Published in: Proceedings of ISSI 2015, Istanbul; 15th International Society of Scientometrics and Informetrics Conference (Editors: Albert Ali Salah, Yaşar Tonta, Alkım Almıla Akdağ Salah,Cassidy Sugimoto, Umut Al), p. 150-159.

The most-cited articles of the 21st century

Elias Sanz-Casado[1,2], Carlos García-Zorita[1,2] and Ronald Rousseau[3,4]

[1,2] *elias.sanz@uc3m.es; czorita@bib.uc3m.es*
[1] Department of Library and Information Science. Laboratory of Metric Studies on Information (LEMI). Carlos III University of Madrid. C/Madrid 126, Getafe, 28903 Madrid, Spain
[2] Research Institute for Higher Education and Science (INAECU).
Carlos III University of Madrid-Autonomous University of Madrid. C/Madrid 126, Getafe, 28903 Madrid, Spain

[3,4] *ronald.rousseau@kuleuven.be*
[3] KU Leuven, Department of Mathematics, Celestijnenlaan 200B, B-3000 Leuven (Heverlee), Belgium
[4] Universiteit Antwerpen, IBW, Venusstraat 35, B-2000 Antwerpen, Belgium

Abstract
The aim of this paper is to collect the most-cited articles of the 21st century and to study how this group changed over time. Here the term “most-cited” is operationalized by considering yearly h-cores in the Web of Science. These h-cores are analysed in terms of authors, research areas, countries, institutions, journals and average number of authors per paper. We only consider publications of article or proceedings type. The research of some of the more prolific authors is on genetics and genomes publishing in multidisciplinary journals, such as *Nature* and *Science*, while the results show that writing a software tool for crystallography or molecular biology may help collecting large numbers of citations. English is the language of all articles in any h-core. The core institutions are largely those best placed in most rankings of world universities. Some attention is given on the relation between h-core articles and the information sciences. We conclude by stating that the notion of an h-core provides a new perspective on leading countries, articles and scientists.

Conference Topic
Citation and co-citation analysis

1. Introduction
The objective of this paper is to collect the most-cited articles of the 21st century and to study how this group changed over time. The term “most-cited” is operationalized by considering the h-core (Hirsch, 2005; Rousseau, 2006) in the Web of Science (WoS) for each period of time, starting with the period 2001-2005, continuing with 2001-2006 and ending with 2001-2013. These periods refer to the publication and the citation window. We recall that the h-core at a given moment in time, for instance on January 1, 2009, consists of the set of articles which at that time received a number of citations at least equal to their rank among all articles published during the period 2001-2008. This approach is different from the one taken in (Van Noorden et al., 2014) where a fixed number, concretely 100, of articles is considered. Furthermore, we study the papers making up the corresponding h-cores in terms of authors, research areas, countries, institutions, journals and average number of authors per paper.

2. Methodology
We have to point out that the 21st century starts on January 1, 2001. This implies that we only consider publications from 2001 on. Moreover, we only consider publications in Thomson Reuters’ Web of Science (WoS) and we restrict ourselves to publications of article or proceedings type.

Although finding today's h-core for a set of articles in the Web of Science is easy, finding an h-core in the past needs some specific knowledge of the tools available in the WoS. First one retrieves the set for which one wants to determine the h-core (ending in the year Y). Its articles are ranked from most cited to least cited. These are collected as a marked list. This is possible for at most 5,000 items. Clicking on Marked List shows this list and now, on this page, the system can provide a Citation Report, which is downloaded as an Excel file showing yearly citations for each of these records. Now we add the same data for the next 5,000 items (more was not necessary for our investigation). In this Excel file, we remove the columns corresponding to the year Y+1 and all later ones. In a next step we sum all remaining citations of each article. Sorting these sums from highest to lowest and comparing with a column of natural numbers leads to the h-index and the h-core. More details of this procedure are provided in (Rousseau & Zhang, 2014).

3. Results

3.1 The most-cited papers

The most-cited articles over the period 2001-2013 (the latest h-core) are shown in Table 1. It is clear that writing a software tool for crystallography or molecular biology may give one's paper a huge boost. The article by the National Cholesterol Education Program Expert Panel (2001) was the most-cited one from 2005 till 2008. From the year 2009 on Sheldrick's became the most-cited one.

Table 1. Most-cited articles over the period 2001-2013.

