



Publication Patterns in the Humanities

Generational Shifts and Changing Research Agendas

Co-authors: N. Robinson-Garcia, W. Arroyo-Machado, E. González-Salmón & D. Torres-Salinas





Daniel Torres-Salinas



Victoria Di Césare



Wenceslao Arroyo-Machado



François van Schalkwyk



Elvira Mª González Salmón



Manuel Escabias



Márcia R. Ferreira



Gabriela F. Nane



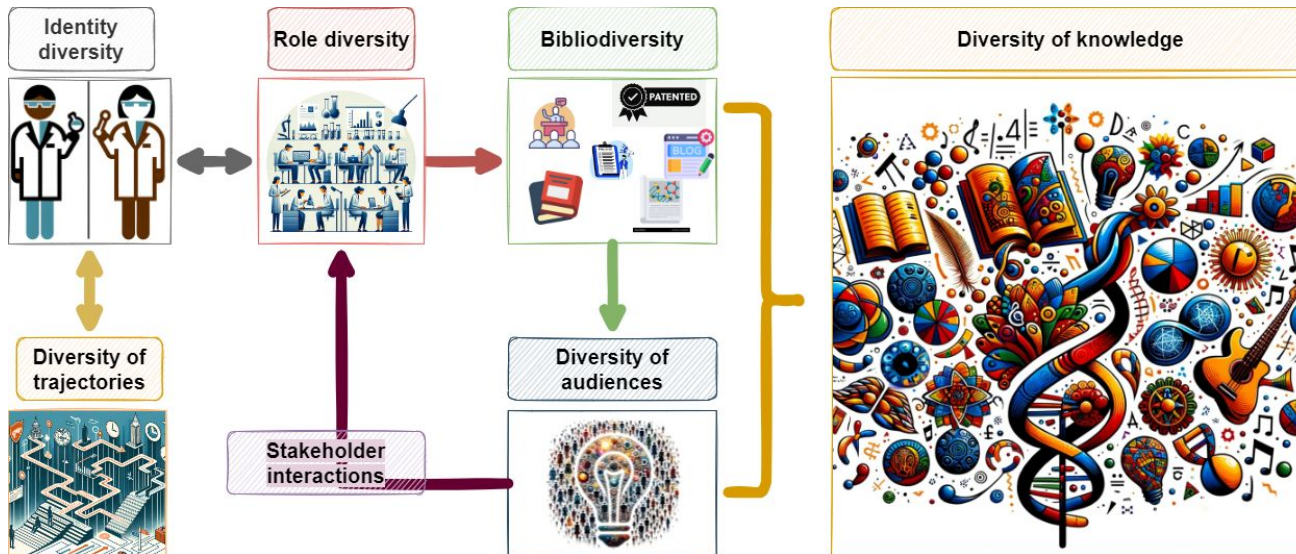
Julia Melkers



Rodrigo Costas

Motivation for the project

Studying diversity within the scientific workforce



Develop scientometric methods to better understanding **different forms of diversity** within the **scientific workforce**

Motivation for the study



Historical mistreatment of humanities in bibliometrics

- Citation windows
- Books and chapters
- **The role of language**
- Database indexing
- Research orientation

Setting | Some antecedents

Scientometrics (2010) 83:337–354
DOI 10.1007/s11192-009-0088-9

Why with bibliometrics the Humanities does not need to be the weakest link

Indicators for research evaluation based on citations, library holdings, and productivity measures

A. J. M. Linmans

Chapter 21

THE FOUR LITERATURES OF SOCIAL SCIENCE

Diana Hicks

*School of Public Policy, Georgia Institute of Technology, GA, USA
E-mail: diana.hicks@pubpolicy.gatech.edu*

What makes Humanities different?

- Diversity of audiences
- Locality of research
- Non-cumulative production knowledge
- The role of citations

To what extent are humanists different?

How extended are these practices within fields?

