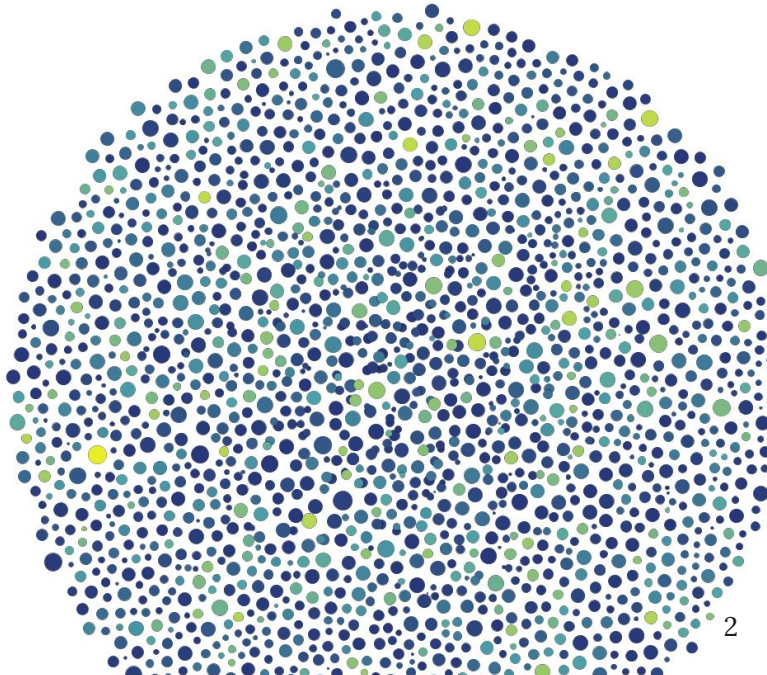
Identifying communities of interest in social media: Microbiology as a case study

Wenceslao Arroyo-Machado
Daniel Torres-Salinas
Nicolas Robinson-Garcia

Table of contents



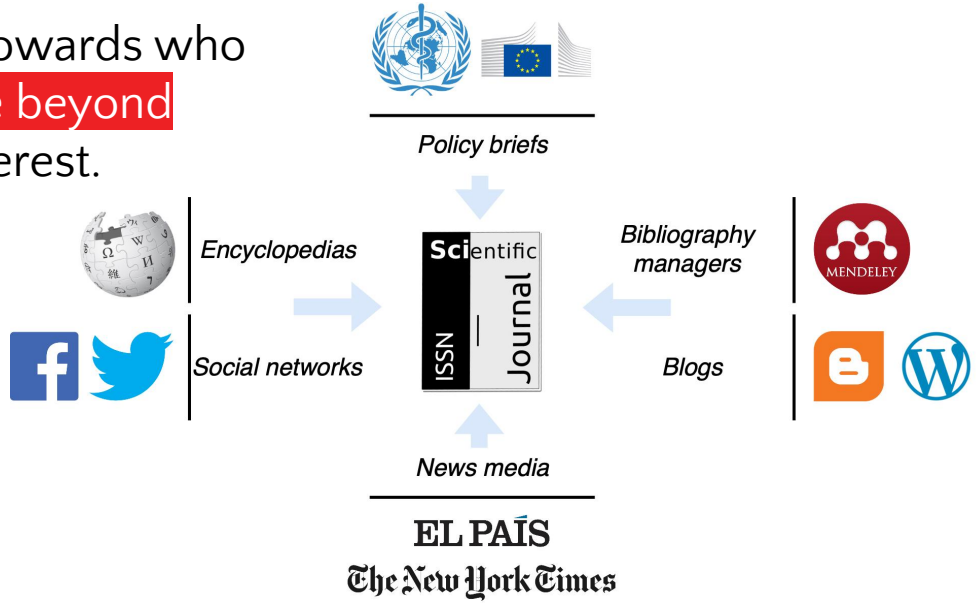
1 Introduction & objectives

In which way the mentions to scientific publications can be used to contextualize communities of attention?

