Preprint of:
Yannis Tzitzikas and Giorgos Dovas, How Co-authorship Affects the H-index?,
Scientometrics, 2024 (accepted for publication).
The Version of Record is available with DOI 10.1007/s11192-024-05088-y
and URL https://link.springer.com/article/10.1007/s11192-024-05088-y

How Co-authorship Affects the H-index?

Yannis Tzitzikas[1,2*] and Giorgos Dovas[2]

[1*]Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH), Voutes, Heraklion, 711 10, Crete, Greece.
[2]Computer Science Department, University of Crete, Panepistimioupoli Vouton, Heraklion, 70013, Greece.

*Corresponding author(s). E-mail(s): tzitzik@ics.forth.gr;
Contributing authors: georgez13@hotmail.gr;

Abstract

H-Index is a widely used metric for measuring scientific output. In this paper we showcase the weakness of this index as regards co-authorship. By ignoring the number of co-authors, each author gets the full credit of a joint work, something that is not fair for evaluation purposes. For this purpose we report the results of simulation scenarios that demonstrate the impact that co-authorship can have. To tackle this weakness, and achieve a more fair evaluation, we propose a few simple variations of H-index that consider the number of co-authors, as well as the active time period of a researcher. In particular we propose using HI/co and HI/(coy), two metrics that are simple to understand and compute, and thus they are convenient for decision making. The simulation shows that they can tackle well co-authorship. Subsequently we report measurements over real data of researchers coming from five universities (Cambridge, Crete, Harvard, Oxford and Ziauddin), as well as other datasets, that reveal big variations in the average number of co-authors. In total, we analyzed 526 authors, having in total more than 127 thousands publications, and 16.7 million citations. These measurements revealed big variations of the number of co-authors. Consequently, by including the number of co-authors in the measures for scientific output (e.g. through the proposed HI/co) we get rankings that differ significantly from the rankings obtained by citations, or by the plain H-Index. The normalized Kendall's tau distance of these rankings ranged from 0.28 to 0.46, which is quite high.

Keywords: H-Index, Research Output Evaluation

1 Introduction

H-Index [1] is a widely used bibliographic metric. It is a single positive integer that aims at considering both the quantity (number of publications), and quality (number of citations). It is heavily used in hiring, promotion, funding, and recognitions. However, a serious limitation of that metric is that it ignores the co-authors of the papers. By ignoring the number of co-authors, each author gets the full credit of a joint work, something that is unfair for evaluation purposes. Indeed, if F in number persons decide to jointly work and write a paper, then (a) they can create a better paper in the sense that more work can be dedicated to the paper (and thus it can have higher changes for acceptance), and (b) in case of acceptance this paper and its citations (and consequently its contribution to H-Index) will be added to the list of publications and citations of each of the F persons. In addition, quite likely this paper will get more citations not only through self-citations but also by citations from the network of the authors. Overall, and quite paradoxically, each of the F co-authors will get the full credit of that paper!

In this paper we showcase this weakness through simulations. Then we propose, and experimentally compare through simulation, a few simple variations of H-index that can tackle this weakness and lead to more fair evaluations. Apart from considering the number of co-authors, we include variations that include the time dimension (a quite overlooked one) since the active time period of a researcher does not only affect his cumulative scientific output (number of publications), but it also affects his citations. A rising question is how much co-authorship varies today. In order to answer this question we collect, extract and process the bibliographic data of the top-researchers (with respect to citations) of five universities (Cambridge, Crete, Harvard, Oxford and Ziauddin). In addition, and to avoid analyzing only top-researchers, we analyzed roughly all active researchers of two universities (University of Crete and University of Ziauddin), we analyzed the researchers of the schools of one university, all faculty members of a single department, as well as a list of famous scientists of the past. These measurements revealed big variations of the number of co-authors. Consequently, by including the number of co-authors in the measures for scientific output (e.g. through the proposed HI/co) we get rankings that differ significantly from the rankings obtained by citations, or by the plain H-Index. We quantify the differences in the obtained rankings using the normalized Kendall's tau distance. The distance values that we get are in most cases bigger than 0.3, meaning that 30% of the relative rankings are different if we consider the number of co-authors, something that is quite high.

The rest of this paper is organized as follows. Section 2 describes the background, i.e. H-Index, for short HI, Section 3 describes simulation scenarios that showcase the problems of HI. Section 4 introduces alternative metrics for tackling the problems of HI. Section 5 evaluates the behaviour of these metrics over the simulation scenarios. Section 6 reports results over researchers coming from five universities, as well over a list of famous researchers of the past. Finally, Section 7 summarizes and concludes the paper.

2 Background

2.1 The H-Index

H-Index was proposed by Hirsch in 2005 [1]. Let A be the set of authors, P be the set of papers, and for an author $a \in A$, let $papers(a)$ be the set of papers, where a occurs as author. For a paper $p \in P$, let $cits(p)$ be the number of papers that cite p. The H-index of an author a, let's denote it by $HI(a)$, is the maximum integer value K such that there are K papers of a each having at least K citations, i.e. $HI(a) = \max_K(|\{p \in papers(a) \mid cits(p) \geq K\}| = K)$.

2.2 Related Work

Several subsequent works are related to H-Index. Just indicatively, [2] compares the H-index for different fields of research, [3] elaborates on the growth of the H-index, through simulations (however it does not focus on co-authorship), the robustness of H-Index to self-citations is described in [4], while [5] focuses on H-index manipulation through self-citation.

As regards the importance of H-index in academy, [6] reported increasing H-index with consecutive academic rank throughout all medical fields. Analogously, [7] evaluates the association of H-index with academic ranking in the interventional radiology community, and found that H-index correlates significantly with faculty position.

In general, the pros and cons of H-Index have been described extensively in the literature, e.g. in [8]. Several metrics have been proposed as alternatives of H-Index, e.g. [9] proposed the g-index (and various extensions of that index have been proposed, e.g. in [10]). The paper [11] also favors the "fractional H-index", proposed in [12], that considers the number of authors, i.e. it gives to an author of an m-authored paper only a credit of c/m if the paper received c citations. The paper [13] proposes an alternative metric, $\bar{h}$, in order to quantify an individual's scientific research output that takes into account the effect of multiple co-authorship, however that metric, as noted also in [14] is harder to calculate (it requires co-authors' H-index), it penalises articles published with collaborative efforts, and may decrease after some time. Some interesting findings about the number and order of authors, as well as discipline-specific measurements over time, are given in [15]. A recent, and quite detailed, review of H-index and its alternative indices is available at [14].

Of course, quantitative metrics is not a panacea, they do have weak points and limitations, e.g. as described in [15]. That work analyzes the trends in current academic publishing and also mentions the issue of longer author lists, shorter papers, and surging publication numbers. However, as stated in [16], the uncritical dismissal of quantitative metrics may aggravate injustices and inequities, especially in nonmeritocratic environments; quantitative metrics could help improve research practices if they are rigorous, field-adjusted, and centralized [16].

Despite the aforementioned efforts and proposals, the *number of citations* and the *H-index*, remain to be the dominant methods which are used for the evaluation of research impact, and this is also evidenced by the default ranking that it is provided by the various bibliographic sources (e.g. Google Scholar). Indeed, two important merits of a metric, that affect its adoption, are *simplicity* and *easy computation*, in the sense that a community is hard to trust a metric that is not clear to all, or a metric that is difficult to compute.

In the current paper we focus on co-authorship. We demonstrate with simulations the importance of the problem, i.e. how co-authorship can affect the H-index, To investigate to what extend co-authorship varies in real data, we perform measurement over the researchers of five different universities. In comparison to [12] and [11], [12] elaborates on the mathematical lower and upper bounds of two versions of the fractional H-index and g-index. That work does not show the effect of co-authorship, and it does not report measurements. Also [11] favors fractional indexes, however it reports very few measurements (over 12 Nobel laureates).