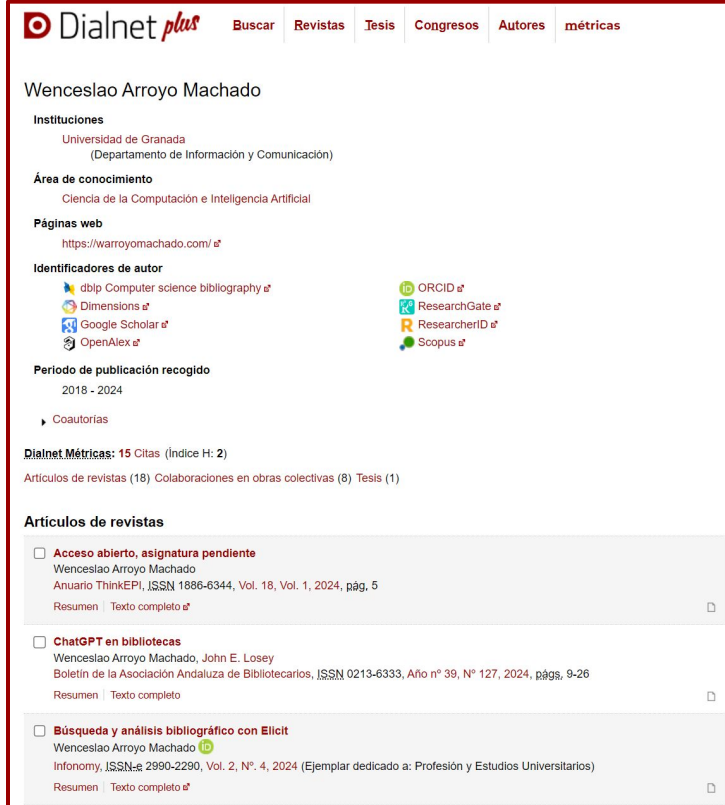
Which are the reasons behind such differences?

Setting | **Goals of the study**

NOT OUTPUTS

1. Identify publication patterns of humanists across and within fields
2. Test two potential explanations for differences:
 - H1.** Differences on publication patterns relate to generational shifts on publication habits
 - H2.** Differences on publication patterns are topic specific

Data and Methods | Dialnet



The screenshot shows the profile of Wenceslao Arroyo Machado on the Dialnet platform. At the top, there is a navigation bar with links for 'Buscar', 'Revistas', 'Tesis', 'Congresos', 'Autores', and 'métricas'. The profile includes the following information:

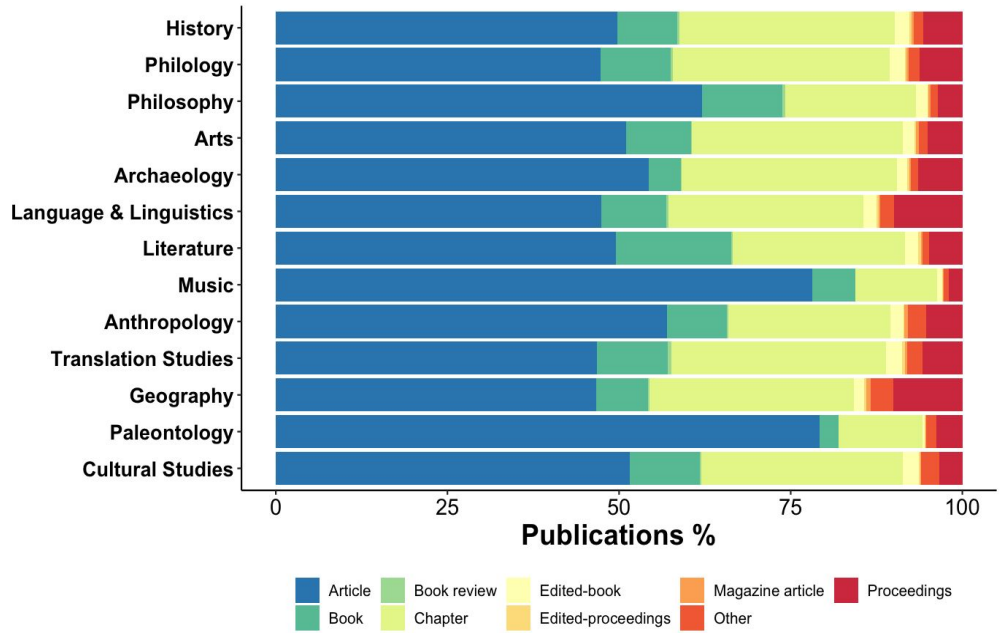
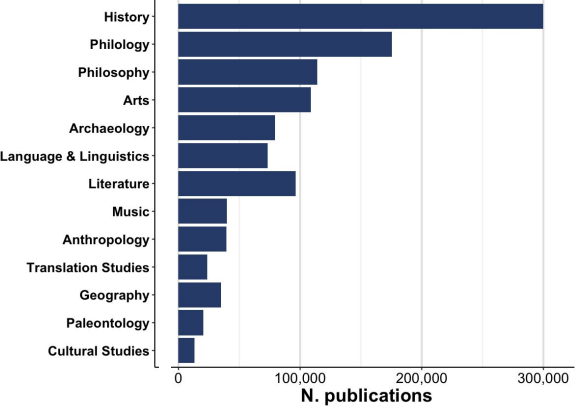
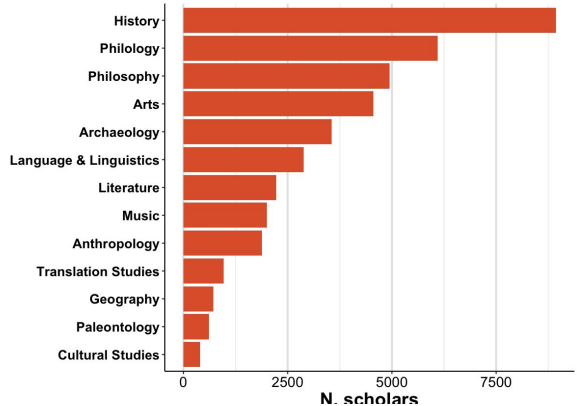
- Instituciones:** Universidad de Granada (Departamento de Información y Comunicación)
- Área de conocimiento:** Ciencia de la Computación e Inteligencia Artificial
- Páginas web:** <https://warroyomachado.com/>
- Identificadores de autor:** dblp Computer science bibliography, Dimensions, Google Scholar, OpenAlex, ORCID, ResearchGate, ResearcherID, and Scopus.
- Periodo de publicación recogido:** 2018 - 2024
- Coautorías:** (link to view)
- Dialnet Métricas:** 15 Citas (Índice H: 2)
- Artículos de revistas (18):** Colaboraciones en obras colectivas (8) Tesis (1)

