



Open Research Seminar. Competence Network Bibliometrics

June 7. 2023

The position of France and French institutions in world scientific publications: The issue of database(s)

Frédérique Sachwald
Science and Technology Observatory
- Hcéres



OUTLINE

1. Examples of analyses by the Science and Technology Observatory (OST)
2. The issue of the database(s)
3. Observation of French scientific publications in the national archive and WoS
4. Discussion

1. EXAMPLES OF OST ANALYSES

Data source in this part : OST publication database. in house WoS



Different types of analyses

Focus on publications here.

- Part of Hcéres evaluation processes : universities' research, public research organizations, academic hospitals
- National disciplinary or thematic evaluations (maths, archaeology)
- Thematic analyses (cancer, cybersecurity).

Geographical or institutional perimeters.

WORLD SHARE AND SHARE OF TOTAL PUBLICATIONS IN AI: 3 CORPORA

	Large AI including applications*			Focused AI*			World share of AI major conferences proceedings**
	World share. %		AI in country publications %	World share. %		AI in country publications %	
	2018	2021	2021	2018	2021	2021	
China	24.8	30.1	12.7	25.5	33.1	4.7	26.1
USA	16.8	15.6	9.1	15.1	12.8	2.5	27.7
India	6.1	5.4	12.1	7.8	7.5	5.6	2.0
S. Korea	2.5	3.3	11.9	2.1	3.4	4.0	3.6
UK	3.8	3.7	8.8	3.8	3.2	2.5	4.6
Germany	3.6	3.4	8.2	3.4	3.1	2.5	5.7
Japan	2.9	2.3	6.7	3.3	2.5	2.4	4.3
Italy	2.3	2.5	7.9	2.1	2.2	2.4	1.7
France	2.4	2.0	8.4	2.5	2.2	3.0	2.9
Canada	2.4	2.5	9.6	2.1	2.1	2.7	3.4
Australia	2.1	2.2	8.9	1.9	2.0	2.7	2.2
Spain	2.2	2.0	8.2	2.2	1.9	2.5	0.8
Iran	2.0	2.0	11.1	1.3	1.6	2.8	0.2
Taiwan	1.5	1.5	12.7	1.5	1.5	4.2	1.2
Brazil	1.8	1.5	6.8	2.1	1.3	2.0	0.4
EU27	18.8	17.3	8.1	18.09	15.8	2.5	16.7
World	100.0	100.0	9.8	100.00	100.0	3.3	100.0

** CVPR. GECCO. ICASSP. ICCV. IJCAI. IROS. ISWC. NeurIPS

* The large corpus has been built on the basis of keywords with help from experts. Publications have at least one of those keywords in their title, abstract or authors' keywords. Publications in the focused corpus are also indexed in the following WoS categories: AI, Cybernetics, Computer hardware and architecture, Information systems, Software engineering, Computer theory and methods, Applied mathematics

0.2 % world pub.

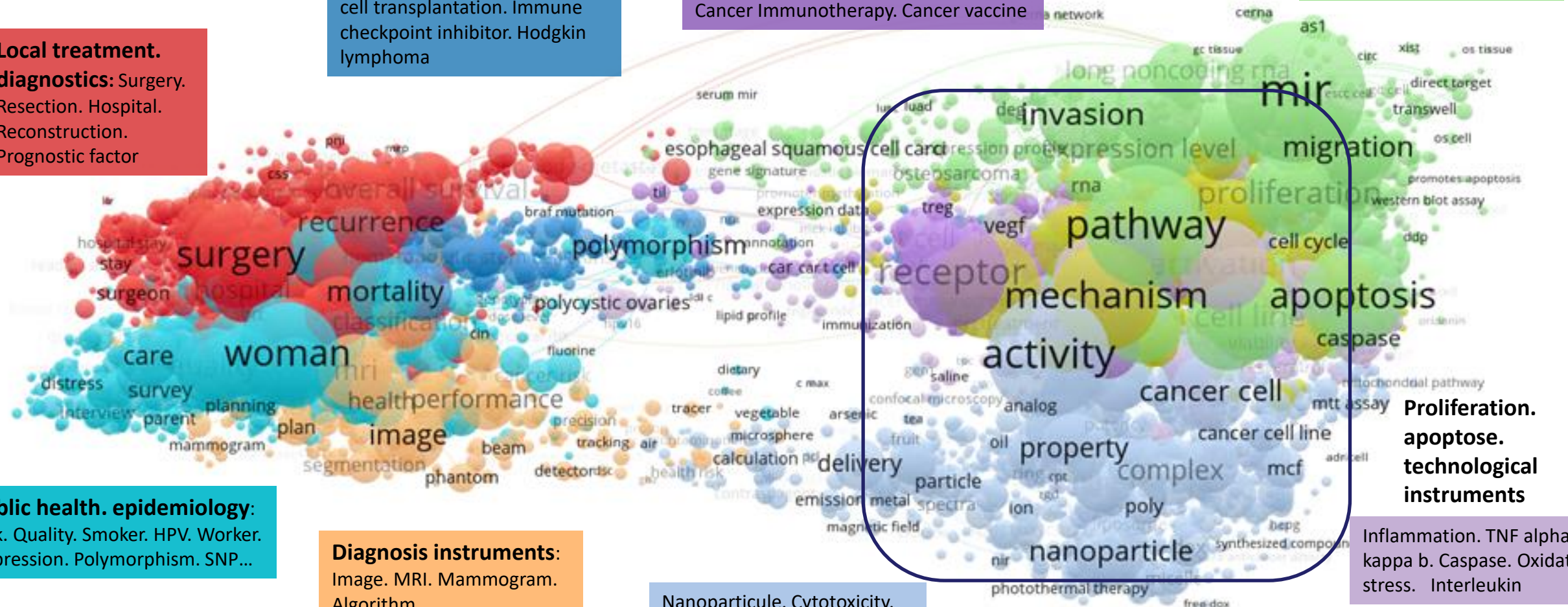
Cancer corpus, world 2016-2020, 940,416 publications