Rank	Article cited	Times cited
1	Sheldrick, G.M. (2008). A short history of SHELX. *Acta Crystallographica Section A*, 64, 112-122.	34,533
2	Livak, K.J. & Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods*, 25(4), 402-408.	24,796
3	Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24(8), 1596-1599.	17,049
4	Novoselov, K.S., Geim, A.K., Morozov, S.V., Jiang, D., Zhang, Y., Dubonos, S.V., Grigorieva, I.V. & Firsov, A.A. (2004). Electric field effect in atomically thin carbon films. *Science*, 306(5696), 666-669.	12,512
5	Ronquist, F. & Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12), 1572-1574.	11,185
6	National Cholesterol Education Program Expert Panel (Group author; includes 28 members). (2001). Executive summary of the Third Report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III). *JAMA-Journal of the American Medical Association*, 285(19), 2486-2497.	11,160
7	Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D – Biological Crystallography*, 60(special issue 1), 2126-2132.	10,392
8	Huelsenbeck, J.P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17(8), 754-755.	10,317
9	Spek, A.L. (2003). Single-crystal structure validation with the program PLATON. *Journal of Applied Crystallography*, 36, 7-13.	9,920
10	Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, 5(2), 150-163.	9,175

3.2. Time evolution of h-index and h-cores

The difference between the h-index and the number of items in the h-core is due to the possible existence of more than one document with the same number of citations as the h-index, as illustrated in Table 2. For the year 2005, for example, there were five articles with 359 citations.

Table 2. H-indices and h-cores for the periods 2001-2005 till 2001-2013

End year	h-index	# articles in the h-core
2005	359	363
2006	441	442
2007	526	527
2008	614	616
2009	704	704
2010	800	800
2011	902	902
2012	1014	1014
2013	1122	1122

It is obvious that only a small percentage of articles included in the WoS belongs to the h-core of a specific period. In order to show the evolution of the ratio of the h-core with respect to all articles we put their values for the period 2001-2004 equal to 100. Figure 1 shows the total number of papers in each period and the number of papers in each h-core when this rescaling has been performed. Linear regression is almost perfect for the two lines: all publications (R^2= 0,9982) and h-core (R^2= 0,9967). For this reason we can forecast the 21st century h-index for, at least, the next years to come. This would lead to an h-core of 1195 documents in 2014 and 1290 in the year 2015.

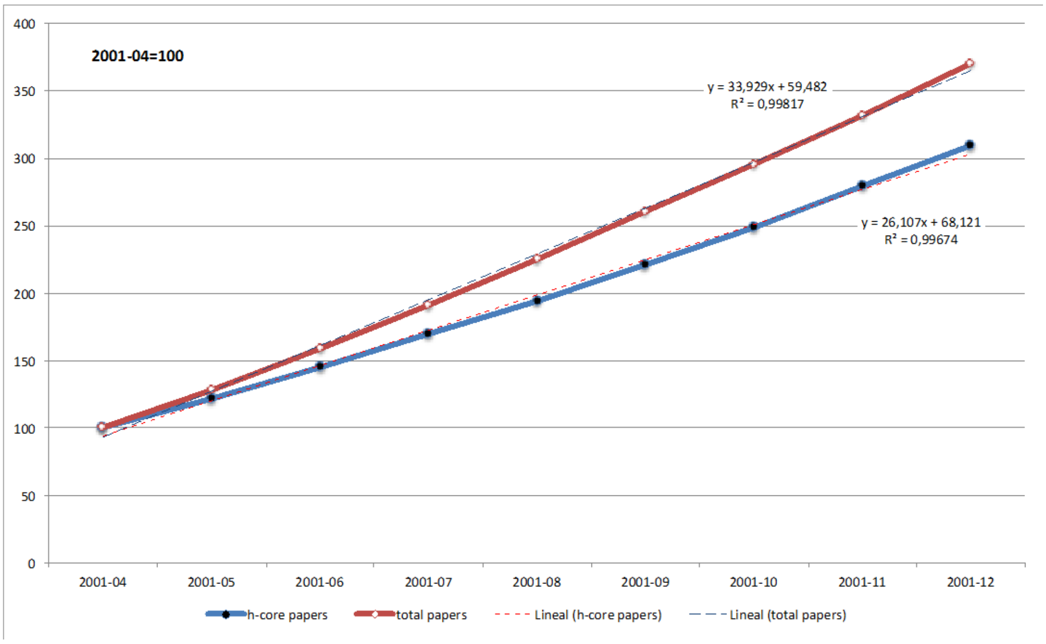

Figure 1. Evolution of the h-core

In Table 3, we show the number of articles published in the years 2001 to 2011 included in each of the h-cores. For each h-core these numbers follow the order of publication, i.e. most articles are published in the year 2001 and least in the latest year included in the core. Core13 has exactly the same number of articles published in 2001 as in 2002 (209 articles), while it does not contain articles published in 2013.

