

European Summer School | September 21, 2017

Bibliometric Solutions for Identifying Potential Collaborators

Nicolas Robinson-Garcia
Georgia Institute of Technology



Daniel Torres-Salinas
EC3Metrics SL
Universidad de Granada y Navarra



[DISCLAIMER]

Bibliometric solutions
inform decisions but do
not *take* decisions

Why seek for collaborators?

Partners for EU projects

Infrastructure

Hiring staff

Writing a paper

Applying novel methods

Who searches for collaborators?

Partners for EU projects

Infrastructure

Hiring staff

Writing a paper

Applying novel methods

**Institutional
level**

**Research
group /
Department**



Data sources and methods

Data sources

Citation indexes and databases



Social media and digital profiles

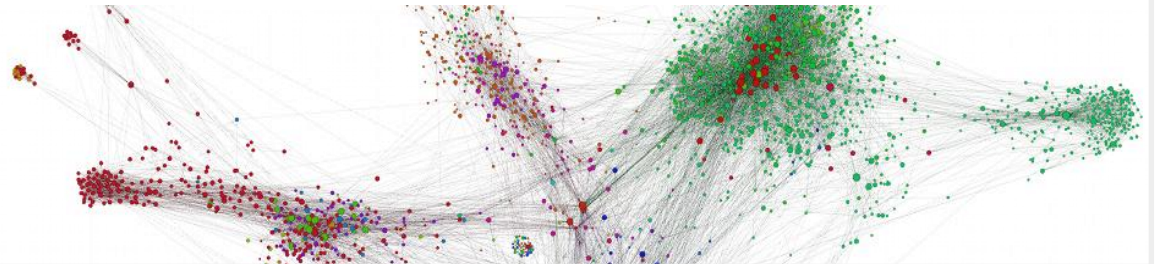


Methods

Refined search queries

```
-----  
ComputerName      SerialNumber      Manufacturer  
-----  
SCCMD1            8692-9122-6370-8325-5256-7599-01      American Megatrends Inc.  
SCCMD1            2791-6406-6583-7238-6564-4758-59      American Megatrends Inc.  
ADD1              6663-2506-1701-2674-1458-7082-52      Microsoft Corporation  
SCC001            2931-2178-6430-9153-0875-4370-14      American Megatrends Inc.  
-----  
PS C:\tmp> .\GetSerialNumberInfo.ps1 -SerialNumber 8692-9122-6370-8325-5256-7599-01  
1  
-----  
ComputerName      SerialNumber      Manufacturer  
-----  
SCCMD1            8692-9122-6370-8325-5256-7599-01      American Megatrends Inc.
```

Visualization techniques



Outline

Sources and digital profiles

Bibliometric approaches

Visualization techniques

An example – InCites

Sources and digital profiles

Academic networks and digital profiles

- ResearchGate
- Google Scholar profiles
- Specialized databases

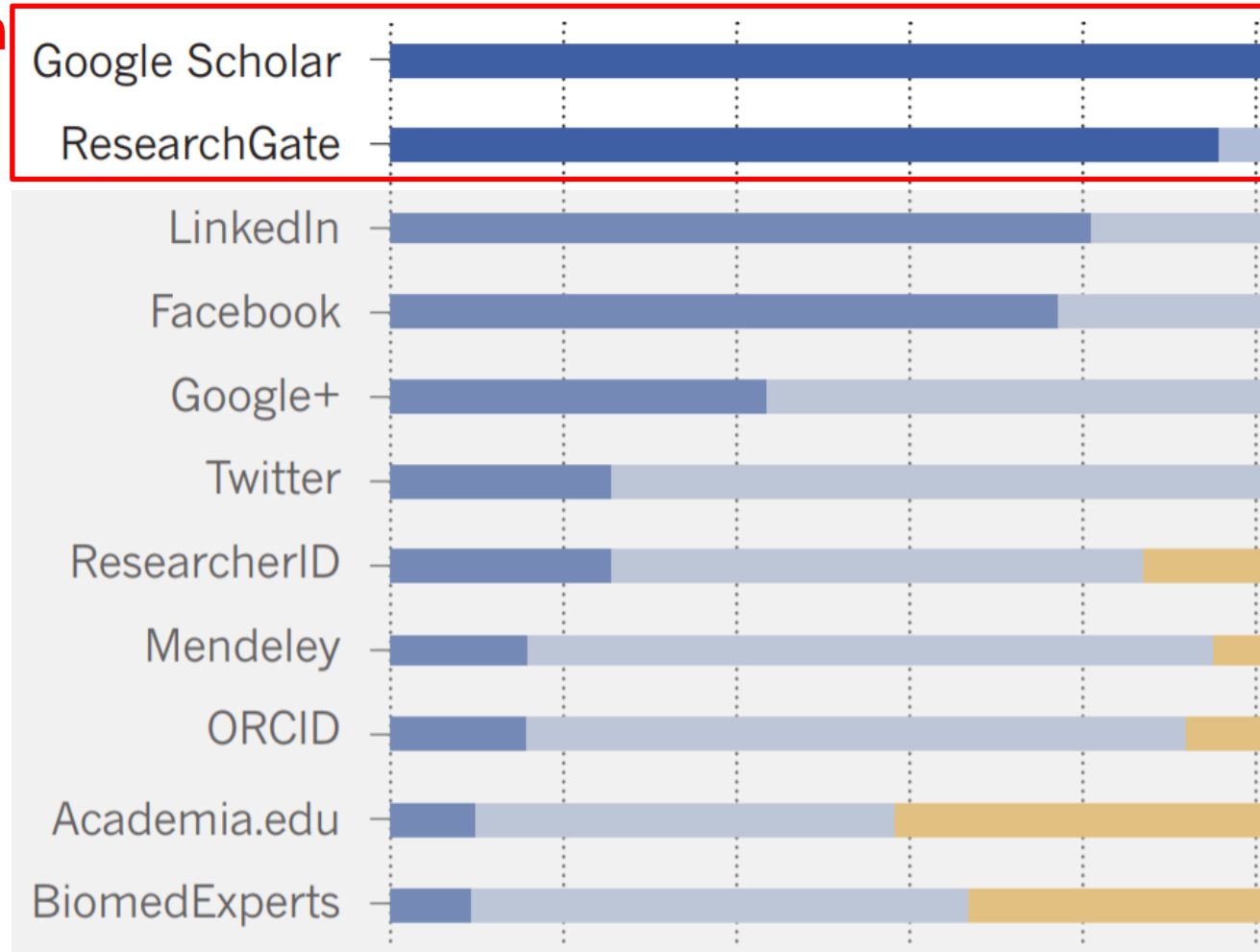
Academic Networks VS Digital Profiles

ACADEMIC NETWORKS. A digital platform where scientists share their work (ResearchGate, Academia.edu) or data (figshare).

Important: contact info, aggregation networks, personal interactions

DIGITAL PROFILES. Online scientific cv with complementary information, (affiliation, research lines...). No interactions between users.

Important: bibliometric indicators



¿Why use social networks or bibliometric profiles?

● Advantages

1. No subscription
2. Access to papers
3. Useful at individual level
4. Good for a quick identification
5. Linking options and directory
6. Scientific collaborators
7. Complementary indicators
8. Diverse profiles (practitioners, librarians,)
9. Not only the main stream science

Disadvantages

1. Accuracy of information no good
2. Need verification
3. Limited search bot
4. Country / discipline bias

“WE HAVE TO BUILD BETTER FILTER SYSTEMS TO EXPLAIN WHAT RESEARCH YOU CAN TRUST.”

Social Networks: ResearchGate

Searching a collaborator
in “bibliometrics”

Barbara Żogała-Siudem

·iI 2.35 · Instytut Badań Systemowych Polskiej Akademii Nauk

Department

Department of Stochastic Methods

[Follow](#) [Message](#)

1 Publication · 1 Citation



Katherine Mccain

·iI 31.76 · Drexel University

Department

College of Computing and Informatics

[Follow](#) [Message](#)

95 Publications · 3745 Citations



Katherine Mccain ·iI 31.76

PhD

Professor

Drexel University , Philadelphia · College of Co...

[Overview](#)

[Contributions](#)

[Info](#)

[Scores](#)

[Research Interests](#)

Research

Sorted by: **Newest**

Projects

Search by publication title or keyword

Research · 95

Article · 88

Chapter · 2

Conference Paper · 5

Full-texts · 12

Mining full-text journal articles to assess obliteration by incorporation: Herbert A. Simon's concepts of bounded rationality and satisficing in economics, management,...

Article

May 2015 · Journal of the Association for Information Science an...

Katherine Mccain

[Request full-text](#)

[Recommend](#)

[Follow](#)

Questions

Answers

Followers · 122

Citations · 3751

Open Reviews

Quick data about impact

Social Networks: ResearchGate

Recruit scientists and researchers

ResearchGate

Jobs

Search jobs

Advanced search

Choose countries

Information Science

- Abnormal Psychology
- Abrahamic Religions
- Accelerator Physics
- Accounting Scholarship
- Acoustic Engineering
- Acoustics
- Addiction Medicine
- Adult Education
- Aerobiology
- Aeronautical Engineering

Jobs / Job post



Lead Cybersecurity Architect (m/w)

Posted on 20 Sep 2017

Carl Zeiss AG

Germany, Oberkochen

5 Job openings at this institution

Go to application page
30 days left to apply

1 of your skills matches this job. [Update skills](#)

Zur Verstärkung unseres Teams im Bereich Konzernfunktion Forschung und Technologie am Standort Oberkochen oder Jena suchen wir ab sofort einen

Lead Cybersecurity Architect (m/w)

Keep up to date with the latest jobs in your field

I'd like to receive jobs via email once a week

Share this job with someone you know



Recommend this job to colleagues, researchers from your network, or someone you know.

Share this job

Similar jobs

[View more](#)

Scientific Evaluation Lead

Galapagos NV
Mechelen, Flanders, Belgium



Bibliometric profiles: Google Scholar



Juan Gorraiz

University of Vienna

Bibliometrics, Library & Information Sciences, Altmetrics, Open Access, Usage Metrics

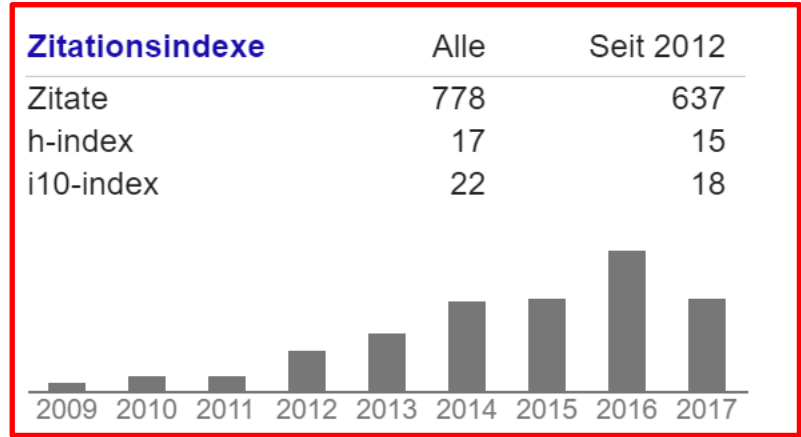
Bestätigte E-Mail-Adresse bei univie.ac.at

Folgen

Google Scholar

Eigenes Profil erstellen

Titel	1-20	Zitiert von	Jahr
A bibliometric analysis of pharmacology and pharmacy journals: Scopus versus Web of Science		66	2008
J Gorraiz, C Schloegl Journal of Information Science 34 (5), 715-725			
Comparison of citation and usage indicators: the case of oncology journals		60	2010
C Schloegl, J Gorraiz Scientometrics 82 (3), 567-580			
Bibliometric practices and activities at the University of Vienna		53	2012
C Gumpenberger, M Wieland, J Gorraiz Library Management 33 (3), 174-183			



- Koautoren** [Alle anzeigen...](#)
- Christian Gumpenberger
 - Christian Schloegl
 - Wolfgang Glänzel
 - Peter Kraker

Bibliometric profiles: Google Scholar

Scientist bibliometric rankings



Wolfgang Glänzel

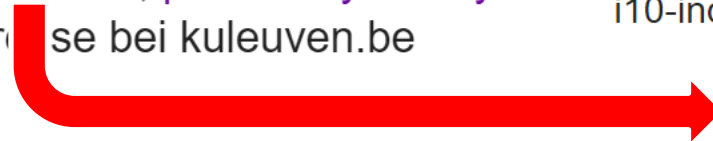
Professor, **KU Leuven**

scientometrics, bibliometrics, probability theory

Bestätigte E-Mail-Adresse bei kuleuven.be

Zitationsindexe


	Alle	Seit 2012
Zitate	14795	7540
h-index	64	47
i10-index	185	143




Research line level

University level


Katholieke Universiteit, Leuven [Weitere Informationen](#)




Jan Staessen
University of Leuven
Bestätigte E-Mail-Adresse bei med.kuleuven.be
Zitiert von: 91622
Hypertension and Cardiovascular ...




Bernd Nilius
KU Leuven
Bestätigte E-Mail-Adresse bei med.kuleuven.be
Zitiert von: 39529
life sciences




De Moor Bart
KU Leuven ESAT-Stadius
Bestätigte E-Mail-Adresse bei esat.kuleuven.be
Zitiert von: 89465
Numerical Linear Algebra System Theory Control Theory




Bart De Strooper
gewoon hoogleraar KU Leuven en onderzoeksdirecteur VIB
Bestätigte E-Mail-Adresse bei cme.vib.kuleuven.be
Zitiert von: 38844
Alzheimer Parkinson mitochondria microRNA secreta




Luc Van Gool
professor computer vision @ ETH Zurich ee.ethz.ch & @ KU Leuven
Bestätigte E-Mail-Adresse bei esat.kuleuven.be
Zitiert von: 78367
computer vision computer graphics machine learning auto



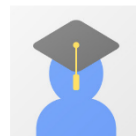
Catherine Verfaillie
KU Leuven
Bestätigte E-Mail-Adresse bei med.kuleuven.be
Zitiert von: 34619
stem cell biology



Joos Vandewalle
Professor at Electrical Engineering Dept (ESAT-SCD) KU Leuven
Bestätigte E-Mail-Adresse bei esat.kuleuven.be
Zitiert von: 41097
mathematical engineering signal processing system theory



Jean Poesen
Professor of Geography, KU Leuven, Belgium
Bestätigte E-Mail-Adresse bei ees.kuleuven.be
Zitiert von: 34200
geography geomorphology soil erosion sediment yield



Wolfgang Glänzel
Professor, KU Leuven
Bestätigte E-Mail-Adresse bei kuleuven.be
Zitiert von: 14795
scientometrics bibliometrics probability theory



Francis Narin
Retired President, CHI Research
Zitiert von: 14066
Science Policy Bibliometrics



David I. Stern
Professor, Crawford School of Public Policy, The Australian National University
Bestätigte E-Mail-Adresse bei anu.edu.au
Zitiert von: 13429
energy economics economic growth climate change bib



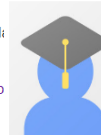
Chaomei Chen
Professor of Informatics, College of Computing and Informatics, Drexel University
Bestätigte E-Mail-Adresse bei drexel.edu
Zitiert von: 13277
Information visualization visual analytics scientometrics



Judit Bar-Ilan
Professor of Information Science, Bar-Ilan University, Israel
Bestätigte E-Mail-Adresse bei biu.ac.il
Zitiert von: 6175
Internet research informetrics information retrieval informa



Nicole Haeflner
DRE Inserm
Bestätigte E-Mail-Adresse bei inserm.fr
Zitiert von: 5744
Immunology - HIV Bibliometrics



Katherine W. McCain
Drexel University
Bestätigte E-Mail-Adresse bei drexel.edu
Zitiert von: 5690
Bibliometrics Scientometrics



Thed van Leeuwen
Leiden University
Bestätigte E-Mail-Adresse bei cwts.nl
Zitiert von: 5609
Bibliometrics science studies research evaluation

Specialized databases

PsycInfo, PubMed, GEOBASE, LISTA,... > Precise + Exhaustive

Results 1–10 of 25 for "Transgender" and "Keita, Gwendolyn Puryear"

Refine Your Search

Topic

Lesbian, gay, bisexual, transgender (9)

Health disparities (4)

HIV & AIDS (4)

Race (4)

Education (3)

11 more... [+]

Precise research profile

« Previous Next »

Relevance ▾

1. Letter of Support for the Student Non-Discrimination Act (SNDA)

Letter to members of congress to support the Student Non-Discrimination Act of 2011 (H.R. 528/S. 555).