Local treatment. diagnostics: Surgery. Resection. Hospital. Reconstruction. Prognostic factor

Systemic treatments: Stem cell transplantation. Immune checkpoint inhibitor. Hodgkin lymphoma

Immunity response: T cell. Cytokine. Tumour microenvironment. Cancer Immunotherapy. Cancer vaccine

Apoptosis. MicroRNA. Cell proliferation. Tumor growth



Public health. epidemiology: Risk. Quality. Smoker. HPV. Worker. Depression. Polymorphism. SNP...

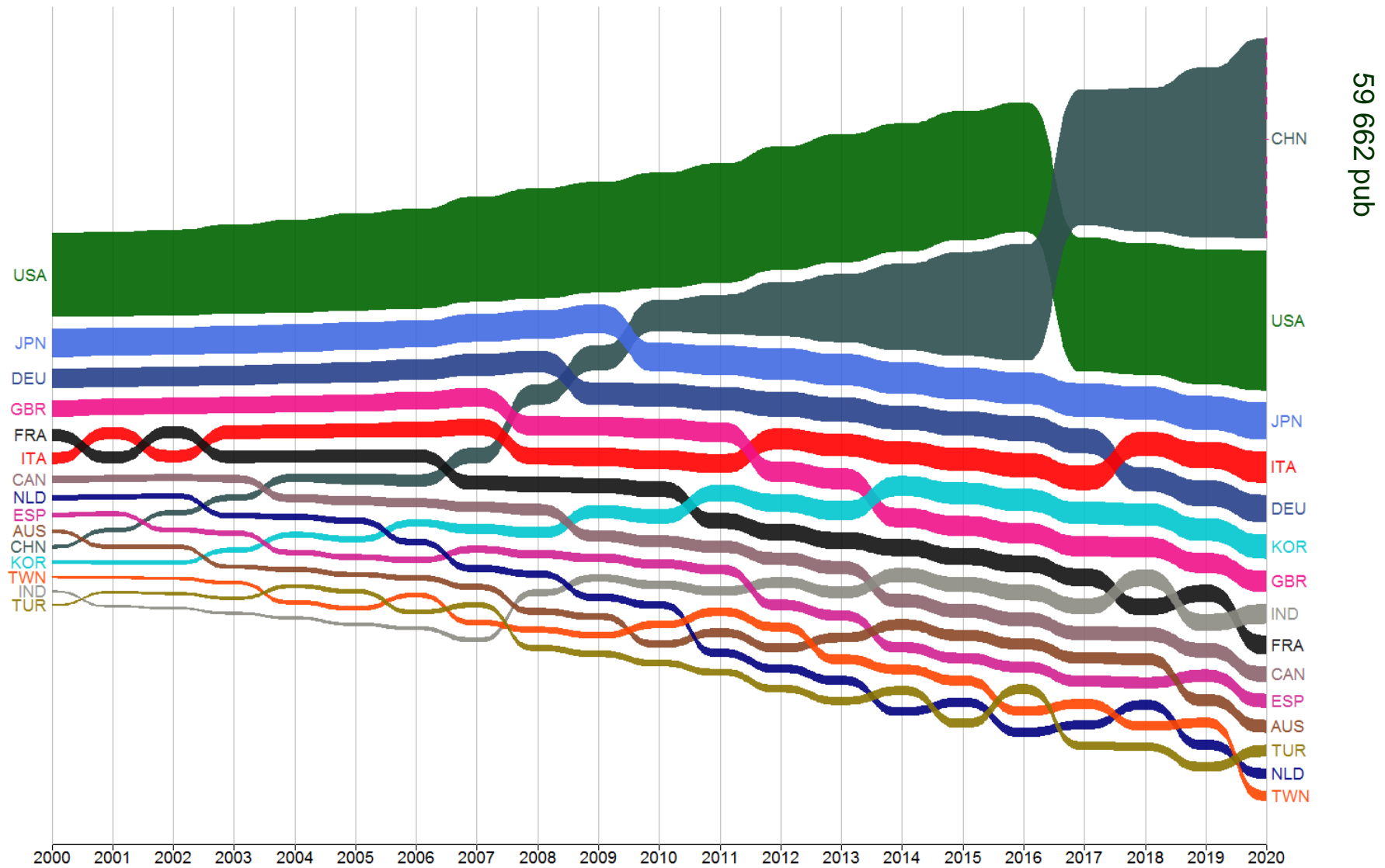
Diagnosis instruments: Image. MRI. Mammogram. Algorithm...