Table 3. Evolution of h-cores

Year of Publication	Core-05	Core-06	Core-07	Core-08	Core-09	Core-10	Core-11	Core-12	Core-13
2001	196	210	218	217	217	213	213	209	209
2002	116	137	158	173	187	197	201	205	209
2003	43	72	96	120	138	151	159	163	169
2004	7	21	41	62	82	99	117	138	146
2005	1	2	11	31	49	74	93	110	121
2006			3	9	17	35	56	70	95
2007				3	10	23	36	47	58
2008				1	3	6	19	39	54
2009					1	2	6	21	32
2010							2	9	19
2011								3	8
2012									2
Total	363	442	527	616	704	800	902	1014	1122

Table 4 shows the number of articles in the h-core (on the diagonal) and on the last line the number of unique articles in the union of all h-cores until the year indicated on top of the column. The other numbers refer to the number of articles originally belonging to the core referred to on the left, but which do not anymore belong to the h-core. We note that there is one article that left the core (in 2007) but re-entered (in 2008) and from then on stayed in the core. This paper is:

Minokoshi, Y., Kim, Y., Peroni, O., Fryer, L., Muller, C., Carling, D., & Kahn, B. (2002). Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *NATURE*, *415* (6869), 339–343. doi:10.1038/415339a

Table 4. H-cores and h-core losses

	Core-05	Core-06	Core-07	Core-08	Core-09	Core-10	Core-11	Core-12	Core-13
Core-05	363	9	9	9	9	9	9	9	9
Core-06		442	13	12	12	12	12	12	12
Core-07			527	17	17	17	17	17	17
Core-08				616	15	15	15	15	15
Core-09					704	26	26	26	26
Core-10						800	27	27	27
Core-11							902	24	24
Core-12								1014	22
Core-13									1122
Total	363	451	549	654	757	879	1008	1144	1274

3.3. H-cores characteristics

All articles in any h-core are written in English. We note that the 2001-2005 h-core contains one article that was later retracted (Chang and Roth, published in *Science*, which has now 533 citations and had 359 citations by the end of 2005, being the last one in the 2005 core). Some of the more prolific authors (E.S. Lander, M.J. Daly, R.A. Gibbs, J. Wang), see Table 5, perform research on genetics and genomes publishing in multidisciplinary journals, such as *Nature* and *Science*, often in hyper co-authored papers (with dozens and even hundreds of authors). A. Jemal and E. Ward publish yearly statistics on cancer, which all enter the h-core. R. Collins and R. Peto work on internal medicine and publish almost exclusively in *Lancet*. The fields of nanotechnology and graphene research are represented by C.M. Lieber and Nobel prize winners A.K. Geim and K.S. Novoselov (Table 5).

Table 5. Authors with highest number of papers in the h-core (Authors with more than 7 papers in the latest core).

Author	Core-05	Core-06	Core-07	Core-08	Core-09	Core-10	Core-11	Core-12	Core-13
Lander, ES	11	13	14	15	16	17	17	19	18
Wang, J	7	8	8	9	9	10	10	14	14
Jemal, A	4	4	5	6	7	8	9	10	12
Collins, R	5	6	7	8	9	11	11	11	10
Daly, MJ	4	5	6	6	7	10	10	12	10
Peto, R	4	5	7	8	8	9	9	9	10
Lieber, CM	5	6	7	7	7	8	8	9	10
Ward, E	3	3	4	5	6	7	8	9	10
Gibbs, RA	2	3	3	3	3	4	4	11	10
Geim, AK				3	3	5	6	8	10
Novoselov, KS				3	3	5	6	8	10
Thun, MJ	5	5	6	7	8	9	9	9	9
Altshuler, D	4	4	5	5	6	8	8	10	9
Abecasis, GR	2	2	2	2	2	4	4	9	9
Golub, TR	4	5	6	8	8	9	9	8	8
Murray, T	5	5	6	7	8	8	8	8	8
Gabriel, SB	3	3	3	3	4	5	5	9	8
Li, Y	1	2	3	3	4	7	7	8	8
Bartel, DP	1	1	1	3	3	4	4	7	8

The multidisciplinary, Science & Technology areas (which include journals such as *Nature*, *Science* and *PNAS*), and the ones related to General and Internal Medicine (such as *Lancet* or the *New England Journal of Medicine*) occur the most in each of the cores, as illustrated in Table 6.