Introduction & objectives

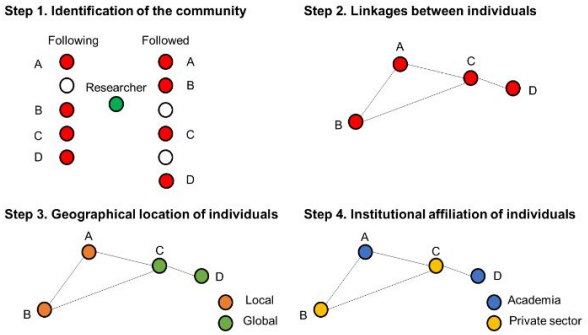
The simple counting of altmetric mentions has been criticized as very limited

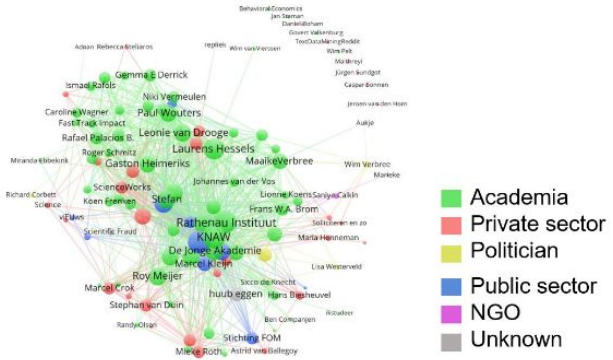
But this mentions can point us towards who is consuming scientific literature beyond academia and their topics of interest.



Introduction & objectives

To analyse these altmetric mentions one approach is to characterise the interaction community network of researchers.





Robinson-Garcia, N., van Leeuwen, T. N., & Rafols, I. (2018). Using altmetrics for contextualised mapping of societal impact: From hits to networks. Science and Public Policy, 45(6), 815-826.

Introduction & objectives

Our main goal is to build a methodology by which we can better understand which research topic and to whom it can be of interest.

Following up from a recent study we expand our analysis by identifying and characterising users with discussions in the field of Microbiology.

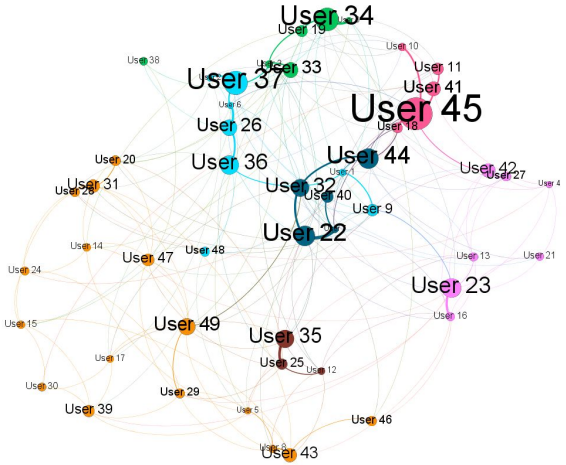
2

Materials & methods

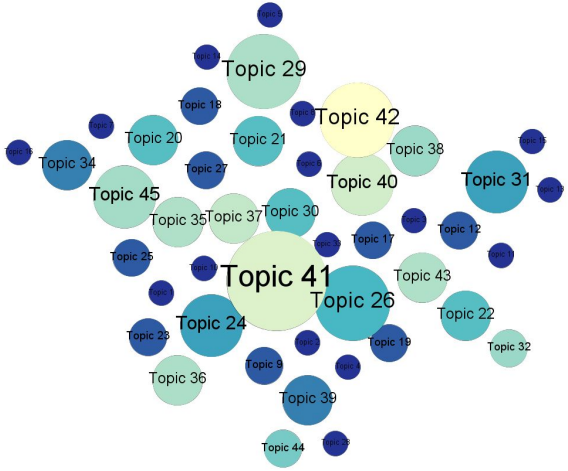
Data collect and visualizations creation

Materials & methods

Network analysis and visualization techniques have been used successfully to visualize the different interactions between actors and scientific publications.



Users' network



Overlay mapping of topics

Materials

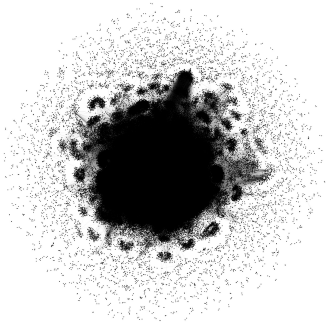
1	Download records of publications indexed in WoS (2012-2018) in the subject categories of *Microbiology* and *Biotechnology and Applied Microbiology*	We downloaded a total of 382,998 records	
2	Recover all papers which at least one mention by any of the Altmetric.com sources using DOIs resources	We downloaded 174,799 papers	Mentions dataset
3	Download mentions to these papers from all Altmetric.com sources	We downloaded 1,594,856 mentions	
4	Filter mentions sources to Twitter, news media and policy briefs	1,252,822 mentions	Publications dataset