2. Building on a successful year

The Public Interest Directorate applies the science and practice of psychology to the fundamental problems of human



Bibliometric approaches

- Defining potential collaborators
- Combining different definitions
- Finding the most suitable match

How to define a potential collaborator?

- A. An institution active in the same field of interest
- B. An institution active in the same field of interest and with high impact research
- C. An institution in the same field and could be inclined to collaborate
- D. Combining the three previous approaches

Option A. Active in the same field

UNIV GRANADA – 2012-2016 – FOOD SCI TECHNOL

TOP 10 COLLABORATORS

- 1 CSIC (Spain)
- 2 CNRS (France)
- 3 Autonomous Univ Barcelona
- 4 Univ Jaen
- 5 Universite Paris Saclay
- 6 Univ Barcelona
- 7 Ghent Univ
- 8 AGROPARISTECH
- 9 Univ California System
- 10 Univ Almeria

TOP 10 IN FOOD SCI TECHNOL

- 1 CSIC (Spain)
- 2 US Department of Agriculture
- 3 Inst Nat de la Recherche Agronomique
- 4 Jiangnan Univ
- 5 Wageningen Univ System
- 6 Chinese Academy of Sciences
- 7 China Agricultural Univ
- 8 CSIR (India)
- 9 CONICET (Argentina)
- 10 Univ California System

Option B. Active + High impact

FILTERING BY HIGHLY CITED PAPERS

TOP 10 COLLABORATORS

- 1 CSIC (Spain)
- 2 CNRS (France)
- 3 Autonomous Univ Barcelona
- 4 Univ Jaen
- 5 Universite Paris Saclay
- 6 Univ Barcelona
- 7 Ghent Univ
- 8 AGROPARISTECH
- 9 Univ California System
- 10 Univ Almeria

TOP 10 HCP IN FOOD SCI TECHNOL

- 1 Univ of Massachussetts System
- 2 CSIC (Spain)
- 3 Wageningen Univ System
- 4 TEAGASC
- 5 US Department of Agriculture
- 6 Inst Nat de la Recherche Agronomique
- 7 Nanchang Univ
- 8 South China Univ of Technol
- 9 Ghent Univ
- 10 King Abdulaziz Univ

Option C. Active + possibly interested

LOOKING AT THOSE CITING OUR WORK

TOP 10 COLLABORATORS

- 1 CSIC (Spain)
- 2 CNRS (France)
- 3 Autonomous Univ Barcelona
- 4 Univ Jaen
- 5 Universite Paris Saclay
- 6 Univ Barcelona
- 7 Ghent Univ
- 8 AGROPARISTECH
- 9 Univ California System
- 10 Univ Almeria

TOP 10 CITING INSTITUTIONS

- 1 CSIC (Spain)
- 2 Univ Bologna
- 3 Marche Polytechnic Univ
- 4 Univ Barcelona
- 5 Univ Rovira I Virgili
- 6 CONICET (Argentina)
- 7 Univ Almeria
- 8 Wageningen Univ System
- 9 Univ de Sfax
- 10 Autonomous Univ Madrid

Option D. Combining the different options

TRYING DIFFERENT COMBINATIONS TO FILTER

A+B

- 1 US Department of Agriculture
- 2 Inst Nat de la Recherche Agronomique
- 3 **Wageningen Univ System**

A+C

- 1 **Wageningen Univ System**
 - 2 CONICET (Argentina)
-

B+C

- 1 **Wageningen Univ System**

Visualization techniques

- Institutional profiles and rankings
- Research focus
- Thematic affinity within fields
 - Research groups
 - Institutional level

Visualization techniques

Institutional level

- Which are the missions of our institution and how they match other potential collaborating institutions?
- What is their research focus and how similar it is to ours?

Research group level

- How similar to a given group is the profile of potential collaborators?

Institutional profiles

- What is the orientation of your university?
- How does it differ from other potential collaborators?

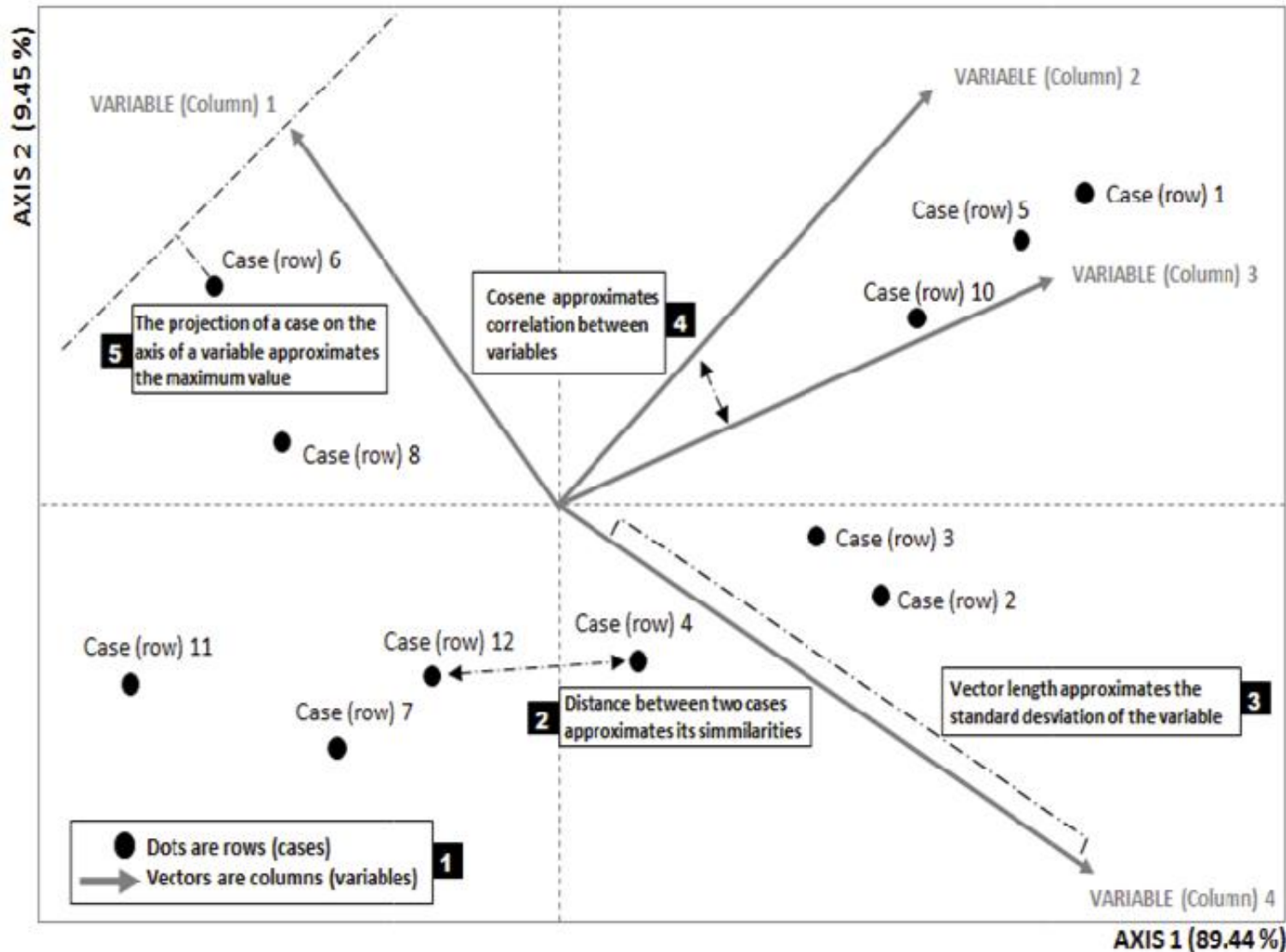
“[Rankings] do present a series of indicators, and institutions can be ranked by each of these separately.”

“A system should not merely present a series of separate rankings in parallel, but rather a dataset and tools to observe patterns in multi-faceted data.”

H.F. Moed, 2017

Institutional profiles

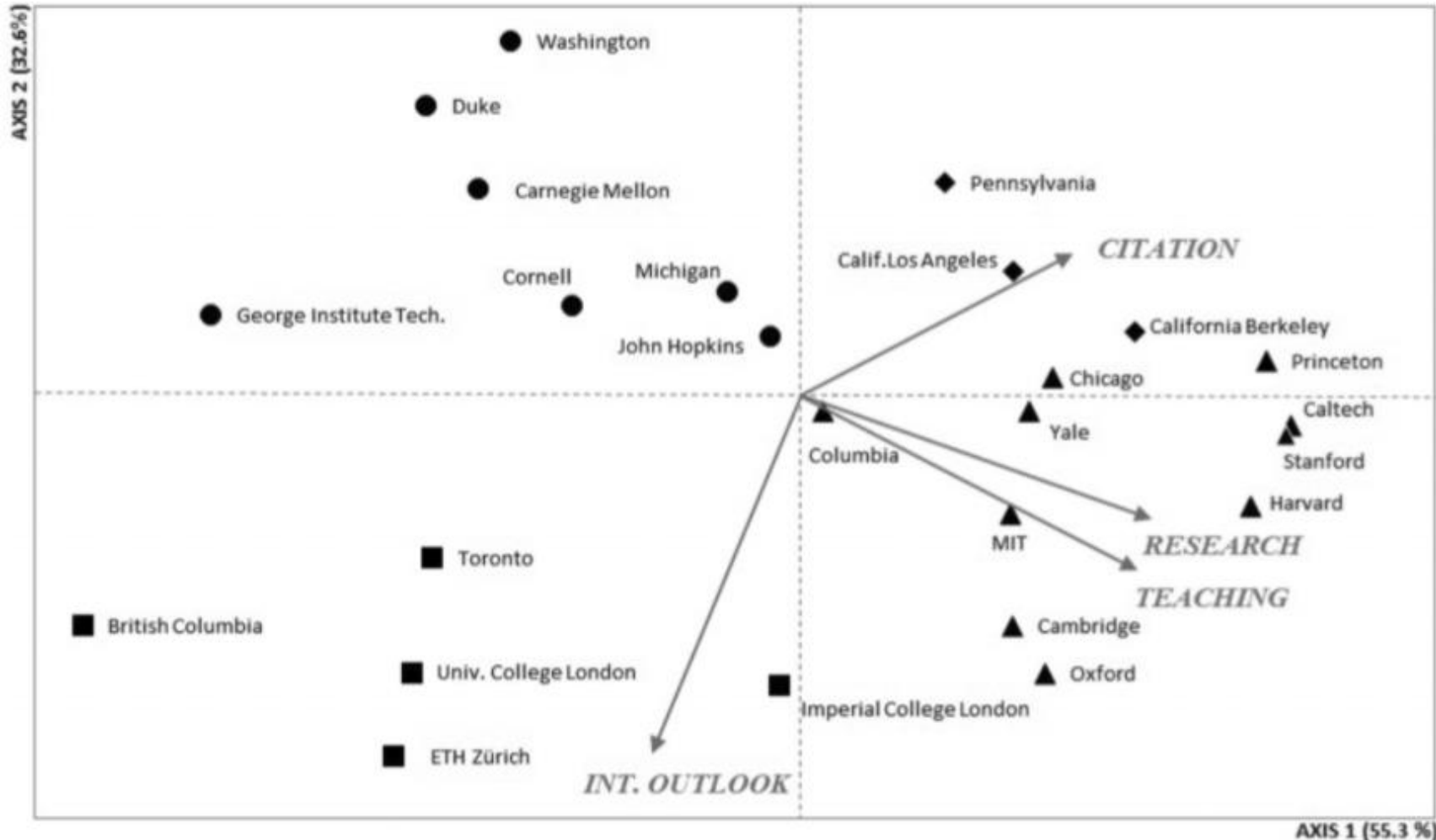
Biplot Analysis



- Visual representation of the relation between both cases and variables
- Arrows represent variables
- Dots represent cases

Institutional profiles

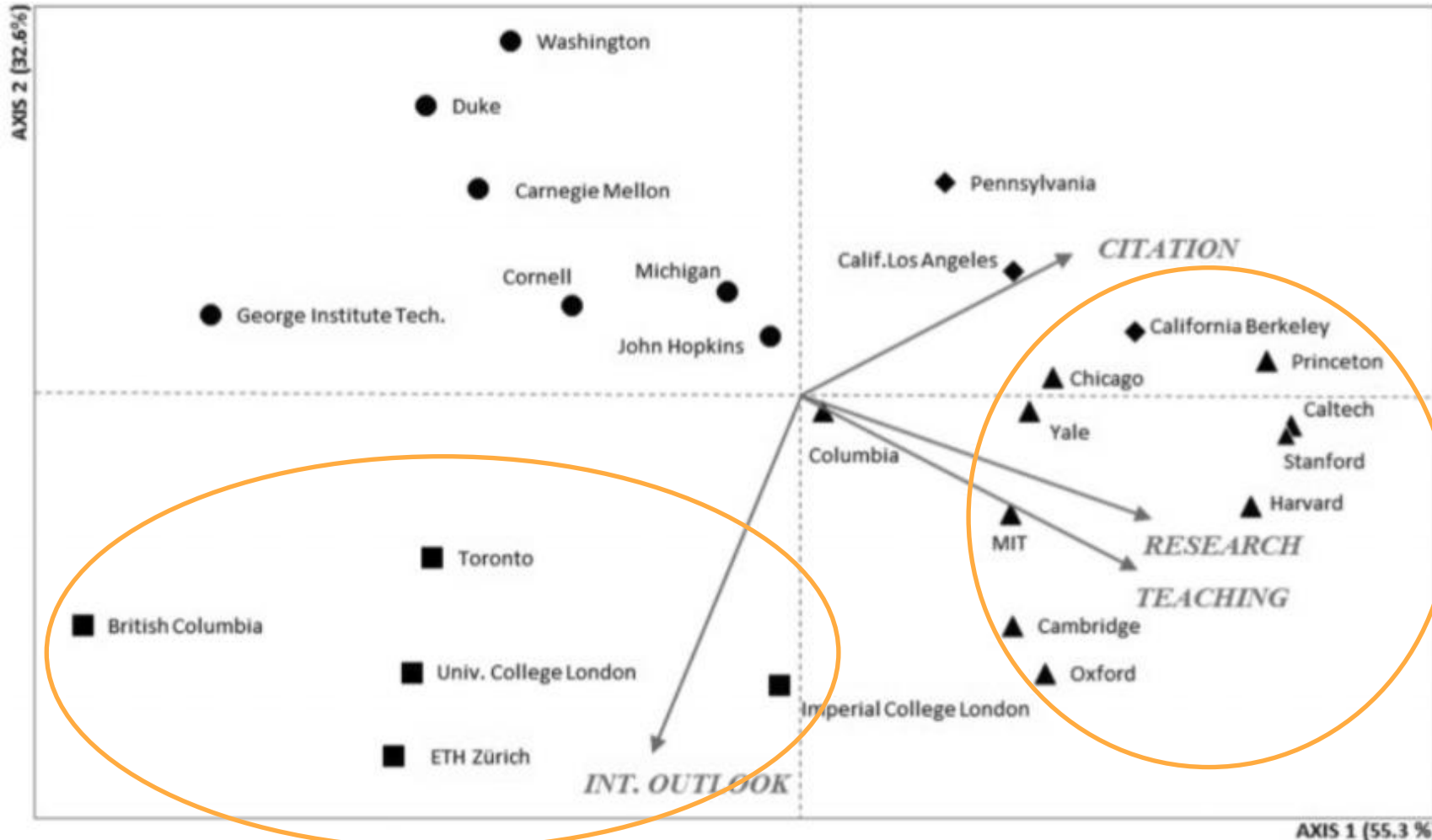
Biplot Analysis



- Top 25 univs. THE Ranking
- We learn about universities' profiles
- We also learn about the variables of the ranking itself