Nanoparticle. Cytotoxicity. Drug delivery. Cancer therapy

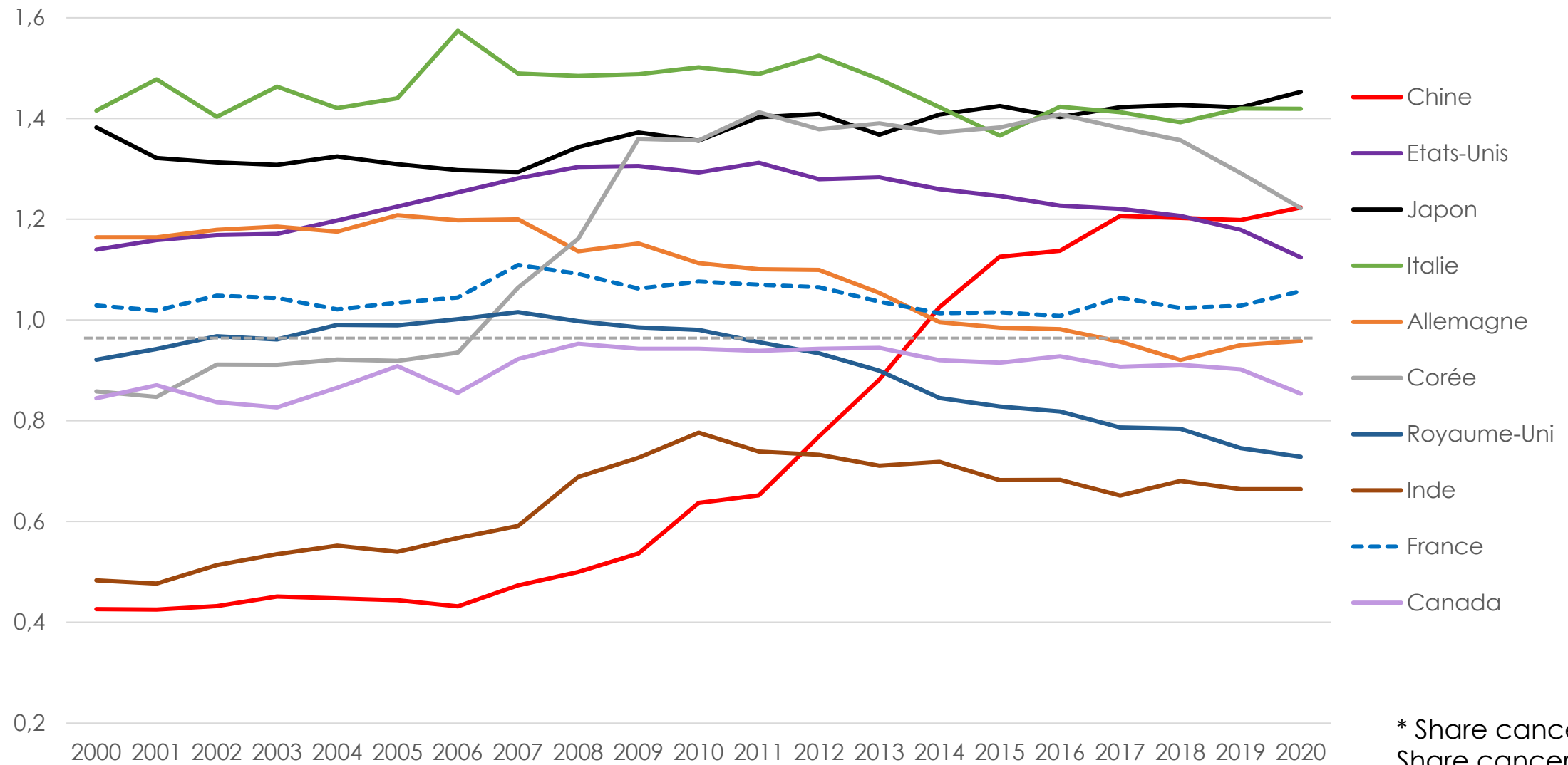
Kinase. Autophagy. Cell death. Phosphorylation. DNA damage