Novelty. To the best of our knowledge, this paper is the first that showcases the impact of co-authorship through simulations, and reports the ranges of co-authorship encountered today. In particular, the measurements (over more than 127 thousand publications) revealed big variations of the number of co-authors, and big variations of the rankings obtained if we consider the number of co-authors (The normalized Kendall's tau distance of these rankings ranged from 0.28 to 0.46).

3 Simulation Scenarios

Suppose that each researcher can dedicate a fixed amount of effort per year, corresponding to the effort required for writing E in number papers. Let assume a modest value for E, e.g. $E = 3$. Now consider the following publication "policies" that a researcher can follow:

- R_{solo}: Here our researcher writes papers alone (and probably with one or two students).
- R_{fK}: Here our researcher has K friends, and whenever he writes a paper he adds his K *friend researchers* to the list of authors. His friends behave the same, i.e. they also add our researcher in the papers that they write. For example, R_{f2} means that our researcher has two friends, R_{f3} means that our researcher has three friends, and so on.

Number of Publications. It follows that each year R_{solo} appears as author of E papers, while R_{fF} (i.e. if the number of friends is F) appears as author in $E(1 + F)$ papers. In Table 1, we can see the cumulative publications per year ($Y = 1, \ldots, 10$), assuming $E = 3$, for R_{solo} and R_{fF} for $F = \{1, 3, 5\}$. Recall that each of these 4 researchers has dedicated exactly the same amount of effort. The first, R_{solo}, in 10 years appears in 30 papers, while R_{f5} appears in 180 papers! In an ideal evaluation system they should get the same score.

Citations. Suppose that every year a paper p receives $Cext$ external citations (not self citations), and $Cself$ self-citations by each of the authors. Therefore we can assume that the total citations of a paper p of a researcher R_{fF}, after Y years is given by $citations(p, Y, R_{fF}) = (Y - 1) * (Cext + (F + 1) * Cself)$. In Table 2 we can see

Year	R_{solo}	R_{f1}	R_{f3}	R_{f5}
1	3	6	12	18
2	6	12	24	36
3	9	18	36	54
4	12	24	48	72
5	15	30	60	90
6	18	36	72	108
7	21	42	84	126
8	24	48	96	144
9	27	54	108	162
10	30	60	120	180

Table 1 The impact of F on the number of papers published

Year	R_{solo}	R_{f1}	R_{f3}	R_{f5}
1	0	0	0	0
2	3	4	6	8
3	6	8	12	16
4	9	12	18	24
5	12	16	24	32
6	15	20	30	40
7	18	24	36	48
8	21	28	42	56
9	24	32	48	64
10	27	36	54	72

Table 2 The impact of F on the number of citations of a single paper

the citations of a paper assuming $Cext$=2 and $Cself$=1, for R_{solo}, R_{f1}, R_{f3} and R_{f5}. Notice the difference between 27 and 72. The first is the number of citations of a paper written by R_{solo} after 10 years, while the second is the number of citations of a paper written by R_{f5} after 10 years.

Let us now compute the total citations of our researchers. We can compute them using the algorithm shown in Alg. 1. In that algorithm we use $PpYear$ for the factor E.

Algorithm 1 Computation of Citations

```
function ComputeCitations(Years, F, PpYear, CExt, CSelf)
    Citations ← 0
    for y ← 1 to Years do
        pastPubls ← (y-1)*PpYear*(1+F)
        yearCits ← pastPubls * (CExt + F*CSelf);
        Citations ← Citations + yearCits
    end for
    return Citations;
end function
```

Years	PpYear	Friends	CExt	CSelf	Papers	Citations	H-Index
10	3	0	2	1	30	270	12
10	3	1	2	1	60	810	18
10	3	2	2	1	90	1620	27
10	3	3	2	1	120	2700	35
10	3	4	2	1	150	4050	42
10	3	5	2	1	180	5670	49

Table 3 The impact of F on papers, citations and H-index in a period of 10 years

This gives the numbers shown in Table 3. Notice, that even if these researchers have dedicated the same effort the last 10 years, R_{solo} has in total 270 citations while R_{f5} has 5670 citations! The difference is tremendous. However, we should note that the simulation is not very precise in the sense that we have not considered any limit to the number of references than a paper can have. However, some conferences/journals do not impose any limit to the number of references.

H-index. In order to compute the H-Index, for short HI, of our researchers, we first compute the citations of each of their papers (i.e. for each year we compute the citations to the publications published the previous years), and then we apply the formula given in §2.1. The results are shown in Table 3. Notice that R_{solo} has HI 12, while R_{f5} has HI 49. Again, the difference is tremendous.

Synopsis. In brief, we have seen very big differences in the number of publications, number of citations and HI when F is greater than one.

4 Towards more Fair Measures

Here we introduce and comparatively evaluate measures that can tackle the aforementioned weakness of HI. In particular we will comparatively evaluate the following measures:

1. HI as defined before, i.e. $HI(a) = \max_K(|\{p \in papers(a) \mid cits(p) \geq K\}| = K)$.
2. HI/co: We compute the HI as before, and we divide it by the average number of co-authors, i.e. $HI/co(a) = HI(a)/avgCoAuthors$, where $avgCoAuthors = avg\{\ authors(p) \mid p \in papers(a)\}$ This means that in the simulation scenarios of §3 we divide the HI by $F+1$.
3. HIdivCit: When computing HI we consider as number of citations of a paper its citations divided by the number of paper's authors, i.e.:
$HIdivCit(a) = \max_K(|\{p \in papers(a) \mid cits(p)/|authors(p)| \geq K\}| = K)$.
4. HI/(coy): Since publications and citations increase over the years, to compare two researchers of different scientific age, i.e. with different periods of research production, we have to consider the time dimension too. For this reason it makes sense to divide publications, citations, and HI/co, by the number of years y. In particular, we propose dividing HI/co by the active research years, i.e.: $HI/coy(a) = HI/co(a)/Y$ where Y is the active years of researcher a.

Years	PpYear	F	CExt	CSelf	Papers	Cits	HI	HI/co	HIdivCit	HI/(coy)
1	3	0	2	1	3	0	0	0.00	0	0.00
1	3	1	2	1	6	0	0	0.00	0	0.00
1	3	2	2	1	9	0	0	0.00	0	0.00
1	3	3	2	1	12	0	0	0.00	0	0.00
1	3	4	2	1	15	0	0	0.00	0	0.00
1	3	5	2	1	18	0	0	0.00	0	0.00
2	3	0	2	1	6	6	2	2.00	2	1.00
2	3	1	2	1	12	18	3	1.50	1	0.75
2	3	2	2	1	18	36	4	1.33	1	0.67
2	3	3	2	1	24	60	5	1.25	1	0.63
2	3	4	2	1	30	90	6	1.20	1	0.60
2	3	5	2	1	36	126	7	1.17	1	0.58
3	3	0	2	1	9	18	3	3.00	3	1.00
3	3	1	2	1	18	54	6	3.00	3	1.00
3	3	2	2	1	27	108	8	2.67	2	0.89
3	3	3	2	1	36	180	10	2.50	2	0.83
3	3	4	2	1	45	270	12	2.40	2	0.80
3	3	5	2	1	54	378	14	2.33	2	0.78
4	3	0	2	1	12	36	4	4.00	4	1.00
4	3	1	2	1	24	108	6	3.00	4	0.75
4	3	2	2	1	36	216	9	3.00	4	0.75
4	3	3	2	1	48	360	12	3.00	3	0.75
4	3	4	2	1	60	540	15	3.00	3	0.75
4	3	5	2	1	72	756	18	3.00	3	0.75
5	3	0	2	1	15	60	6	6.00	6	1.20
5	3	1	2	1	30	180	9	4.50	6	0.90
5	3	2	2	1	45	360	12	4.00	5	0.80
5	3	3	2	1	60	600	15	3.75	5	0.75
5	3	4	2	1	75	900	18	3.60	4	0.72
5	3	5	2	1	90	1260	21	3.50	4	0.70

Table 4 Comparative Results (Part A)

5 Comparative Evaluation

Here we compare the four metrics of §4 over our simulated researchers. Table 4 and 5 show their values for various time periods, from 1 year to 10 years.

