&q=Chile`. The search bar contains the text "institutions" and "Chile". The page displays 69 results, with the following institutions listed:

Institution Name	Address	Papers
U. Chile, Santiago Universidad de Chile Av. Libertador Bernardo O'Higgins 1058, Santiago de Chile	website	38 papers
Central U., Chile Central University of Chile Toesca 1783, 8320000 Santiago, Chile	website	18 papers
SPACE, Chile San Pedro de Atacama Celestial, Explorations, Casilla 21, San Pedro de Atacama CHILE	website	3 papers
IBM, Chile IBM de Chile SAC, Av. Providencia 655, Providencia, Santiago	website	2 papers
U. Autonoma Chile, Santiago Universidad Autonoma de Chile Av. Pedro de Valdivia 425, Providencia, Región Metropolitana, Santiago, Chile	website	2 papers

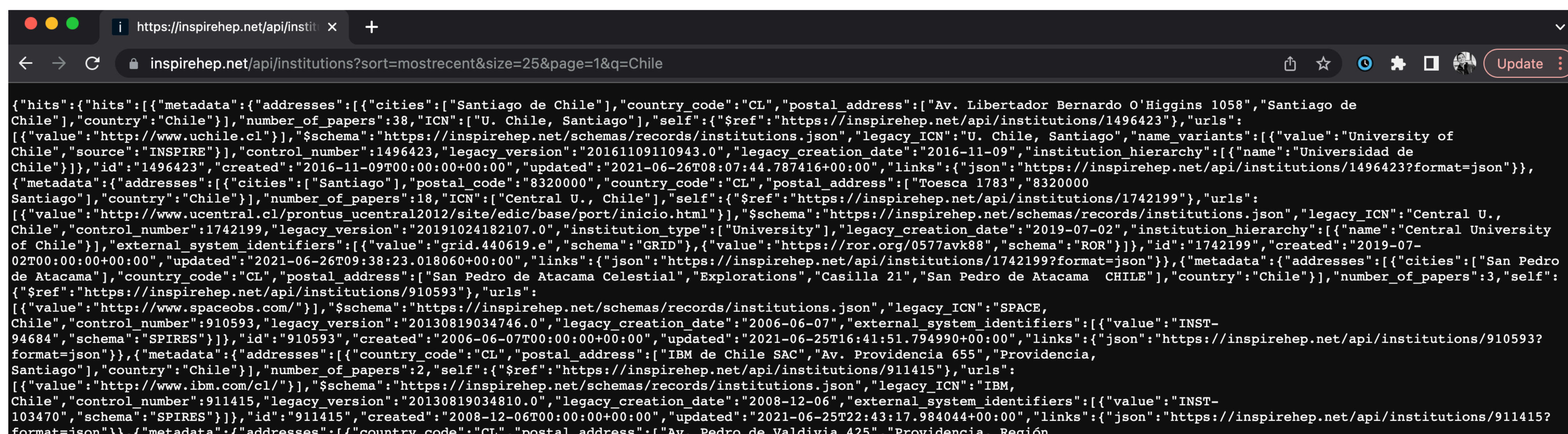
A vertical "Feedback" button is visible on the right side of the page.

A SHORT EXERCISE

2 - Access the API

```
https://inspirehep.net/institutions?sort=mostrecent&size=25&page=1&q=Chile
```

```
https://inspirehep.net/api/institutions?sort=mostrecent&size=25&page=1&q=Chile
```



The screenshot shows a web browser window with the URL `https://inspirehep.net/api/institutions?sort=mostrecent&size=25&page=1&q=Chile`. The page content is a JSON response from the API, listing several institutions in Chile. The JSON structure includes fields for `hits`, `metadata`, `addresses`, `country_code`, `postal_address`, `number_of_papers`, `ICN`, `self`, `urls`, `value`, `schema`, `legacy_ICN`, `name_variants`, `source`, `control_number`, `legacy_version`, `legacy_creation_date`, `institution_hierarchy`, `id`, `created`, `updated`, `links`, and `external_system_identifiers`. The first institution listed is the University of Chile (ICN: U. Chile, Santiago) with 38 papers. The second is Central U. Chile (ICN: Central U., Chile) with 18 papers. The third is San Pedro de Atacama (ICN: SPACE) with 3 papers. The fourth is IBM de Chile (ICN: IBM) with 2 papers. The fifth is Pedro de Valdivia (ICN: IBM) with 1 paper.

A SHORT EXERCISE

3 - Extract the list of institutions of Chile

```
In [35]: 1 list_legacyes_ICN
```

```
Out[35]: ['U. Chile, Santiago',  
'Central U., Chile',  
'IBM, Chile',  
'SPACE, Chile',  
'European Southern Obs., Chile',  
'U. Autonoma Chile, Santiago',  
'San Sebastian U., Chile',  
'Santiago de Chile U.',  
'Unlisted, CL',  
'Pontificia U. Catol. Chile, Santiago',  
'Chile U., CMM',  
'Chile U., Beauchef',  
'Chile U., Santiago',  
'Chile Austral U., Valdivia',  
'Chile U., Catolica',  
'U. Arturo Prat, Iquique',  
'INI, Santiago',  
'CIDCA, Santiago',  
'Valparaiso U.',  
'Cerro Calan Observ.',  
'Concepcion Catolica U.',  
'Catolica del Norte U.',  
'Valparaiso U., Catolica',  
'CECS, Valdivia',  
'Temuco U., Catolica',
```


A SHORT EXERCISE

4 - Access the papers published by an institution. Then, iterate over institutions

<https://inspirehep.net/api/literature?sort=mostrecent&page=1&q=aff+U.+Chile,+Santiago+and+ac+1->+10>

