Tracking content updates in Scopus (2011-2018): a quantitative analysis of journals per subject category & subject categories per journal

Frédérique Bordignon
Ecole des Ponts ParisTech, Direction de la Documentation, Champs-sur-Marne, France

17th INTERNATIONAL CONFERENCE ON SCIENTOMETRICS & INFORMETRICS
September 2-5, 2019 - Sapienza University, Rome, Italy

Study context and design

Journal classification schemes

- Many systems used for evaluation purposes
- No international standard
- No single classification scheme widely adopted

–> Difficult to make comparative analyses because of the dual problem of matching categories and delineating journals comparable sets.

Mono *vs* multi-disciplinary classification systems

- Systems with a great level of details in a restricted research area
  -> JEL (Journal of Economic Literature), MeSH (Medical Subject Headings)

- Systems appended to multidisciplinary databases
  -> WoS, Scopus...

- Journal level classification systems:
  - sometimes too fuzzy
  - at least not so accurate as article level classifications

Multiaffectation classification systems

- Many journals are assigned to multiple categories
- Many journals are assigned to too many categories

Wang, Q., & Waltman, L. (2016). **Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus**. *Journal of Informetrics*, 10(2), 347–364. 10.1016/j.joi.2016.02.003

- Classifications are generally not well documented

Archambault, É., Beauchesne, O. H., & Caruso, J. (2011). **Towards a multilingual, comprehensive and open scientific journal ontology**. In *Proceedings of the 13th international conference of the international society for scientometrics and informetrics*. E. Noyons, P. Ngulube, J. Leta (Eds.) (pp. 66–77).

Very few studies dedicated to changes over time

- "*Comparing like with like*" (Martin & Irvine, 1983)

Martin, B. R., & Irvine, J. (1983). **Assessing basic research: Some partial indicators of scientific progress in radio astronomy**. *Research Policy*, 12(2), 61–90. https://doi.org/10.1016/0048-7333(83)90005-7

- About the problem of fitting new journals into existing schemes: "*if one does not systematically account for redelineation in the groupings over time but uses "fixed journal sets" instead, one risks making a prediction of performance with reference to an outdated unit*."

Leydesdorff, L. (2002). **Dynamic and evolutionary updates of classificatory schemes in scientific journal structures**. *Journal of the American Society for Information Science and Technology*, 53(12), 987–994. https://doi.org/10.1002/asi.10144

Reliability of analyses based on data extracted from Scopus

Are Scopus coverage updates substantial?

Do they affect categories per journal and/or journals per category?

?

ASJC journal classification system

- Confusing subfield labels

Wang, Q., & Waltman, L. (2016). **Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus**. *Journal of Informetrics*, 10(2), 347–364. 10.1016/j.joi.2016.02.003

- Strong imbalanced distribution of journals and therefore documents among the fields

Jacsó, P. (2013). **The need for end-user customization of the journal-sets of the subject categories in the SCImago Journal Ranking database for more appropriate league lists. A case study for the Library & Information Science field**. *El Profesional de La Informacion*, 22(5), 459–473. 10.3145/epi.2013.sep.12

- The nomenclature structure itself is stable.

Table 1. ASJC journal classification system

Supergroups	*Fields*	*No. of Subfields*
–	Multidisciplinary	1
Health Sciences	Medicine	48
Health Sciences	Nursing	23
Health Sciences	Veterinary	4
Health Sciences	Dentistry	6
Health Sciences	Health Professions	16
Life Sciences	Agricultural and Biological Sciences	11
Life Sciences	Biochemistry, Genetics and Molecular Biology	15
Life Sciences	Immunology and Microbiology	6
Life Sciences	Neuroscience	9
Life Sciences	Pharmacology, Toxicology and Pharmaceutics	5
Physical Sciences	Chemical Engineering	8
Physical Sciences	Chemistry	7
Physical Sciences	Computer Science	12
Physical Sciences	Earth and Planetary Sciences	13
Physical Sciences	Energy	5
Physical Sciences	Engineering	16
Physical Sciences	Environmental Science	12
Physical Sciences	Materials Science	8
Physical Sciences	Mathematics	14
Physical Sciences	Physics and Astronomy	10
Social Sciences	Arts and Humanities	13
Social Sciences	Business, Management and Accounting	10
Social Sciences	Decision Sciences	4
Social Sciences	Economics, Econometrics and Finance	3
Social Sciences	Psychology	7
Social Sciences	Social Sciences	22
4 supergroups	**26 fields**	**307 subfields**