Observations.
HI, HI/co and HIdivCit. We can see that HI/co behaves well, almost all researchers get the same score (recall that they have worked equally hard). R_{solo} has a bit higher HI/co, something that is reasonable. HIdivCit also behaves well, we observe small variations. These are evident also from Figure 1 that shows these values for the case Y=10.

HI/(coy). As expected, we can see that according to this metric years do not matter, therefore it can be used to fairly compare researchers of different academic age. To see that Hi/(coy) manages to normalize over time, Figure 2 shows the values of HI/co and HI/(coy) for various combinations of Years and F, in particular for the following

Years	PpYear	F	CExt	CSelf	Papers	Cits	HI	HI/co	HIdivCit	HI/(coy)
6	3	0	2	1	18	90	6	6.00	6	1.00
6	3	1	2	1	36	270	12	6.00	6	1.00
6	3	2	2	1	54	540	16	5.33	6	0.89
6	3	3	2	1	72	900	20	5.00	6	0.83
6	3	4	2	1	90	1350	24	4.80	6	0.80
6	3	5	2	1	108	1890	28	4.67	5	0.78
7	3	0	2	1	21	126	8	8.00	8	1.14
7	3	1	2	1	42	378	12	6.00	7	0.86
7	3	2	2	1	63	756	18	6.00	8	0.86
7	3	3	2	1	84	1260	24	6.00	7	0.86
7	3	4	2	1	105	1890	30	6.00	7	0.86
7	3	5	2	1	126	2646	35	5.83	7	0.83
8	3	0	2	1	24	168	9	9.00	9	1.13
8	3	1	2	1	48	504	15	7.50	9	0.94
8	3	2	2	1	72	1008	20	6.67	9	0.83
8	3	3	2	1	96	1680	25	6.25	8	0.78
8	3	4	2	1	120	2520	30	6.00	8	0.75
8	3	5	2	1	144	3528	36	6.00	8	0.75
9	3	0	2	1	27	216	10	10.00	10	1.11
9	3	1	2	1	54	648	18	9.00	10	1.00
9	3	2	2	1	81	1296	24	8.00	9	0.89
9	3	3	2	1	108	2160	30	7.50	10	0.83
9	3	4	2	1	135	3240	36	7.20	9	0.80
9	3	5	2	1	162	4536	42	7.00	9	0.78
10	3	0	2	1	30	270	12	12.00	12	1.20
10	3	1	2	1	60	810	18	9.00	12	0.90
10	3	2	2	1	90	1620	27	9.00	10	0.90
10	3	3	2	1	120	2700	35	8.75	11	0.88
10	3	4	2	1	150	4050	42	8.40	10	0.84
10	3	5	2	1	180	5670	49	8.17	10	0.82

Table 5 Comparative Results (Part B)

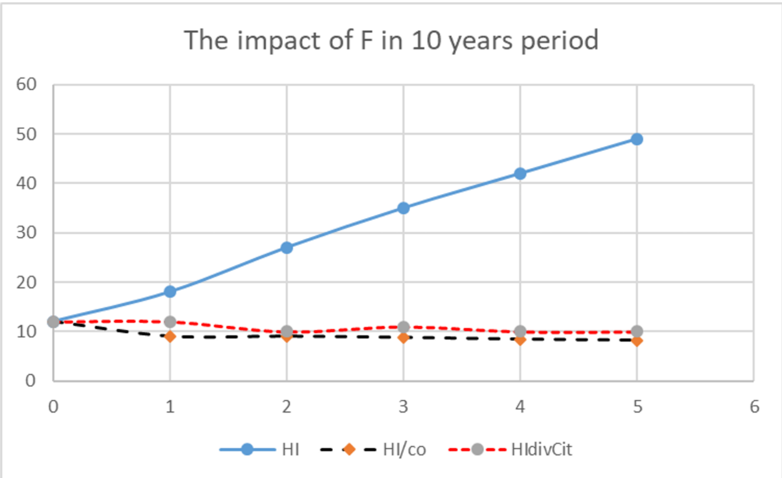

Fig. 1 The impact of F in 10 years at HI, HI/co and HdivCit

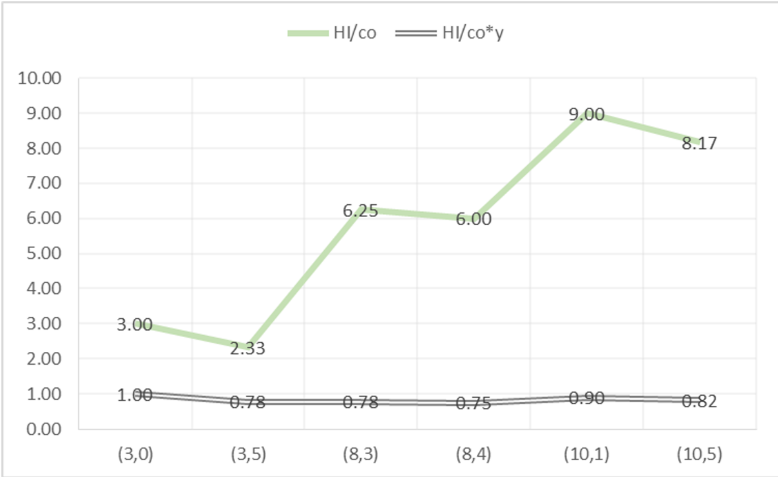

Fig. 2 HI/co and HI/(coy) for various (Y,F) combinations

cases: (Y=3, F=0), (Y=3, F=5), (Y=8, F=3), (Y=8, F=4), (Y=10, F=1), (Y=10, F=5). We can see that even if their HI/co varies a lot, HI/(coy) ranges 0.78 to 1.

6 Evaluation over Real Data

At first we describe our methodology (in §6.1) and how we implemented it (in §6.2). Then we analyze the top-10 profiles from five universities (in §6.3) where we also summarize our findings (at Section 6.3.6). Subsequently we report the results of a more thorough analysis that comprises roughly all researchers from two universities (in §6.4). Then we include an analysis of the faculty members of one department (in §6.5), as well as an analysis at school level (in §6.6). Subsequently, in §6.7, we analyze the bibliographic profiles of a few famous scientists of the past (and compare them with those of §6.3). Finally, in §6.8 we provide suggestions for the community and bibliographic sources, while in §6.9 we provide suggestions on how to compare researchers.