Institutional profiles

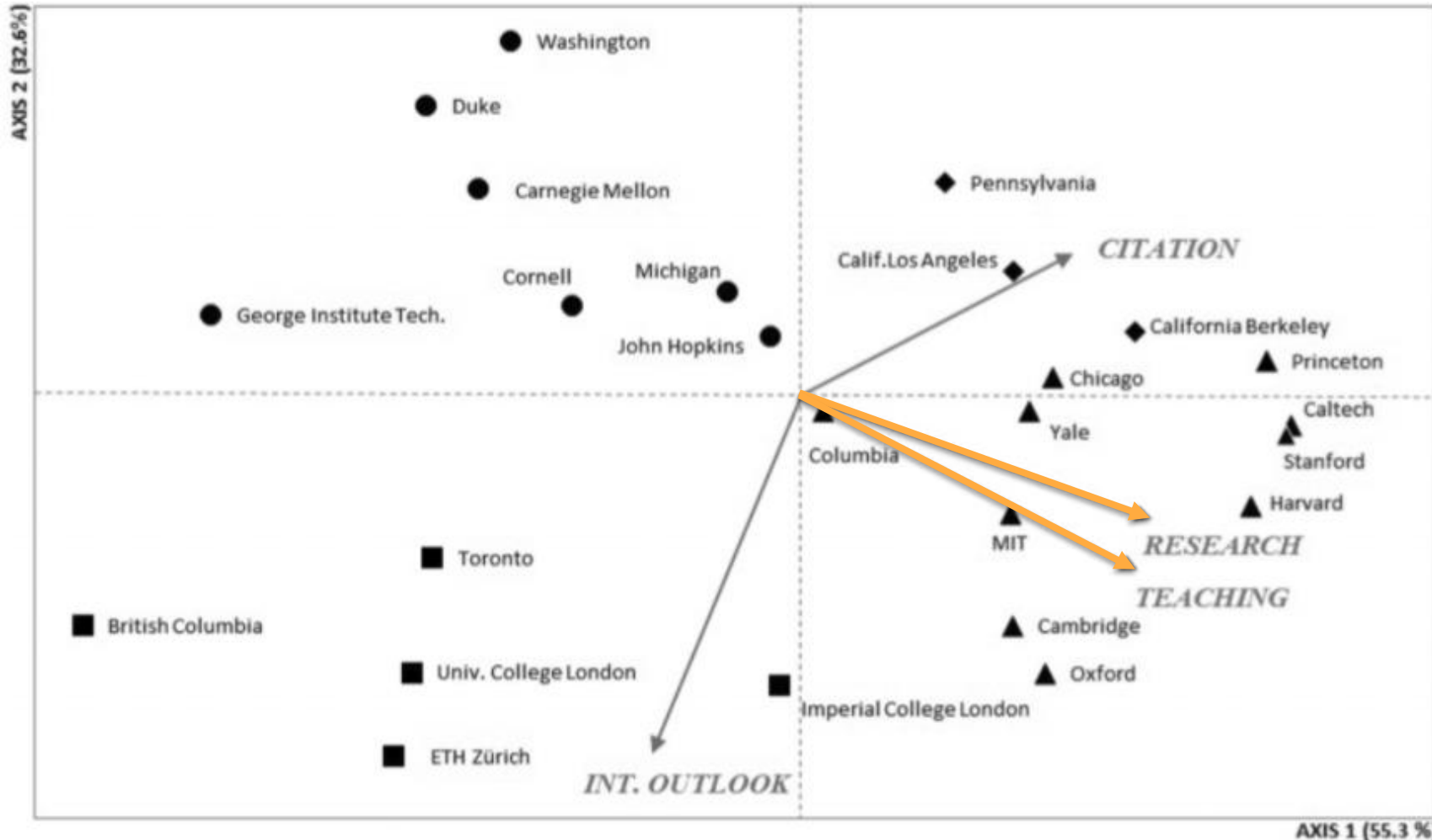
Biplot Analysis



- Top 25 univs. THE Ranking
- We learn about universities' profiles
- We also learn about the variables of the ranking itself

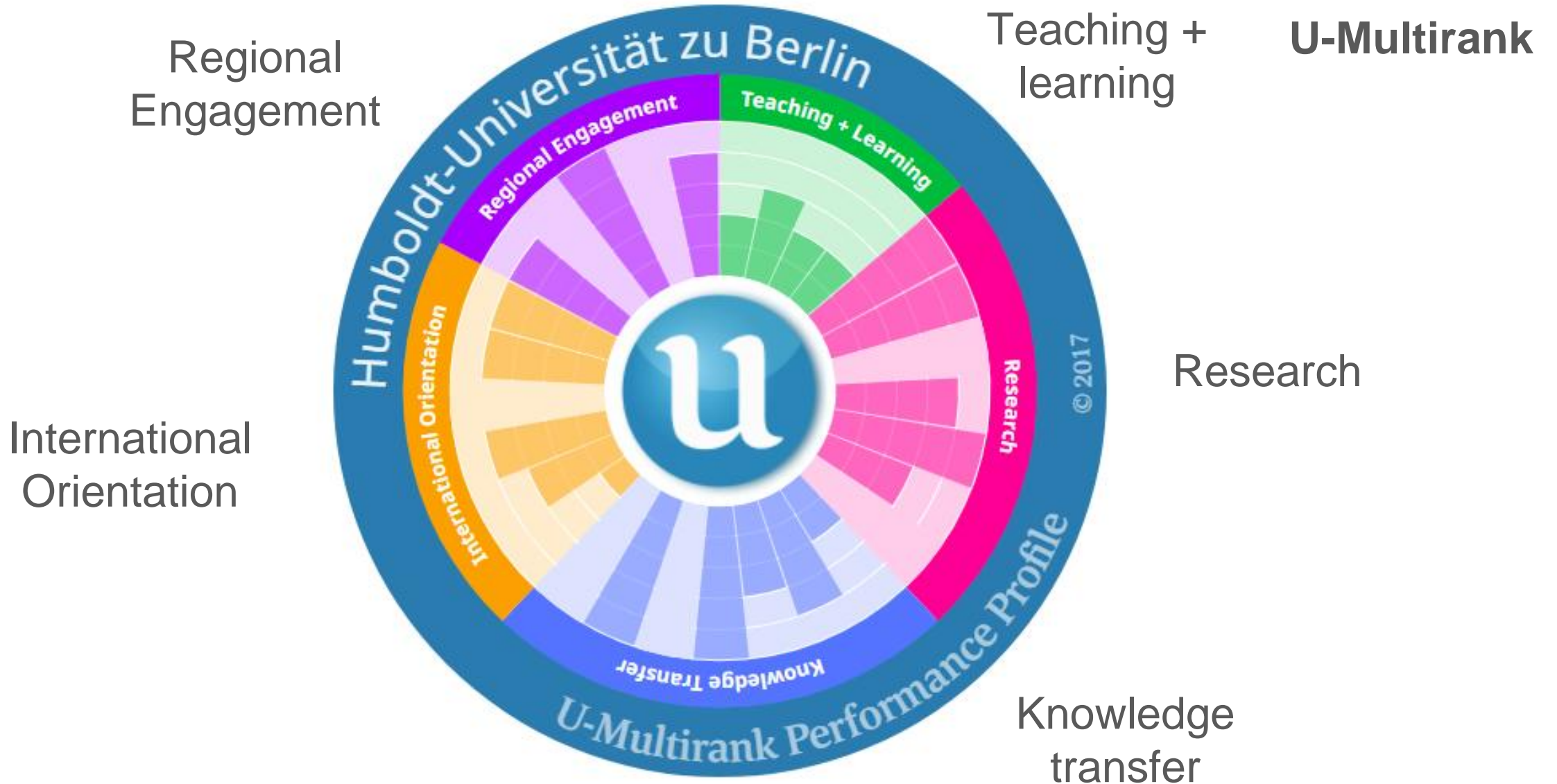
Institutional profiles

Biplot Analysis



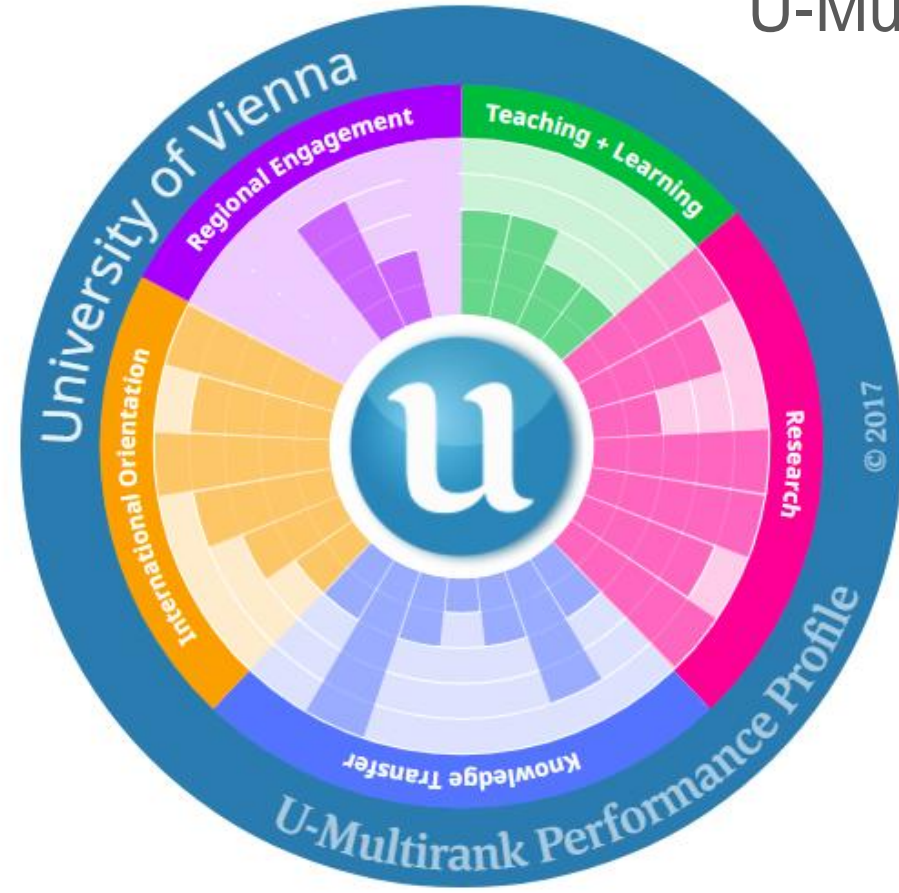
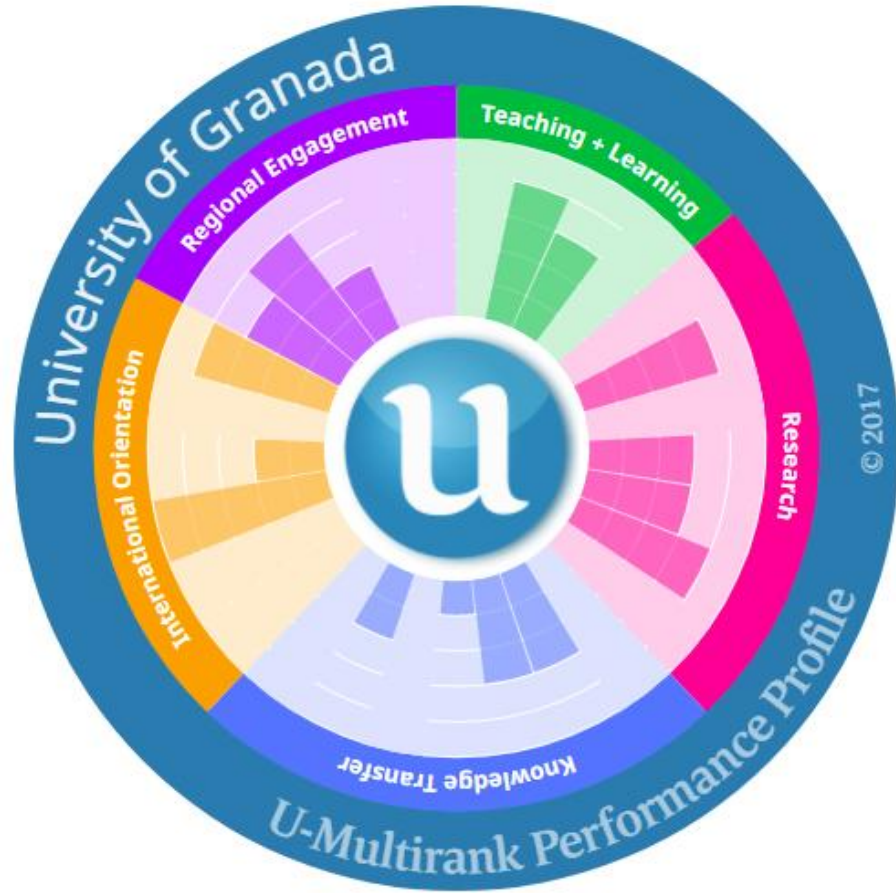
- Top 25 univs. THE Ranking
- We learn about universities' profiles
- We also learn about the variables of the ranking itself