Table 6. H-cores in different research areas (Areas with more than 10 papers in the last core)

Research area	Core-05	Core-06	Core-07	Core-08	Core-09	Core-10	Core-11	Core-12	Core-13
Science & Technology - Other Topics	39,1%	38,0%	35,3%	34,9%	32,8%	33,4%	32,7%	32,0%	31,9%
General & Internal Medicine	27,8%	26,2%	26,4%	25,0%	24,6%	23,1%	21,6%	20,4%	20,0%
Biochemistry & Molecular Biology	8,3%	9,0%	8,3%	9,7%	10,1%	10,6%	11,4%	12,8%	13,3%
Physics	5,5%	5,0%	4,9%	4,5%	5,0%	5,5%	6,4%	6,9%	7,0%
Chemistry	0,8%	1,4%	2,1%	1,9%	3,0%	3,9%	5,3%	6,0%	6,1%
Computer Science	2,5%	3,6%	4,7%	4,2%	4,5%	4,5%	5,1%	5,3%	5,5%
Cell Biology	4,1%	4,3%	4,0%	4,5%	4,5%	5,0%	5,2%	5,3%	5,1%
Engineering	1,4%	1,6%	3,0%	3,4%	3,6%	3,5%	3,8%	3,6%	3,9%
Biotechnology & Applied Microbiology	2,2%	3,4%	2,8%	3,1%	3,4%	3,1%	3,3%	3,8%	3,8%
Materials Science	0,6%	0,7%	0,8%	0,8%	1,7%	2,1%	3,0%	3,4%	3,8%
Oncology	2,8%	2,3%	2,3%	2,4%	2,7%	2,9%	2,5%	2,6%	2,9%
Genetics & Heredity	3,6%	3,4%	3,4%	3,2%	3,7%	3,4%	3,2%	3,3%	2,8%
Mathematics	0,8%	1,8%	1,7%	1,8%	1,7%	1,8%	2,0%	2,5%	2,7%
Mathematical & Computational Biology	0,8%	2,0%	1,7%	1,9%	1,8%	1,8%	1,8%	2,4%	2,4%
Research & Experimental Medicine	3,0%	3,2%	3,4%	3,2%	3,1%	2,9%	2,5%	2,5%	2,2%
Crystallography	0,8%	0,7%	0,9%	1,1%	1,3%	1,5%	1,6%	1,8%	2,0%
Neurosciences & Neurology	0,3%	0,7%	0,8%	0,8%	0,9%	1,4%	1,4%	1,9%	2,0%
Astronomy & Astrophysics	2,5%	2,9%	2,5%	2,3%	2,1%	2,1%	1,9%	1,9%	1,6%
Cardiovascular System & Cardiology	1,4%	1,8%	1,9%	1,8%	1,6%	1,4%	1,4%	1,5%	1,5%
Evolutionary Biology	0,0%	0,2%	0,2%	0,5%	0,7%	0,9%	1,2%	1,4%	1,5%
Immunology	2,8%	3,2%	3,2%	2,4%	2,7%	2,1%	1,8%	1,6%	1,3%
Biophysics	0,0%	0,0%	0,4%	0,5%	0,9%	1,0%	1,0%	1,1%	1,3%
Environmental Sciences & Ecology	0,3%	0,5%	0,4%	0,2%	0,4%	0,9%	0,9%	1,1%	1,3%
Radiology, Nuclear Medicine & Medical Imaging	0,0%	0,2%	0,4%	0,5%	0,6%	0,8%	1,0%	1,2%	1,2%
Endocrinology & Metabolism	1,4%	1,1%	1,1%	1,5%	1,4%	1,3%	1,1%	1,0%	1,1%

Table 7 shows a list of most used sources, where we observe, together with the mentioned multidisciplinary journals, the presence of medicine-related journals, including the specialized journal, *CA-A Cancer Journal for Clinicians*, whose presence is due to the systematic publication of the highly-cited annual statistics on cancer (all of them are in core 13). Other journal in the top positions, such as *Physical Review Letters* or *Nature Materials* occur less frequently.

Table 7. Journals of h-core publications (sources with 10 or more papers)