Under 'Artículos de revistas', three entries are visible:

- Acceso abierto, asignatura pendiente**
Wenceslao Arroyo Machado
Anuario ThinkEPI, ISSN 1886-6344, Vol. 18, Vol. 1, 2024, pág. 5
Resumen | Texto completo
- ChatGPT en bibliotecas**
Wenceslao Arroyo Machado, John E. Losey
Boletín de la Asociación Andaluza de Bibliotecarios, ISSN 0213-6333, Año nº 39, Nº 127, 2024, págs. 9-26
Resumen | Texto completo
- Búsqueda y análisis bibliográfico con Elicit**
Wenceslao Arroyo Machado
Infonomiy, ISSN-e 2990-2290, Vol. 2, Nº. 4, 2024 (Ejemplar dedicado a: Profesión y Estudios Universitarios)
Resumen | Texto completo

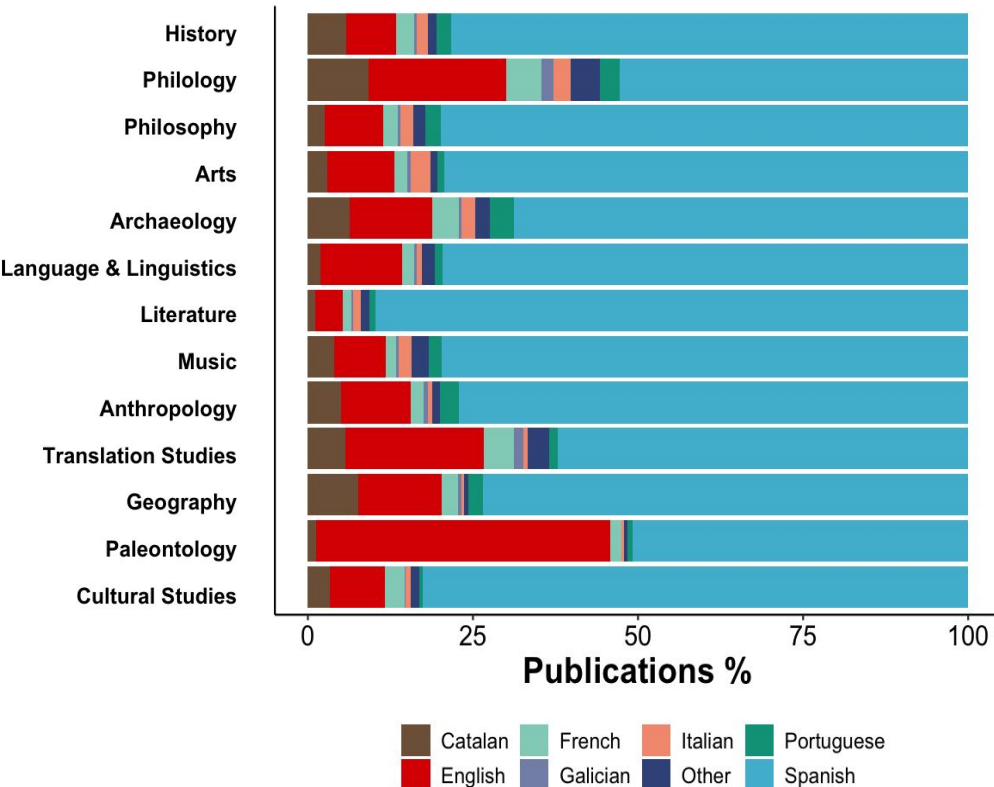
- Dialnet is a **free bibliographic database** created and maintained by the Dialnet Foundation from University of La Rioja.
- It contains mainly Spanish-language SSH literature **~2,500 journals + books, proceedings...**
- It is curated by university librarians (through a consortium) and contains author profiles with their publication records