Process of collection and the evolution of samples size

Methods

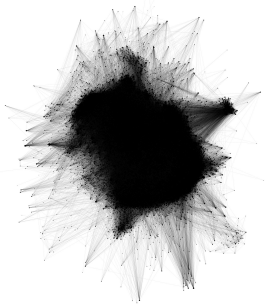
Mentions dataset

1. Full two-mode network



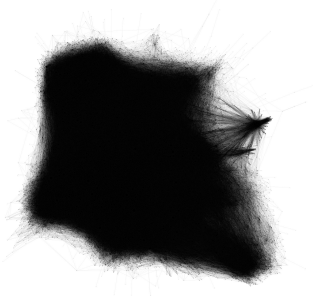
1,252,822 mentions

2. Network filtering



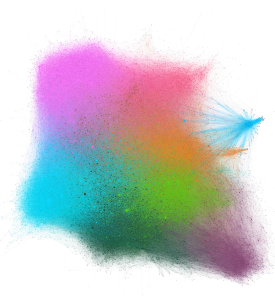
The main component of only nodes with a minimum indegree or outdegree of 10

3. Network layout



Nodes distributed using the ForceAtlas2 layout algorithm

4. Communities detection



15 communities detected using the Louvain method (Q=0.5)

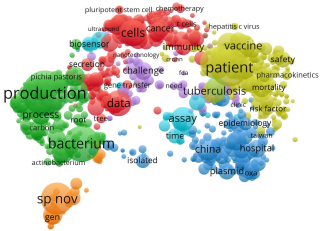
Methods

Publications dataset

1	Cefpodoxime vs ciprofloxacin for short-course treatment of acute uncomplicated cystitis: a randomized trial.
2	Tedizolid Phosphate vs Linezolid for Treatment of Acute Bacterial Skin and Skin Structure Infections
3	A Culture-Independent Sequence-Based Metagenomics Approach to the Investigation of an Outbreak of Shiga-Toxigenic Escherichia coli O104:H4
4	Supreme Court Rules Against Gene Patents
5	Pustulonodular Lesion on the Nose
6	Oral, Capsulized, Frozen Fecal Microbiota Transplantation for RelapsingClostridium difficileInfection
7	Association Between Vancomycin Minimum Inhibitory Concentration and Mortality Among Patients With Staphylococcus aureus Bloodstream Infections: A Systematic Review and Meta-analysis
8	Rational Use of Antibiotics in the ICU: Balancing Stewardship and Clinical Outcomes
9	Drugs for Urinary Tract Infections
10	The Human Microbiome and the Future Practice of Medicine

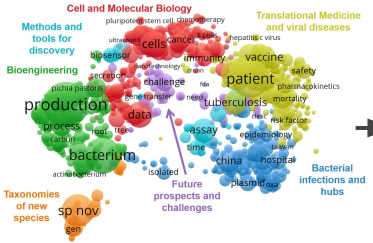
5. Extract titles

6. Create thematic landscape



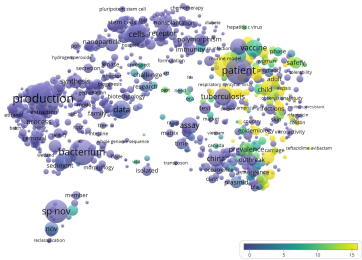
The network was filtered to a minimum of 50 co-occurrences and only top 60% most relevant terms

7. Identify communities

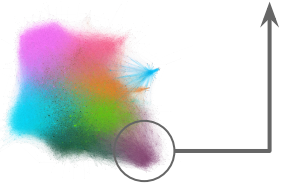


Experts in Microbiology corroborated

8. Overlay terms map



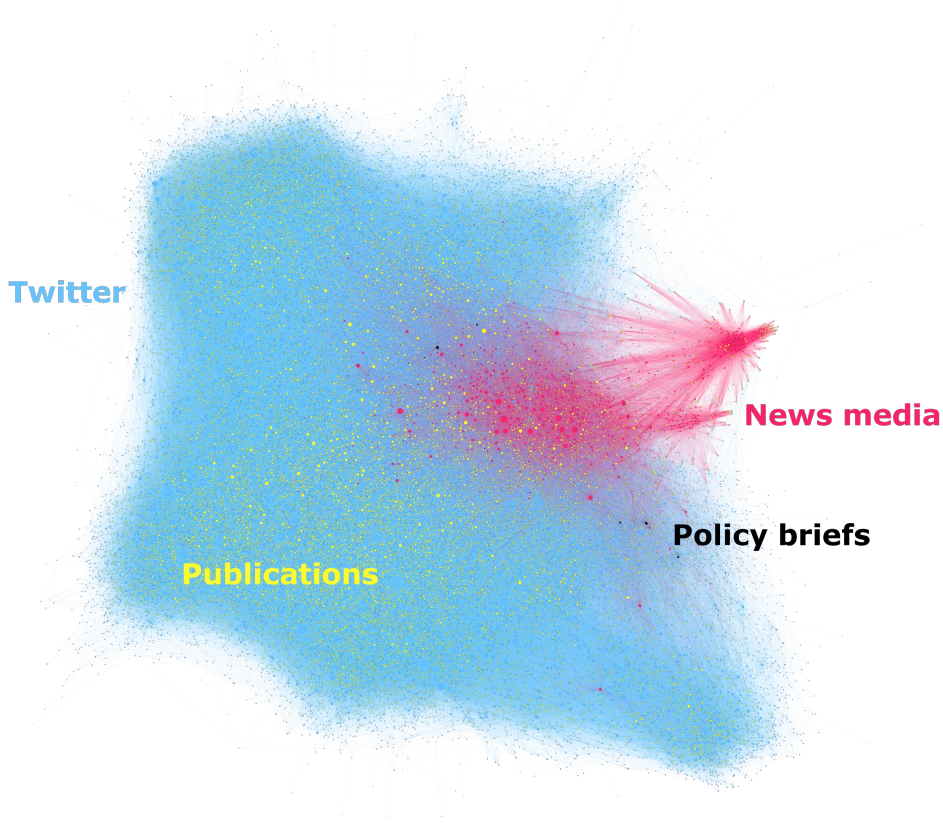
The score applied depends on the binary presence of the publications title terms mentioned by each cluster