Source Titles	Core-05	Core-06	Core-07	Core-08	Core-09	Core-10	Core-11	Core-12	Core-13
NATURE	19,6%	17,4%	15,6%	15,9%	14,6%	14,9%	14,6%	14,4%	13,9%
SCIENCE	15,2%	16,1%	15,6%	15,1%	14,1%	14,0%	13,4%	12,9%	12,7%
NEW ENGLAND JOURNAL OF MEDICINE	16,5%	15,4%	15,2%	14,9%	14,8%	14,0%	13,1%	12,1%	11,9%
LANCET	5,2%	5,0%	5,1%	4,5%	4,4%	4,4%	4,1%	3,7%	3,6%
JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION	5,5%	5,2%	5,1%	4,7%	4,4%	3,9%	3,5%	3,3%	3,1%
PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA	3,6%	3,8%	3,6%	3,4%	3,3%	3,5%	3,4%	3,2%	3,1%
CELL	0,8%	0,7%	0,9%	1,5%	1,8%	2,4%	2,7%	2,9%	2,9%
NUCLEIC ACIDS RESEARCH	3,3%	2,9%	2,5%	2,8%	2,4%	2,1%	2,2%	2,5%	2,6%
BIOINFORMATICS	0,8%	1,6%	1,3%	1,1%	1,1%	1,1%	1,1%	1,5%	1,6%
PHYSICAL REVIEW LETTERS	3,6%	2,5%	2,3%	1,8%	1,6%	1,1%	1,4%	1,5%	1,4%
CA-A CANCER JOURNAL FOR CLINICIANS	1,4%	1,1%	1,3%	1,3%	1,3%	1,3%	1,2%	1,2%	1,2%
NATURE MATERIALS	0,0%	0,0%	0,0%	0,2%	0,6%	0,9%	1,2%	1,4%	1,2%
ACTA CRYSTALLOGRAPHICA SECTION D-BIOLOGICAL CRYSTALLOGRAPHY	0,0%	0,0%	0,4%	0,5%	0,7%	0,9%	0,9%	1,0%	1,2%
NATURE MEDICINE	1,7%	1,8%	1,5%	1,6%	1,6%	1,5%	1,4%	1,3%	1,2%
CIRCULATION	1,1%	1,4%	1,5%	1,3%	1,1%	1,0%	0,9%	1,0%	1,1%
IEEE TRANSACTIONS ON INFORMATION THEORY	0,8%	0,7%	1,3%	1,1%	1,1%	1,1%	1,0%	0,9%	1,0%
JOURNAL OF CLINICAL ONCOLOGY	0,8%	0,7%	0,6%	0,8%	0,9%	1,0%	0,8%	0,7%	0,9%
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY	0,0%	0,2%	0,6%	0,5%	0,6%	0,6%	0,9%	1,0%	0,9%
NATURE GENETICS	2,5%	2,0%	1,9%	1,6%	1,8%	1,6%	1,4%	1,3%	0,9%

We observe that the shares of the top journals such as *Nature*, *Science* and the *NEJM* are slowly declining over the years, while the share of *Cell* is increasing. This corresponds with recent findings (Lozano et al., 2012; Larivière et al., 2014; Acharya et al., 2014) that more and more highly-cited publications are published in journals that do not have the highest impact factors, say "non-elite journals". Of course, this is as such not surprising as the number of publications world-wide increases faster than the publication opportunities provided by so-called elite journals.

In table 8 we show the distribution of countries in the h-cores, where an article is classified as belonging to a country if at least one author has an address in this country. The first place goes to the USA. If, however, we consider the European Union (EU-28) as one entity then it leads the rankings in all except one year. Our results correspond to those obtained by King (2004) for the percentage of documents published by USA in the 1% most cited papers. Our results are also similar to those found by Leydesdorff et al. (2014). In their work the EU-28 gains gradually in the top-10% segment at the expense of the USA, and one can expect a cross-over between the EU28 and the USA in the near future within the top-10% segment. However, the distance between the U.S. and the EU is much larger in the top-1% segment.

Also here we see that the top performers (USA, EU-28 and Germany) lose in the share of h-core articles. This observation also holds for the Netherlands and most Scandinavian countries. England and Scotland consolidate their share, while Brazil and New Zealand show an increase. Although China's share in publications shows an exponential growth (Jin & Rousseau, 2005; Zhou & Leydesdorff, 2006, 2008) its share in h-core papers is much lower and shows at best a small increase in the latest years, after a decrease in the period 2008-2009.

Table 8. Countries of publication (with 10 or more papers in the latest core).