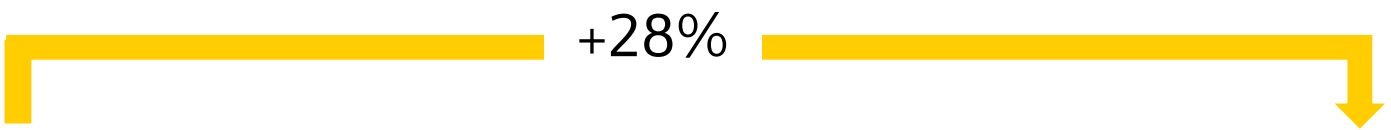
Data

- Title list files Elsevier has released twice a year since 2011
- Focus on journals only
- Aggregated data is available for reuse and further investigation (SNIP values and Open Access status are included in the dataset even if they are not analyzed in the study).

Bordignon, Frederique (2019), “Scopus source title list: aggregated data (2011-2018)”, Mendeley Data, V1, doi: 10.17632/855x2zwjd2.1

Results

Inclusion and withdrawal of journals at the category level



	2011	2012	2013	2014	2015	2016	2017	2018
No. of journals	28 335	29 561	31 154	32 332	33 058	33 810	34 772	36 189
Annual growth	–	**+4,3%**	**+5,4%**	+3,8%	+2,2%	+2,3%	+2,8%	+4,1%

Table 2. Number of journals included in Scopus and annual growth

- Very few journals are merely dropped
- Even inactive journals are sometimes added

Unpredictable fluctuations

- Contrasted evolution from one field to another

- Mainly additions

newly included journals or journals "coming from" other fields?