6.1 Methodology

To evaluate the impact of these metrics on real data, we analyzed various data. We have chosen 3 prestigious universities, one from US (Harvard) and two from Europe (Cambridge, Oxford), as well as, a relatively new one (University of Crete, founded less than 50 years ago), as well as a younger one, University of Ziauddin, and from these five universities we analyzed the profiles of the top-10 researchers, and the relative rankings as produced by various metrics. For not restricting our analysis to the top researchers only, for two universities (University of Crete and Ziauddin), we analyzed essentially all profiles. To test also the metrics on scientists of the same domain, we analyzed all profiles of one department (Computer Science Department, University of Crete). To test if co-authorship varies in all schools of a university, we analyzed all

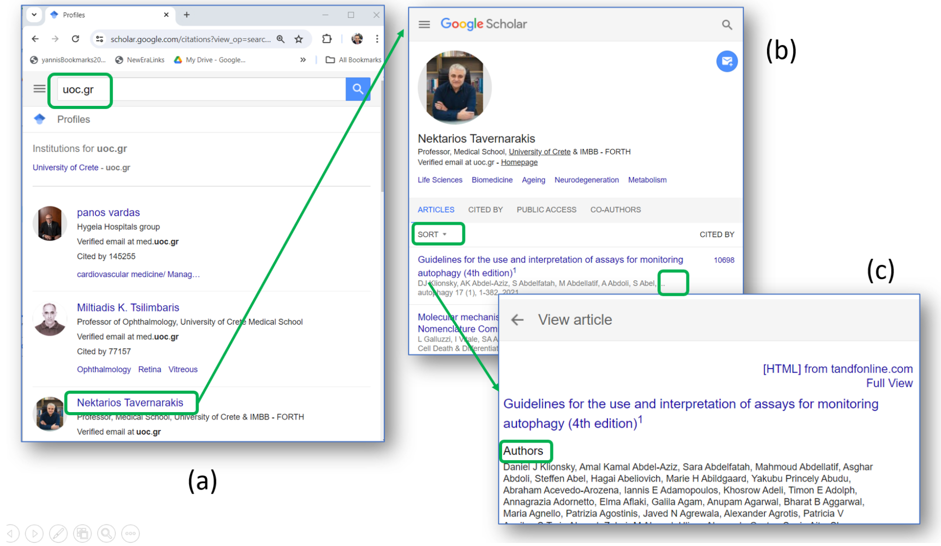

Fig. 3 Getting information from Google Scholar

schools of the University of Crete. To test whether the metrics can reveal the different publishing policies of different eras, we also report results about a few famous scientists of the past. Finally, we should clarify that we focus on the fair evaluation of persons, not on the evaluation of departments nor of universities.

6.2 Implementation

Bibliographic Source. We used Google Scholar[1] as the source of our data, which provides the ranking of researchers with respect to citations. In particular, if one searches using the name of one university and then from the top-left menu selects the option "Profiles", he gets a list of all profiles associated with that university, ranked according to the number of citations. An example is shown in Figure 3(a). By clicking on a profile, the user can get the list of all publications of that author, sorted either by the number of citations or by date, as shown in Figure 3(b). In that list each publication is shown as an item of the list, however, in case of papers with many authors, only the first few authors are shown and a symbol "...". To overcome this inaccuracy, and thus get the complete list of authors of a paper, we have to click on the particular paper, as shown in Figure 3(c).

Method. For each author that we examined we downloaded all of his/her papers and for each paper the number of its authors and the number of citations. This allowed us to compute the avgCoAuthors for each author, and thus to compute HI/co. In this way, we can compute the rank of each these profiles with respect to each metric, for investigating how co-authorship affects ranking.

[1] https://scholar.google.com/

Automating the extraction. Obviously, the above process cannot be done manually. To automate the extraction process from Google Scholar, according to our methodology, we developed a *playwright*[2] script (playwright is a tool for scraping the web, or testing a web application). It takes as input a domain name (e.g. "uoc.gr"), and the number of top author profiles to analyze. It can also take as input a text file containing the URLs of Google Scholar profiles. The script is about 700 lines of code, and uses the playwright tool. It took about 70 hours to write and test the program. Note that Google Scholar, might temporary ban the IP address used if multiple calls happen in very short amount of time. To overcome this obstacle our script support custom delays. With the timeout used, to fetch all data of hundreds of profiles, requires a few days. The fetched data are then stored in an output file in JSON format. The size of the output file depends on the number of articles scraped from every user, it can be from 4MB to more than 50MB; it depends on how many articles the author has. Note that some profiles, especially those corresponding to the most cited researchers, can have 1000 or more articles. To fully scrape such a profile it takes around 35 minutes[3] Most of the time is spent waiting to avoid IP bans. All data that are presented below were downloaded on November and December of 2023.

Limitations. The entire process is automatic, so the results of our analysis depend on the correctness of the data provided by Google Scholar. As mentioned earlier, the program that scrapes the data from each article checks the "Author" section in the Google Scholar page, in order to get the full list of authors. However, we have spotted a few cases where for some articles the number of authors listed is inaccurate, i.e. the author might be an "association" or some of the authors might be missing. Moreover we checked manually all 300 profiles of the University of Crete. We spotted only two cases where one researcher had two profiles (each corresponding to a different spelling of his/her name). This corresponds to around 1% inaccuracy in the analyzed profiles. Overall, these errors were too few, and we cannot say that they affect the main results of our analysis, i.e. that (as we shall see) the degree of co-authorship varies a lot. Moreover, to ensure transparency and enable reproducibility, the paper includes a link to a public folder where all data are placed.

6.3 Analyzing the Top-10 Researchers of Universities

Below we analyze the researchers of five universities, in lexicographic order, specifically University of Cambridge (in §6.3.1), University of Crete (in §6.3.2), University of Harvard (in §6.3.3), a University of Oxford (in §6.3.4), and University of Ziauddin (in §6.3.5) For each university we downloaded the top-10 profiles according to citations. Then we analyzed all the papers of these authors, and then we computed the relative rankings of these top-10 persons with respect to the other criteria.

[2] https://playwright.dev/

[3] Based on our experiments, to download $authorArticles$ in number articles, it takes about $2(authorArticles{+}1)$ seconds to scrape them, and around 7-8 minutes to scrap 100 profiles.

6.3.1 University of Cambridge

The top-10 researchers of University of Cambridge[4], according to citations, plus the extra information that we scratched and computed for them, are shown in Table 6. We observe big variations in the average number of co-authors: from 7.44 to 41.98. Now the relative rankings of these top-10 researchers with respect to the other criteria and metrics, are shown in Table 7. We observe significant changes. For instance the 1st with respect to citations (Nicholas Wareham) is 5th according to HI/co. The 2nd with respect to HI/co (RH Friend) is 6th with respect to Pubs and HI. None of the top-3 with respect to HI/co, belongs to the top-3 with respect to HI.

	Name	Citations	Publications	H-index	avgCo-Authors	HI/co
1	Nicholas Wareham	328987	1875	245	29.30	8.36
2	Richard Durbin	294900	466	153	20.15	7.59
3	John Danesh	229053	606	170	39.01	4.35
4	Prof. TW Robbins	225421	1739	261	7.65	34.09
5	Simon Baron-Cohen	215843	1423	225	8.81	25.53
6	RH Friend	208419	1678	204	7.44	27.41
7	Douglas Easton	198117	1424	206	41.98	4.9
8	Gregory Hannon	196773	549	178	8.03	22.16
9	James Jackson	176357	809	156	38.62	4.03
10	Sir Stephen O'Rahilly	166882	960	205	29.34	6.98

Table 6 The top-10 profiles of University of Cambridge with respect to Citations, and the extracted and computed information

	Name	wrt Citations	wrt Pubs	wrt H-index	wrt HI/co
1	Nicholas Wareham	1	1	2	**5**
2	Richard Durbin	2	10	10	6
3	John Danesh	3	8	8	9
4	Prof. TW Robbins	4	2	1	1
5	Simon Baron-Cohen	5	5	3	3
6	RH Friend	6	3	6	**2**
7	Douglas Easton	7	4	4	8
8	Gregory Hannon	8	9	7	4
9	James Jackson	9	7	9	10
10	Sir Stephen O'Rahilly	10	6	5	7

Table 7 Ranking of the top-10 profiles of University of Cambridge with respect to various criteria

To quantify the difference between the rankings obtained, we can use one distance function between rankings [17]. We have selected the *normalized Kendall tau distance.* that counts the number of pairwise disagreements between two ranking lists. Let N be a universe of elements. Let S_N be the set of all permutations on N. The Kendall's tau distance between two rankings s and t (where $s, t \in S_N$) measures the total number of pairwise inversions. In particular, the Kendall's tau distance is given by $K_d(s,t) =$

[4]The profiles retrieved with the query "University of Cambridge".