aff: use `legacy_ICN`

and: logical operator

ac: defines number
of authors

```

{"hits":{"hits":[{"id":"2136681","metadata":{"citation_count":3,"citation_count_without_self_citations":3,"publication_info":[{"journal_volume":"666","artid":"A36","year":2022,"journal_record":{"$ref":"https://inspirehep.net/api/journals/1214902"},"page_start":"A36","journal_title":"Astron.Astrophys."},"journal_volume":"666","pubinfo_freetext":"A&A 666, A36 (2022)","artid":"A36","material":"publication","year":2022,"page_start":"A36","journal_title":"Astron.Astrophys."}],authors":[{"raw_affiliations":[{"value":"Finnish Centre for Astronomy with ESO, University of Turku, Vesilinnantie 5, 20014 Turku, Finland"},"full_name_unicode_normalized":"liodakis, i.", "full_name":"Liodakis, I.", "record":{"$ref":"https://inspirehep.net/api/authors/1891882"},"ids":{"schema":"INSPIRE BAI","value":"I.Liodakis.2"},"last_name":"Liodakis","affiliations":[{"record":{"$ref":"https://inspirehep.net/api/institutions/903664"},"value":"Turku U."},"bai":"I.Liodakis.2","signature_block":"LADACi","first_name":"I.", "uuid":"62bfa9ea-b060-488e-8549-d9fedf8235c1","recid":1891882},"raw_affiliations":[{"value":"Finnish Centre for Astronomy with ESO, University of Turku, Vesilinnantie 5, 20014 Turku, Finland"},"value":"Aalto University Metsähovi Radio Observatory, Metsähovintie 114, 02540 Kylmälä, Finland"},"full_name_unicode_normalized":"hovatta, t.", "full_name":"Hovatta, T.", "record":{"$ref":"https://inspirehep.net/api/authors/1029923"},"ids":{"schema":"INSPIRE BAI","value":"T.Hovatta.1"},"last_name":"Hovatta","affiliations":[{"record":{"$ref":"https://inspirehep.net/api/institutions/903664"},"value":"Turku U."},"record":{"$ref":"https://inspirehep.net/api/institutions/911903"},"value":"Aalto U."},"bai":"T.Hovatta.1","signature_block":"HAVATt","first_name":"T.", "uuid":"3f3fa50b-b8f7-423b-956c-578407d2a379","recid":1029923},"raw_affiliations":[{"value":"Department of Physics and Institute of Theoretical and Computational Physics, University of Crete, 71003 Heraklion, Greece"},"value":"Department of Physics, Univ. of Crete, 70013 Heraklion, Greece"},"full_name_unicode_normalized":"pavlidou, v.", "full_name":"Pavlidou, V.", "record":{"$ref":"https://inspirehep.net/api/authors/1019414"},"ids":{"schema":"INSPIRE BAI","value":"V.Pavlidou.1"},"last_name":"Pavlidou","affiliations":[{"record":{"$ref":"https://inspirehep.net/api/institutions/912413"},"value":"IESL, Heraklion"},"record":{"$ref":"https://inspirehep.net/api/institutions/911754"},"value":"IFW, Dresden"},"record":{"$ref":"https://inspirehep.net/api/institutions/903969"},"value":"Crete U."},"bai":"V.Pavlidou.1","signature_block":"PAVLADv","first_name":"V.", "uuid":"ec5c7b46-1ee7-4ace-b704-d172715627ec","recid":1019414},"raw_affiliations":[{"value":"Department of Physics and Institute of Theoretical and Computational Physics, University of Crete, 71003 Heraklion, Greece"},"value":"Department of Physics, National and Kapodistrian University of Athens, University Campus Zografos, 15783 Athens, Greece"},"full_name_unicode_normalized":"readhead, a.c.s.", "full_name":"Readhead, A.C.S.", "record":{"$ref":"https://inspirehep.net/api/authors/1027796"},"ids":{"schema":"INSPIRE BAI","value":"A.C.S.Readhead.1"},"last_name":"Readhead","affiliations":[{"record":{"$ref":"https://inspirehep.net/api/institutions/912413"},"value":"IESL, Heraklion"},"record":{"$ref":"https://inspirehep.net/api/institutions/911754"},"value":"IFW, Dresden"},"record":{"$ref":"https://inspirehep.net/api/institutions/903969"},"value":"Crete U."},"record":{"$ref":"https://inspirehep.net/api/institutions/902711"},"value":"Caltech"},"bai":"A.C.S.Readhead.1","signature_block":"RADADa","first_name":"A.C.S.", "uuid":"bcd0eec3-b29b-4645-9a2c-7161ef650c90","recid":1027796},"raw_affiliations":[{"value":"Institute of Astrophysics, Foundation for Research and Technology-Hellas, 71110 Heraklion, Greece"},"full_name_unicode_normalized":"blandford, r.d.", "full_name":"Blandford, R.D.", "record":{"$ref":"https://inspirehep.net/api/authors/1020799"},"ids":{"schema":"INSPIRE BAI","value":"R.D.Blandford.1"},"last_name":"Blandford","affiliations":[{"record":{"$ref":"https://inspirehep.net/api/institutions/909388"},"value":"KIPAC, Menlo Park"},"bai":"R.D.Blandford.1","signature_block":"BLANDFADr","first_name":"R.D.", "uuid":"3a2b27f1-6525-4c41-ad9c-6dce815fbc6","recid":1020799},"raw_affiliations":[{"value":"Department of Physics and Institute of Theoretical and Computational Physics, University of Crete, 71003 Heraklion, Greece"},"value":"Owens Valley Radio Observatory, California Institute of Technology, Pasadena, CA 91125, USA"},"full_name_unicode_normalized":"kiehlmann, s.", "full_name":"Kiehlmann, S.", "record":{"$ref":"https://inspirehep.net/api/authors/1930033"},"ids":{"schema":"INSPIRE BAI","value":"S.Kiehlmann.2"},"last_name":"Kiehlmann","affiliations":[{"record":{"$ref":"https://inspirehep.net/api/institutions/912413"},"value":"IESL, Heraklion"},"record":{"$ref":"https://inspirehep.net/api/institutions/911754"},"value":"IFW, Dresden"},"record":{"$ref":"https://inspirehep.net/api/institutions/903969"},"value":"Crete U."},"bai":"S.Kiehlmann.2","signature_block":"CALNANs","first_name":"S.", "uuid":"990fb285-4588-463b-af51-3ead2b172c4f","recid":1930033},"raw_affiliations":[{"value":"Finnish Centre for Astronomy with ESO, University of Turku, Vesilinnantie 5, 20014 Turku, Finland"},"full_name_unicode_normalized":"lindfors, e.", "full_name":"Lindfors, E.", "record":{"$ref":"https://inspirehep.net/api/authors/1026783"},"ids":{"schema":"INSPIRE BAI","value":"E.Lindfors.1"},"last_name":"Lindfors","affiliations":[{"record":

```

A SHORT EXERCISE

5 - For each article of each institution, extract the profile URL of the authors

In each entry:



```

Out[152]: [['https://inspirehep.net/api/authors/1973718',
'https://inspirehep.net/api/authors/1403047',
'https://inspirehep.net/api/authors/1477164',
'https://inspirehep.net/api/authors/2013953',
'https://inspirehep.net/api/authors/1413660',
'https://inspirehep.net/api/authors/999090'],
['https://inspirehep.net/api/authors/1993758',
'https://inspirehep.net/api/authors/1993759',
'https://inspirehep.net/api/authors/995374',
'https://inspirehep.net/api/authors/1037240'],
['https://inspirehep.net/api/authors/1017586',
'https://inspirehep.net/api/authors/1875153',
'https://inspirehep.net/api/authors/996257',
'https://inspirehep.net/api/authors/993295',
'https://inspirehep.net/api/authors/1287459',
'https://inspirehep.net/api/authors/985110'],
['https://inspirehep.net/api/authors/1037240',
'https://inspirehep.net/api/authors/1987711',
'https://inspirehep.net/api/authors/995374']]
  
```

We have obtained all the authors from Chile



SOME COMMENTS

- In this way we have obtained the list of all papers of a given country with up to 10 authors
- This means that the papers of experimental collaborations are left out of this analysis (we have another database for them!)
- The database was obtained in June 2022, and has to be updated after some time
- The actual nationality of the author is irrelevant. It counts as an author of a given country by being affiliated to an institution of that country

Now some results ...

COUNTRIES IN THE DATABASE

We include the following countries in our study:

Country	Population
Argentina	45,606,000
Bolivia	11,833,000
Brazil	213,993,000
Chile	19,212,000
Colombia	51,266,000
Costa Rica	5,139,000
Cuba	11,318,000
Ecuador	17,888,000
Guatemala	18,250,000
Honduras	10,063,000
Mexico	130,262,000
Paraguay	7,220,000
Peru	33,359,000
Uruguay	3,485,000
Venezuela	28,705,000

The population of the countries is useful to find “metrics per capita”.

They have been obtained from <https://data.un.org/en/index.html>
(checked in September 2022)

Why are some countries missing?

They simply do not have indexed entries on INSPIRE-HEP

METRICS FOR EACH COUNTRY

After generating the database we have obtained:

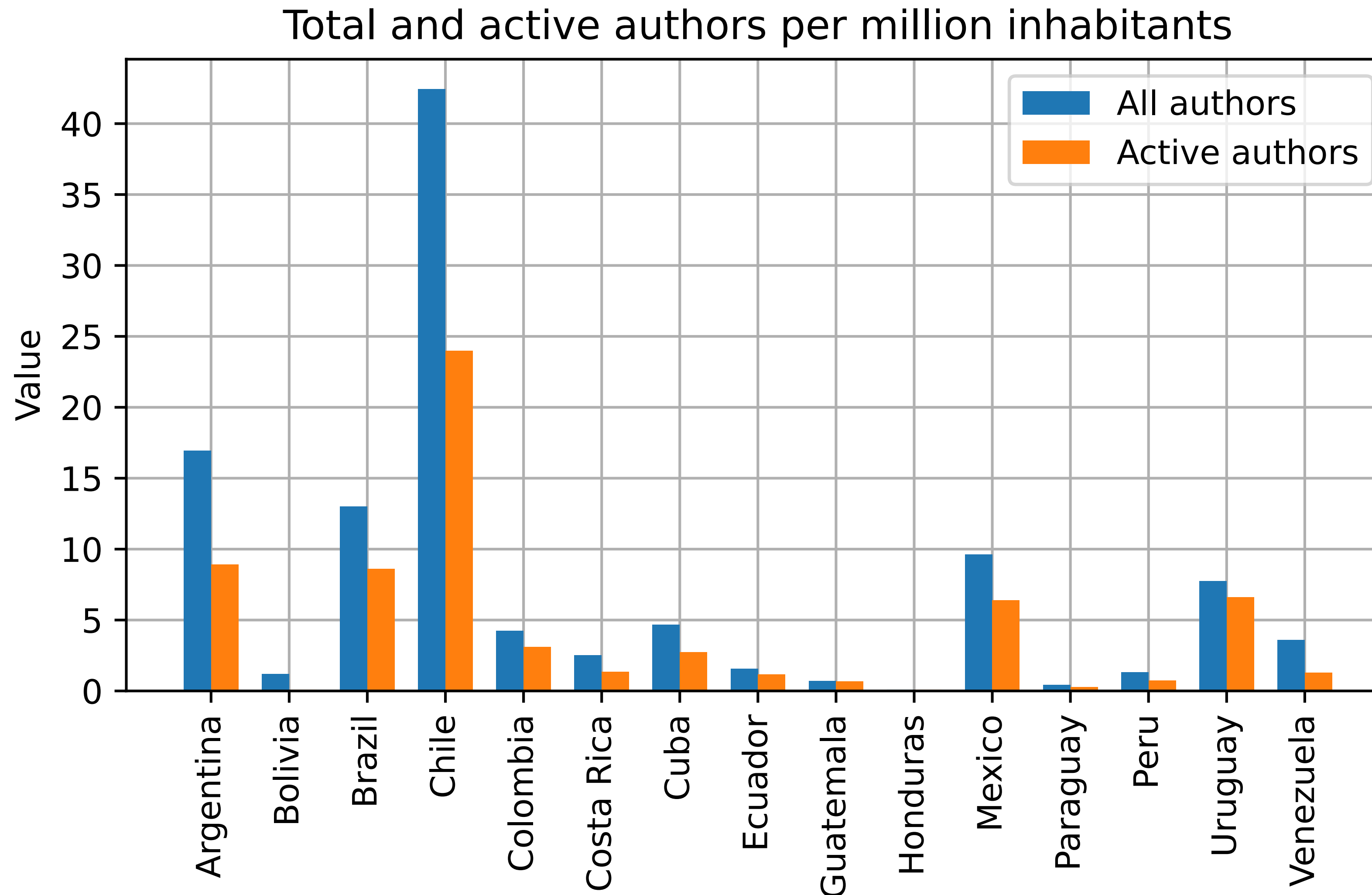
- Name and email of all active authors (with publications in the last 5 years)
- Number of total and active authors
- Number of publications
- Number of citations
- H-index
- Collaborations with other countries in Latin America
- And more...

TOTAL AND ACTIVE AUTHORS

(p.m.i. = per million inhabitants)

Country	Total authors	Active authors	Total authors p.m.i.	Active authors p.m.i.
Argentina	773	406	17	9
Bolivia	14	1	1	0
Brazil	2782	1843	13	9
Chile	815	461	42	24
Colombia	218	159	4	3
Costa Rica	13	7	3	1
Cuba	53	31	5	3
Ecuador	28	21	2	1
Guatemala	13	12	1	1
Honduras	1	1	0	0
Mexico	1253	832	10	6
Paraguay	3	2	0	0
Peru	44	24	1	1
Uruguay	27	23	8	6
Venezuela	103	37	1	1