Data and Methods | The dataset



Our dataset contains ~1,2M publications from ~40k scholars from 13 ≠ fields - 1950-2021 period

Data and Methods | The dataset



- Spanish is the main publication language in all fields
- Despite the preponderance of Spanish, there is a significant share of publications in other languages
- Philology and Paleontology have the highest shares of non-Spanish publications
- English language is the second language

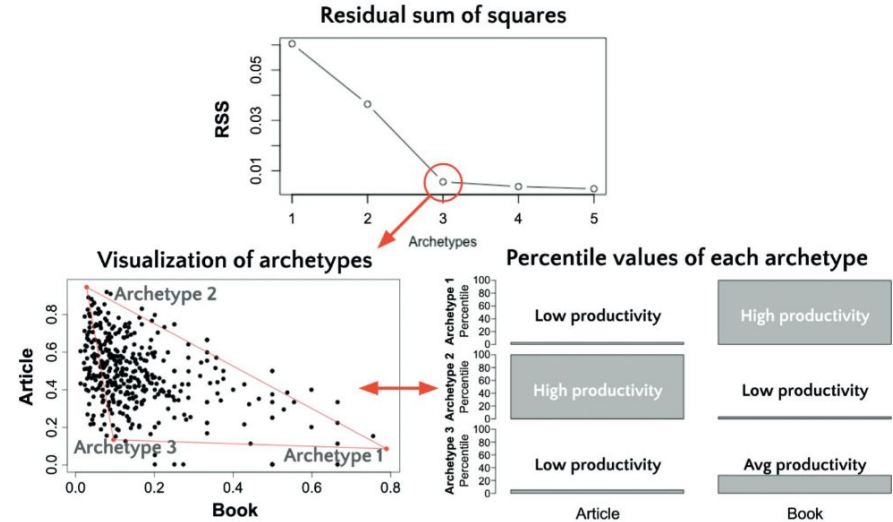
Data and Methods | Variables

For each humanist we compute the following variables:

● No. of books	● No. of conf. papers
● No. of chapters	● No. edited proceedings
● No. of journal articles	● % indexed in WoS/Scopus
● No. of edited books	● % non-English pubs

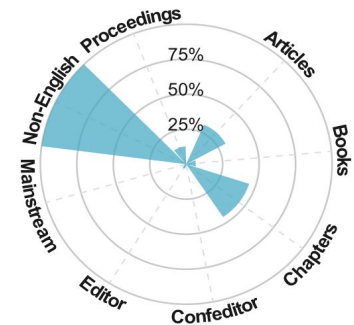
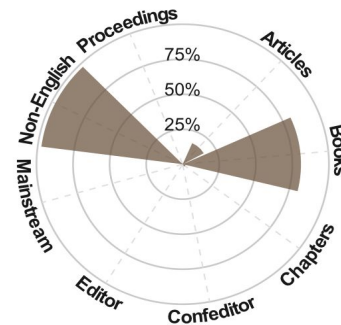
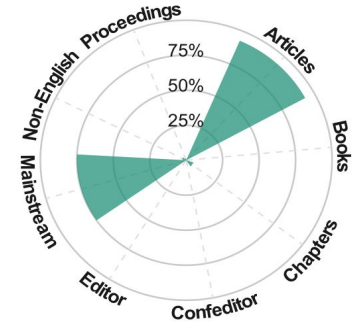
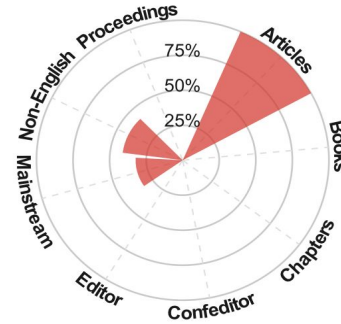
Data and Methods | Methodological design