$\sum_{\{i,j\}\in P, i<j} dist_{i,j}(s,t)$ where P is the set of unordered pairs N and $dist_{i,j}(s,t) = 0$ if i and j are in the same order in s and t, while $dist_{i,j}(s,t) = 1$ if they are in opposite order in s and t. It follows that $K_d(s,t)$ is equal to 0 if s and t are identical, and is equal to $\frac{1}{2}|N|(|N|-1)$ if one is the reverse of the other. The normalized Kendall tau distance, K_n, is defined as $K_n = \frac{K_d}{\frac{1}{2}|N|(|N|-1)}$ and therefore lies in the interval [0,1].

Hereafter we shall use R_X to denote the ranking obtained by a metric X, and $dist(R_X, R_Y)$ to denote the the normalized Kendall tau distance between R_X and R_Y, e.g. with $dist(R_{cits}, R_{HI/co})$ we will denote the normalized Kendall tau distance between R_{cits} and $R_{HI/co}$. Below we show the distance values obtained:

$$\begin{aligned}
dist(R_{cits}, R_{HI}) &= 0.42\\
dist(R_{cits}, R_{HI/co}) &= 0.40\\
dist(R_{cits}, R_{Pubs}) &= 0.42\\
dist(R_{HI/co}, R_{HI}) &= 0.28\\
dist(R_{HI/co}, R_{Pubs}) &= 0.33
\end{aligned}$$

Notice that $dist(R_{HI/co}, R_{HI}) = 0.28$ meaning that the relative rankings of 28% of the pairs of researchers, is different in HI and HI/co.

6.3.2 University of Crete

The top-10 researchers of UoC (University of Crete)[5], according to citations, plus the extra information that we scratched and computed, are shown in Table 8. We observe variations in the average number of co-authors: from 3.88 to 18.24. Now the rankings of these top-10 researchers of UoC, according to each criterion are shown in Table 9. We observe significant changes. For instance, the 1st with respect to HI/co (E. Economou), is 10th with respect to citations (!), and 7th with respect to HI.

	Name	Citations	Publications	H-index	avgCo-Authors	HI/co
1	P. Vardas	135826	1213	109	12.65	8.61
2	M Tsilimbaris	72105	386	48	13.04	3.67
3	N. Tavernarakis	62851	583	100	7.67	13.03
4	A Tsatsakis	50561	1159	100	8.80	11.35
5	C Stoumpos	44442	307	86	9.23	9.30
6	C Lionis	39251	738	63	9.15	6.87
7	D Mavroudis	30627	743	89	18.24	4.87
8	N Komodakis	29206	151	56	4.67	11.97
9	N Mihalopoulos	27669	691	92	11.57	7.94
10	E Economou	27334	512	75	3.88	19.32

Table 8 The top-10 profiles of UoC with respect to citations and the extracted/computed information

[5]The profiles retrieved with the query "uoc.gr".

	Name	wrt Citations	wrt Pubs	wrt H-index	wrt HI/co
1	P. Vardas	1	1	1	6
2	M Tsilimbaris	2	8	10	10
3	N. Tavernarakis	4	6	2*	2
4	A Tsatsakis	4	2	2*	4
5	C Stoumpos	5	9	6	5
6	C Lionis	6	4	8	8
7	D Mavroudis	7	3	5	9
8	N Komodakis	8	10	9	3
9	N Mihalopoulos	9	5	4	7
10	E Economou	**10**	**7**	7	**1**

Table 9 Ranking of the top-10 profiles of UoC with respect to various criteria

Below we show the distances of these rankings:

$$
\begin{aligned}
dist(R_{cits}, R_{HI}) &= 0.37 \\
dist(R_{cits}, R_{HI/co}) &= 0.57 \\
dist(R_{cits}, R_{Pubs}) &= 0.37 \\
dist(R_{HI/co}, R_{HI}) &= 0.42 \\
dist(R_{HI/co}, R_{Pubs}) &= 0.60
\end{aligned}
$$

Notice that $dist(R_{HI/co}, R_{HI}) = 0.42$

6.3.3 University of Harvard

The top-10 researchers of University of Harvard[6], according to citations, plus the extra information that we scratched, are shown in Table 10. We observe variations in the average number of co-authors: from 1.94 to 48.22. The relative rankings are shown in Table 11. Here we observe that the 1st with respect to citations (Michael E. Porter) is also 1st with respect to HI/co, but 9th with respect to HI. In general we observe significant changes in the rankings, for instance the 2nd with respect to Hi/co (Andrei Shleifer) is 10th with respect to HI, and 8th with respect to citations.

Below we show the distances of these rankings:

$$
\begin{aligned}
dist(R_{cits}, R_{HI}) &= 0.31 \\
dist(R_{cits}, R_{HI/co}) &= 0.28 \\
dist(R_{cits}, R_{Pubs}) &= 0.22 \\
dist(R_{HI/co}, R_{HI}) &= 0.42 \\
dist(R_{HI/co}, R_{Pubs}) &= 0.20
\end{aligned}
$$

Notice that $dist(R_{HI/co}, R_{HI}) = 0.42$.

[6]The profiles retrieved with the query "Harvard".

	Name	Citations	Publications	H-index	avgCo-Authors	HI/co
1	Michael E. Porter	575772	2338	182	1.94	93.68
2	Ronald C Kessler	538486	1994	332	11.75	28.27
3	Frank B. Hu	476966	1218	308	15.72	19.59
4	Dr. JoAnn E. Manson	460146	3000	314	10.18	30.86
5	Paul M Ridker, MD, MPH	450479	1825	270	20.23	13.34
6	Stacey Gabriel	435819	585	220	48.22	4.56
7	Mark Daly	414992	1230	237	27.89	8.50
8	Andrei Shleifer	411288	1183	168	3.58	46.90
9	Gad Getz	336904	604	227	32.45	7.00
10	Matthew Meyerson	308764	892	220	24.93	8.82

Table 10 The top-10 profiles of Harvard with respect to Citations and the extracted/computed information

	Name	wrt Citations	wrt Pubs	wrt H-index	wrt HI/co
1	Michael E. Porter	1	2	**9**	1
2	Ronald C. Kessler	2	3	1	4
3	Frank B. Hu	3	6	3	5
4	Dr JoAnn E. Manson	4	1	2	3
5	Paul M Ridker, MD, MPH	5	4	4	6
6	Stacey Gabrie	6	10	7	10
7	Mark Daly	7	5	5	8
8	Andrei Shleifer	8	7	10	**2**
9	Gad Getz	9	9	6	9
10	Matthew Meyerson	10	8	8	7

Table 11 Ranking of the top-10 profiles of Harvard with respect to various criteria

6.3.4 University of Oxford

The top-10 researchers of University of Oxford[7], according to citations, plus the extra information that we scratched, are shown in Table 12. We observe big variations in the average number of co-authors: from 3.75 to 114.45. The relative rankings are shown in Table 13. We observe that the 1st according to HI/co (Robert M. May) is 10th with respect to citations, and 6th with respect to HI.