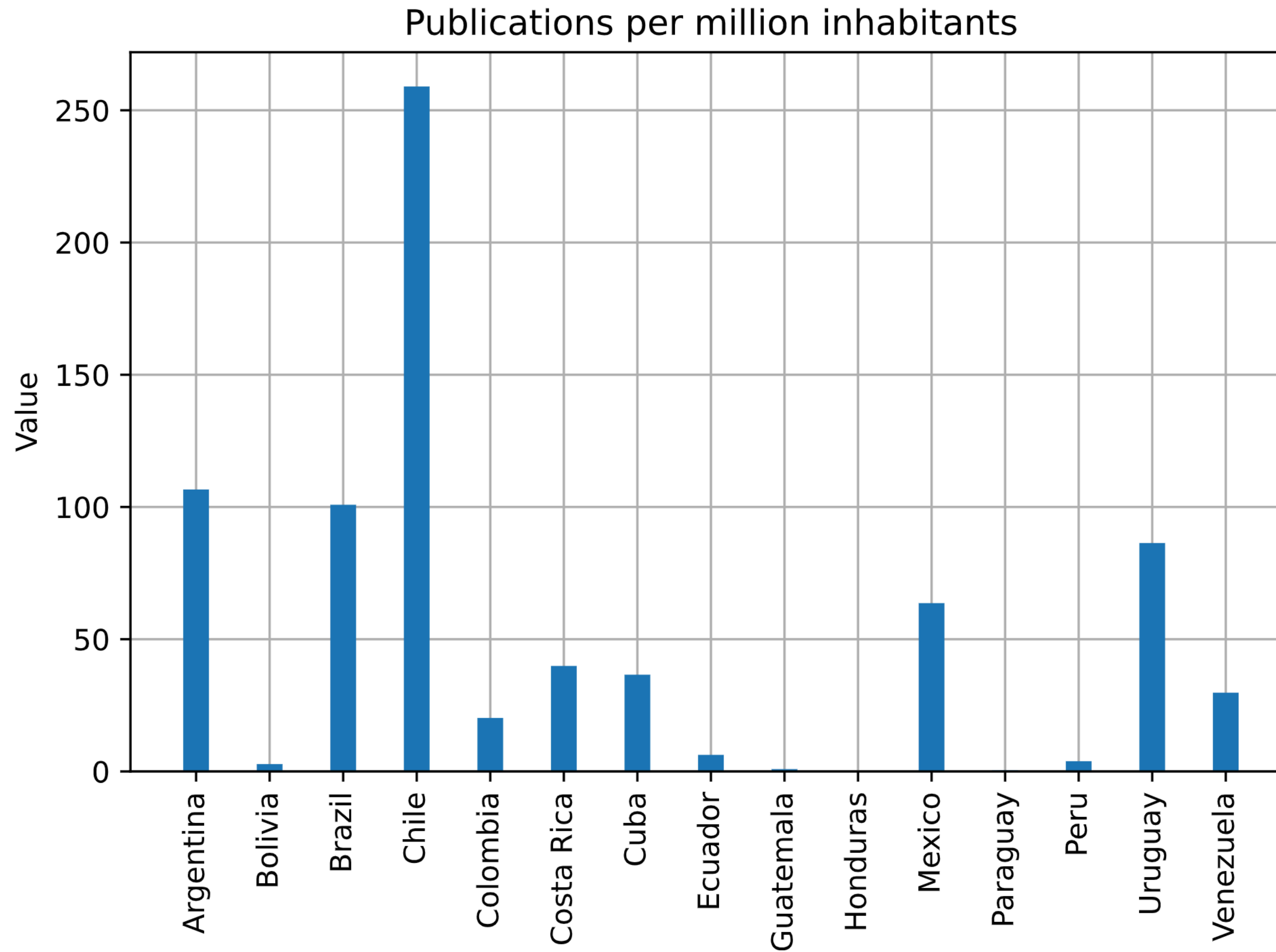
AUTHORS NORMALISED BY POPULATION



PUBLICATIONS

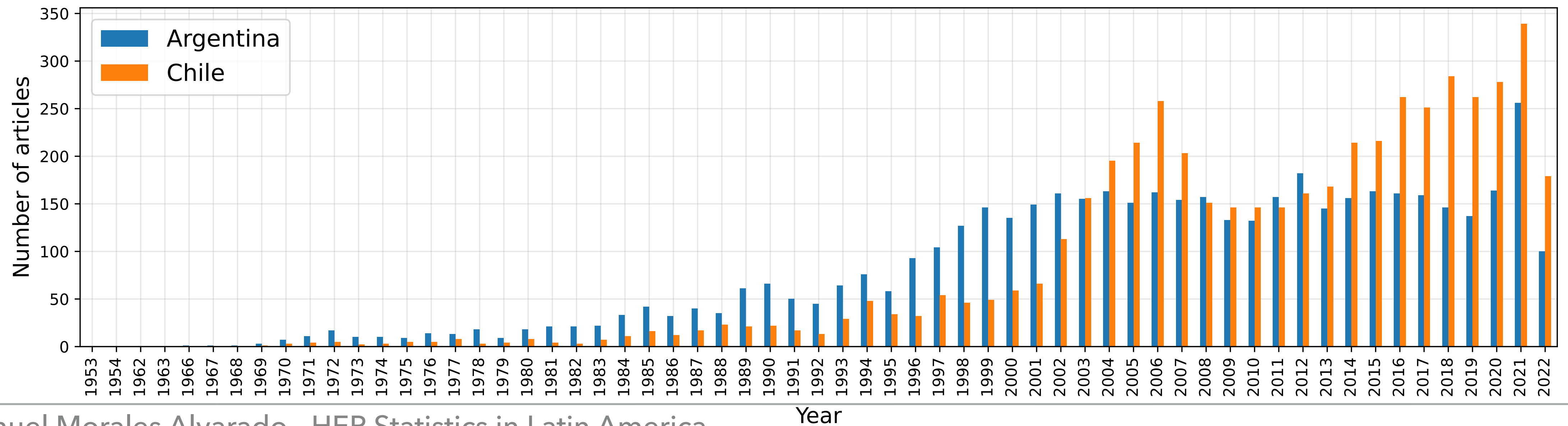
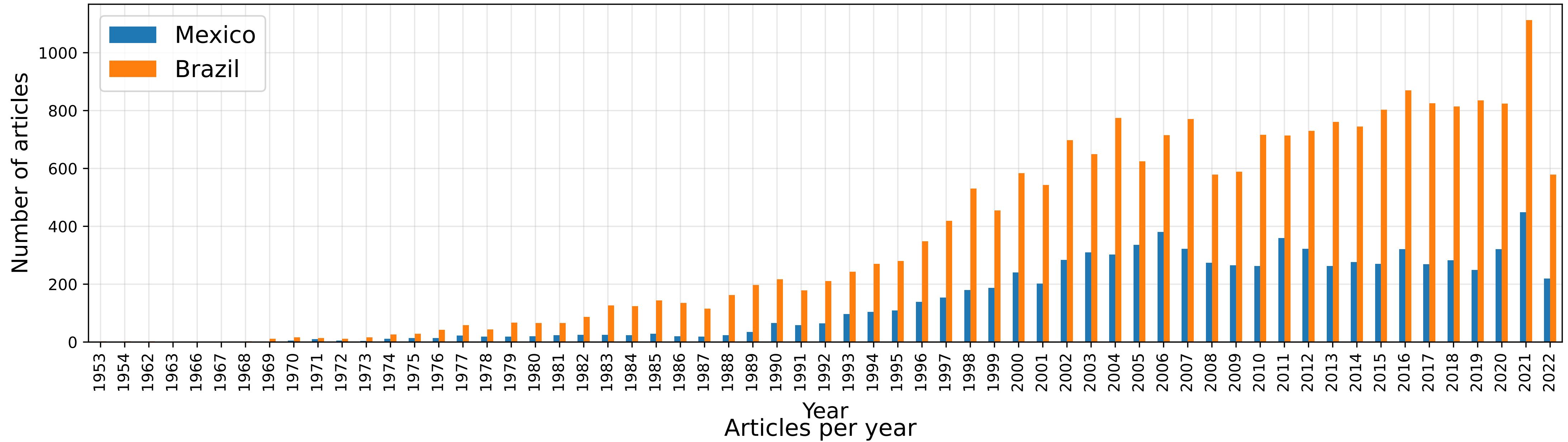
Country	Total publications	Total publications per million inhabitants
Argentina	4862	107
Bolivia	33	3
Brazil	21578	101
Chile	4976	259
Colombia	1037	20
Costa Rica	205	40
Cuba	414	37
Ecuador	112	6
Guatemala	16	1
Honduras	1	0
Mexico	8289	64
Paraguay	3	0
Peru	129	4
Uruguay	301	86
Venezuela	855	30

NORMALISED PUBLICATIONS



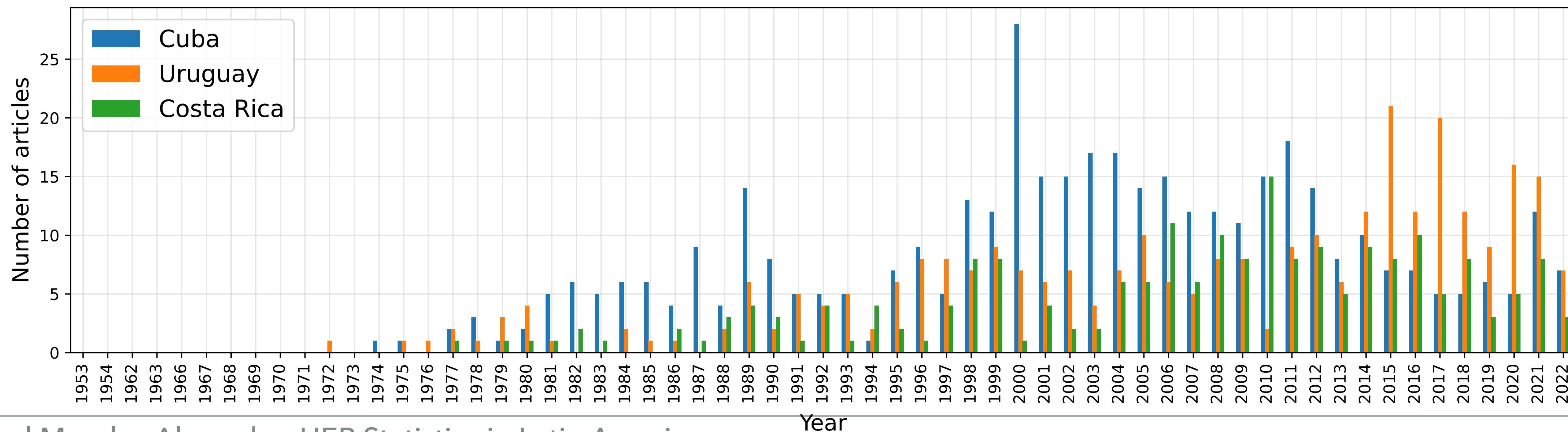
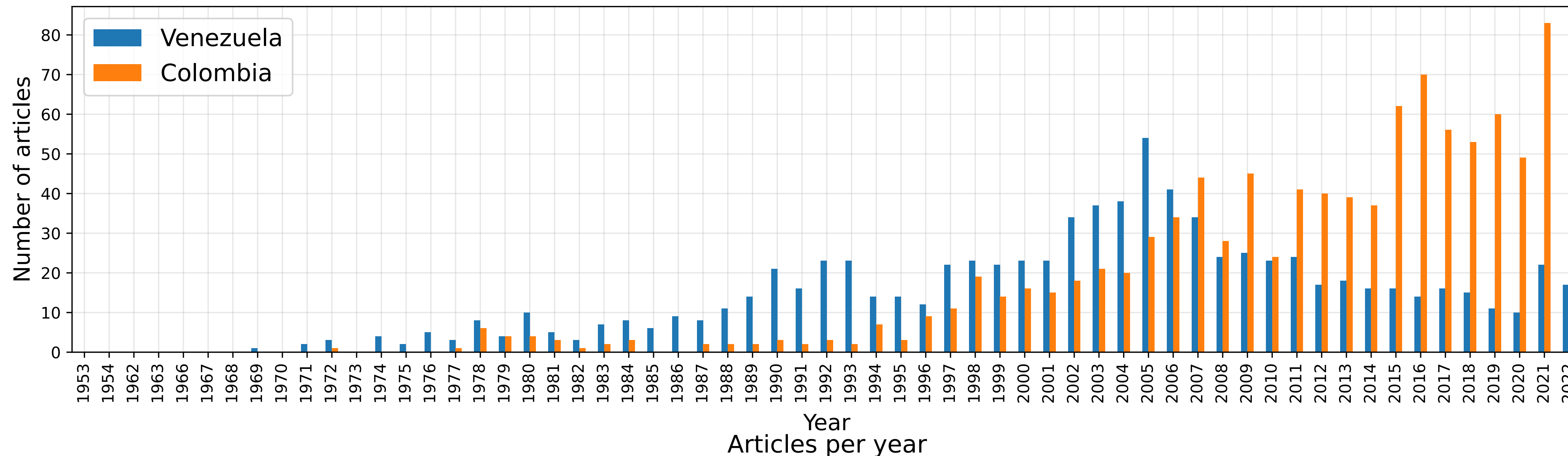
PUBLICATIONS PER YEAR