Institutional profiles



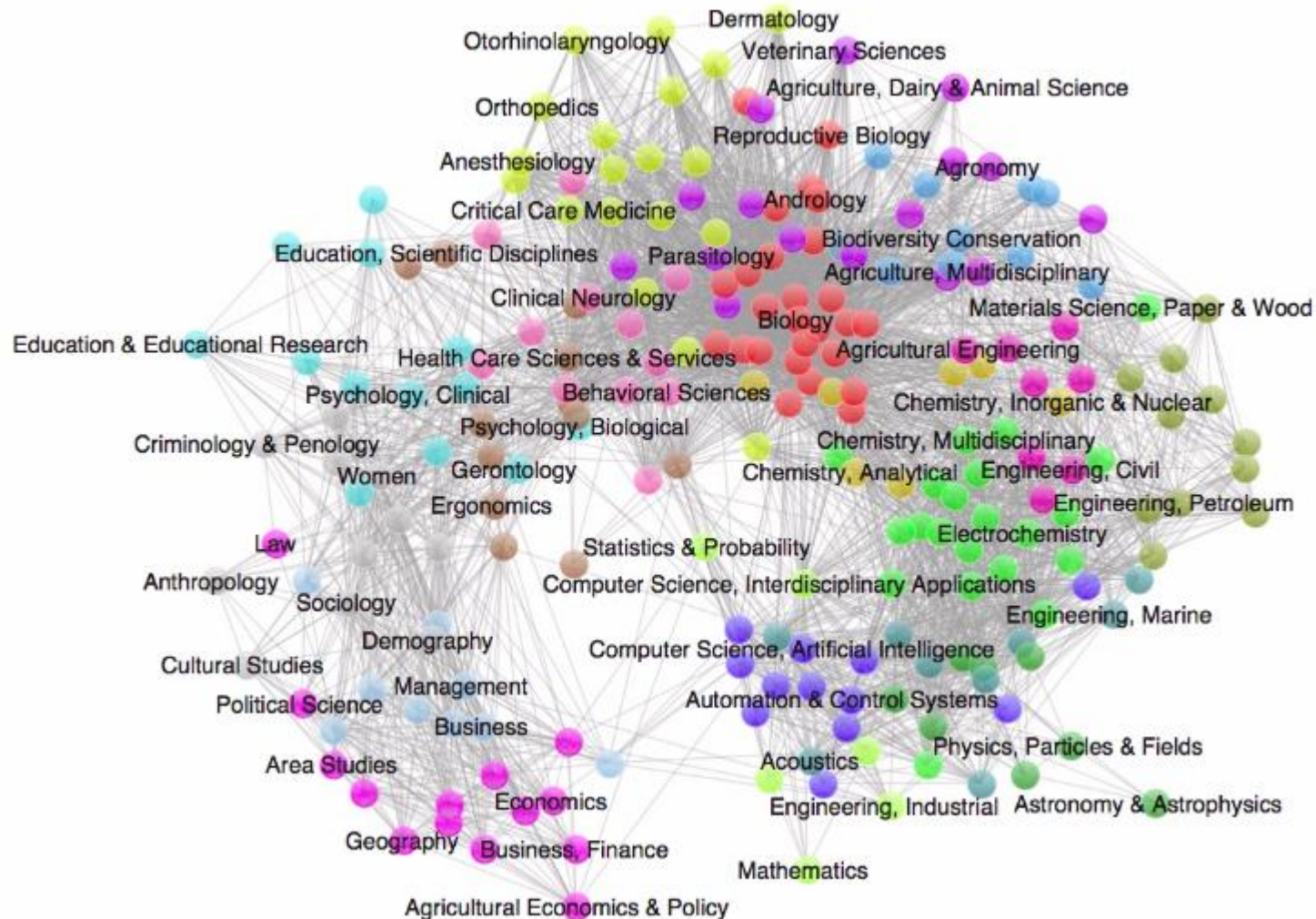
Institutional profiles

U-Multirank



Research focus

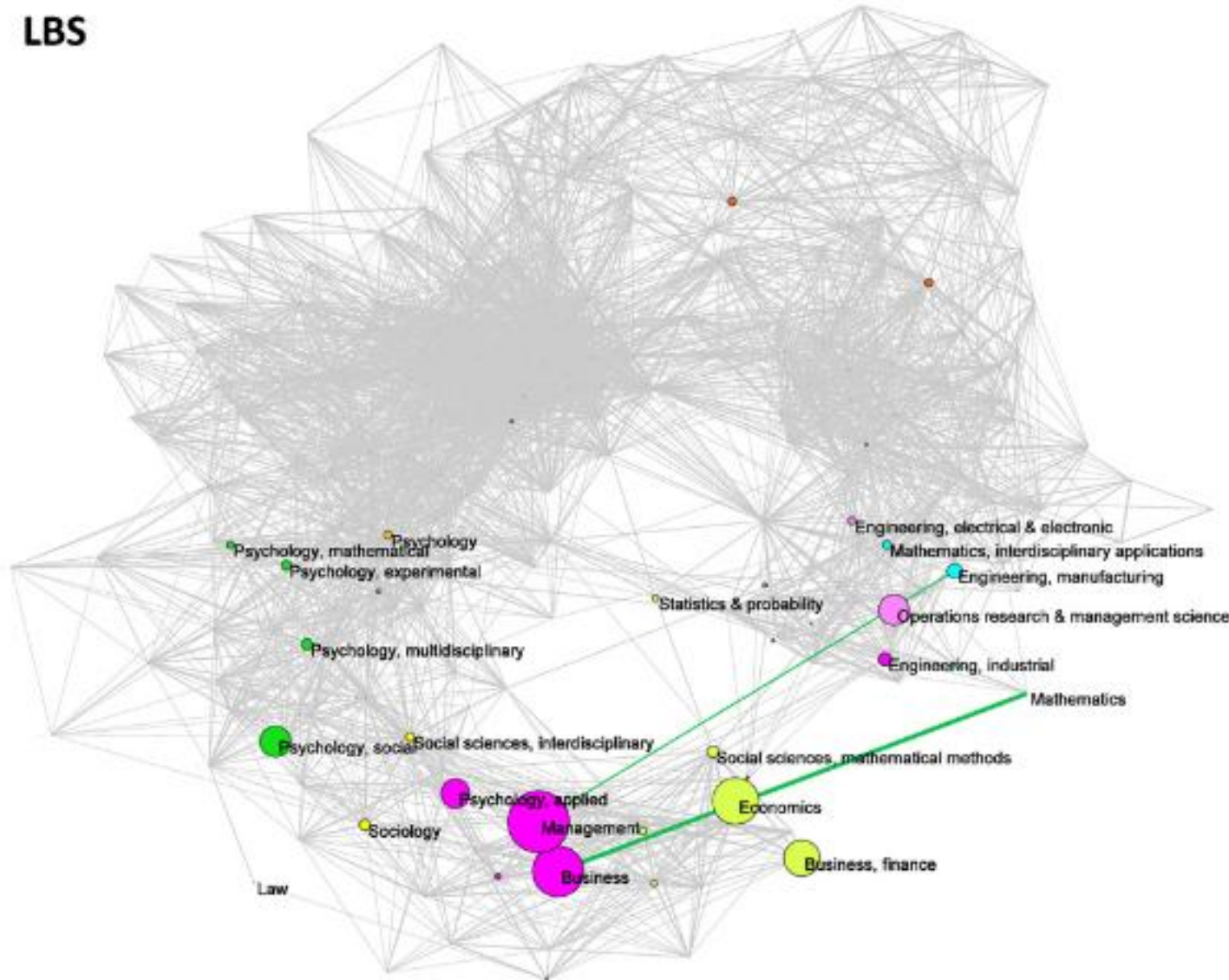
Overlay maps of science



- Basemap of WoS categories
- Overlay of the institutional research profile
- Comparisons between different profiles

Research focus

LBS



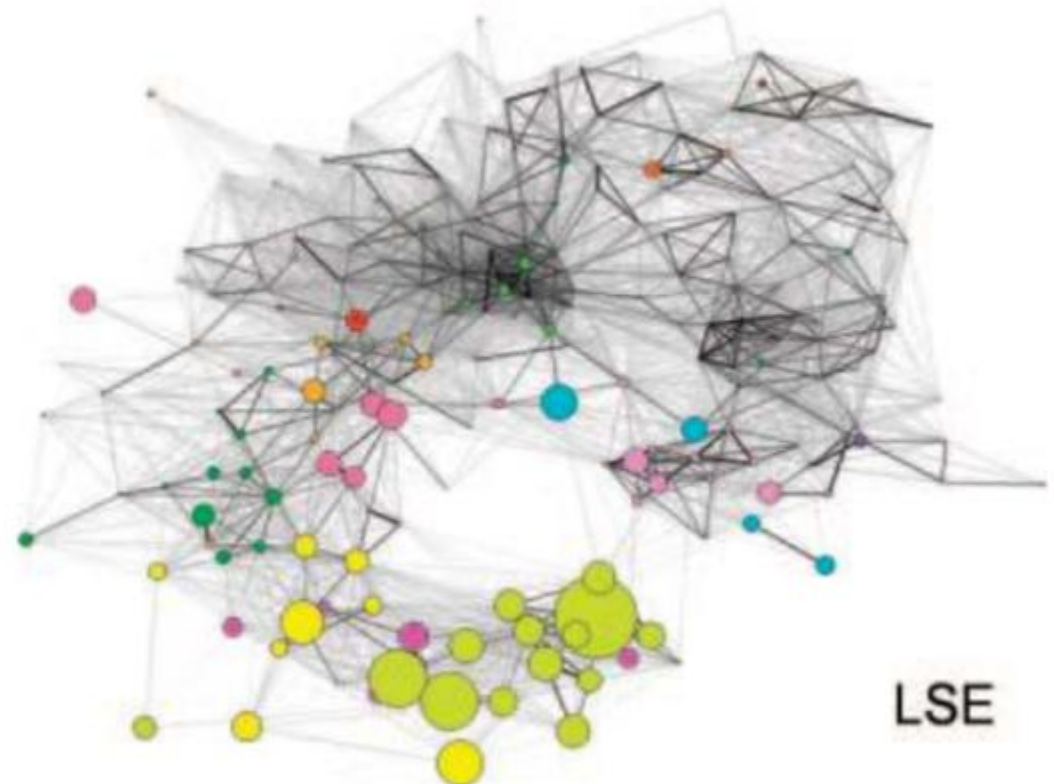
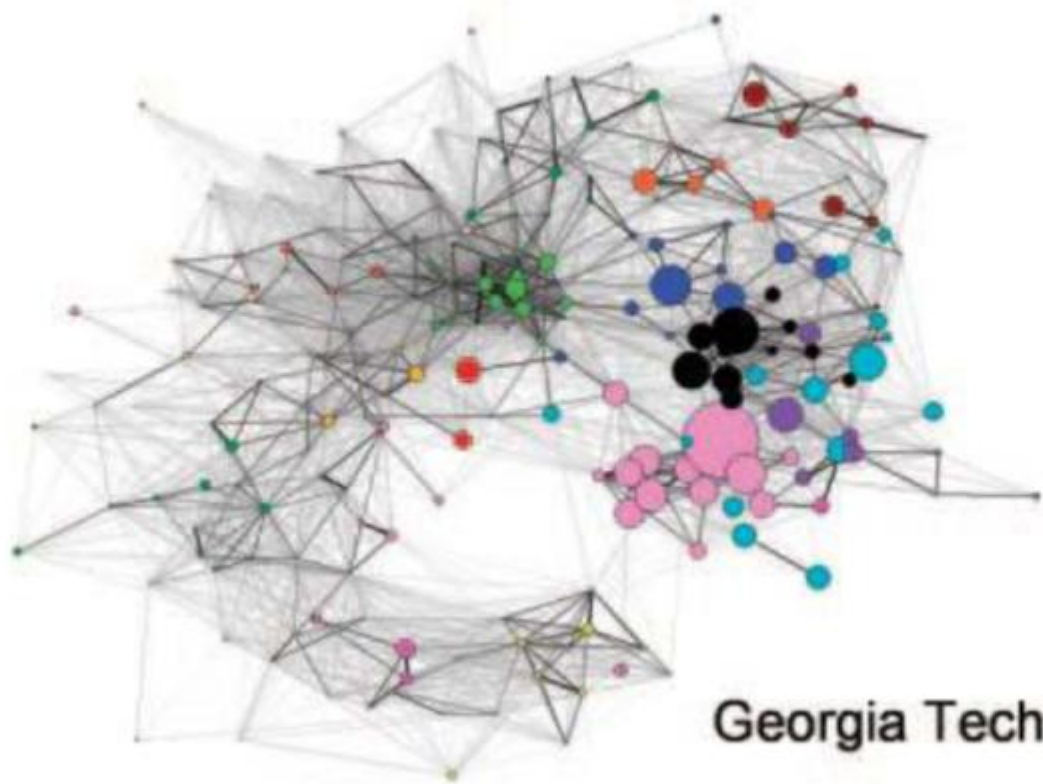
Overlay maps of science

- Basemap of WoS categories
- Overlay of the institutional research profile
- Comparisons between different profiles