Below we show the distances of these rankings:

$$\begin{aligned} dist(R_{cits}, R_{HI}) &= 0.40 \\ dist(R_{cits}, R_{HI/co}) &= 0.60 \\ dist(R_{cits}, R_{Pubs}) &= 0.57 \\ dist(R_{HI/co}, R_{HI}) &= 0.46 \\ dist(R_{HI/co}, R_{Pubs}) &= 0.51 \end{aligned}$$

Notice that $dist(R_{HI/co}, R_{HI}) = 0.46$, a very high value!

[7]The profiles retrieved with the query "University of Oxford"

	Name	Citations	Publications	H-index	avgCo-Authors	HI/co
1	Douglas G Altman	854796	1703	278	7.94	34.98
2	Andrew Zisserman	388991	973	192	4.49	42.69
3	Derrick A. Bennett	282175	472	115	38.18	3.01
4	Amanda Cooper-Sarkar	276023	1809	245	114.45	2.14
5	Peter Jüni	244322	1034	156	14.78	10.55
6	Cornelia M van Duijn	238866	1890	236	35.23	6.69
7	Stephen M. Smith	218639	744	153	8.72	17.52
8	Adrian Vivian Hill	215443	3000	211	9.60	21.97
9	Peter Rothwell	203576	1352	169	14.60	11.57
10	Robert M May	200289	1111	171	3.75	45.55

Table 12 The top-10 profiles of Oxford with respect to citations and the related extracted and computed information

	Name	wrt Citations	wrt Pubs	wrt H-index	wrt HI/co
1	Douglas G Altman	1	4	1	3
2	Andrew Zisserman	2	8	5	2
3	Derrick A. Bennett	3	10	10	9
4	Amanda Cooper-Sarkar	4	3	2	10
5	Peter Jüni	5	7	8	7
6	Cornelia M van Duijn	6	2	3	8
7	Stephen M. Smith	7	9	9	5
8	Adrian Vivian Hill	8	1	4	4
9	Peter Rothwell	9	5	7	6
10	Robert M May	**10**	6	6	**1**

Table 13 Ranking of the top-10 profiles of Oxford with respect to various criteria

6.3.5 University of Ziauddin

Ziauddin University is a relative new university from Pakistan, founded in 1995, whose position in the ranking for 2024 produced by THE (Times Higher Education)[8] is very low[9]. It has seven academic faculties, i.e. health sciences, law, liberal arts and human sciences, eastern medicine and natural sciences, engineering science technology and management, pharmacy, and the nursing and midwifery, and some of these departments offer both undergraduate and postgraduate courses. The top-10 researchers of University of Ziauddin[10], according to citations, plus the extra information that we scratched, are shown in Table 14. Again we observe big variations in the average number of co-authors: from 2.33 to 24.39.

The relative rankings of these researchers are shown in Table 15. We observe that the 1st according to HI/co (Fauzia Shamim) is 9th according to H-Index. Below we show the distances of these rankings:

$$\begin{aligned} dist(R_{cits}, R_{HI}) &= 0.06 \\ dist(R_{cits}, R_{HI/co}) &= 0.46 \\ dist(R_{cits}, R_{Pubs}) &= 0.2 \end{aligned}$$

[8] https://www.timeshighereducation.com/world-university-rankings/2024/world-ranking?page=106#
[9] In particular its position it resides at position 2,651 of the 2,671 universities that were evaluated.
[10] The profiles were retrieved with the query "zu.edu.pk".

	Name	Citations	Publications	H-index	avgCo-Authors	HI/co
1	dr sehrish ahmed	23254	1131	38	7.01	5.42
2	Madiha Hashmi	3626	122	28	24.39	1.15
3	Shaikh Ziauddin Ahammad	2701	132	29	5.02	5.77
4	Murtaza Ziauddin	2543	87	26	4.71	5.52
5	ALI ASGHAR	2354	132	24	6.50	3.69
6	Haroon Rashid Baloch	2070	74	17	3.24	5.24
7	Haider Naqvi	1821	153	23	6.31	3.65
8	Dr. Saif Ullah	1814	53	20	3.68	5.44
9	Fauzia Shamim	1526	46	16	2.33	6.88
10	Muhammad Mustafa Swaleh	1247	12	4	4.92	0.81

Table 14 Ranking of the top-10 profiles of Ziauddin wrt Citations

	Name	wrt Citations	wrt Pubs	wrt H-index	wrt HI/co	wrt Pubc/co
1	dr sehrish ahmed	1	1	1	5	1
2	Madiha Hashmi	2	5	3	9	9
3	Shaikh Ziauddin Ahammad	3	3	2	2	2
4	Murtaza Ziauddin	4	6	4	3	7
5	ALI ASGHAR	5	4	5	7	5
6	Haroon Rashid Baloch	6	7	8	6	4
7	Haider Naqvi	7	2	6	8	3
8	Dr. Saif Ullah	8	8	7	4	8
9	Fauzia Shamim	**9**	**9**	**9**	**1**	6
10	Muhammad Mustafa Swaleh	10	10	10	10	10

Table 15 Ranking of the top-10 profiles of Ziauddin wrt various criteria

$$\begin{aligned} dist(R_{HI/co}, R_{HI}) &= 0.44 \\ dist(R_{HI/co}, R_{Pubs}) &= 0.53 \end{aligned}$$

Again we can see very big differences in the rankings, mainly for those pairs that include HI/co.

6.3.6 Summary of Findings by Analyzing the Top Researchers of 5 Universities

We have observed big variations of the number of co-authors. The ranges and the median encountered by analyzing only the top-10 profiles with respect to citations, are shown in Table 16. As regards co-authors, note that [15] that analyzed more than 120 million papers, from several domains, reports mean number of authors from 2.83 to 6.14. However, they also state that the maximal number of authors for a single paper in each year increased sharply over time, and have spotted some recent papers with more than 3,000 authors. In our case, we have seen much larger average number of co-authors. This provides evidence that the highly cited researchers, mainly have more co-authors in average.

Table 16 also shows shows the ranges and medians of HI and HI/co. These ranges are illustrated as interval plots in Figure 4. Notice that UofCrete, has considerably lower range of HI, in comparison to Cambridge, Harvard and Oxford, but since it also has lower range of co-authors, its score in HI/co is considerably better.

	avgCo-authors		HI		HI/co	
University	Range	Median	Range	Median	Range	Median
University of Cambridge	7.44 to 41.98	24.72	153 - 261	205	4.03 - 34.09	7.97
University of Crete	4.67 to 18.24	9.19	48 - 109	87.5	3.67 - 19.32	8.95
University of Harvard	1.94 to 48.22	22.11	168 - 332	232	4.56 - 93.68	16.46
University of Oxford	3.75 to 114.45	12.1	115 - 278	181.5	2.14 - 45.55	14.54
University of Ziauddin	2.33 to 24.39	4.97	4 - 38	23.5	0.81 - 6.88	5.33

Table 16 Range and medians of average co-authors, HI and HI/co of the top-10 profiles

We have observed that by considering the number of co-authors (through HI/co), the ranking of researchers changes significantly. The normalized Kendall's tau distance between the rankings obtained by HI and HI/co, ranges from 0.28 to 0.46, meaning that more than one third of researcher pairs have different relative ranking in HI and HI/co.

Finally, we should clarify that we focus on the fair evaluation of individual researchers, not universities, so the role of the aforementioned data, is to provide information about the scale of co-authorship and to highlight that metrics that consider co-authorship affect the obtained rankings.