Proliferation. apoptose. technological instruments
Inflammation. TNF alpha. nf kappa b. Caspase. Oxidative stress. Interleukin

Number of publications, first 15 countries, 2000-20

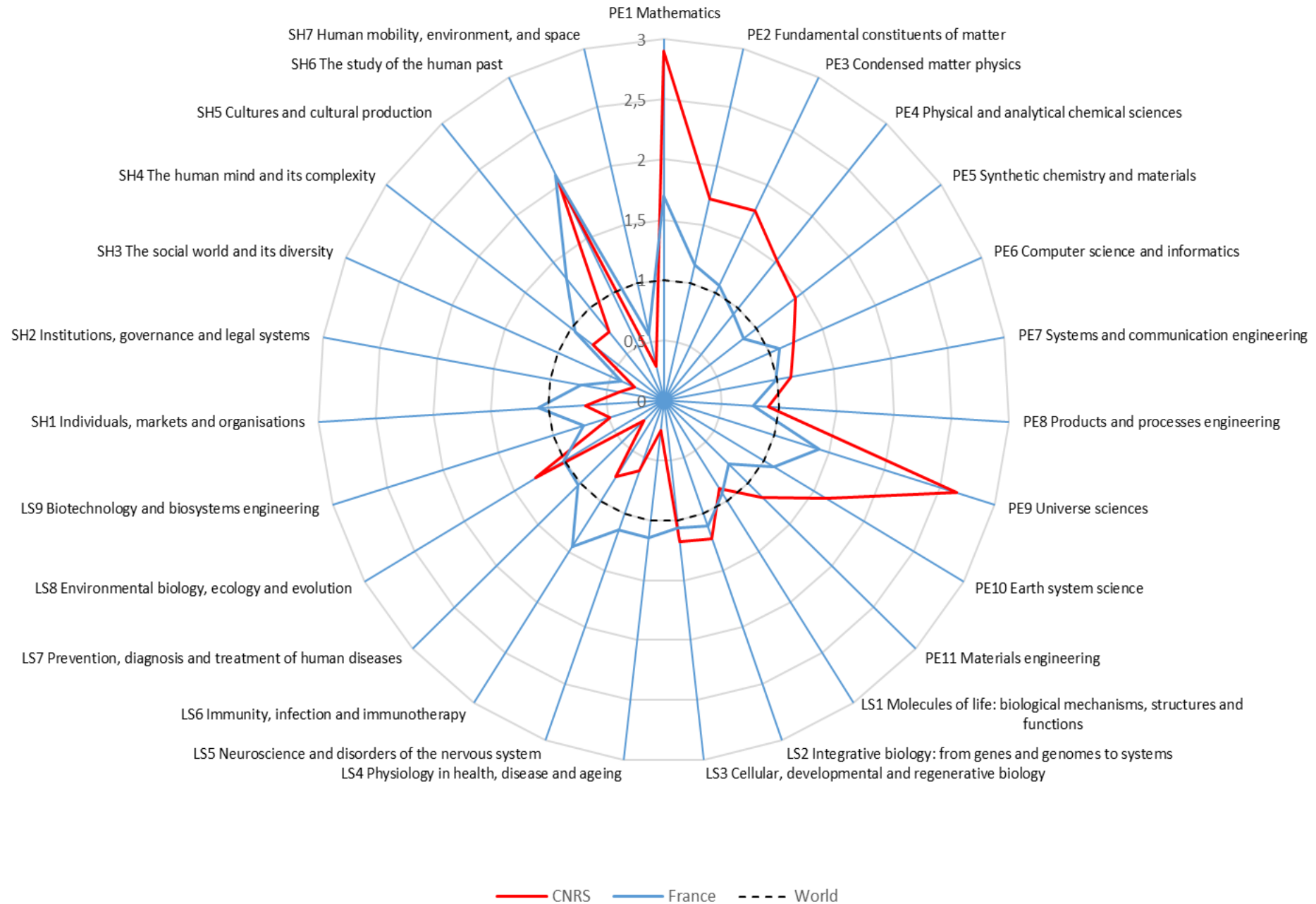


Specialisation* in the field of cancer, first 10 countries

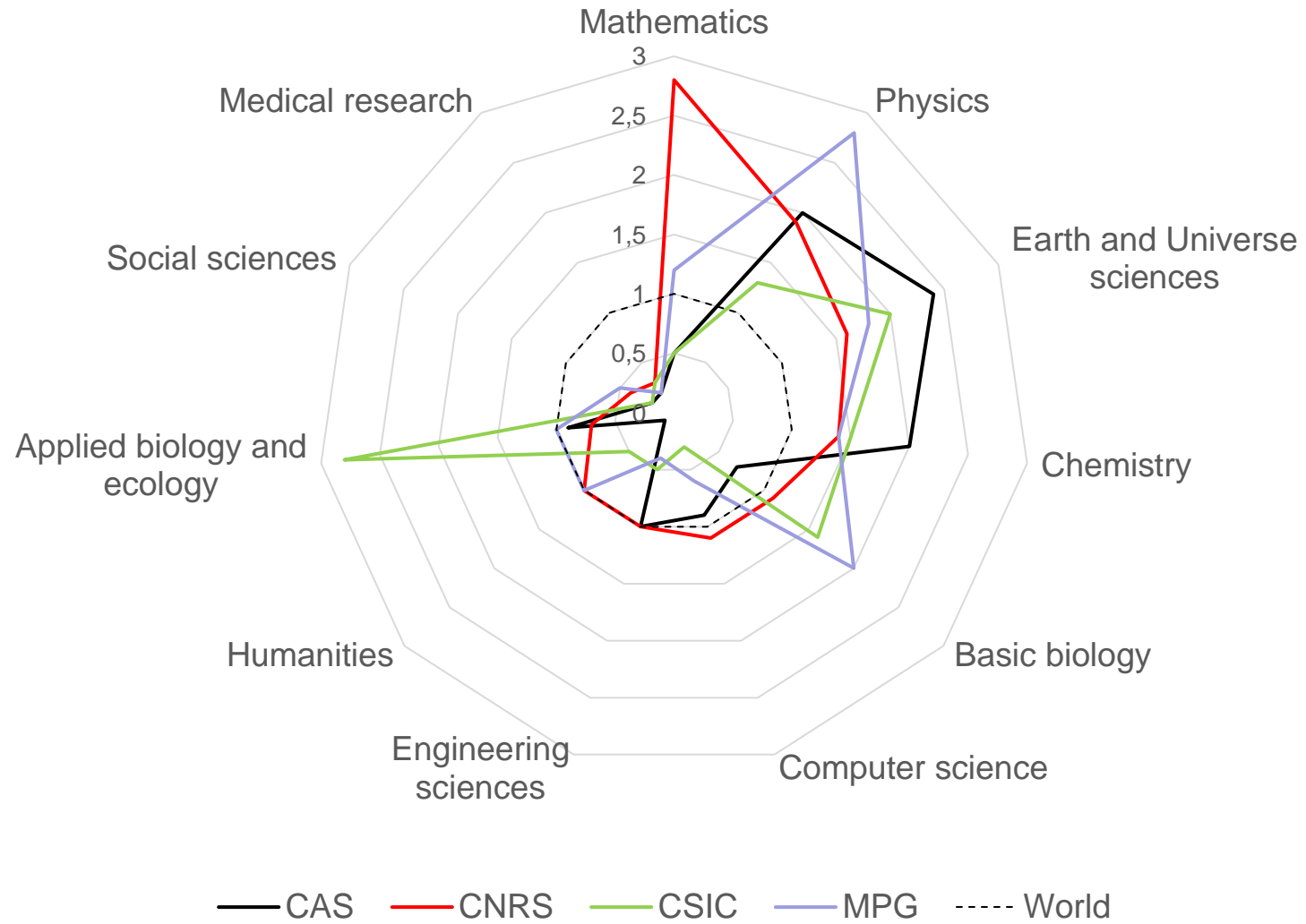


* Share cancer-country / Share cancer-world (fractional count)