Research focus

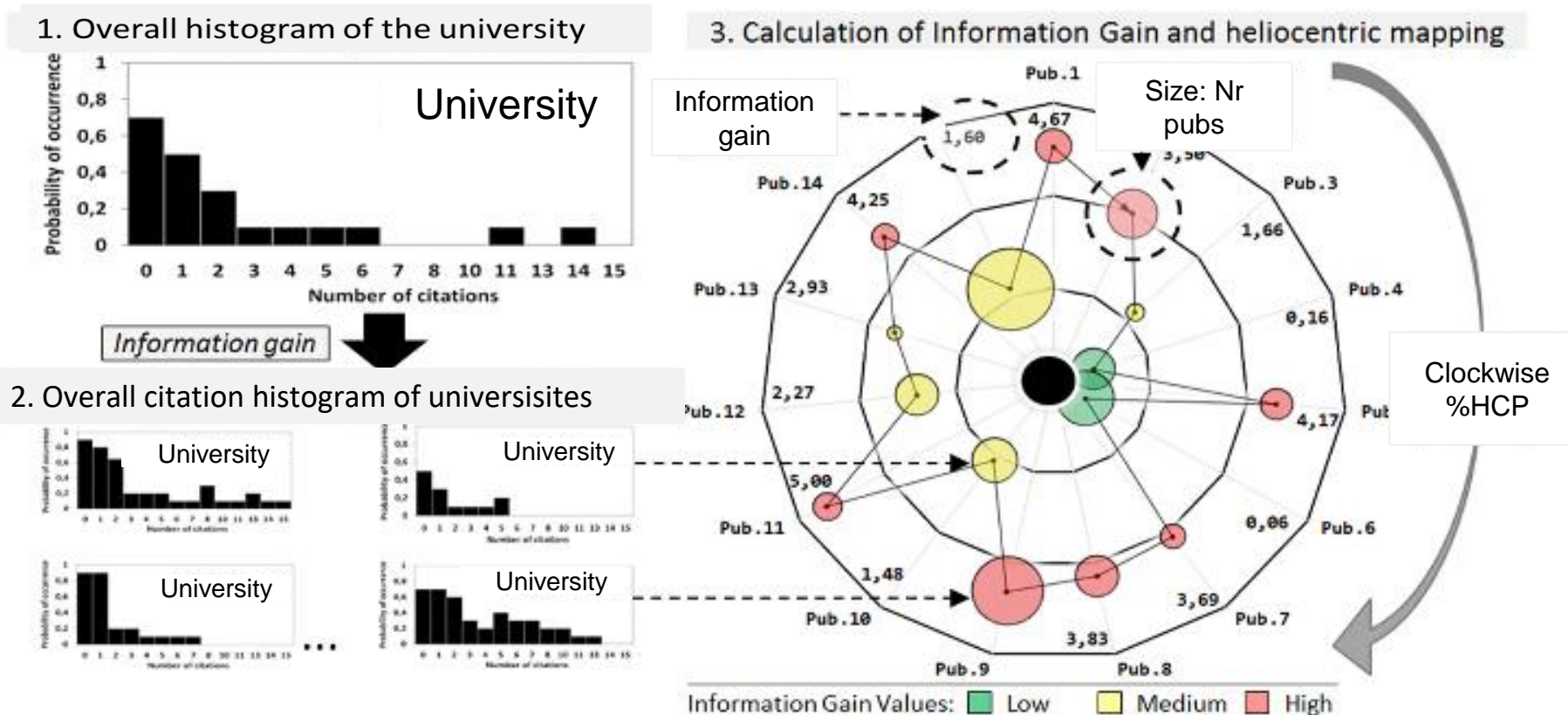
Overlay maps of science

- Comparisons between different profiles



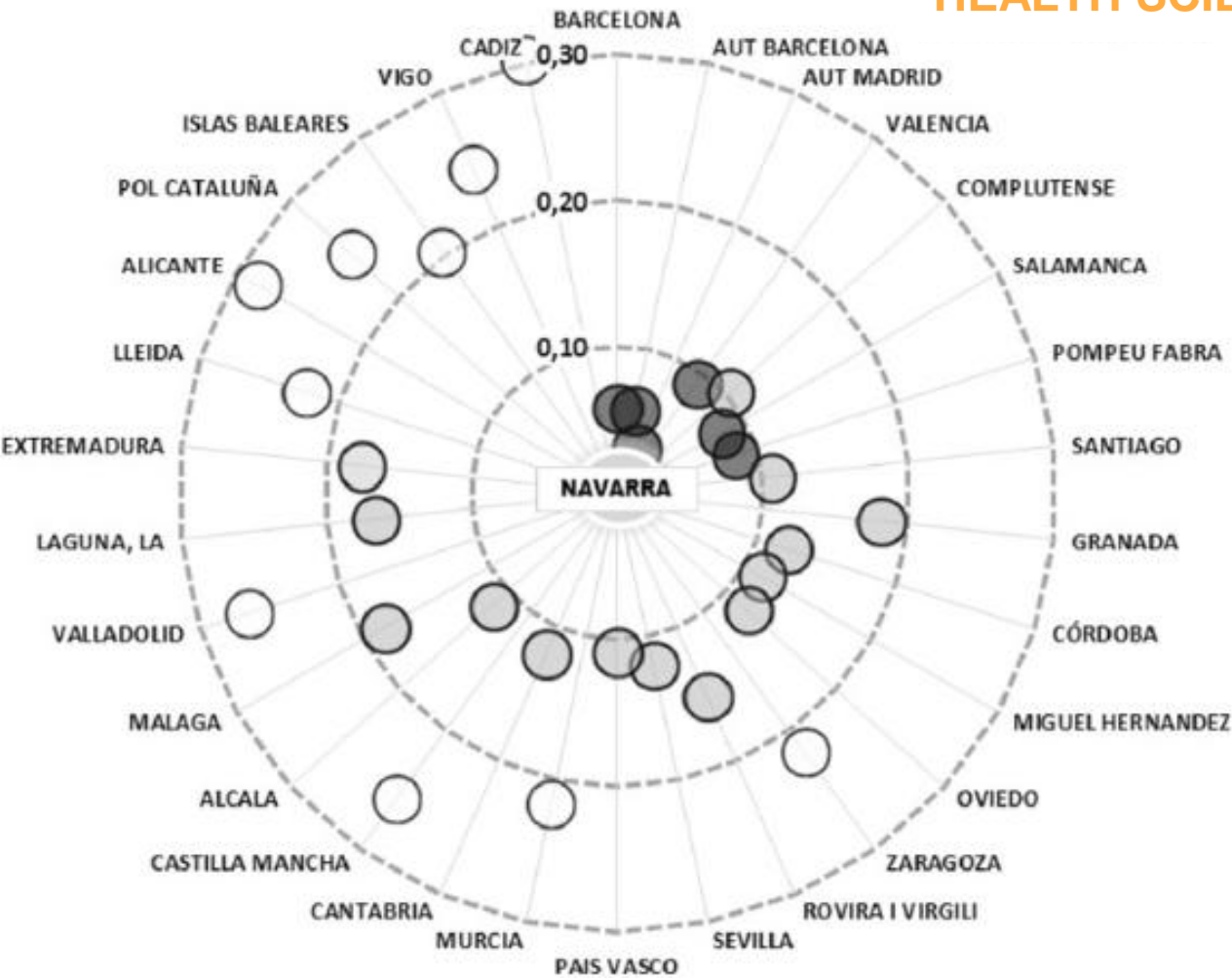
Thematic affinity within fields – Institutional level

- Information gain compares similarity of distribution
- Here we use citation distributions and benchmark our institution with others



Thematic affinity within fields – Institutional level

HEALTH SCIENCES



- Information gain is used to measure similarity between institutions
- The closer to the center the most similar to our institution
- Univs are organized clockwise according to their share of HCP

Thematic affinity within fields – Institutional level

SPANISH NATIONAL RANKING

INFORMATION & COMMUNICATION TECHNOLOGIES

1. University of Granada
2. Polytechnic University of Catalonia
3. University of Jaen
4. Polytechnic University of Valencia
5. ...

Another approach
looking at national
thematic rankings

Thematic affinity within fields – Institutional level

SPANISH NATIONAL RANKING

INFORMATION & COMMUNICATION TECHNOLOGIES

1. **University of Granada**
2. Polytechnic University of Catalonia
3. **University of Jaen**
4. Polytechnic University of Valencia
5. ...

- **Co-authorship analysis** shows that these two universities are close collaborators
- These two universities are competitors / potential collaborators

Thematic affinity within fields – Institutional level

SPANISH NATIONAL RANKING

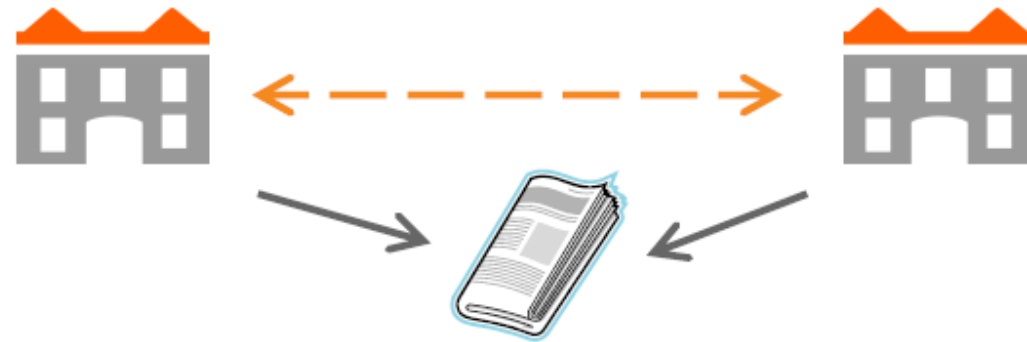
INFORMATION & COMMUNICATION TECHNOLOGIES

1. **University of Granada**
2. **Polytechnic University of Catalonia**
3. **University of Jaen**
4. **Polytechnic University of Valencia**
5. ...

- **Co-authorship analysis** shows that these two universities are close collaborators
- These two universities are competitors / potential collaborators

Thematic affinity within fields – Institutional level

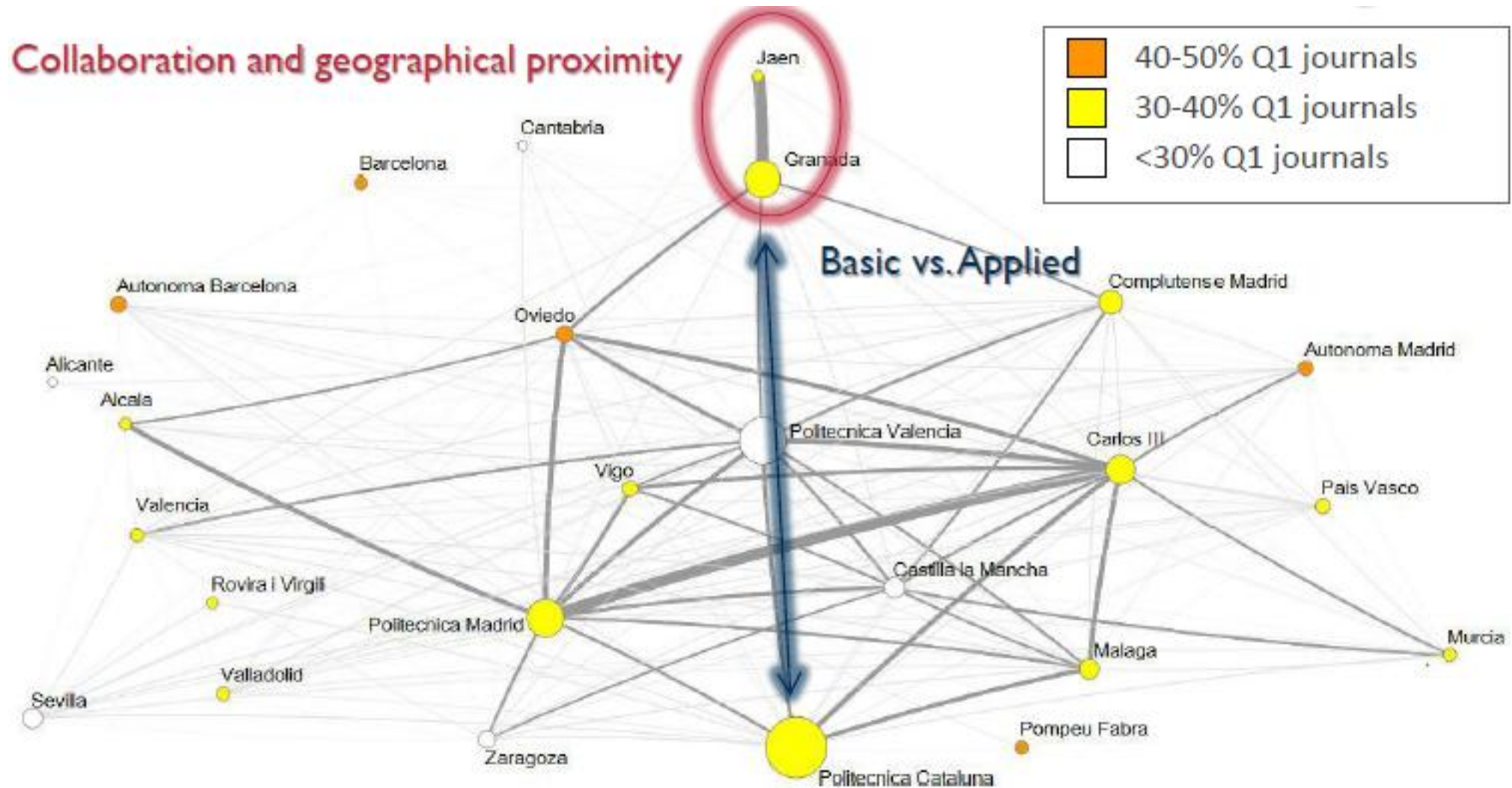
Lets take a closer look
at those disciplines...



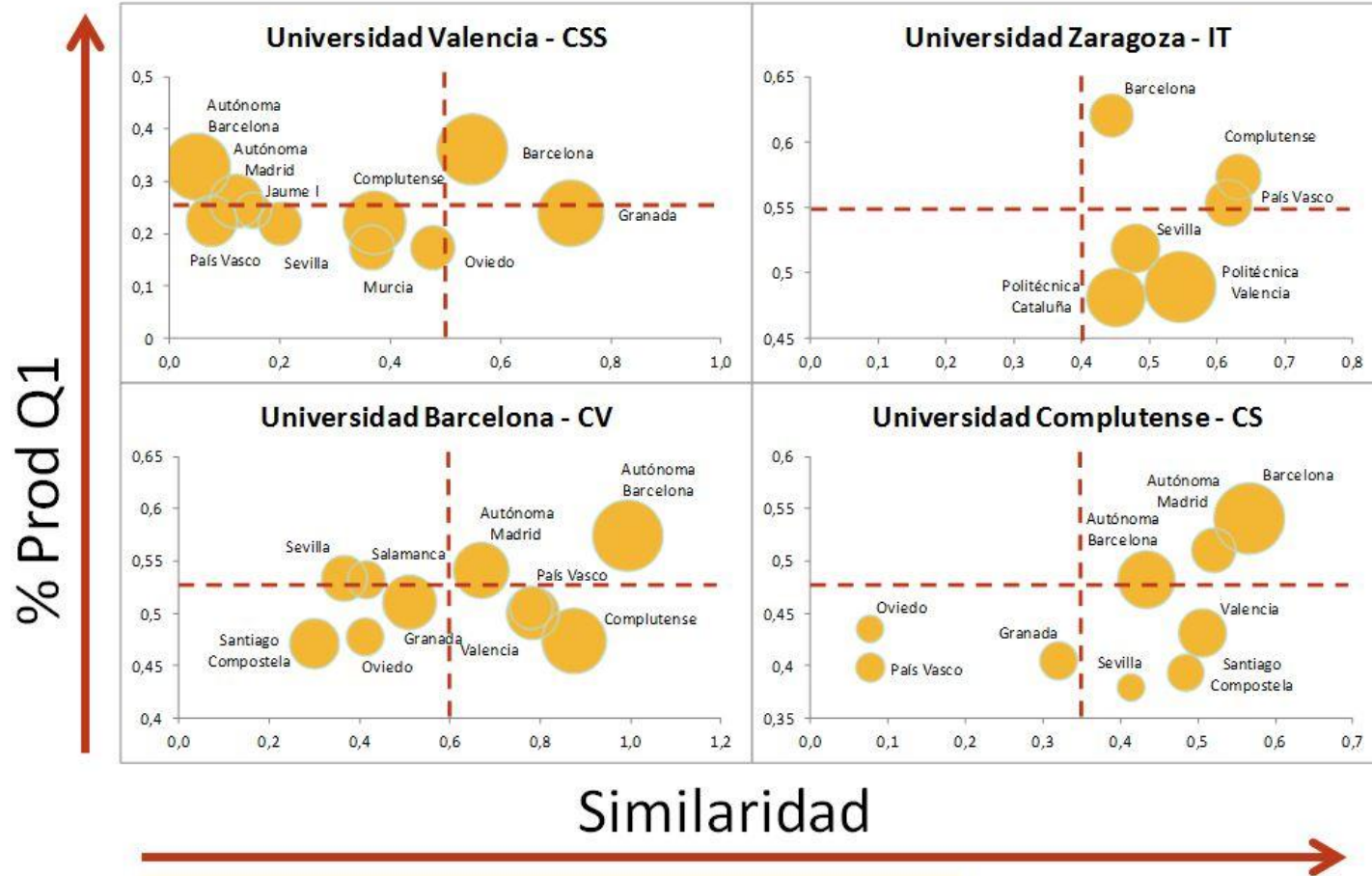
Hypothesis

*Universities with a similar research profile
should publish in the same scientific journals*