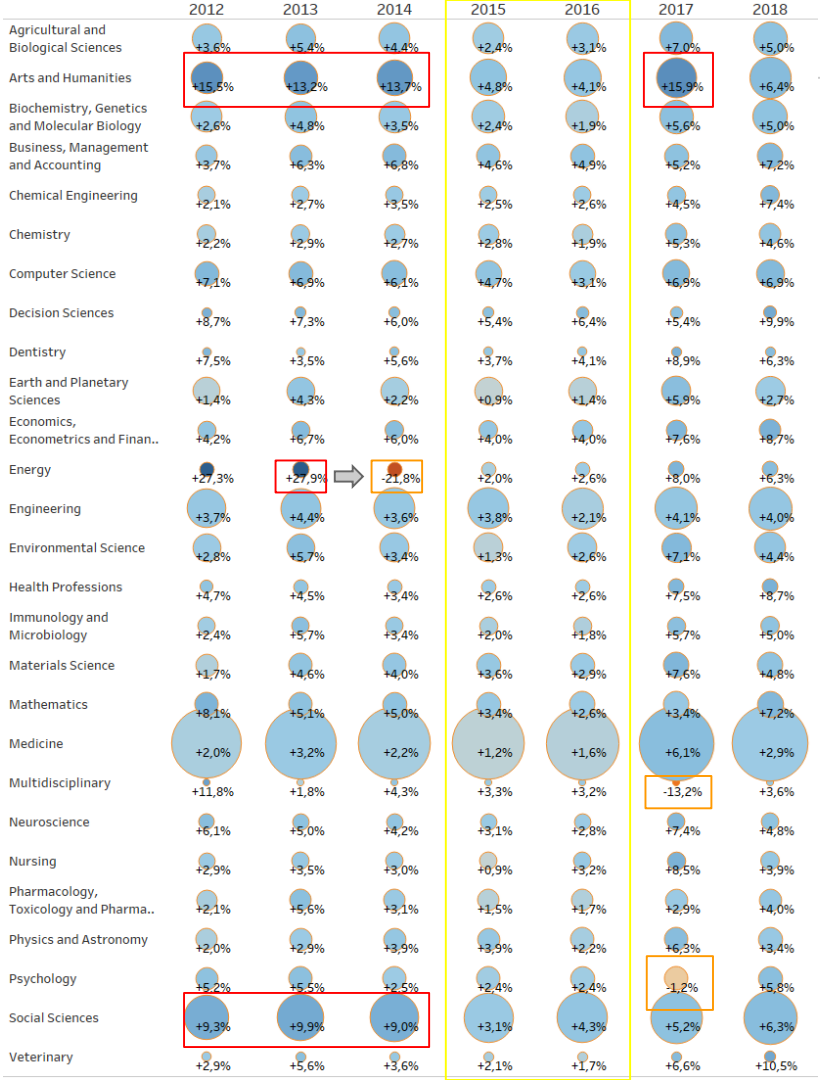


Figure 1. Number of journals per field and annual growth

Field or subfield shifts are very unusual

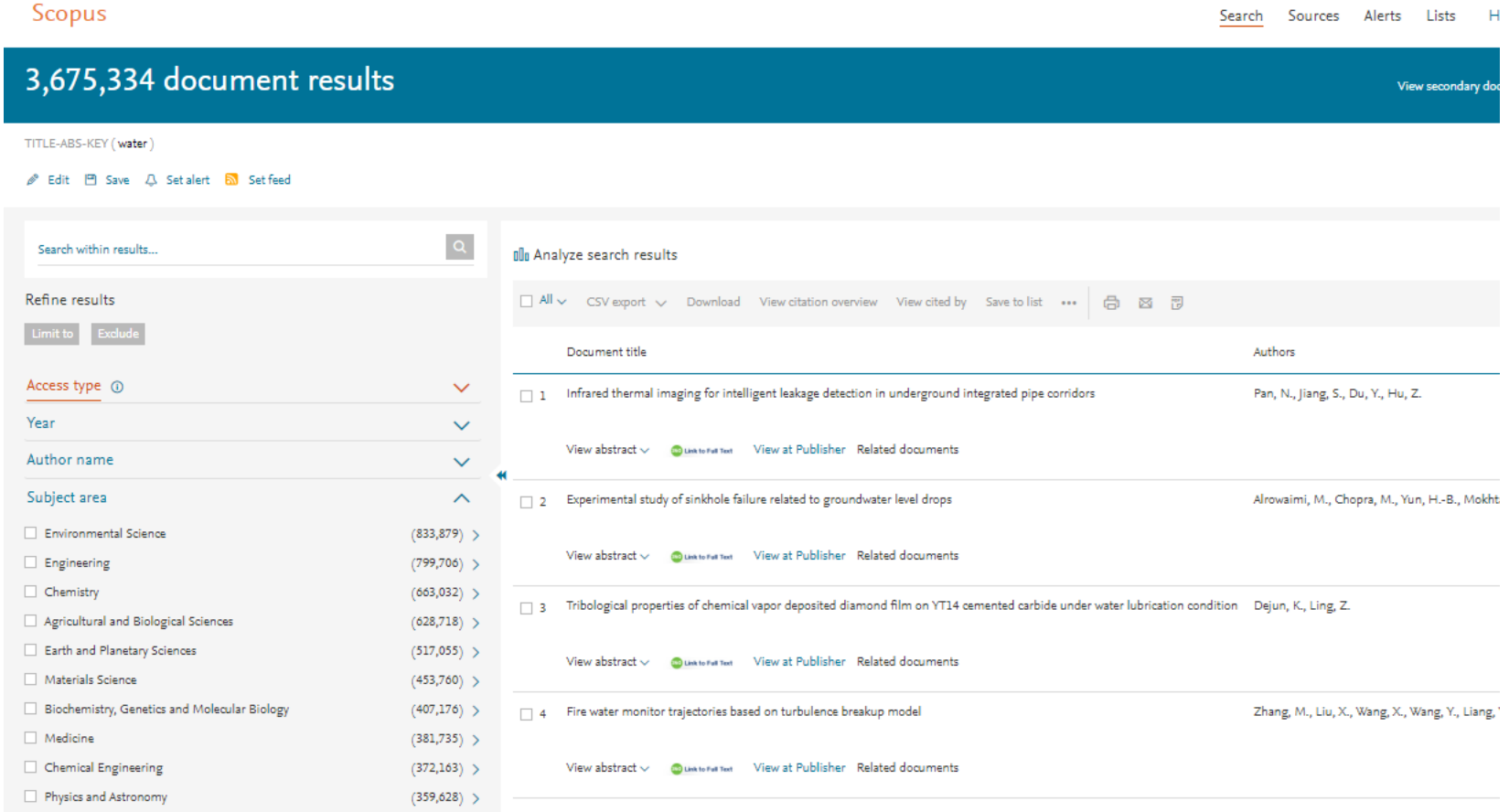
Table 3. Annual percentage of journals whose assignment to fields and subfields has been updated