Specialization index of CNRS and France by ERC panel, fractional count. 2017-21

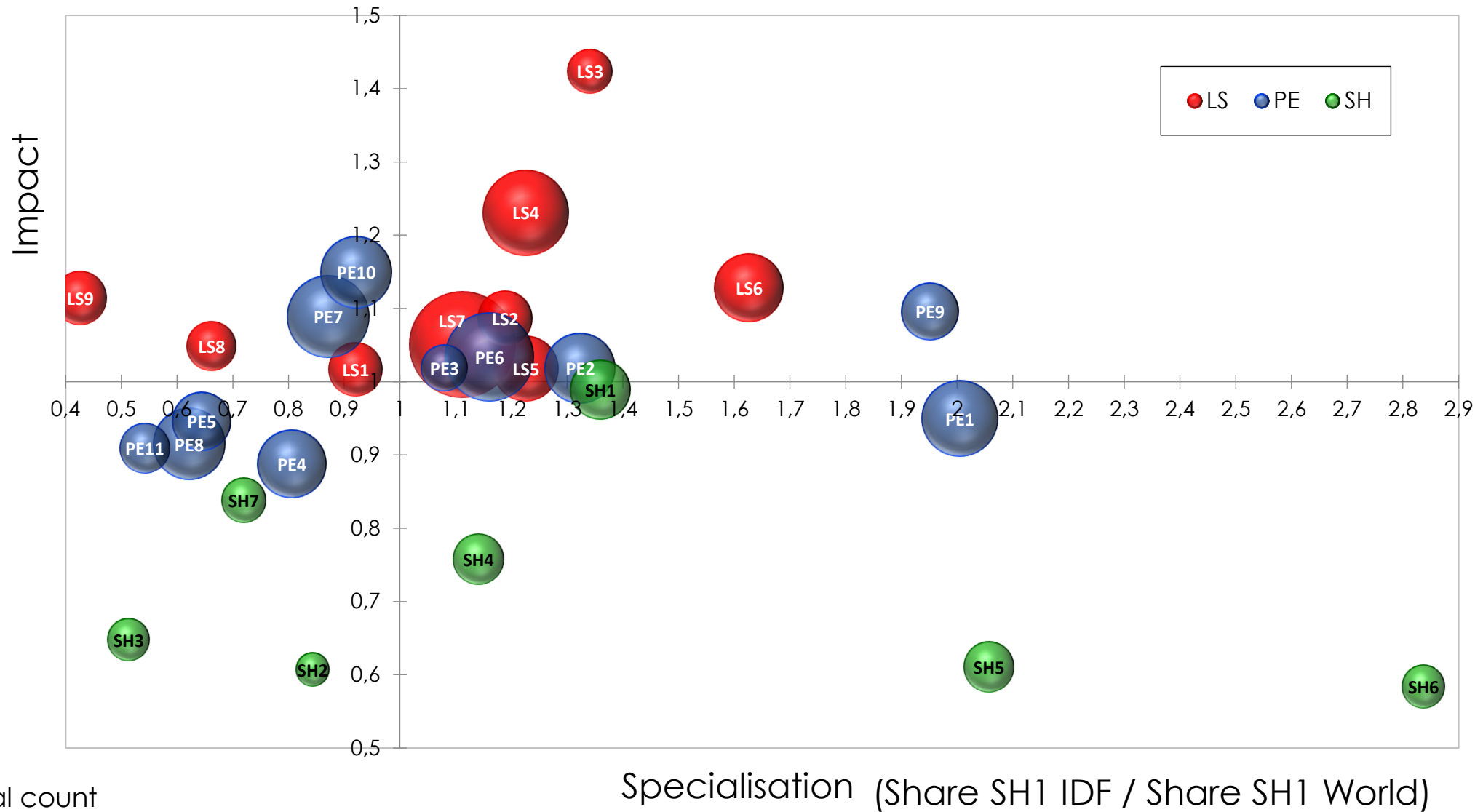


DISCIPLINARY PROFILE: CNRS, MPG, CSIC, CAS, 2020



Fractional count

Specialisation and impact of Île-de-France publications. ERC panels. 2018-19



Fractional count

Specialisation (Share SH1 IDF / Share SH1 World)

2. THE ISSUE OF THE DATABASE(S)

—



In France the issue has been raised for two types of disciplines

- Computer science
Validation of the WoS list of journals and proceedings
(Maths)
- Social sciences and humanities
 - Literature review : WoS + ESCI \approx Scopus < national repositories
 - Larger databases, such as Dimensions or OpenAlex
 - National archives/repositories

3. OBSERVATION OF THE FRENCH SCIENTIFIC PUBLICATIONS IN HAL AND THE WOS



CORPUS HAL

Extraction from HAL 14 sept. 2022: 3.1 M notices, 2.8 M uniques.