Articles per year

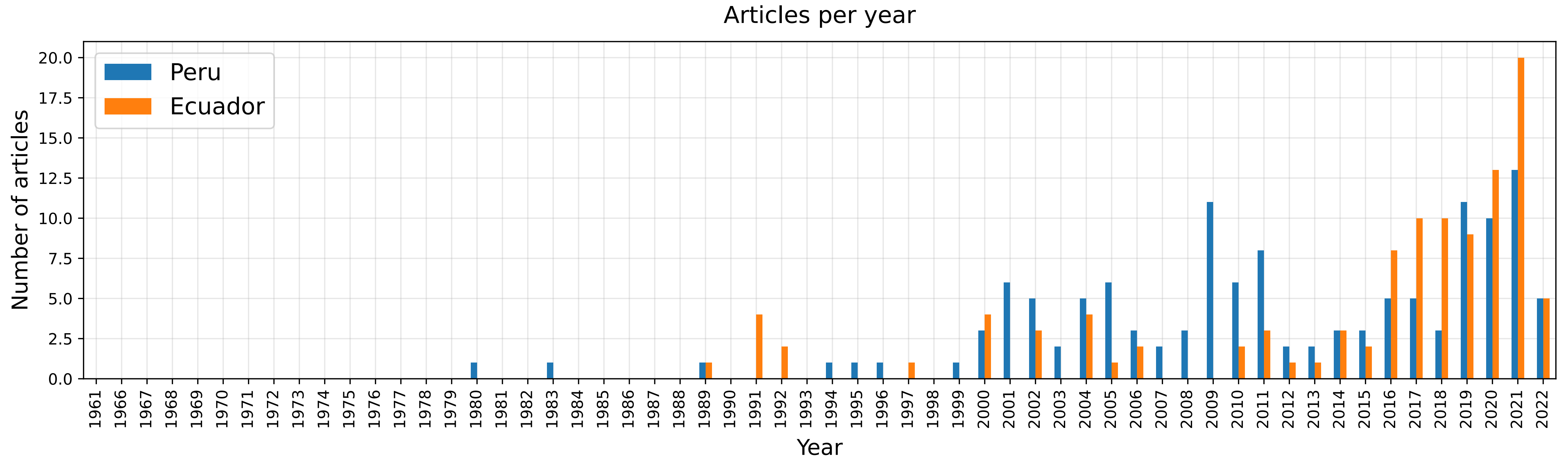


PUBLICATIONS VS YEAR

Articles per year



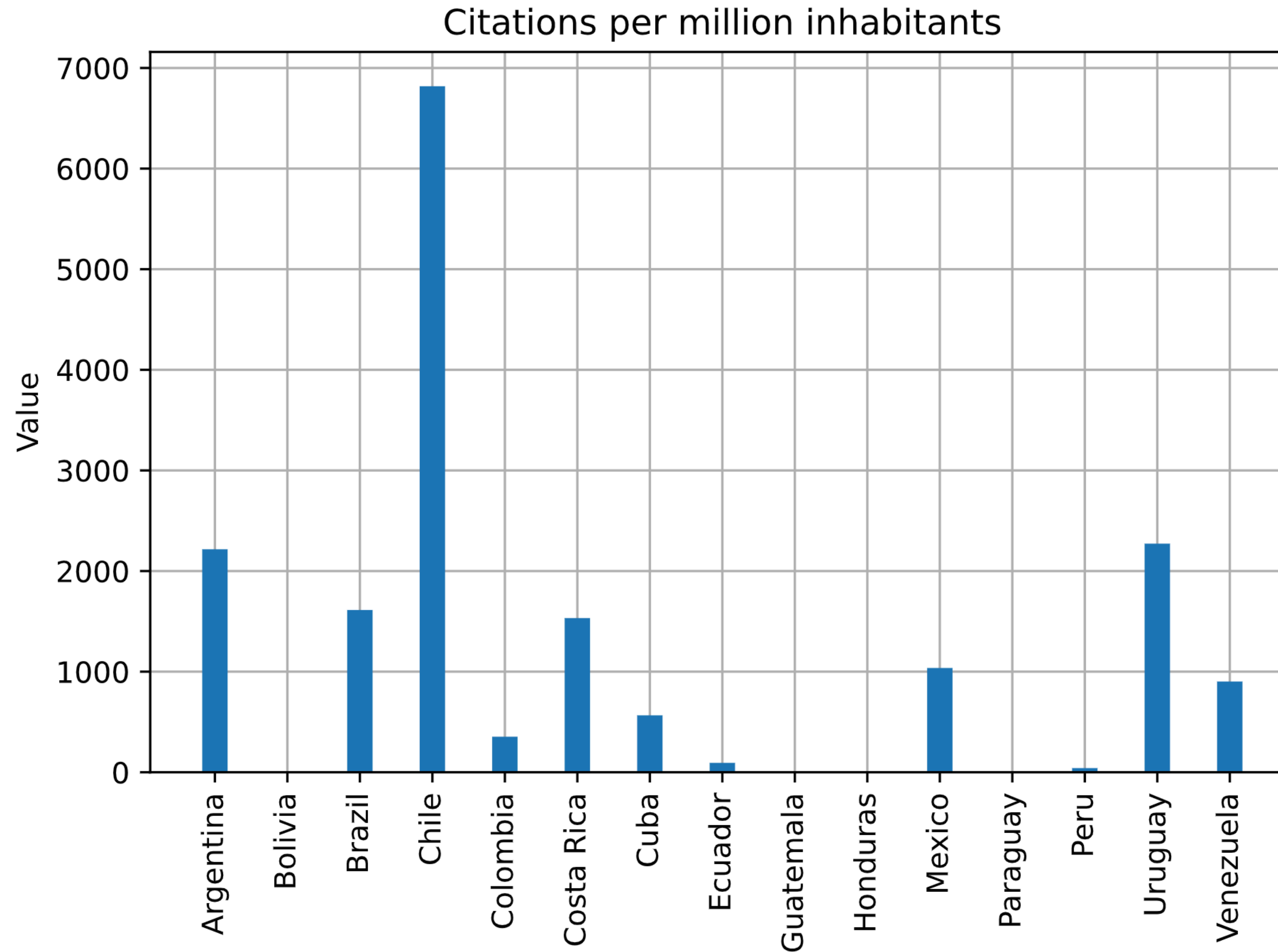
PUBLICATIONS VS YEAR



And many more countries ...