Countries	Core-05	Core-06	Core-07	Core-08	Core-09	Core-10	Core-11	Core-12	Core-13
European Union	78,8%	76,9%	76,5%	76,8%	73,6%	75,3%	73,8%	75,8%	76,0%
USA	75,2%	75,1%	75,5%	74,8%	75,1%	74,5%	73,1%	72,0%	71,7%
England	18,2%	19,0%	17,5%	17,9%	17,3%	17,6%	17,1%	17,9%	17,8%
Germany	14,0%	13,6%	13,5%	12,5%	11,9%	12,0%	12,2%	12,2%	11,7%
France	8,5%	8,8%	9,1%	9,3%	8,9%	8,8%	8,2%	8,3%	8,5%
Canada	9,9%	9,0%	8,7%	8,6%	8,1%	8,6%	8,0%	8,4%	8,3%
Japan	7,4%	8,8%	8,3%	8,3%	7,7%	7,9%	7,3%	7,8%	7,7%
Italy	5,8%	5,7%	6,1%	5,8%	5,5%	6,4%	6,3%	6,4%	6,1%
Switzerland	5,5%	4,8%	5,1%	5,2%	4,8%	5,1%	5,2%	5,6%	6,0%
Netherlands	6,9%	6,3%	5,7%	5,7%	5,5%	5,5%	5,1%	5,7%	5,8%
Australia	5,0%	5,2%	5,1%	5,4%	5,3%	5,4%	5,5%	5,3%	5,7%
Sweden	5,2%	5,4%	5,1%	5,4%	5,3%	5,3%	5,4%	5,5%	5,3%
Spain	3,6%	3,4%	3,6%	3,4%	3,0%	3,1%	3,2%	3,5%	3,8%
Belgium	4,1%	3,8%	3,6%	4,1%	4,0%	4,0%	3,7%	3,7%	3,7%
Scotland	2,8%	2,7%	3,2%	3,4%	3,3%	3,5%	3,3%	3,3%	3,1%
Denmark	3,6%	3,2%	3,0%	3,1%	2,8%	3,1%	3,0%	2,7%	2,8%
Finland	3,3%	2,7%	2,8%	2,3%	2,1%	2,3%	2,3%	2,6%	2,6%
Peoples R China	2,2%	1,8%	1,9%	1,5%	1,4%	1,8%	1,8%	2,5%	2,4%
Austria	2,2%	1,8%	1,9%	1,9%	1,8%	2,0%	2,2%	2,1%	2,1%
Israel	1,4%	1,6%	1,9%	1,6%	1,6%	1,6%	1,6%	1,7%	1,6%
Norway	1,7%	1,6%	2,3%	2,1%	1,8%	1,9%	1,8%	1,6%	1,5%
Russia	1,4%	0,7%	0,9%	1,1%	1,1%	1,3%	1,4%	1,5%	1,5%
South Korea	1,1%	0,9%	0,8%	1,0%	0,9%	0,9%	1,1%	1,4%	1,5%
Poland	1,1%	0,9%	0,8%	1,1%	1,4%	1,3%	1,3%	1,3%	1,4%
Ireland	1,4%	1,4%	1,5%	1,3%	1,3%	1,6%	1,3%	1,3%	1,3%
Brazil	0,8%	0,7%	0,8%	1,0%	1,0%	1,0%	1,1%	1,1%	1,2%
New Zealand	0,3%	0,5%	0,6%	0,6%	0,9%	1,0%	1,0%	1,0%	1,2%
Taiwan	1,1%	0,9%	0,9%	1,0%	1,0%	0,9%	0,7%	0,7%	0,9%

Core institutions are shown in Table 9. Leading institutions are those that one can find in most rankings of world universities, although The University of Texas (Austin) is only 39th in the latest ARWU ranking.

Table 9. Core institutions restricted to those with 25 or more papers in the latest core

Institution	Core-05	Core-06	Core-07	Core-08	Core-09	Core-10	Core-11	Core-12	Core-13
Harvard Univ	37	47	52	63	69	80	86	97	106
MIT	16	18	23	29	33	41	43	53	56
Univ Calif Berkeley	17	22	28	34	39	39	49	54	54
Univ Texas	11	16	20	25	30	35	39	41	45
Johns Hopkins Univ	12	17	19	26	29	34	33	40	43
Univ Washington	21	25	30	36	38	38	38	39	42
Univ Michigan	10	12	18	20	20	27	27	35	41
Univ Cambridge	11	13	16	20	22	26	29	34	39
Univ Oxford	15	14	16	18	19	24	27	34	39
Stanford Univ	15	21	24	24	26	26	33	37	38
Brigham & Womens Hosp	13	18	24	29	32	32	31	34	35
Univ Calif Los Angeles	13	19	19	20	21	24	26	28	35
Univ Calif San Diego	9	12	13	15	18	23	25	29	32
Columbia Univ	3	4	8	13	15	19	22	28	31
Massachusetts Gen Hosp	9	11	13	15	18	24	25	27	31
Univ Calif San Francisco	13	14	18	21	22	23	25	28	29
Univ Penn	13	13	14	15	17	19	19	25	26
Duke Univ	8	9	11	12	17	18	18	23	25
NCI	12	14	16	20	21	24	25	27	25
Univ Pittsburgh	7	9	11	16	16	18	19	22	25

In table 10 we have calculated average co-authorship values of articles in h-cores by research areas. For several research areas these values are higher than the co-authorship values of all publications: for example, in Clinical Medicine the co-authorship value for all publications was 4.5 authors per document and 5 in Bioscience and Biomedical Research (Bordons & Gómez 2000; Glänzel & Schubert, 2005). For several research areas these values are higher than the co-authorship values expected from previous research. For example, in Clinical Medicine the co-authorship value for all publications was 4.5 authors per document and 5 in Bioscience and Biomedical Research (Bordons & Gómez 2000; Glänzel & Schubert, 2005).