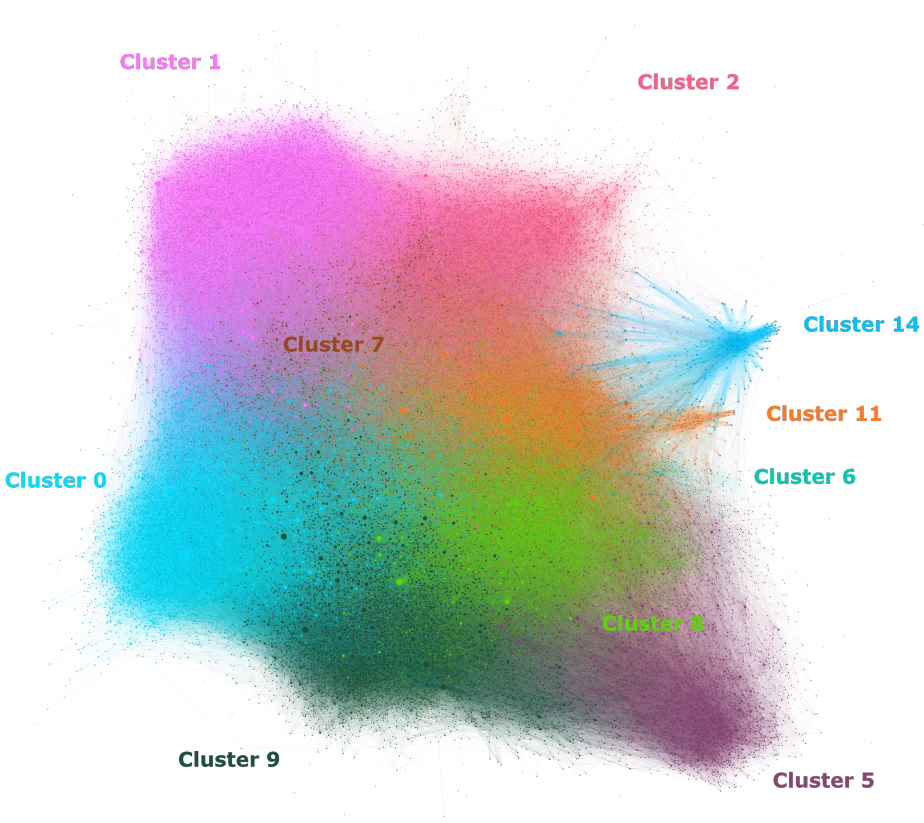
3

Results

Are there specific communities discussing high impact research?