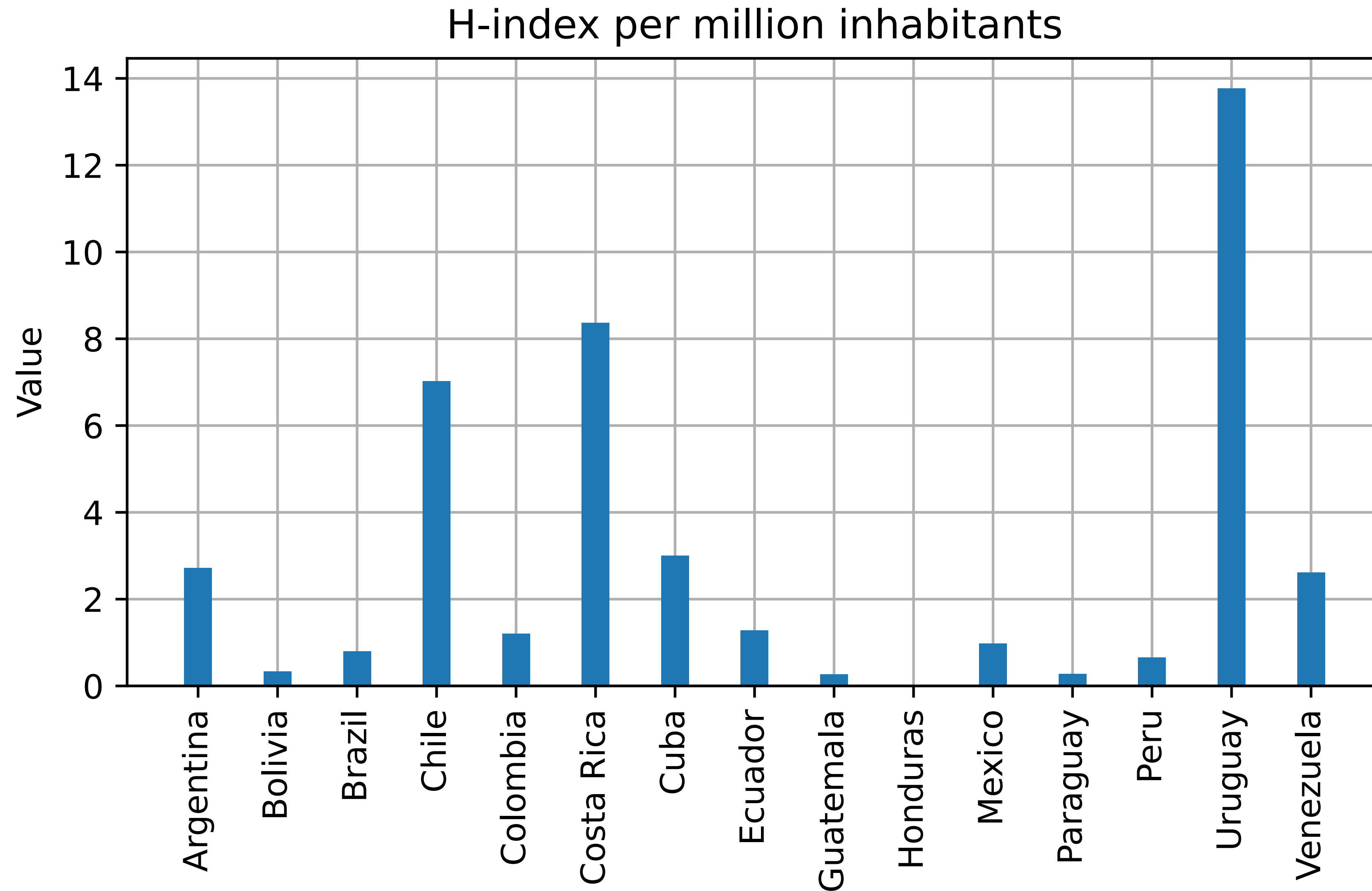
Country	Total citations	Total citations per million inhabitants
Argentina	101086	2217
Bolivia	50	4
Brazil	345902	1616
Chile	130995	6818
Colombia	18118	353
Costa Rica	7876	1533
Cuba	6406	566
Ecuador	1675	94
Guatemala	109	6
Honduras	0	0
Mexico	135046	1038
Paraguay	43	6
Peru	1389	42
Uruguay	7919	2272
Venezuela	25893	902

NORMALISED CITATIONS



Country	h-index	h-index per million inhabitants
Argentina	124	3
Bolivia	4	0
Brazil	172	1
Chile	135	7
Colombia	62	1
Costa Rica	43	8
Cuba	34	3
Ecuador	23	1
Guatemala	5	0
Honduras	0	0
Mexico	127	1
Paraguay	2	0
Peru	22	1
Uruguay	48	14
Venezuela	75	3

NORMALISED H-INDEX



CORRELATIONS WITH GDP

Country	Percentage of GDP
Argentina	0.46
Bolivia	0.28 (2002)
Brazil	1.21
Chile	0.34
Colombia	0.32
Costa Rica	0.37 (2018)
Cuba	0.55
Ecuador	0.44 (2014)
Guatemala	0.03
Honduras	0.04 (2004)
Mexico	0.28
Paraguay	0.14
Peru	0.16
Uruguay	0.48
Venezuela	0.34 (2014)

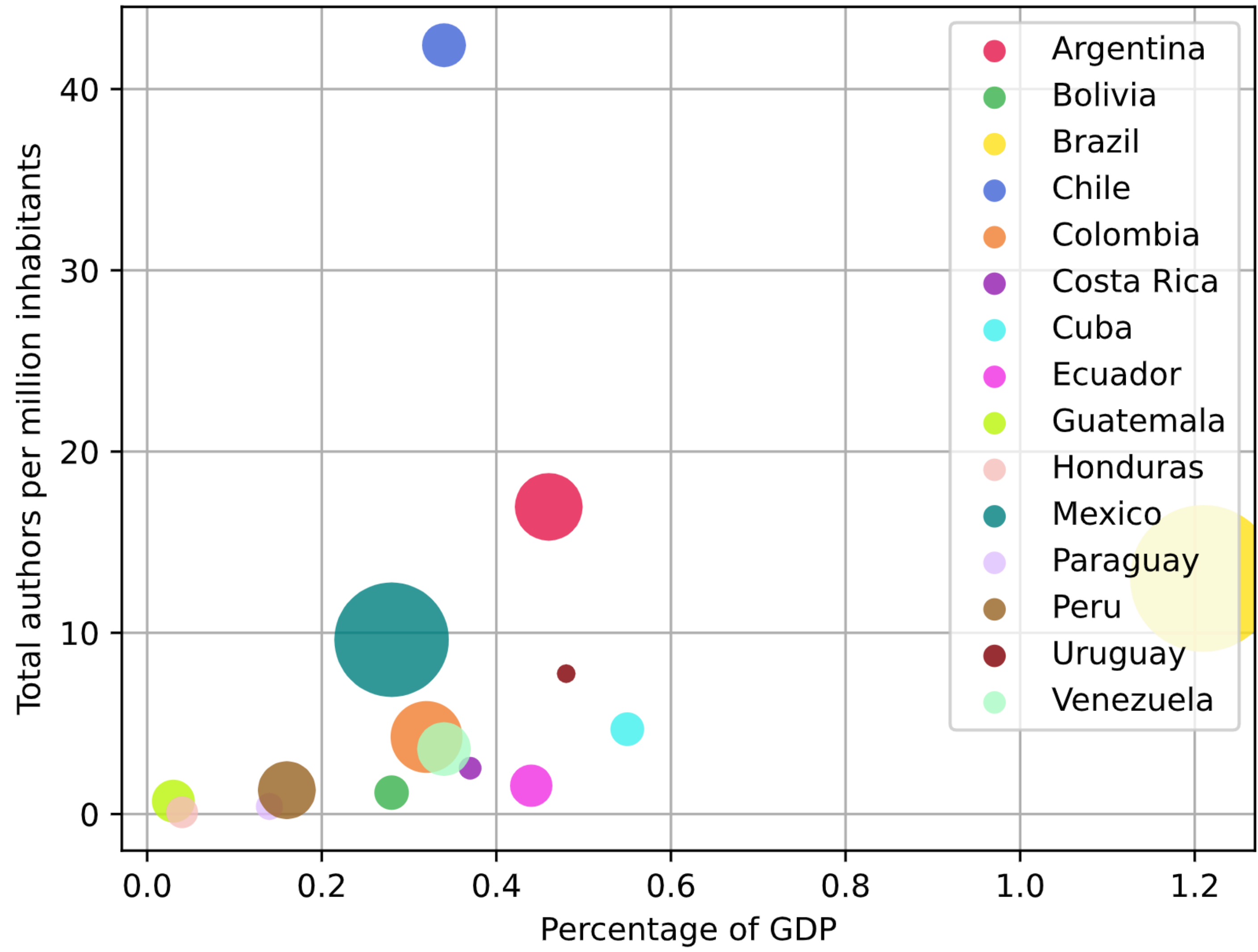
GDP = gross domestic product (invested in research and development)

Data obtained from <https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS> (checked in September 2022)