Table 10. Average numbers of authors for papers in the h-cores by research areas (areas with more than 10 papers in 2013)

Research Area	Core-05	Core-06	Core-07	Core-08	Core-09	Core-10	Core-11	Core-12	Core-13	Average
Science & Technology - Other Topics	15,5	16,1	14,6	13,9	14,7	14,5	14,5	17,0	15,9	15,3
General & Internal Medicine	19,8	20,4	23,4	25,6	24,2	25,9	22,7	22,1	22,1	23,1
Biochemistry & Molecular Biology	8,2	8,6	8,3	8,5	8,4	7,9	7,3	7,5	7,4	7,8
Physics	52,2	45,0	40,4	37,9	31,3	19,4	15,3	13,6	49,6	31,0
Chemistry	4,0	3,8	4,5	4,4	4,8	5,4	5,2	5,2	5,2	5,1
Computer Science	3,6	3,3	3,0	3,0	3,2	3,1	3,0	3,1	3,0	3,1
Cell Biology	11,4	11,8	11,7	10,9	10,8	10,7	10,2	11,1	11,1	10,9
Engineering	3,8	3,6	3,1	2,9	2,8	2,9	2,8	2,8	2,8	2,9
Biotechnology & Applied Microbiolog	6,8	5,9	7,0	7,4	7,4	6,5	6,0	5,6	5,4	6,2
Materials Science	4,5	3,3	6,5	5,6	5,0	5,2	5,6	5,8	6,3	5,7
Oncology	10,6	10,6	9,8	10,1	10,8	11,2	11,1	11,2	11,1	10,8
Genetics & Heredity	7,1	6,7	8,4	8,0	7,5	7,0	6,5	6,2	5,9	6,9
Mathematics	3,3	3,9	3,9	3,5	4,3	3,9	3,8	3,7	3,6	3,8
Mathematical & Computational Biolo	3,3	3,8	3,8	4,7	5,3	5,0	4,7	4,3	4,3	4,5
Research & Experimental Medicine	11,5	12,1	11,6	11,6	11,0	11,5	11,8	11,4	11,4	11,5
Crystallography	3,3	3,3	3,0	2,6	2,6	3,4	3,1	4,2	5,1	3,7
Neurosciences & Neurology	16,0	10,7	8,8	8,6	8,7	8,5	8,3	7,6	7,8	8,3
Astronomy & Astrophysics	41,8	30,7	30,7	37,3	35,9	37,5	38,8	46,5	45,8	38,8
Cardiovascular System & Cardiology	12,6	10,5	8,8	10,1	10,1	9,7	9,8	11,9	13,5	10,9
Evolutionary Biology		2,0	2,0	3,0	3,4	3,1	2,6	2,7	2,9	2,8
Immunology	8,1	7,3	7,2	7,4	7,5	7,7	7,6	7,6	7,8	7,6
Biophysics			2,5	2,3	3,3	4,0	3,8	5,1	5,9	4,5
Environmental Sciences & Ecology	7,0	4,0	4,0	7,0	5,0	3,1	2,9	2,6	2,8	3,2
Radiology, Nuclear Medicine & Medical Imagin		6,0	4,5	5,7	6,3	5,0	4,3	5,0	5,5	5,1
Endocrinology & Metabolism	7,2	7,2	6,8	8,4	6,9	6,9	6,9	6,6	5,9	6,9

Areas with an average of less than 5 authors (in 2013) are: computer science, engineering, mathematics, mathematical and computational biology, crystallography, evolutionary biology, biophysics and environmental sciences & ecology. Areas with an average larger than 15 are: science & technology – other topics, general & internal medicine, physics and astronomy & astrophysics.

3.4 The 21st century h-core (2001-2013) and the information sciences

Only one article classified by Thomson Reuters as *Information science and library science* belongs to this h-core, namely Venkatesh, V., Morris, M.G., Davis, G.B. et al. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, 27(3), 425-478 (cited 2261 times in total).

Yet, other ones were used and cited in *Information science and library science* articles. We list those that were cited at least 30 times by ILS researchers (on December 25, 2014).