Nodes colored by type



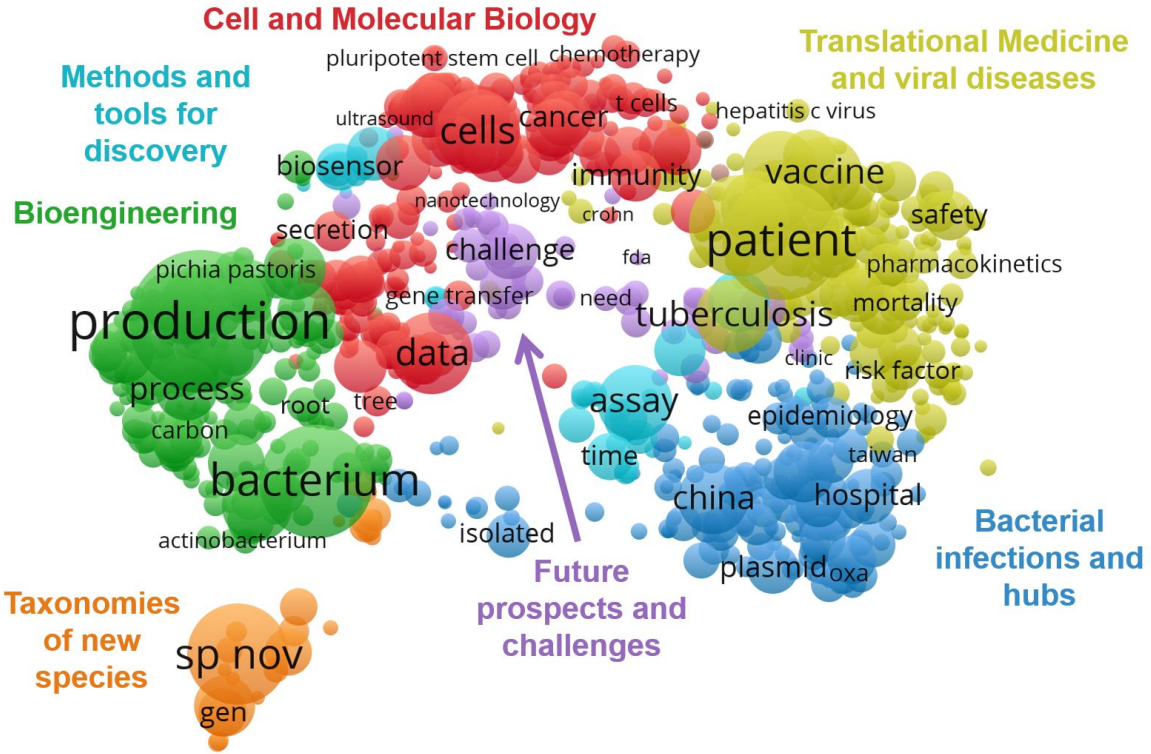
Nodes colored by cluster

Giant component of two-mode top 10% network

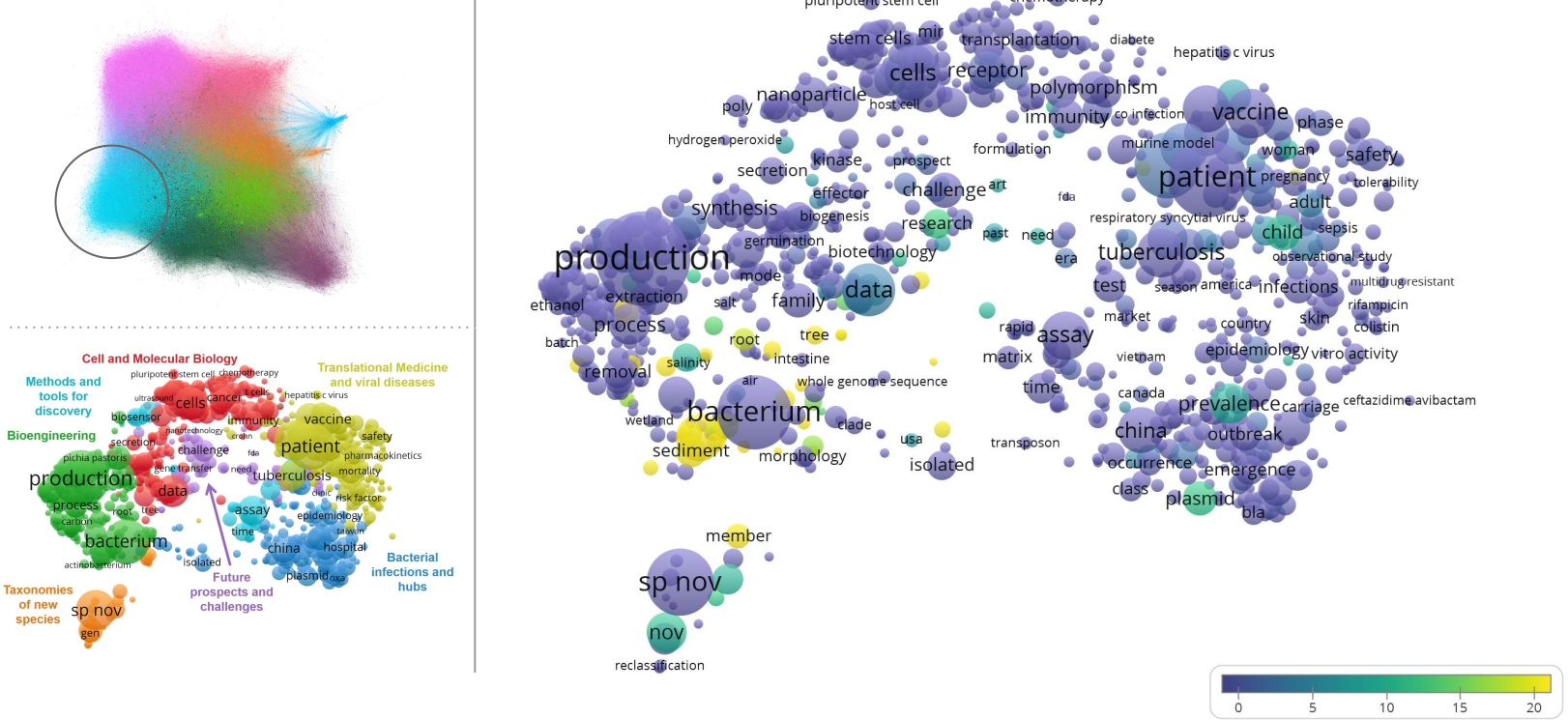
Results

Cluster	Twitter actors (tweets)	News story actors (mentions)	Policy document actors (mentions)	Papers (mentions received)	Total internal mentions	Total actors (mentions)
0	**2836 (131707)**	5 (139)	0 (0)	5205 (138728)	89768	8046 (131846)
1	**2794 (113782)**	1 (87)	0 (0)	4934 (109117)	81277	7729 (113869)
9	**1918 (93533)**	1 (16)	0 (0)	3802 (94906)	53617	5721 (93549)
8	**1953 (69608)**	5 (57)	1 (22)	3439 (64152)	36904	5398 (69687)
11	759 (13113)	**675 (45292)**	**6 (405)**	2291 (61919)	32870	3731 (58810)
2	**2366 (58248)**	19 (490)	4 (38)	2276 (58235)	34821	4665 (58776)
5	1367 (45124)	11 (827)	**8 (546)**	2343 (45516)	33143	3729 (46497)
7	**548 (13964)**	4 (225)	0 (0)	632 (13696)	6786	1184 (14189)
14	14 (120)	**102 (8202)**	0 (0)	179 (9567)	6457	295 (8322)
6	**115 (1984)**	2 (33)	0 (0)	112 (1522)	843	229 (2017)
13	**38 (1016)**	0 (0)	0 (0)	72 (813)	531	110 (1016)
12	**4 (138)**	0 (0)	0 (0)	10 (240)	13	14 (138)
3	**5 (118)**	0 (0)	0 (0)	25 (424)	30	30 (118)
4	**2 (6)**	0 (0)	0 (0)	1 (4)	2	3 (6)
10	**1 (1)**	0 (0)	0 (0)	1 (2)	1	2 (1)
Total	*14720 (542462)*	*825 (55368)*	*19 (1011)*	*25322 (598841)*	*377063*	*40886 (598841)*