6.4 Analyzing all Profiles of two Universities

In the previous subsections we reported big variations in the rankings if co-authorship is considered. The real difference can be bigger in the sense that previously we have analyzed only the top-10 profiles. Since the scrapping process is slow, we decided to make a more complete analysis of two universities. We selected University of Crete, since it feasible for us to check the validity of the results, in addition the university of Zaudin (whose top researchers were analyzed in §6.3.5)

University of Crete. For the university of Crete, we decided to analyze the top-300 profiles with respect to citations, and for these 300 profiles to provide the relative rankings. In brief, we noticed even bigger changes. Table 17 shows the top-15 researchers according to HI, HI/co, Pubs and Pubs/co. Note that we have also the column *Pubs/co* as indicator of the productivity of each researcher. Analyzing so many profiles enables us to inspect the difference in the rankings obtained by the metrics, not only over of the top researches, but also over the all members of the university (300 essentially contains all active members).

For reasons of space, in Table 18 we show only the top 55 with respect to citations At first we observe quite different researchers in top-10. The number of common elements between the top-10 with respect to Citations and the top-10 with respect to HI is 7 (i.e. 70%). The number common elements between the top-10 with respect to Citations and the top-10 with respect to HI/co is 2 (i.e. 20%). The number common elements between the top-10 with respect to Citations and the top-10 with respect

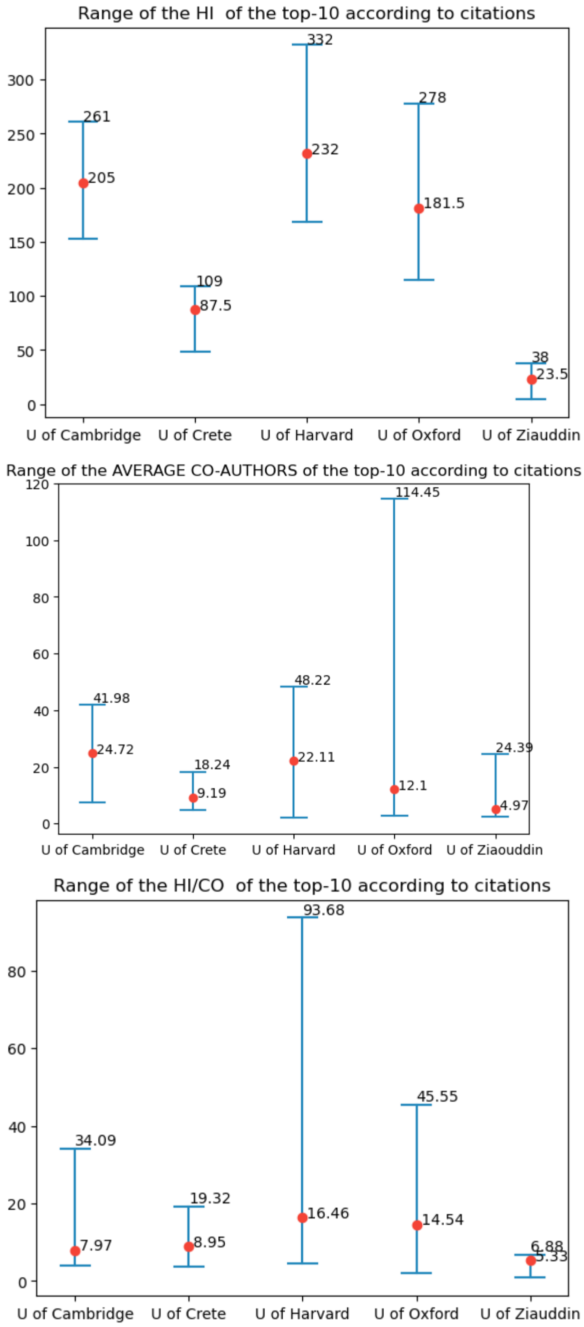

Fig. 4 Range and medians of HI, average co-authors, and HI/co of the top-10 profiles of 5 universities

to Pubs/co is 5 (i.e. 50%). In general we observe very big changes in the rankings. The 1st with respect to citations (Panos Vardas) is 49th with respect to HI/co. The 2nd with respect to citations (M K Tsilimbaris) is 207th with respect to HI/co. The 2nd with respect to HI/co (Eleftherios Zouros) was 27th with respect to to citations, and was not included in Table 9, as well as the 3rd with respect to HI/co (Stamatios Papadakis) who is 55th with respect to citations.

	Wrt HI	Wrt HI/co	Wrt Pubs	Wrt Pubs/co
1	Panos Vardas	**Eleftherios N. Economou**	Panos Vardas	**Eleftherios N. Economou**
2	Nektarios Tavernarakis	Eleftherios Zouros	Aristidis Tsatsakis	Aristidis Tsatsakis
3	Aristidis Tsatsakis	Stamatios Papadakis	MAVROUDIS DIMITRIOS	Ioannis G. Pallikaris
4	Nikos Mihalopoulos	Nikos Papanicolaou	Christos Lionis	Panos Vardas
5	MAVROUDIS DIMITRIOS	Vassilios Makrakis	Nikos Mihalopoulos	Apostolos Karantanas
6	Costas Stoumpos	Elias Kiritsis	A. Zezas	Giorgos Tsironis
7	Elias Kiritsis	Nektarios Tavernarakis	Nektarios Tavernarakis	Christos Lionis
8	Vassilis Charmandaris	Vasilis Niarchos	Vassilis Charmandaris	Nektarios Tavernarakis
9	**Eleftherios N. Economou**	Ioannis G. Tollis	Ioannis G. Pallikaris	JOHN DAMILAKIS
10	George Samonis	Yannis Stylianou	Eleftherios N. Economou	Vassilios Makrakis
11	Dimitris Mavroudis	Kostas Vlassopoulos	Apostolos Karantanas	Yannis Stylianou
12	Maria Kanakidou	Nikos Frantzikinakis	George Samonis	Yannis Tzitzikas
13	Euripides G. Stephanou	Georgios Tziritas	Dimitris Mavroudis	Panagiotis Tsakalides
14	Christos Lionis	Giorgos Tsironis	JOHN DAMILAKIS	A. Zezas
15	Ioannis Koutroubakis	Nikos Komodakis	Nikos Tzanakis	Panos Trahanias

Table 17 The top-15 profiles of UoC with respect to various criteria (based on the analysis of the top-300

Below we show the distances of these rankings of 300 profiles:

$$\begin{aligned}
dist(R_{cits}, R_{HI}) &= 0.12 \\
dist(R_{cits}, R_{HI/co}) &= 0.349 \\
dist(R_{cits}, R_{Pubs}) &= 0.22 \\
dist(R_{HI/co}, R_{HI}) &= 0.342 \\
dist(R_{HI/co}, R_{Pubs}) &= 0.38
\end{aligned}$$

Notice that $dist(R_{HI/co}, R_{HI}) = 0.34$.