Thematic affinity within fields – Institutional level



Thematic affinity within fields – Institutional level



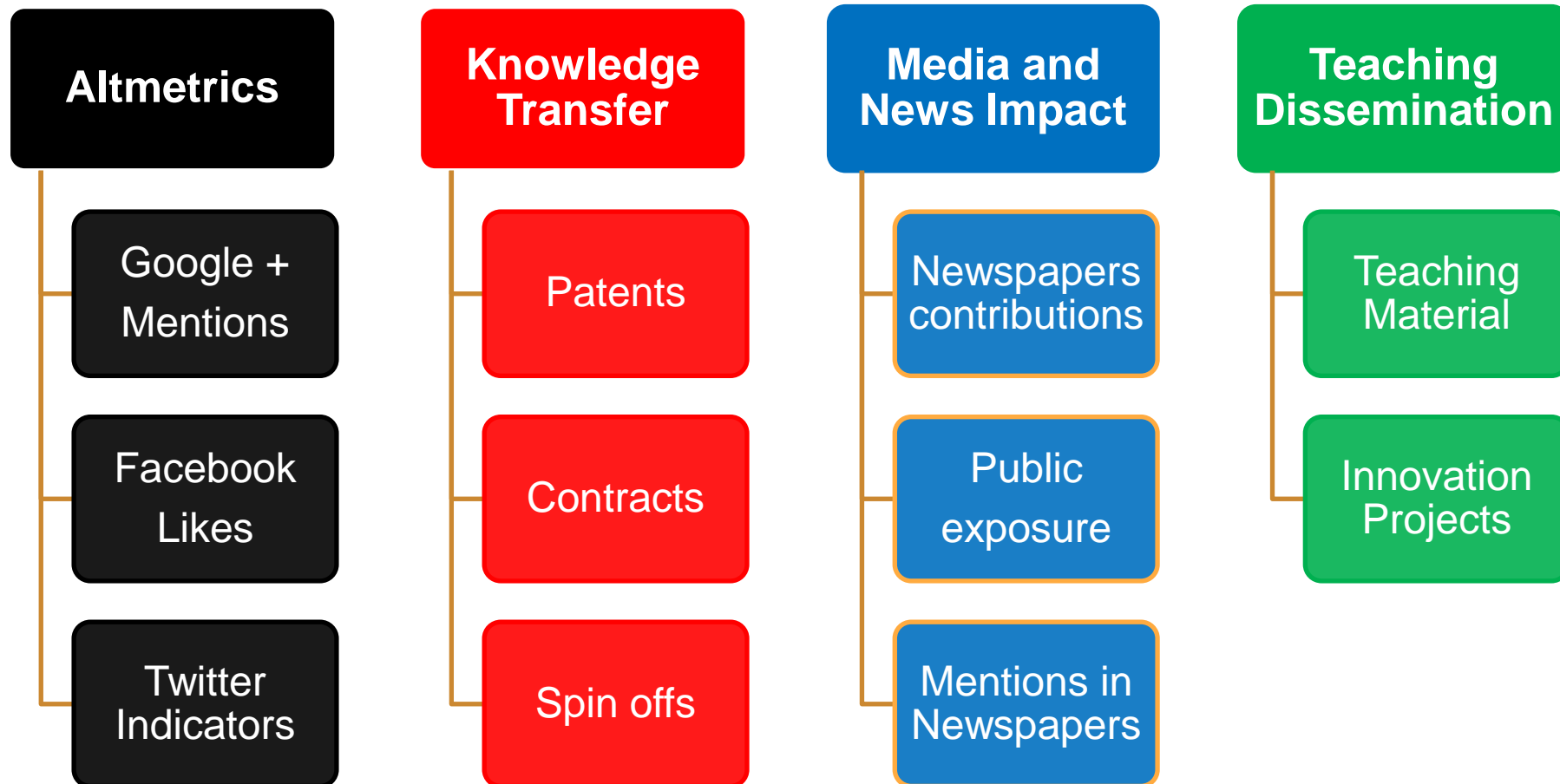
We can delve into the data even more by combining distances in the network between one institution and the rest with other indicators:

- Production
- % Highly Cited Papers
- % Co-authored publications
- ...

CSS: Ciencias Sociales – **IT:** Ingeniería y Tecnología
CV: Ciencias de la Vida - **CS:** Ciencias de la Salud









Thematic affinity within fields – Research group level

Dimensions analyzed



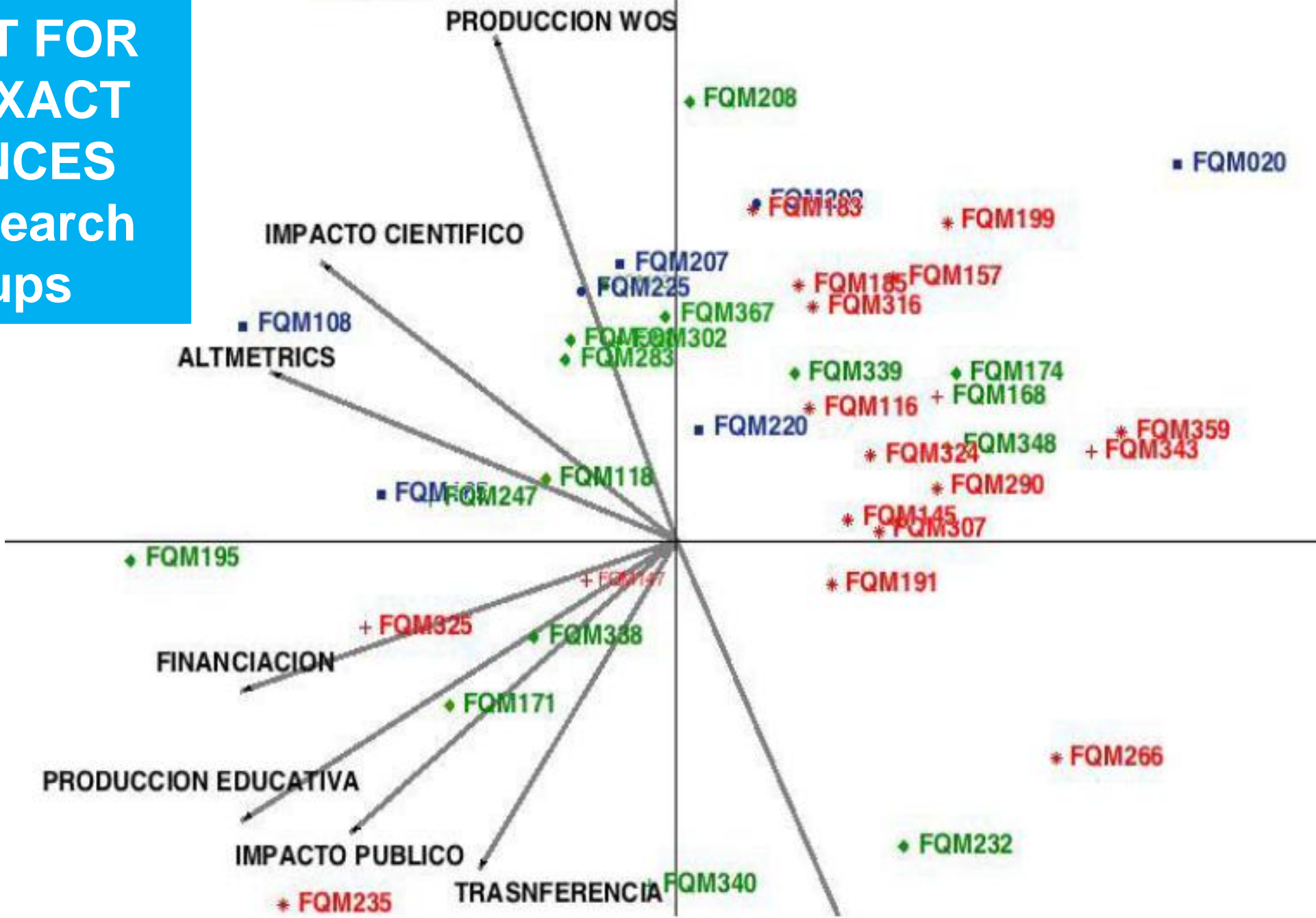
Thematic affinity within fields – Research group level

Some examples of the information sources

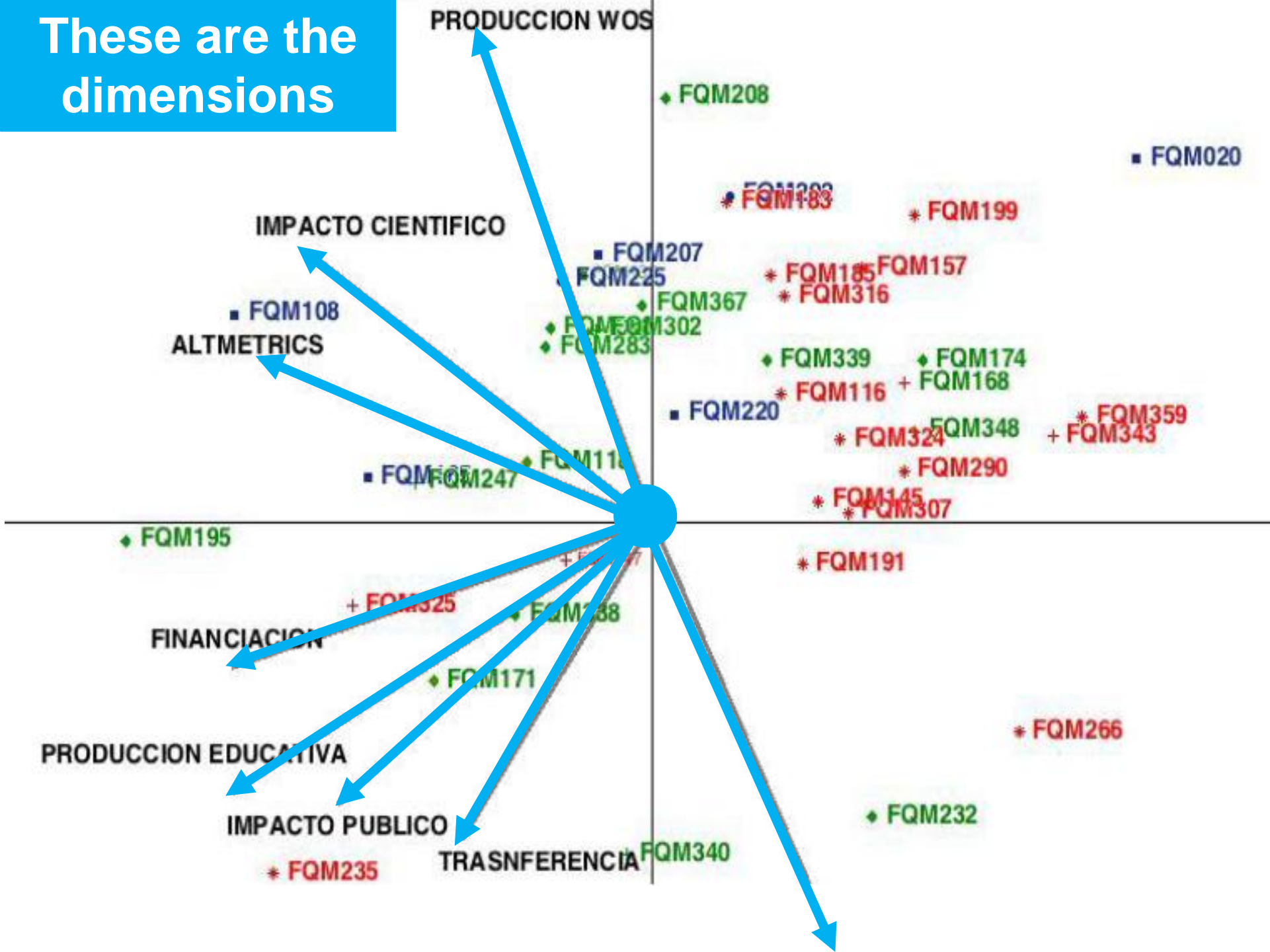
Bibliometrics	Non-bibliometrics
 <p>WEB OF SCIENCE</p>	<p>REGIONAL INFORMATION SYSTEM</p> 
 <p>ESSENTIAL SCIENCE INDICATORS</p>	<p>ALTMETRIC.COM</p> 
 <p>GOOGLE SCHOLAR CITATION</p>	<p>MY NEWS ON LINE</p> 
	<p>OCLC-WorldCat</p> 
	<p>Knowledge Transfer Office</p> 
	<p>INTERNAL DATABASES</p>