1. Archetypal Analysis by field
2. Hierarchical clustering of archetypes
3. Generational year
4. Topic identification



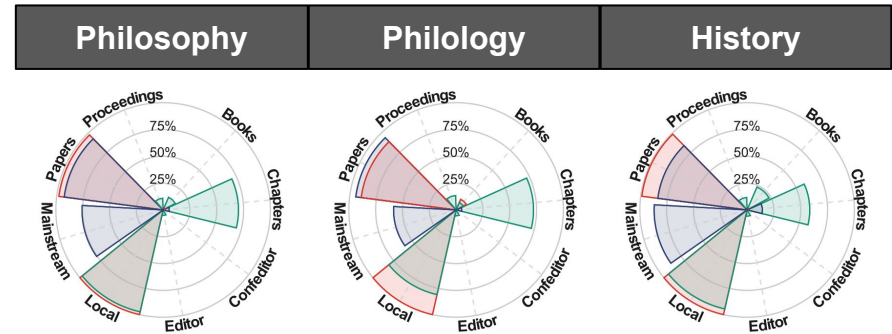
Data and Methods | Methodological design

1. Archetypal Analysis by field
2. Hierarchical clustering of archetypes
3. Generational year
4. Topic identification



Data and Methods | Methodological design

1. Archetypal Analysis by field
2. Hierarchical clustering of archetypes
3. Generational year
4. Topic identification



Philosophy1 ≈ Philology2 ≈ History1

Philosophy2 ≈ Philology3 ≈ History3

Philosophy3 ≈ Philology1 ≈ History2

Data and Methods | Methodological design

1. Archetypal Analysis by field
2. Hierarchical clustering of archetypes
- 3. Generational year**
4. Topic identification



Dialnet *plus* [Buscar](#) [Revistas](#) [Tesis](#) [Congresos](#) [Autores](#) [métricas](#)

Wenceslao Arroyo Machado

Instituciones
Universidad de Granada
(Departamento de Información y Comunicación)

Área de conocimiento
Ciencia de la Computación e Inteligencia Artificial

Páginas web
<https://warroyomachado.com/>

Identificadores de autor

dblp Computer science bibliography	ORCID
Dimensions	ResearchGate
Google Scholar	ResearcherID
OpenAlex	Scopus

Periodo de publicación recogido
2018 - 2024

► Coautorías

Data and Methods | Methodological design

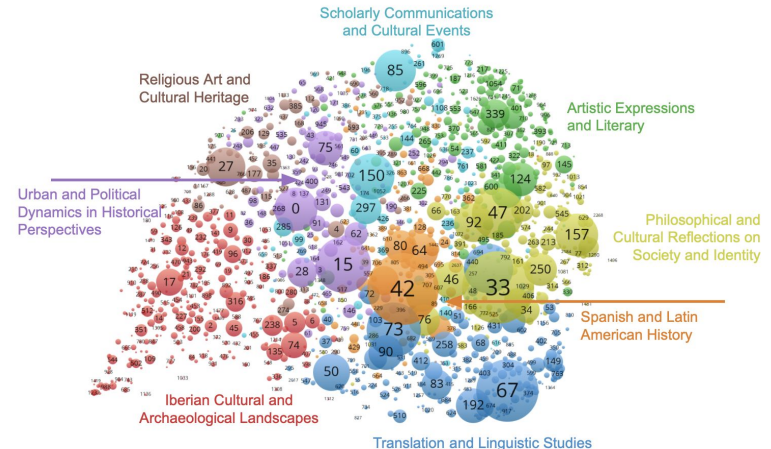
1. Archetypal Analysis by field
2. Hierarchical clustering of archetypes
3. Generational year
4. **Topic identification**