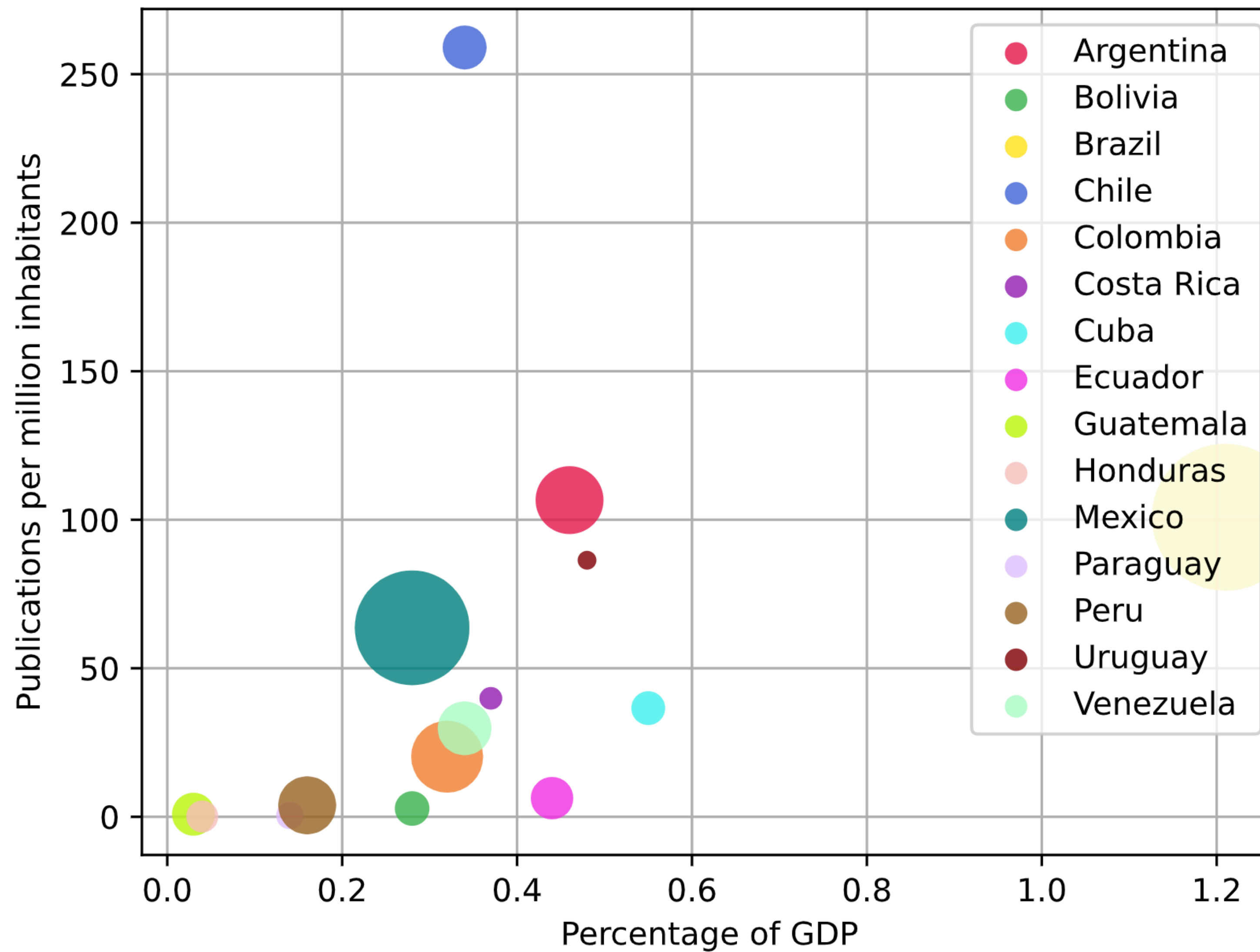
Some references (and issues):

Entity	Percentage of GDP
China	2.24
European Union	2.22
OECD	2.67
United Nations	2.33
United States	3.17