2012		2013		2014		2015		2016		2017		2018	
subfield	field	subfield	field	subfield	field	subfield	field	subfield	field	subfield	field	subfield	field
0,1%	0,1%	0,2%	0,2%	1,0%	1,0%	0,5%	0,5%	0,3%	0,3%	**10,6%**	**6,9%**	1,5%	1,0%

Fields	No. of journals with subfield shifts	% of all journals in the field
Multidisciplinary	33	25,60%
Psychology	274	21,6%
Immunology and Microbiology	132	18,0%
Biochemistry, Genetics and Molecular Biology	439	17,7%
Health Professions	85	17,7%
Pharmacology, Toxicology and Pharmaceutics	172	16,3%
Neuroscience	99	15,9%
Veterinary	33	13,60%
Medicine	1630	13,4%
Nursing	88	12,5%
Environmental Science	239	12,1%
Social Sciences	652	11,2%
Engineering	376	9,5%
Computer Science	136	8,7%
Earth and Planetary Sciences	149	8,1%
Agricultural and Biological Sciences	186	8,0%
Arts and Humanities	201	6,3%

Table 4. Number and percentage of journals per field with subfield shifts between 2016 and 2017

Limited Scopus interface



Categories per journal

Table 5. Average number of fields and subfields per journal for added or previously included ones

fields: +1,49%

		2011	2012	2013	2014	2015	2016	2017	2018
Avg number of fields	Added	1,48	1,51	1,56	1,57	1,61	1,66	1,67	1,66
	Previously included	1,53	1,53	1,53	1,53	1,53	1,54	1,54	1,54
Avg number of subfields	Added	1,91	2	2,02	2,08	2,26	2,37	2,45	2,38
	Previously included	1,99	1,98	1,98	1,99	2	2,01	2,01	2,03

subfields: +3,11%

- Increasing average number of fields and subfields assigned to journals.
- Almost exclusively due to the addition of journals to the database index.
- Newly included sources have been assigned to more fields and subfields (since 2014) on average than those previously indexed.

Conclusion and discussion

Changes are significant

- Scopus coverage updates can be substantial, very irregular, invisible to the average user and unpredictable.

- Large volume changes in some fields certainly do not reflect the scientific reality of the field but rather Elsevier's objectives to increase its coverage.

Consequences can be significant

For example:

- on reports/analyses
- on SNIP values due to an unstable scope of journals and therefore a very unstable citations rate,
- on international thematic university rankings whose evaluation criteria are based partly on the collection of citations and outputs according to subject areas.

What could be done?

- Elsevier should (better) inform the user of Scopus content updates in order to prevent potential impacts on the resulting analyses.

- One possibility is to reflect category changes of a journal only on newly added publications (recently published or not) and not on all publications already present in the database.

Thanks!

Any questions ?

frederique.bordignon@enpc.fr

École des Ponts
ParisTech