Step 1. Translate title to English with GPT-4

Step 2. Extract keywords using RAKE algorithm

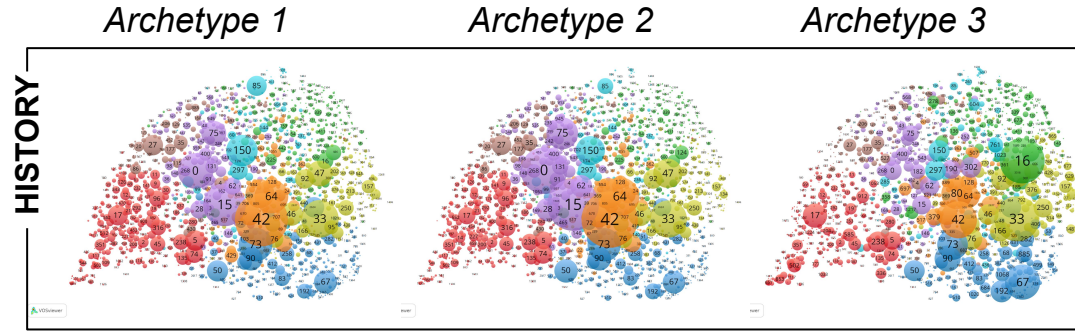
Step 3. Create co-occurrence matrix of keywords

Step 4. Apply the Leiden community detection algorithm to identify topics (1,452) and fields (8)



Data and Methods | Methodological design

1. Archetypal Analysis by field
2. Hierarchical clustering of archetypes
3. Generational year
4. **Topic identification**



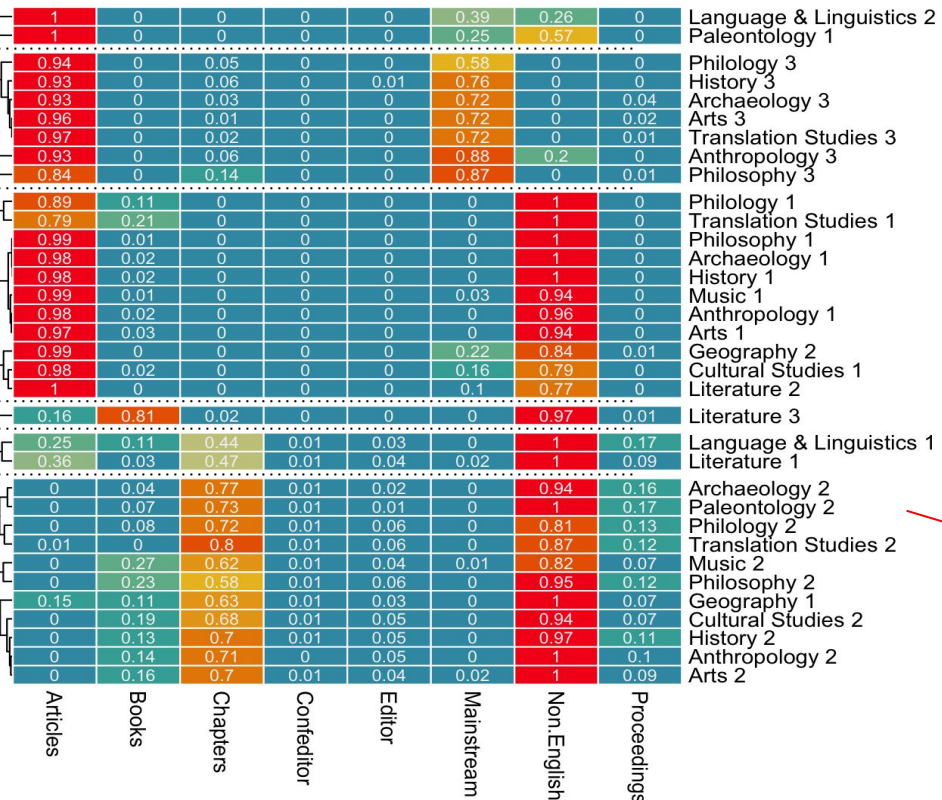
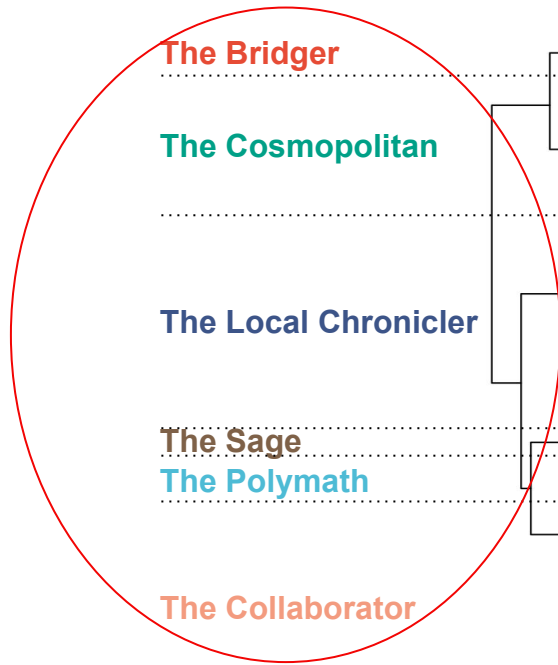
Once reduce to 1 dimension we can see how alike profiles are in terms of topic distribution

Step 5. Each archetype-discipline combination is a distribution of pubs across topics

Step 6. Compute Jaccard distance between vectors within disciplines

Step 7. Reduce to 1 dimension using Multidimensional Scaling.