Analysis 2017-21. 800 K notices

- IDs to match with other data sources
- construction d'une nomenclature disciplinaire commune à HAL et WoS à partir des domaines/spécialités d'indexation dans chacune

SSH Journal articles in HAL, 2020

Data sources that observe the editorial process	# journals of notices in HAL SHS 2020	Share of HAL SHS 2020 journals
Ulrich + « refereed » = Y	3 257	56%
SCOPUS	2 432	42%
WoS (inc. ESCI)	2 265 (487)	39 % (8 %)
BONA FIDE	1 105	19%
Directory of Open Access Journals / DOAJ	774	13%
<i>Indexed by at least 1 source avec év. par les pairs</i>	3 661	63%
# Total journals	5 777	80 % or more in other disciplines

Notices in HAL, by type

	Type of notice	Share 2017-2021
	Articles	
	All Journal Articles	49.1%
	<i>with peer review (PR)¹</i>	38.2%
	Conferences	
	All communications	20.9%
	<i>with proceedings</i>	6.5%
	Poster	2.0%
	Books	
	Book chapters	8.6%
	Book	1.5%
	Book editing	1.5%
	Others	
	Ph. D. thesis	5.2%
	Memo	3.4%
	Others ²	3.8%
	Preprint. Working paper	2.7%
	Total	100%

Publications

Journal art. + proceedings =
83% of publications

Proceedings = 31 % of communications

1. Journals indexed in at least one of the following: DOAJ. Bona Fide. WoS (including ESCI). Scopus. Ulrich "refereed".
2. For example: reports. HDR dissertation (*Habilitation à diriger des recherches*). images. sound. video. patents.

Publications indexed in HAL, WoS or both

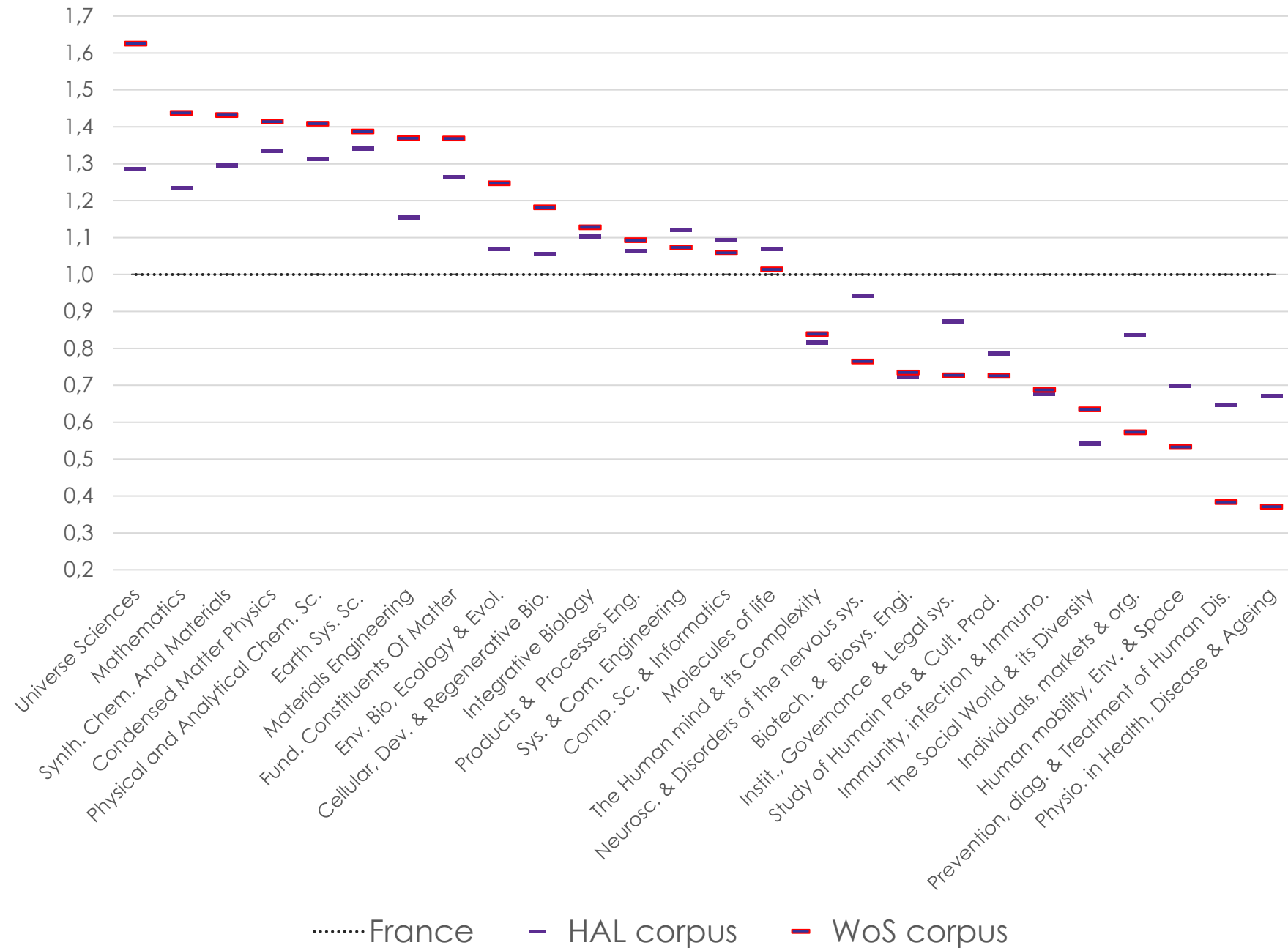
Corpus in HAL	Indicators	France		CNRS	
		HAL	WoS	HAL	WoS
1. All productions	Number	746 628	570 221	446 851	292 265
	Ratio : HAL / WoS	1.31	1.00	1.53	1.00
2. All publications and contributions to conferences	Number	598 023	478 774	369 939	271 897
	Ratio : HAL / WoS	1.25	1.00	1.36	1.00
3. Journal articles and conference proceedings	Number	522 088	478 774	328 373	271 897
	Ratio : HAL / WoS	1.09	1.00	1.21	1.00
4. Journal articles PR and conference proceedings	Number	351 419	478 774	243 462	271 897
	Ratio : HAL / WoS	0.73	1.00	0.90	1.00