BIPLOT FOR THE EXACT SCIENCES

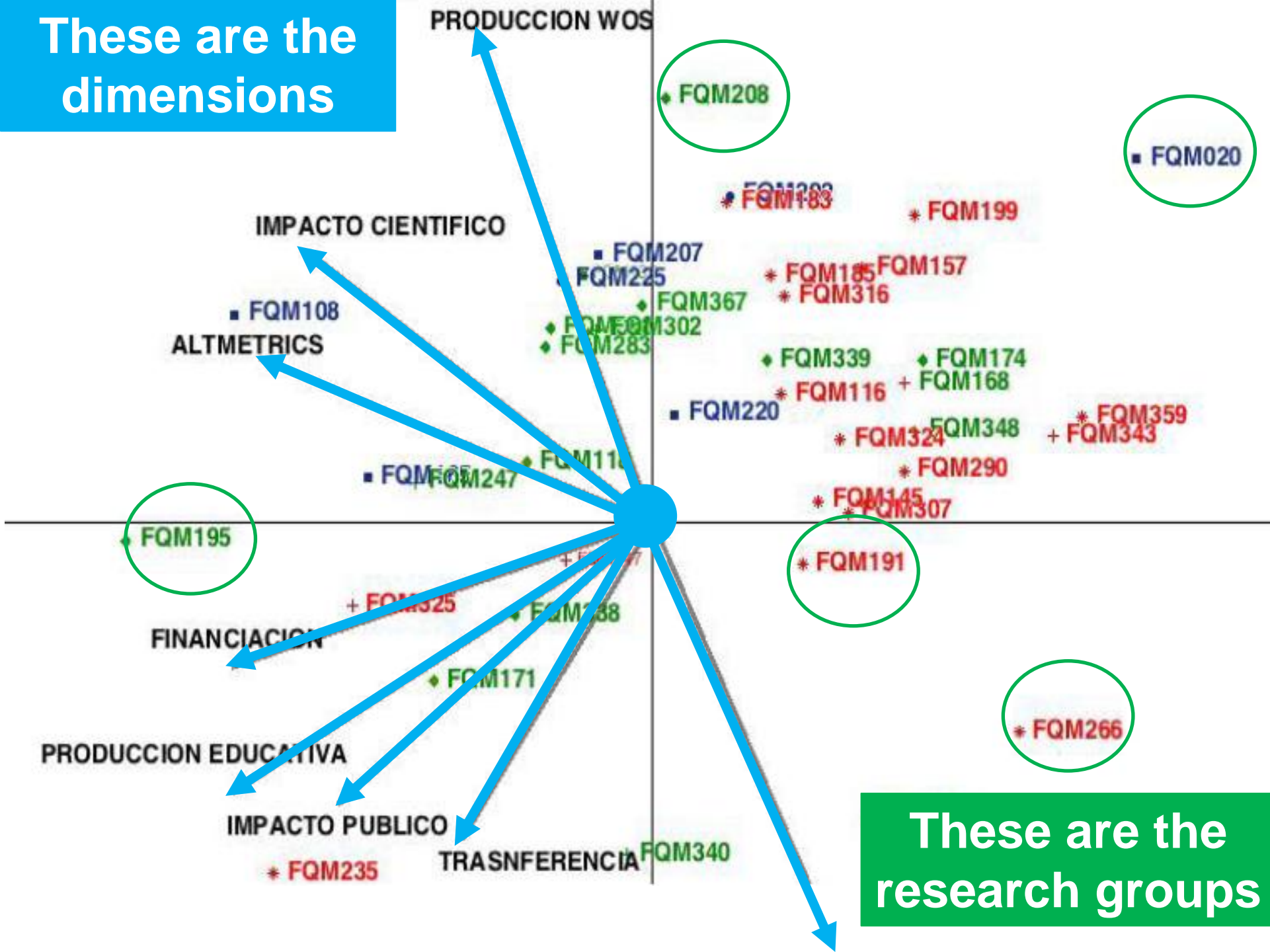
52 Research groups



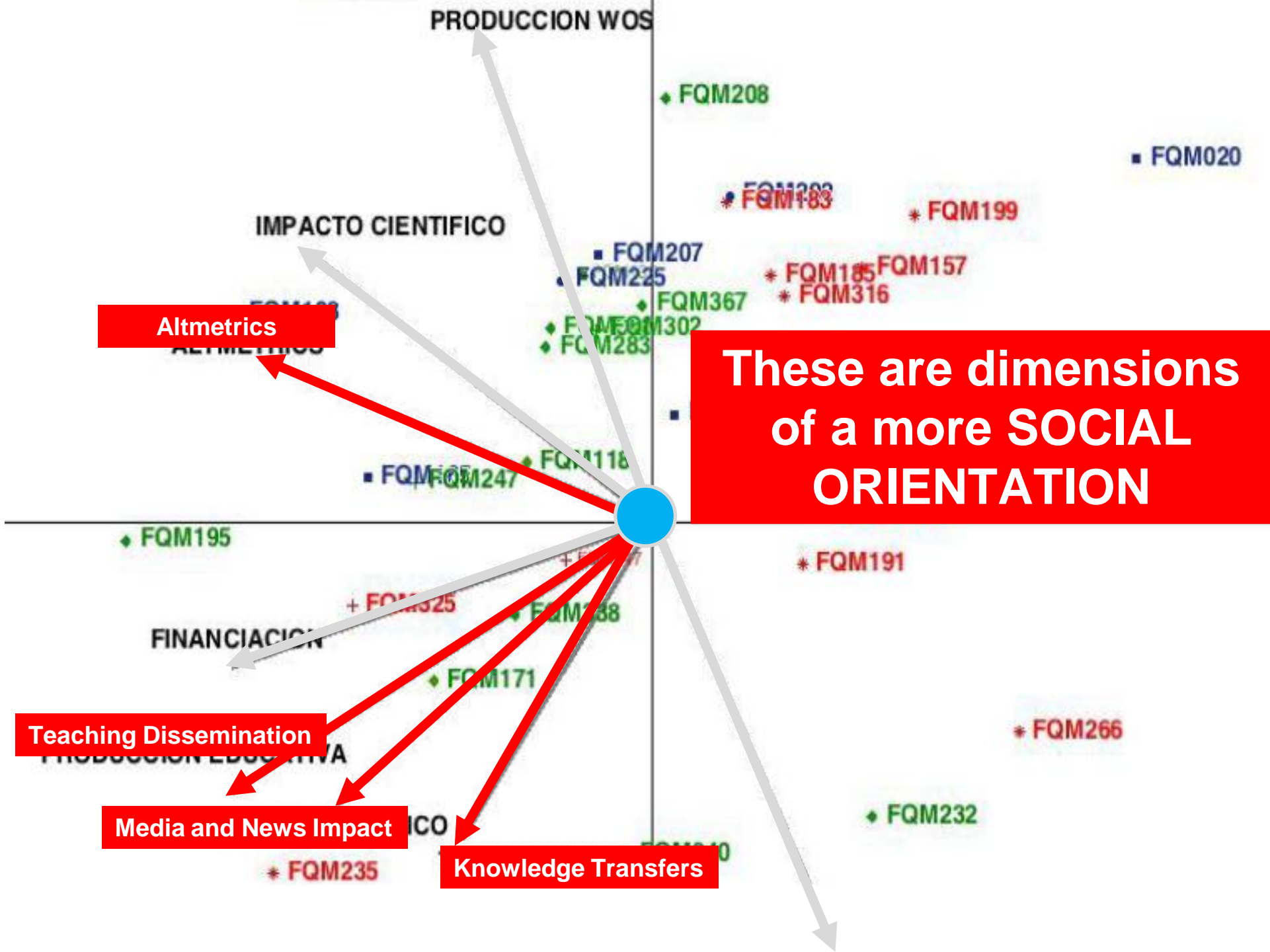
These are the dimensions

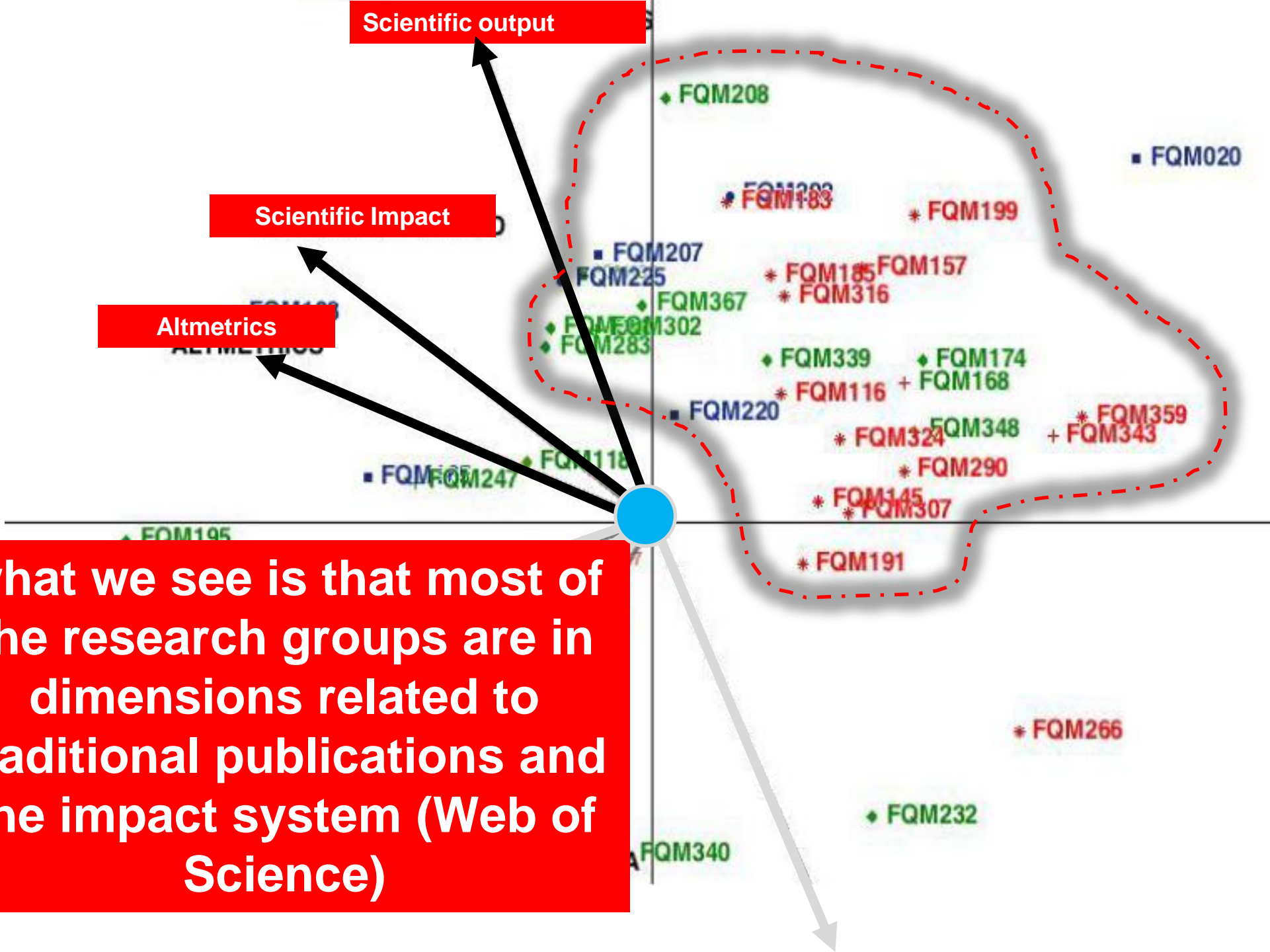


These are the dimensions

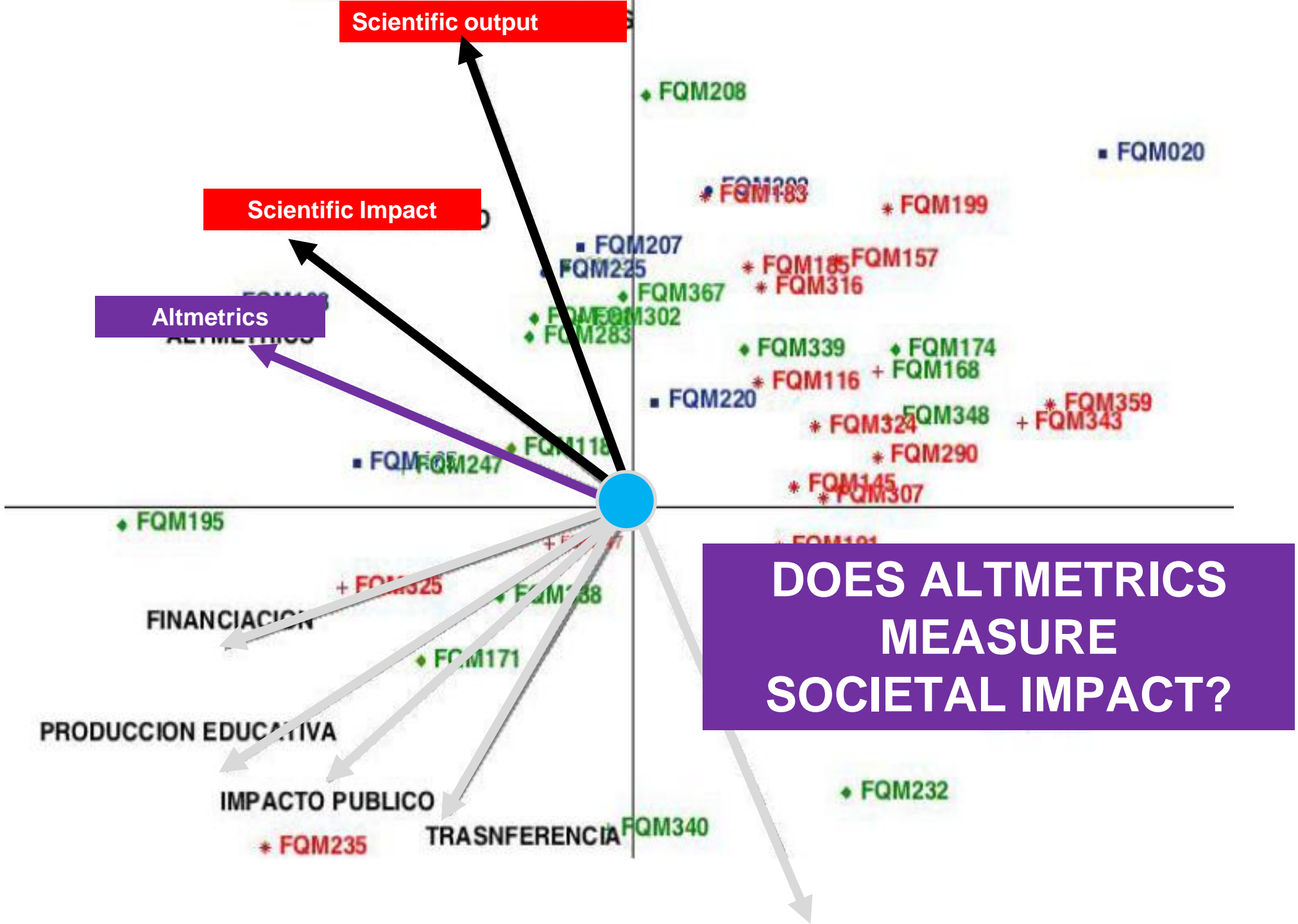


These are the research groups

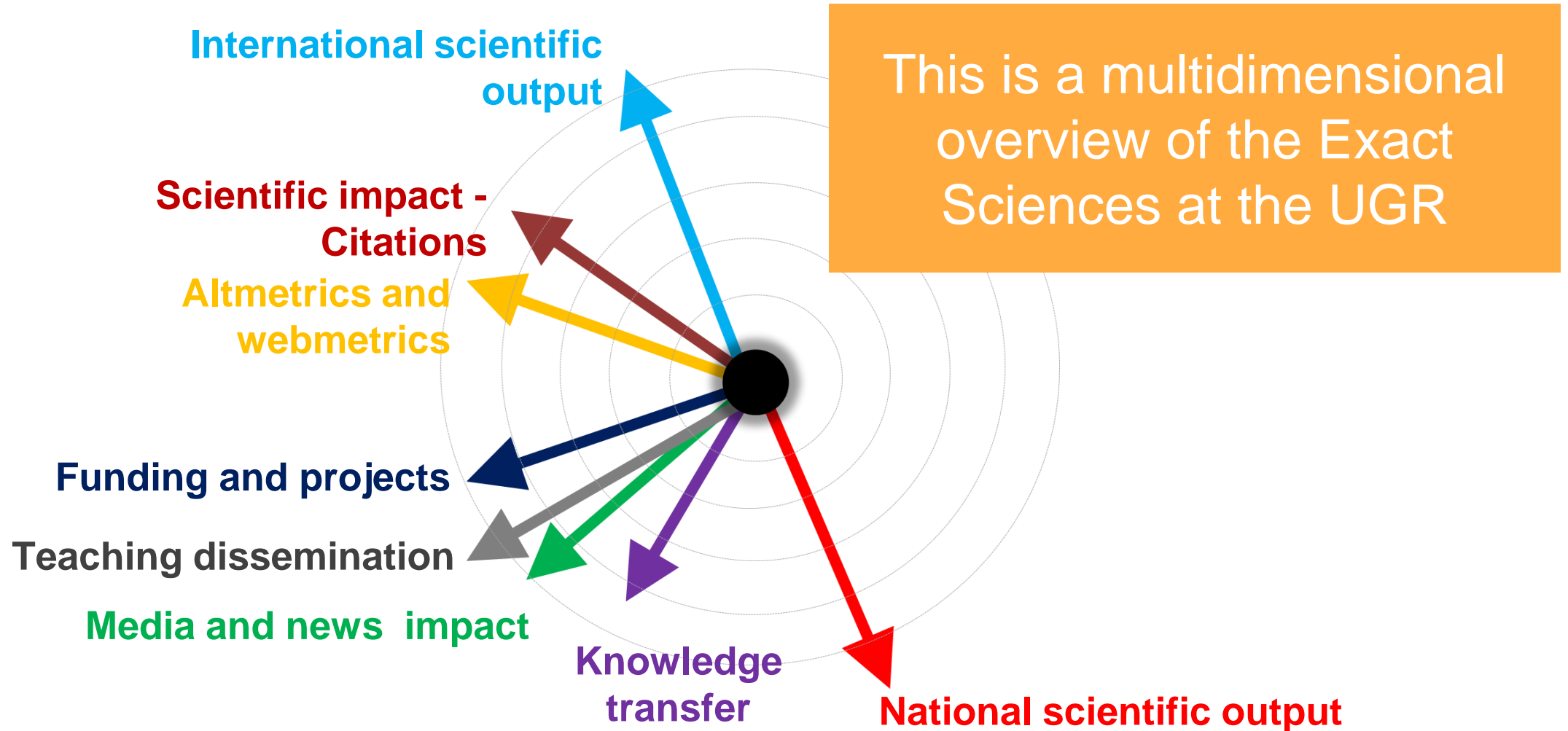




what we see is that most of the research groups are in dimensions related to traditional publications and the impact system (Web of Science)