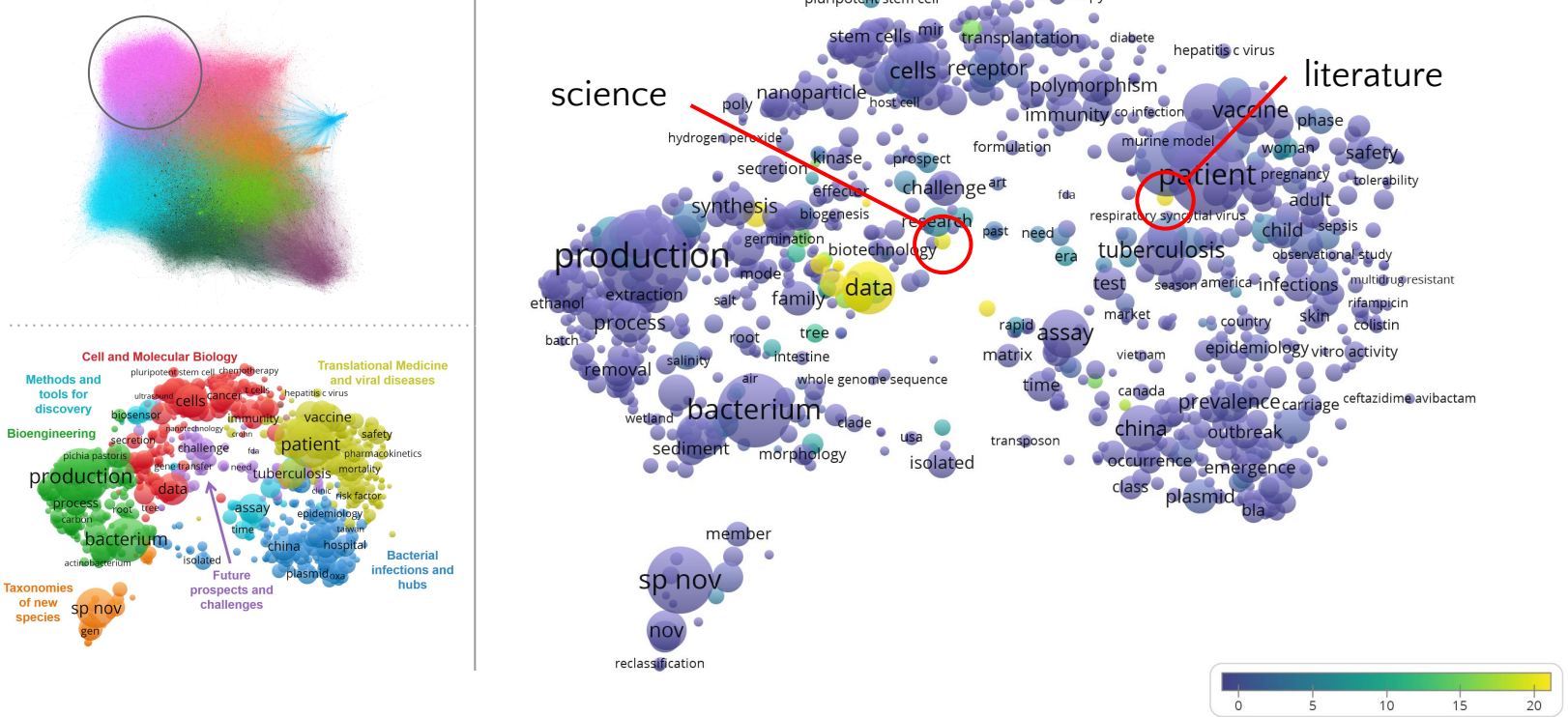
Descriptive of the 15 communities detected and the documents mentioned



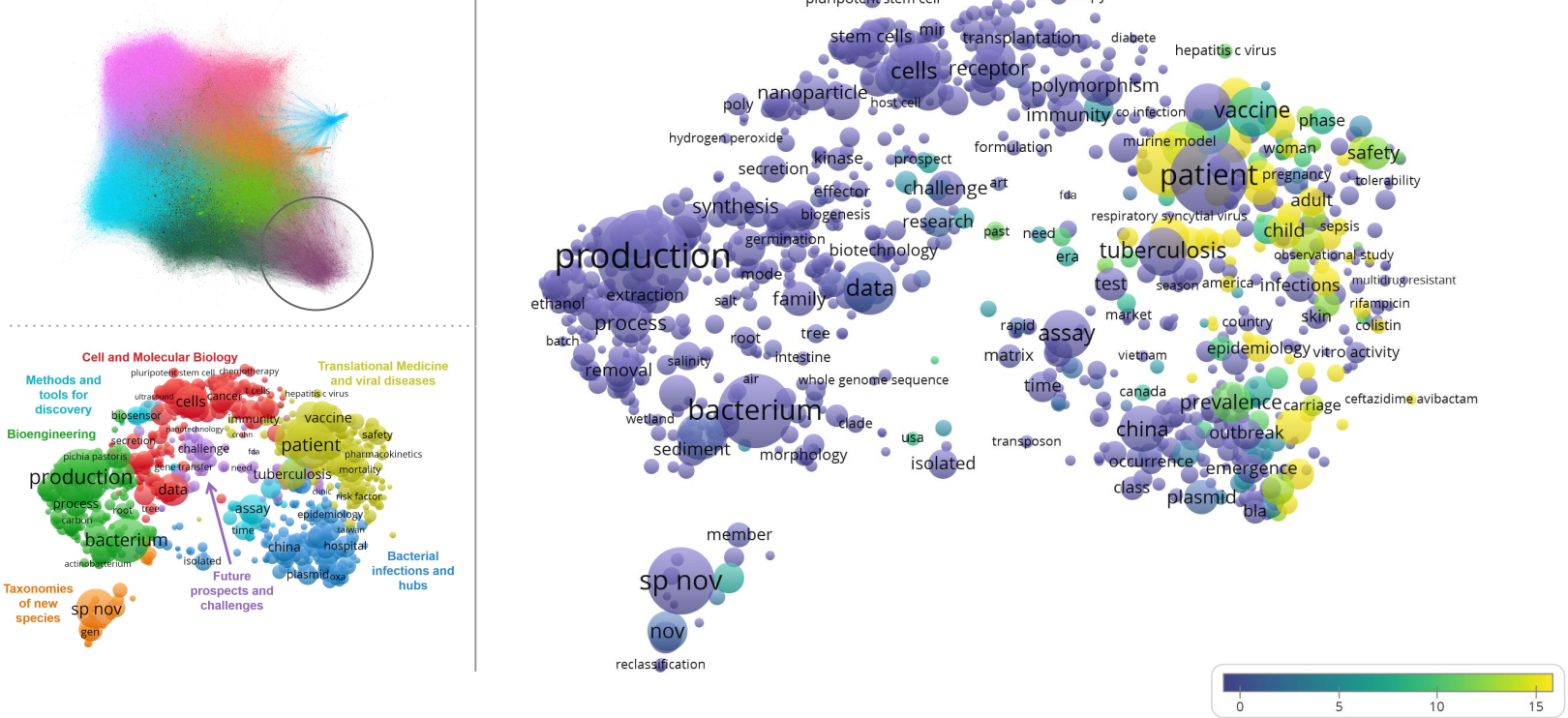
Most relevant terms from titles of all microbial publications indexed in Altmetric.com