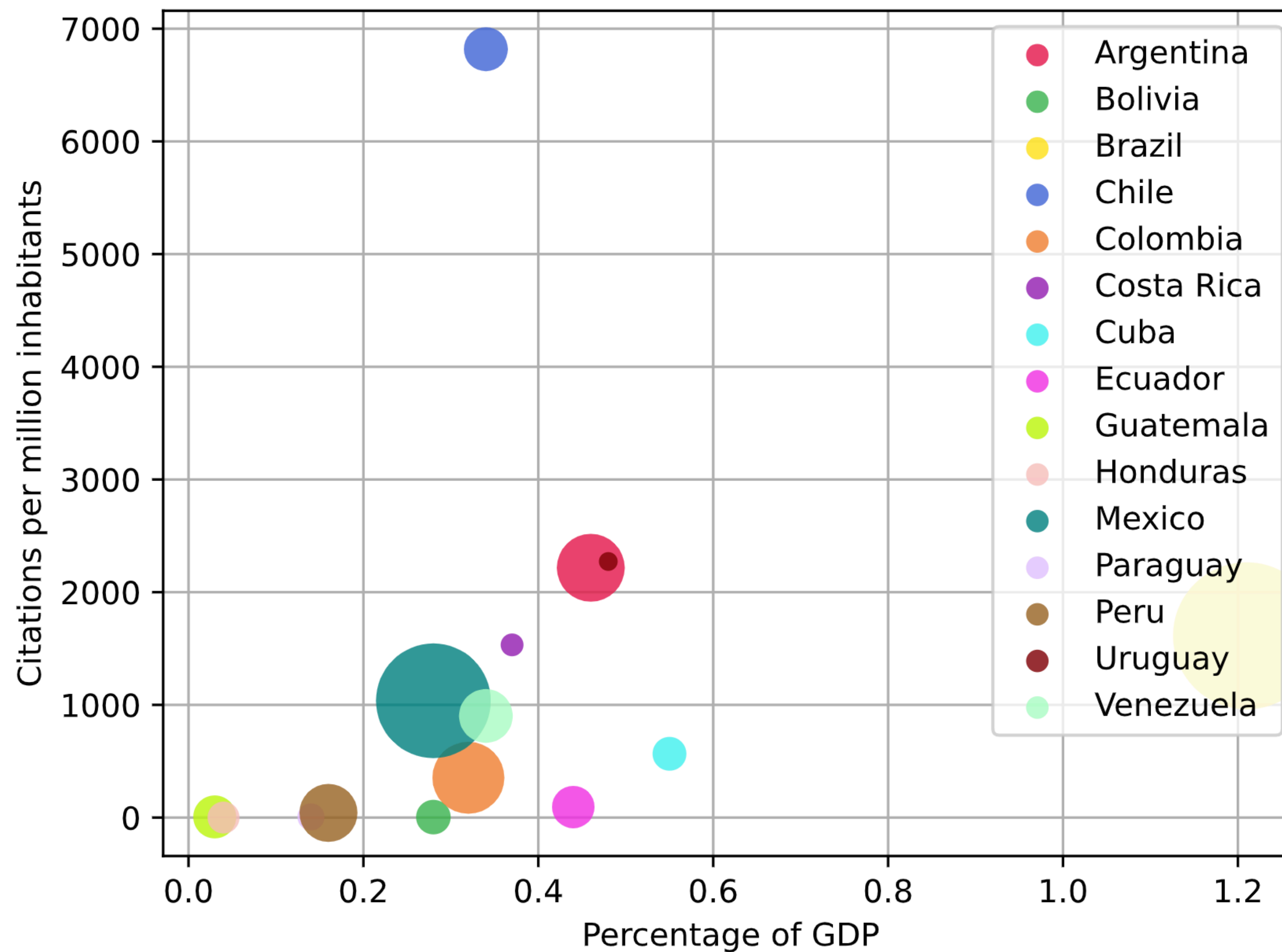
NORMALISED AUTHORS VS GDP



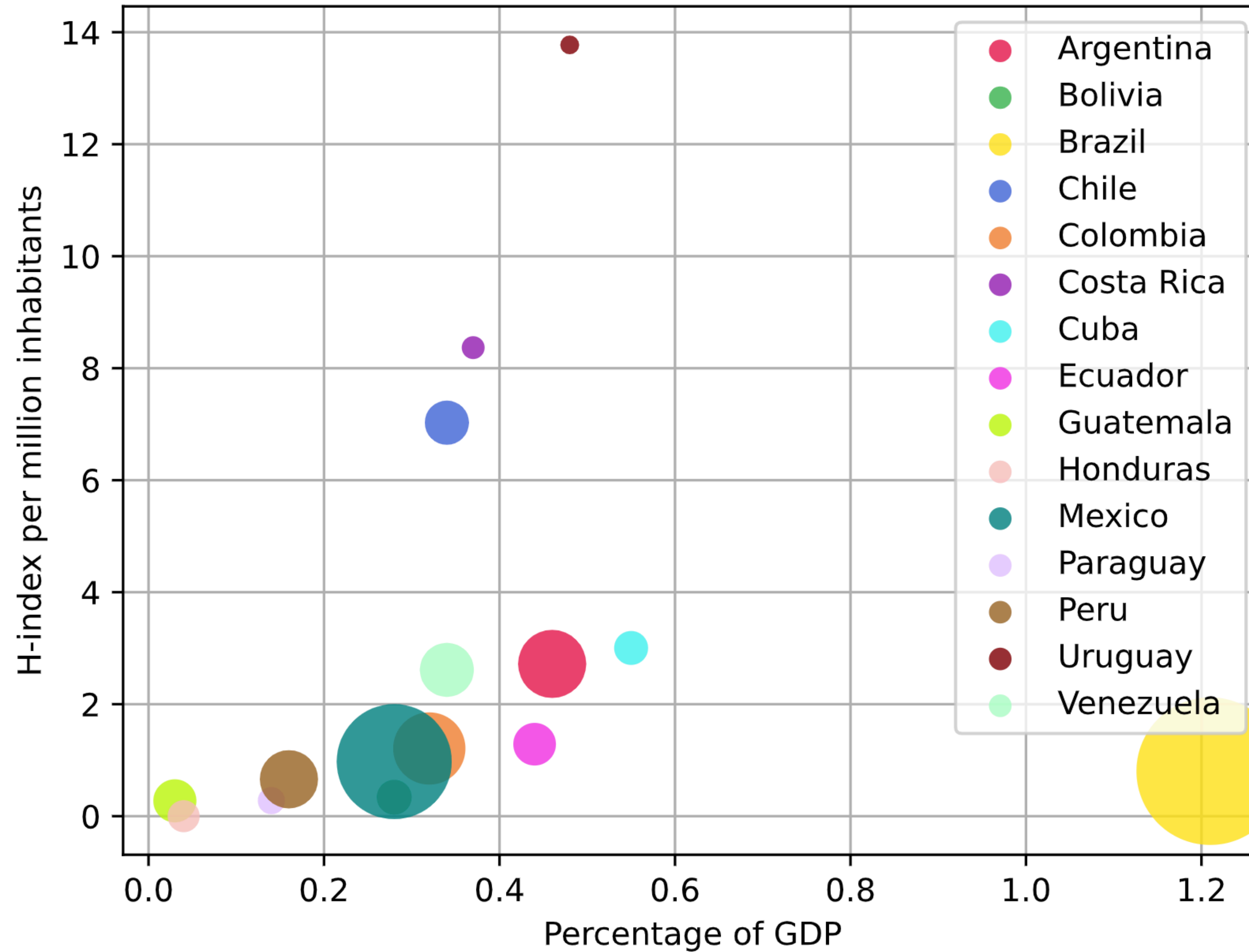
NORMALISED PUBLICATIONS AND GDP



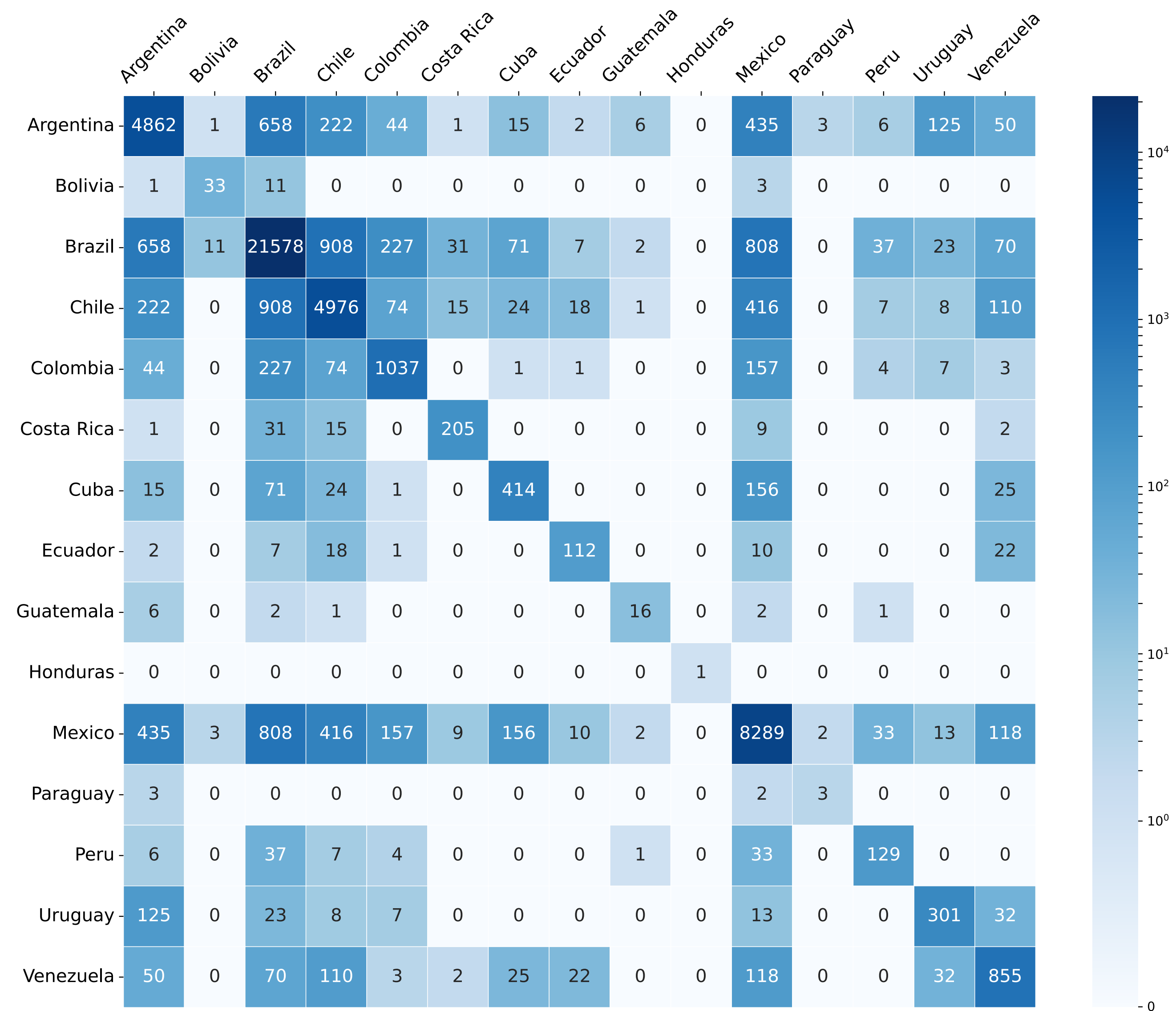
NORMALISED CITATIONS AND GDP



NORMALISED H-INDEX AND GDP



NUMBER OF COLLABORATIONS WITHIN LATIN AMERICA



CAN YOU HELP US?

Yes, by completing your INSPIRE-HEP profile

The image shows a screenshot of an INSPIRE-HEP author profile. The profile is for an author named "Author" (Cambridge U., DAMTP) with an ORCID iD icon. The profile includes a "hep-ph" subject tag, an "Author Identifier" field with the value "Identifier", and a "PhD Advisor" field with the value "Advisor". The profile also features a vertical timeline of education and work experience:

- 2021-present: **POSTDOC**, Cambridge U., DAMTP
- 2017-2021: **PHD**, Louvain U., CP3
- 2014-2017: **MASTER**, U. Bologna, DIFA
- 2011-2014: **UNDERGRADUATE**, U. Bologna, DIFA

At the bottom left, there are icons for email, Twitter, and an "edit" button. At the bottom right, it says "Updated on Feb 10, 2022".

FOR THE FUTURE

- Obtain collaborations with the rest of the world
- Explore the mobility of Latin American researchers
- Extra coordination with other LAAHECAP committees (JEDI, for example)
- Extend this analysis to other areas of physics and/or science in general

CONCLUSIONS AND OUTLOOK

- We discussed HEP Statistics in Latin America

- We have developed a database to study HEP research in Latin America
- We have developed a code to study different metrics (authors, publications, citations, etc)
- We are currently working in an extended report on the subject

- In the future:

- Study the experimental collaborations in Latin America
- Coordinate analysis with other committees in LAAHECAP
- Explore Latin American collaborations with the rest of the world
- ...

Thank you for your attention