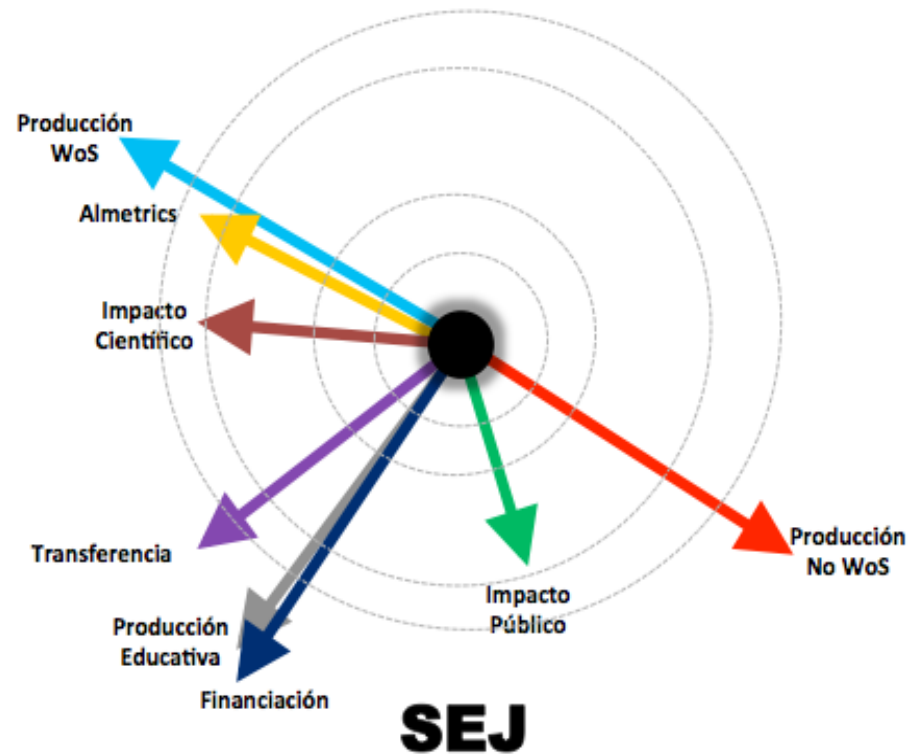


Thematic affinity within fields – Research group level

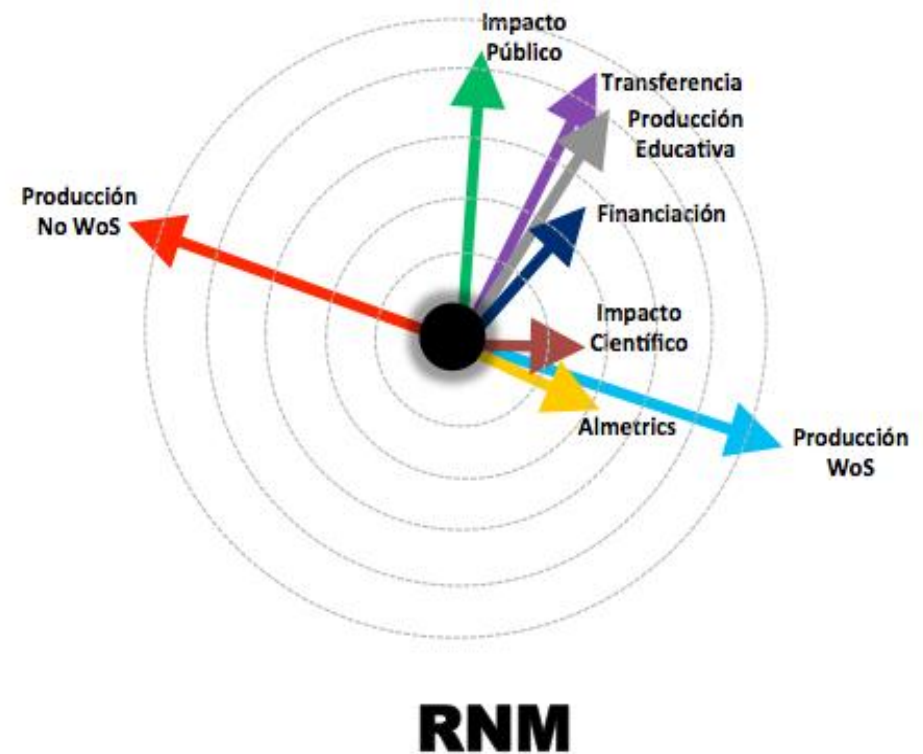


Thematic affinity within fields – Research group level

Social sciences



Natural sciences



An example – InCites

An example – InCites

1. PROFILE

¿ORGANIZATION TYPE?

Non-academic, Academic system, health, lab, publishers, museums, groups, national academies

LOCATION

Country or region level (e.g., States from US, EU)

RESEARCH AREA

ESI, WC, OECD categories...










Extra: By Research Network

2. TIME PERIOD

3. INDICATORS

- Most productive institutes (# of papers)
- Scientific excellence (e.g., Highly Cited Papers)
- Collaboration (e.g., % International collaboration)
- Journal Profile (e.g., % 1Q papers)

Name	Web of Science	Rank	Science Documents
▶ Helmholtz Association			
▶ CARL GUSTAV CARUS UNIVERSITY HOSPITAL		1	4,391
▶ Bayer AG			
▶ Bayer Healthcare Pharmaceuticals			
▶ Goethe University Frankfurt Hospital			
▶ Max Planck Society			
▶ Helmholtz-Zentrum Munchen			
▶ Boehringer Ingelheim	8	355	77.59% 5.0%
▶ Merck KGaA	9	298	84.25% 3.0%
▶ Heinrich Heine University Dusseldorf Hospital	10	256	37.4% 1.5%
▶ Max Delbruck Center for Molecular Medicine	11	238	69.06% 1.2%

-  Collaborating Organizations
-  Collaborating People
-  Collaborating Countries
-  Research Areas
-  Journals
-  Affiliated People
-  Associated Countries
-  Funding Agency
-  Collaborating Countries

Dataset:
InCites Dataset

Organization Type:
NOT Academic

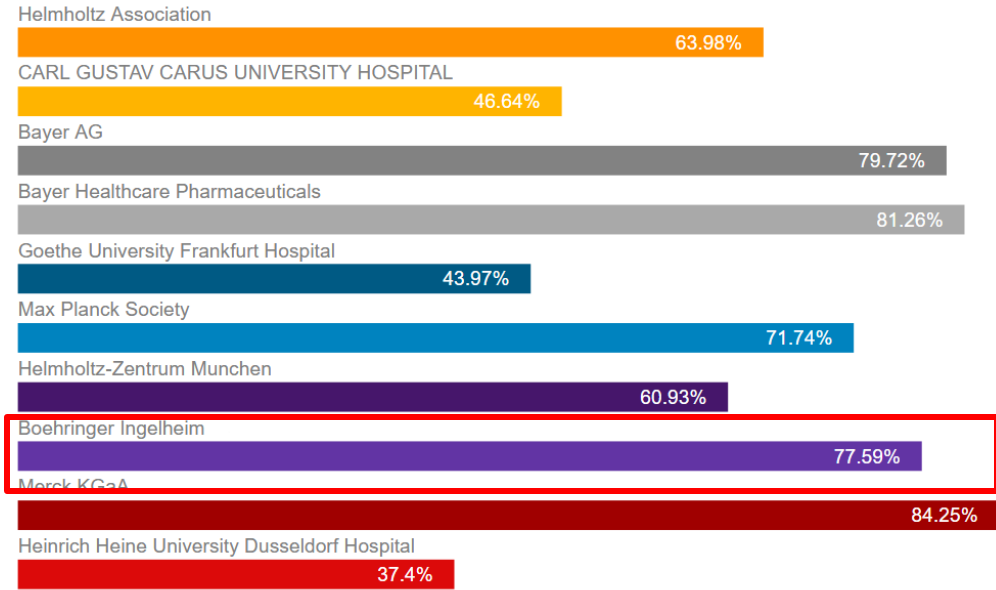
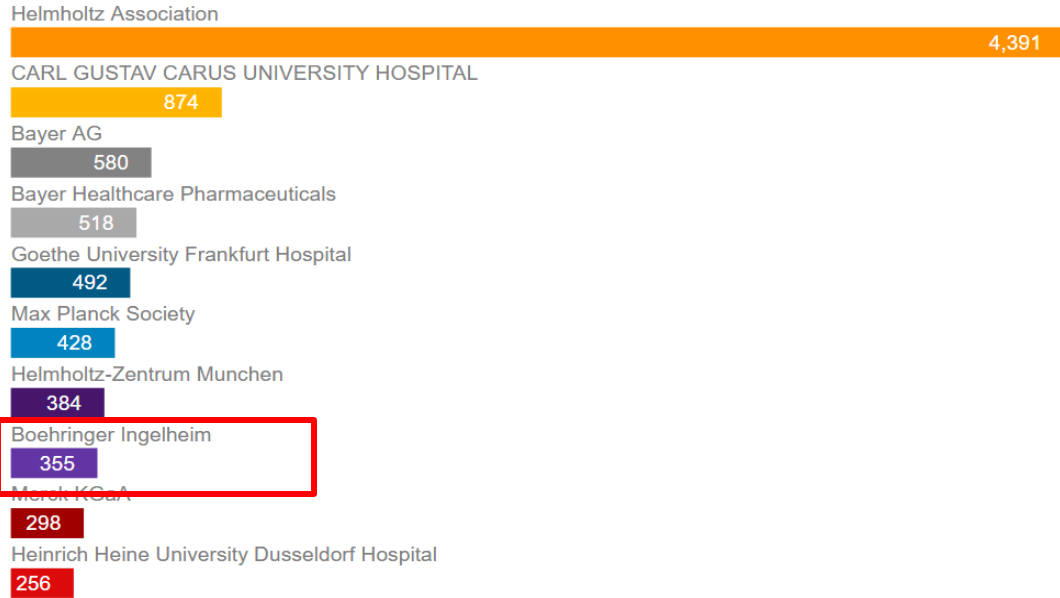
Location:
GERMANY (FED REP GER)

Schema:
Web of Science

Research Area:
ONCOLOGY

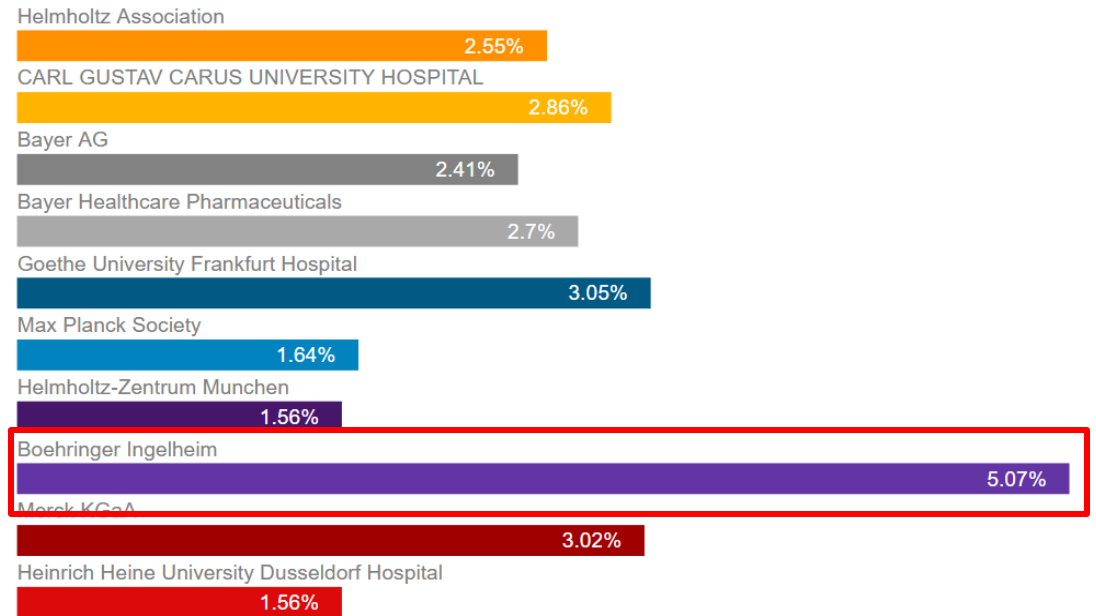
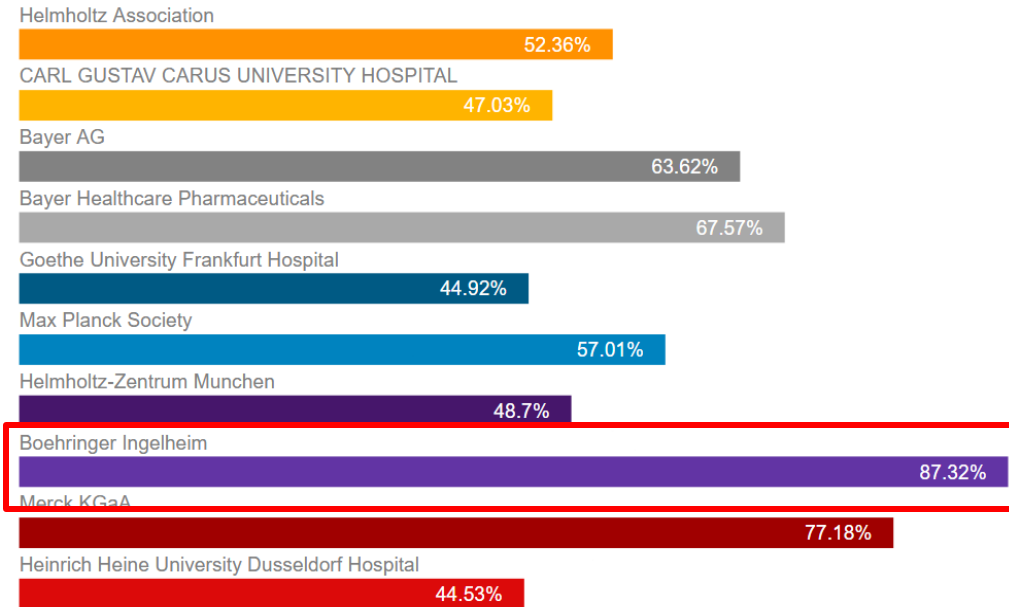
Time Period:
2012-2016

PAPERS



% FIRST Q

% INT. COLL



HIGHLY CITED P

European Summer School | September 21, 2017

THANK YOU!

Nicolas Robinson-Garcia
Georgia Institute of Technology



Daniel Torres-Salinas
EC3Metrics SL
Universidad de Granada y Navarra



References

- García, J. A., Rodríguez-Sánchez, R., Fdez-Valdivia, J., Robinson-García, N., & Torres-Salinas, D. (2013). Benchmarking research performance at the university level with information theoretic measures. *Scientometrics*, 95(1), 435-452.
- Milanés Guisado, Y. Evaluación multidimensional de la investigación. Análisis micro en la universidad de Granada durante el periodo 2009-2013. Granada: Universidad de Granada, 2016. [<http://hdl.handle.net/10481/42894>]
- Moed, H.F. (2017). A critical comparative analysis of five world rankings. *Scientometrics*, 110(2), 823-834
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy*, 41(7), 1262-1282.
- Rafols, I., Porter, A.L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *J Assoc Inform Sci Technol*, 61(9), 1871-1887.
- Robinson-García, N., Rodríguez-Sánchez, R., García, J. A., Torres-Salinas, D., & Fdez-Valdivia, J. (2013). Análisis de redes de las universidades españolas de acuerdo a su perfil de publicación en revistas por áreas científicas. *Rev Esp Doc Cient*, 36(4), 027.
- Rotolo, D., Rafols, I., Hopkins, M. M., & Leydesdorff, L. (2017). Strategic intelligence on emerging technologies: Scientometric overlay mapping. *Journal of the Association for Information Science and Technology*, 68(1), 214-233.
- Torres-Salinas, D., Robinson-García, N., Jimenez-Contreras, E., Herrera, F., Lopez-Cozar, E.D. (2013). On the use of biplot analysis for multivariate bibliometric and scientific indicators. *J Assoc Inform Sci Technol* 64(7), 1468-1479
- Torres-Salinas, D., Rodríguez-Sánchez, R., Robinson-García, N., Fdez-Valdivia, J., García, J.A. (2013). Mapping citation patterns of book chapters in the Book Citation Index. *J informetrics*, 7(2), 412-424.
- Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature*, 512, 126–129.

Recommended readings