The case of SSH

Productions by type and source, France and CNRS, 2017-21

	HAL (specialisation CNRS : 4/3)				WoS (6/5)	
	1. Articles, conferences, books	2. Articles. Proc.	3. Articles PR. Proc	4. Articles PR. Proc CNRS	5. Articles. Proc	6. Articles. Proc CNRS
Human past and cultural prod.	15.9%	11.5%	6.7%	5.3%	1.7%	1.2%
Instit.. gouv. et Judiciary sys.	10.1%	8.9%	2.6%	2.3%	0.3%	0.2%
Immunity, infection. immunotherapy	9.1%	10.1%	13.1%	8.9%	4.3%	2.9%
Material sc.	5.5%	6.1%	7.1%	8.2%	3.1%	4.1%
Individuals, markets & org.	4.6%	4.6%	3.8%	3.2%	2.7%	1.5%
Computer sc. & info. sys.	4.4%	4.8%	5.2%	5.7%	5.7%	6.0%
Sys. Engineering & communication	4.4%	4.9%	5.3%	6.0%	5.9%	6.3%
Synth. Chemicals & materials	3.8%	4.2%	5.2%	6.8%	4.1%	5.8%
Environ. Bio, ecology & evolution	3.8%	4.1%	4.5%	4.8%	3.8%	4.8%
Earth Sc.	3.5%	3.9%	4.9%	6.6%	6.8%	9.4%
Human mob. env. Space	3.4%	3.3%	3.0%	2.1%	1.3%	0.7%
Universe sc.	3.4%	3.8%	5.0%	6.4%	3.3%	5.5%
Human mind and its complexity	3.2%	2.9%	2.3%	1.9%	1.6%	1.3%
Social world and its diversity	3.0%	2.7%	1.6%	0.8%	0.8%	0.5%
Fund. constituents of matter	2.8%	3.2%	3.7%	4.7%	4.5%	6.2%
Mathematics	2.4%	2.5%	3.2%	4.0%	4.6%	6.6%
Products and process engineering	2.1%	2.3%	2.5%	2.6%	5.5%	6.0%
Physio., physiopathology, aging	1.9%	2.2%	2.9%	1.9%	7.2%	2.6%
Physical & anal. chemistry	1.8%	2.1%	2.6%	3.5%	5.4%	7.5%
Physics of condensed matter	1.8%	2.1%	2.4%	3.2%	1.8%	2.5%
Biotech., engineering of biosys.	1.6%	1.8%	1.9%	1.4%	4.6%	3.4%
Prev., diag., treatment of human diseases	1.6%	1.8%	2.4%	1.6%	10.1%	3.9%
Neurosc. Nervous sys. troubles	1.5%	1.7%	2.5%	2.1%	3.7%	2.8%
Integrative biology	1.4%	1.5%	2.0%	2.2%	2.7%	3.0%
Biomolecules	0.9%	1.0%	1.3%	1.4%	2.5%	2.5%
.Cellular, dev. & regenerative bio.	0.7%	0.7%	1.0%	1.0%	1.3%	1.6%

CNRS specialisation indices compared to France in WoS and HAL



Full count : SDU vs Maths

— DISCUSSION



1. Method for evaluating the issue of coverage vs representativeness of WoS in SSH.
2. Issue/discussion in Germany
3. ...



hceres.com

 [@Hceres_](https://twitter.com/Hceres_)

 [Hcéres](https://www.youtube.com/Hceres)