Overlay terms map for cluster 0



Overlay terms map for cluster 1



Overlay terms map for cluster 5

Cluster structure	35.3% - Twitter accounts 0.1% - News media 64.7% - Publications	**CHARACTERISTICS AT THE CLUSTER LEVEL**
Coverage by actor for the complete set of pubs	Twitter - 30.6% pubs News media - 0.11% pubs	
Most discussed topic from term map	Bioengineering	
Most discussed terms from term map	Tree, taxonomy and human gut	

Publication coverage of social media actors	**19.56% of the publications of the entire network are mentioned**	**CHARACTERISTICS BASED ON THE TOP 25 SOCIAL MEDIA ACTORS**
Types of Twitter accounts	13 bots, **9 academics**, 2 scientific communities and 1 journal	
Types of news media	3 specialized media, 1 university press and 1 scientific association	
Types of policy organizations	-	
Most discussed topic from term map	General interest in the field of microbiology	
Most influential actors	 @jcamthrash with 5441 mentions	

Highlights of cluster 0

Cluster structure	36.2% - Twitter accounts 0.0% - News media 63.8% - Publications	**CHARACTERISTICS AT THE CLUSTER LEVEL**
Coverage by actor for the complete set of pubs	Twitter - 29.3% pubs News media - 0.1% pubs	
Most discussed topic from term map	Cell and molecular biology	
Most discussed terms from term map	Human genome, literature and genome assembly	

Publication coverage of social media actors	17.59% of the publications of the entire network are mentioned	**CHARACTERISTICS BASED ON THE TOP 25 SOCIAL MEDIA ACTORS**
Types of Twitter accounts	**14 bots**, 7 academics, 2 professionals, 1 company and 1 journal	
Types of news media	1 journal press	
Types of policy organizations	-	
Most discussed topic from term map	Mostly mentioning publications in the field of bioinformatics	
Most influential actors	 @yeast_papers with 6147 mentions	

Highlights of cluster 1

Cluster structure	36.7% - Twitter accounts 0.3% - News media 0.2% - Policy briefs 62.8% - Publications	CHARACTERISTICS AT THE CLUSTER LEVEL
Coverage by actor for the complete set of pubs	Twitter - 16.7% pubs News media - 1.8% pubs Policy briefs - 1.4% pubs	
Most discussed topic from term map	Translational medicine and viral diseases and bacterial infections and hubs	
Most discussed terms from term map	Infectious diseases society, america and update	

Publication coverage of social media actors	12.46% of the publications of the entire network are mentioned	CHARACTERISTICS BASED ON THE TOP 25 SOCIAL MEDIA ACTORS
Types of Twitter accounts	8 bots, 10 professionals, 2 academics, 2 journals, 1 hospital, 1 society and 1 specialized website	
Types of news media	**7 specialized media**, 2 journals and 1 research centre	
Types of policy organizations	**3 government organizations (national), 5 health organizations (4 national and 1 global)**	
Most discussed topic from term map	Actors are related to hospitals and clinical medicine. Focused on viral diseases and bacterial infections	
Most influential actors	 @AntibioticResis with 6496 mentions	

Highlights of cluster 5

4 Further research

How to improve it

Further research

- Expand the methodology to other research fields
- Include additional altmetric data sources
- Further analyses including novel methods and approaches suggested elsewhere

Leydesdorff, L., & Nerghes, A. (2017). Co-word maps and topic modeling: A comparison using small and medium-sized corpora (N< 1,000). *Journal of the Association for Information Science and Technology*, *68*(4), 1024-1035.

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports*, *9*.

Thanks!

Any questions?

You can find me at

- @Wences91
- wences@ugr.es

Presentation available at https://doi.org/10.5281/zenodo.3384573