Findings | Archetypes and publication profiles

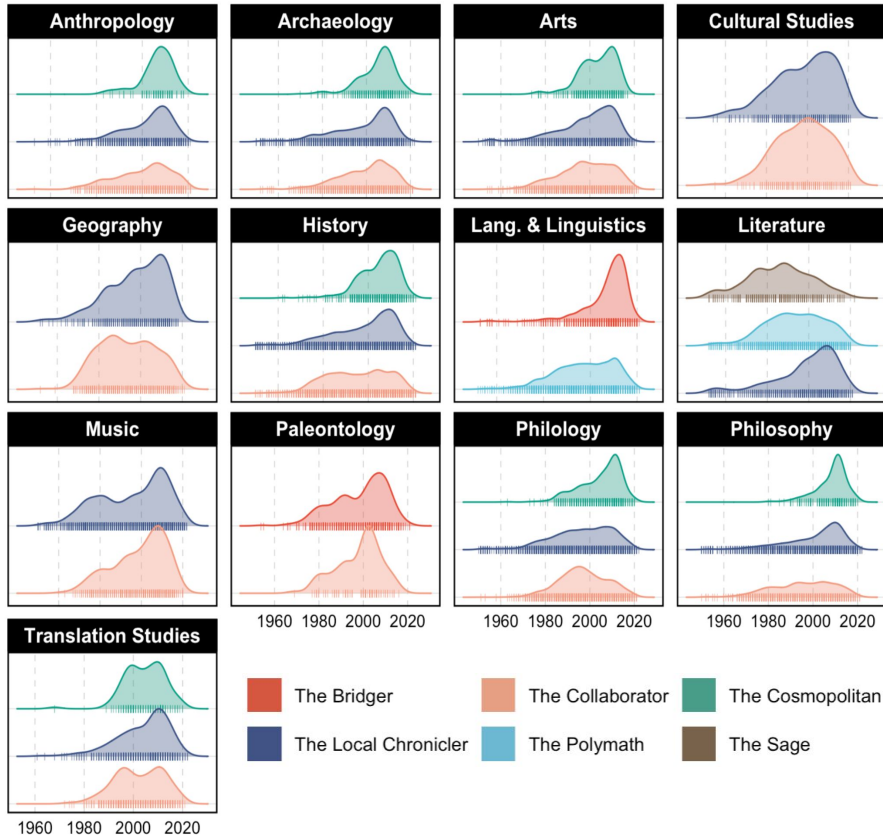


Normalized parameters

Archetypes

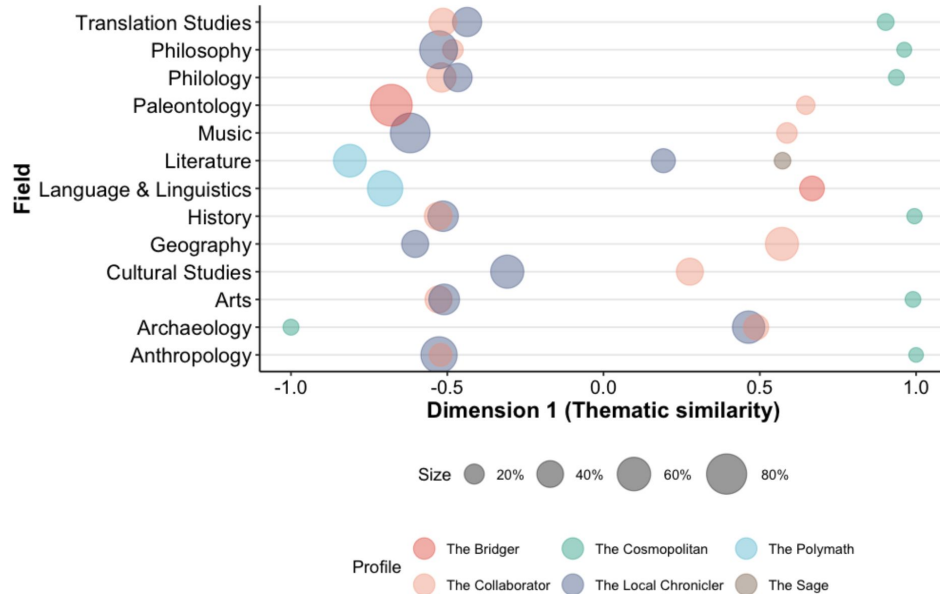
Profiles
(understood as clusters of archetypes)