1. Hirsch, J.E. (2005). An index to quantify an individual's research output. *Proceedings of the National Academy of Sciences of the USA*, 102(46), 16569-16572. Cited 682 times by ILS researchers.
2. Venkatesh, V., Morris, M.G., Davis, G.B. et al. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, 27(3), 425-478. Cited 595 times.
3. Newman, M.E.J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the USA*, 98(2), 404-409. Cited 118 times.
4. Blei, D.M., Ng, A.Y. & Jordan, M/I. (2003). Latent Dirichlet allocation. *Journal of Machine-Learning Research*, 3(4-5), 993-1022. Cited 93 times.
5. Zhara, S.A. & George, G. (2002). Absorptive capacity: a review, reconceptualization, and extension. *Academy of Management Review*, 27(2), 185-203. Cited 91 times.
6. Berners-Lee, T., Hendler, L. & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), 28-37. Cited 64 times.
7. Newman, M.E.J., Strogatz, S.H. & Watts, D.J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 62(2), article number 026118. Cited 60 times
8. Girvan, M. & Newman, M.E.J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the USA*, 99(12), 7821-7826. Cited 50 times.
9. Newmann. M.E.J. & Girvan, M. (2004). Finding and evaluating community structure in networls. *Physical Review E*, 69(2), article number 026113. Cited 36 times
10. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. Cited 30 times.

Besides Hirsch' famous article on the h-index (Hirsch, 2005), we see also Berners-Lee's article on the semantic web (Berners-Lee et al., 2001) and note the fact that Mark Newman occurs four times in this ILS h-core.

4. Conclusions

- Using the notion of an h-core provides a new perspective on leading countries, articles and scientists.
- The scientific contribution to the h-cores by the EU-28 is slightly higher than the USA's.
- The trend of annual h-cores since 2001 can predict future values of this indicator.

Of course, the view provided in this contribution is highly biased in favor of certain research areas such as General & Internal Medicine, or Biochemistry & Molecular Biology, and certain methodologies (writing heavily used software programs). Yet, it is a fact of life that these areas provide today's leading research. One should clearly realize that publishing highly cited research is different from realizing outstanding intellectual achievements.

Acknowledgements.
The authors thank Raf Guns (University of Antwerp) for helpful comments.

References

Acharya, A., Verstak, A., Suzuki, H., Henderson, S., Iakhiaev, M., Chiung, C., Lin, Y. & Shetty, N. (2014). Rise of the rest: the growing impact of non-elite journals. arXiv:1410:2217v1. 8 Oct 2014.

Bordons, M., & Gómez, I. (2000). Collaboration networks in science. In: B. Cronin and H.B. Atkins (eds), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*. Information Today Inc, Medford, NJ, 197–213.

Glänzel, W., & Schubert, A. (2005). Analysing scientific networks through co-Authorship. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 257–276). Dordrecht: Kluwer Academic Publishers.

Hirsch, J.E. (2005). An index to quantify an individual's research output. *Proceedings of the National Academy of Sciences of the USA*, 102(46), 16569-16572.

Jin, BH. & Rousseau, R. (2005). China's quantitative expansion phase: exponential growth but low impact. In: *Proceedings of ISSI 2005*. Edited by P. Ingwersen and B. Larsen. Karolinska University Press, Stockholm, p. 362-370.

King, D. A. (2004). The scientific impact of nations. *Nature*, *430*(6997), 311–316.

Larivière, V., Lozano, G.A., & Gingras, Y. (2014). Are elite journals declining? *Journal of the Association for Information Science and Technology*, 65(4), 649-655.

Leydesdorff, L., Wagner, C.S. & Bornmann, L. (2014). The European Union, China, and the United States in the top-1% and top-10% layers of most-frequently cited publications: Competition and collaborations. *Journal of Informetrics*, 8(3), 606-617.

Lozano, G.A., Larivière, V. & Gingras, Y. (2012). The weakening relationship between the impact factor and papers' citations in the digital age. *Journal of the American Society for Information Science and Technology*, 63(11), 2140-2145.

Rousseau, R. & Zhang, L. (2014). How to determine the h-index of a set of publications in the WoS? *ISSI Newsletter*, 10(3), 63-65.

Van Noorden, R., Maher, B. & Nuzzo, R. (2014). The Top-100 papers. *Nature*, 514(7524), 550-553.

Rousseau, R. (2006). New developments related to the Hirsch index. *Science Focus* 1(4), 23-25 (in Chinese). An English version is available at: http://eprints.rclis.org/7616/

Zhou, P. & Leydesdorff, L. (2006). The emergence of China as a leading nation in science. *Research Policy*, 35(1), 83-104.

Zhou, P. & Leydesdorff, L. (2008). China ranks second in scientific publications since 2006. *ISSI Newsletter*, 4(1), 7–9.