Findings | Generational differences



- Shift across profiles go in multiple directions - **No common trend**
- **History, Anthropology, Archaeology and the Arts** younger cosmopolitan profiled with collaborator and local chronicler remaining cross-generational
- **Philology and Philosophy** younger cohorts of cosmopolitans, but decline of the collaborator
- **Polymath is unique to Lang. & Linguistics and Literature**, rise of bridgers for L&L, decline of sages for Literature
- **Paleontology** distinct from the rest of the fields

Findings | Topic-relatedness



- Except for Geography, Music and Cultural Studies, **no topic differences between the local chronicler and the collaborator profiles**
- **Cosmopolitan, polymaths, sages and bridgers scholars have a topic orientation distinct from the rest of the profiles**
- In the case of Literature, we clearly observe how **each of the three publication profile responds to different topic landscapes**

Discussion | Recap

- We analyze a unique dataset of publication records assigned to authors from 13 fields from the humanities, mainly located in Spain (also Spanish speaking South American countries)
- We applied an archetypal analysis to identify publication profiles based on the type of outputs they produce **by language, publication type and database indexing**
- We cluster archetypes to identify similarities across fields
- We look into differences and similarities based on first year publication of authors and topic distribution

Discussion | Take-away points

- Humanists exhibit **six distinct publication profiles** which **vary across and within fields** in the humanities
- These differences in publication patterns **seem to be related with** *but not only!* **generational changes in publication habits**
- But, there is **no unique trend that could point towards an influence of journal-centric evaluation practices**
- Topic orientation also seems to be determining publication habits, especially for some of these profiles
- **There is a danger on journal/English-centric evaluative practices of suppressing topic diversity** by pushing scholars towards certain publication venues

1 Publication Patterns in the Humanities: Generational
2 Shifts and Changing Research Agendas

3 [Nicolas Robinson-García](#)¹, [Wenceslao Arroyo-Machado](#)^{2*}, [Elvira González-Salmón](#)¹ and
4 [Daniel Torres-Salinas](#)¹

5
6 ¹Unit for Computational Humanities and Social Sciences (U-CHASS), EC3 Research Group,
7 [University of Granada](#), Granada, Spain

8 ²Center for Science, Technology and Environmental Policy Studies (CSTEPS), School of Public
9 Affairs, [Arizona State University](#), Phoenix, AZ 85004, USA

10
11 * Corresponding author: warrayom@asu.edu

12 Abstract

13 The humanities have distinct publication practices compared to the sciences, with books and
14 local language literature being essential. This study aims to identify and analyze the
15 publication patterns of humanities scholars in Spanish-speaking countries, revealing unique
16 publication behaviors and fostering diverse perspectives rather than linear knowledge
17 progression. We analyzed the publication histories of approximately 40,000 scholars from
18 1950 to 2021 using data from the Dialnet database. By identifying archetypal publication
19 profiles, we explored their distribution across generational cohorts and research topics. Our
20 findings reveal substantial generational shifts towards journal-centric publication patterns
21 probably influenced by bibliometric-driven evaluation systems. The also show a relation
22 between publication patterns and research topics. This highlights the need for more inclusive
23 assessment practices that recognize the diverse nature of humanities scholarship. We
24 contribute to ongoing discussions on promoting bibliodiversity in research assessment and
25 the potential impacts of current policies on the humanities.

26 Keywords

27 humanities, publication patterns, bibliodiversity, research evaluation, Spanish-speaking
28 countries, generational shifts
29
30
31
32
33
34
35
36
37
38
39
40

41 _____
42 This is a preprint and has not undergone peer review
43 The final published version may differ
44 [10.5281/zenodo.13989299](https://doi.org/10.5281/zenodo.13989299)



Preprint alert!

<https://doi.org/10.5281/zenodo.13969299>



**Feel free to share your
feedback with us!** 😊



Thank you!

Questions, comments?



Acknowledgements: We thank the **Dialnet Foundation** for facilitating access to the publication and author data.