University of Ziauddin. We downloaded all profiles of the University of Ziauddin,

	Name	wrt Cits	wrt Pubs	wrt H-index	wrt HI/co	wrt Pubs/co
1	Panos Vardas	1	1	1	**49**	4
2	Miltiadis K. Tsilimbaris	2	18	42	**207**	91
3	Nektarios Tavernarakis	3	7	2	7	8
4	Aristidis Tsatsakis	4	2	3	17	2
5	Costas Stoumpos	5	26	6	37	71
6	Christos Lionis	6	4	14	79	7
7	MAVROUDIS DIMITRIOS	7	3	5	151	43
8	Nikos Komodakis	8	115	22	15	75
9	Nikos Mihalopoulos	9	5	4	59	19
10	Eleftherios N. Economou	10	10	9	1	1
11	Vassilis Charmandaris	11	8	8	118	54
12	Euripides G. Stephanou	12	68	13	92	145
13	Maria Kanakidou	13	19	12	96	62
14	George Samonis	14	12	10	43	22
15	Elias Kiritsis	15	30	7	6	23
16	Vasiliki Pavlidou	16	43	29	286	267
17	Konstantinos Makris	17	78	84	64	48
18	Ioannis Koutroubakis	18	16	15	144	77
19	Elias Castanas	19	23	30	74	40
20	Dimitris Mavroudis	20	13	11	120	60
21	Ioannis G. Pallikaris	21	9	19	18	3
22	Maria Kafesaki	22	31	25	67	49
23	A. Zezas	23	6	20	97	14
24	Ioannis Tsamardinos	24	54	49	57	39
25	Panagiotis Simos, Ph.D	25	42	16	58	61
26	Souglakos	26	22	33	201	119
27	Eleftherios Zouros	27	82	17	**2**	27
28	George Froudakis	28	73	34	153	167
29	Katerina Antoniou	29	17	26	170	81
30	Marilena Kampa	30	102	67	140	149
31	Nikos Tzanakis	31	15	18	91	31
32	Elena Anagnostopoulou	32	69	56	34	29
33	Antonis Argyros	33	28	52	88	36
34	Iossif Papadakis	34	36	23	253	212
35	Ioannis G. Tollis	35	56	50	9	16
36	Stelios Tzortzakis	36	45	36	62	44
37	Christos Tsatsanis	37	83	27	78	121
38	Tzanakakis George	38	66	21	51	76
39	Kostas Demadis	39	40	31	24	24
40	Nikolaos K. Efremidis	40	93	88	72	68
41	Evangelos Markatos	41	34	39	23	18
42	JOHN DAMILAKIS	42	14	35	33	9
43	Yannis Stylianou	43	35	51	10	11
44	Achilleas Gikas	44	104	57	233	217
45	Ioannis Karakassis	45	59	40	84	85
46	Prodromos Sidiropoulos	46	20	53	178	70
47	Maria Vamvakaki	47	48	28	32	33
48	Apostolos Karantanas	48	11	37	42	5
49	Sofia Agelaki	49	46	73	216	130
50	KALLIOPI ROUBELAKIS-ANGELAKIS	50	120	38	56	117
51	Achille Gravanis	51	41	32	68	53
52	Georgios Tziritas	52	117	89	13	26
53	Demetrios Anglos	53	94	45	52	79
54	Emmanuel Prokopakis	54	70	58	267	236
55	Stamatios Papadakis	55	118	24	**3**	42

Table 18 The top-55 profiles of UoC and their relative ranking with respect to various criteria in top-300

they are 169. Below we show the distances of induced rankings:

$$\begin{aligned} dist(R_{cits}, R_{HI}) &= 0.14 \\ dist(R_{cits}, R_{HI/co}) &= 0.18 \\ dist(R_{cits}, R_{Pubs}) &= 0.20 \\ dist(R_{HI/co}, R_{HI}) &= 0.15 \\ dist(R_{HI/co}, R_{Pubs}) &= 0.22 \end{aligned}$$

Here we observe smaller differences of the induced rankings i.e. from 0.14 to 0.22, in comparison to those of the University of Crete.

6.5 Analyzing all faculty members of one department

For checking how co-authorship varies if all compared researchers are of the same discipline, and how co-authorship affects the rankings at department level, we analyzed all faculty profiles associated with the Computer Science Department of the University of Crete. The data of these 25 profiles are shown in Table 19. For reasons of discretion, the last members of the list are written anonymously We observe the the avgCo-authors ranges from 2.96 to 8.44. The position of each faculty member with respect to various criteria is shown in Table 20. Again we observe significant changes, just indicatively, the 1st with respect to citations and HI is Nikos Komodakis, however the 1st with respect to HI/co is Ioannis G. Tollis, and the 1st with respect to Pubs/co is Constantine Stephanidis.

Since the members have different academic age, we can use Pubs/(co*y) and HI/(co*y) as a time-normalized version of Pubs/co and HI/co. As expected, the obtained ranking is different from the previous ones. The 1st with respect to HI/(coy) is Xenofontas Dimitropoulos. The range of HI/co is 1.88 to 12.53. Note that in the computation of Y we have considered the publication year of the first and the last paper. Alternatively, instead of considering the year of the first paper, one could consider the year of PhD graduation in order to avoid penalizing those researchers who have started publishing very early, e.g. when they were undergraduate students.

Below we show the distances of these rankings over the 25 faculty members:

$$\begin{aligned} dist(R_{cits}, R_{HI}) &= 0.09 \\ dist(R_{cits}, R_{HI/co}) &= 0.22 \\ dist(R_{cits}, R_{Pubs}) &= 0.23 \\ dist(R_{HI/co}, R_{HI}) &= 0.19 \\ dist(R_{HI/co}, R_{Pubs}) &= 0.32 \\ dist(R_{HI/co}, R_{HI/(coY)}) &= 0.22 \end{aligned}$$

Notice that $dist(R_{HI/co}, R_{HI}) = 0.19$. This is the smallest value that we have observed, probably due to the fact that all scientists here come from the same domain (Computer Science), and because the range of average co-authors is smaller than the ranges measured at university level.

	Name	Citations	Publications	H-index	avgCo-Authors	HI/co
1	Nikos Komodakis	29765	152	56	4.68	11.96
2	Ioannis Tsamardinos	12542	237	45	5.51	8.17
3	Antonis Argyros	10882	292	44	6.61	6.66
4	Ioannis G. Tollis	10657	225	45	3.59	12.53
5	Constantine Stephanidis	9408	669	45	4.09	11.00
6	Evangelos Markatos	9375	279	49	4.64	10.56
7	Yannis Stylianou	9285	278	45	3.61	12.46
8	Georgios Tziritas	8430	151	37	3.09	11.99
9	Dimitris Plexousakis	8286	327	43	4.82	8.92
10	Panagiotis Tsakalides	6337	275	35	3.91	8.96
11	Panos Trahanias	5724	241	34	3.73	9.10
12	Xenofontas Dimitropoulos	5104	120	38	4.09	9.29
13	Manolis G.H. Katevenis	3922	156	31	6.79	4.56
14	Haridimos Kondylakis	3456	197	34	6.99	4.86
15	Angelos Bilas	3126	241	30	8.44	3.55
16	Yannis Tzitzikas	2921	281	26	3.80	6.84
17	George Papagiannakis	2883	157	28	5.61	5.00
18	Anthony Savidis	2660	170	28	2.96	9.46
19	Maria Papadopouli	2293	130	25	6.91	3.62
20	Grigorios Tsagkatakis	1602	119	20	4.69	4.27
21	Panagiota Fatourou	1489	112	20	3.38	5.93
22	Polyvios Pratikakis	1294	56	16	4.09	3.91
23	faculty1	1217	99	20	4.17	4.79
24	faculty2	1043	91	19	7.67	2.48
25	faculty3	679	70	14	7.43	1.88

Table 19 The data of the 25 profiles of faculty members of UoC-CSD

6.6 Analyzing faculty members by School

To check if there are variations at school level, we separated the 300 profiles of the University of Crete (those analyzed in Section 6.4) to three groups: a) School of Sciences and Engineering (that comprises 6 departments) b) Medical School (1 department) c) Schools of Philosophy (3 departments), Education (2 departments), Social Sciences (4 departments), in total 9 departments. Hereafter, we shall use the term Humanities and Social Sciences, for the last group. For each of these three groups below we report measurements.

The first group (School of Sciences and Engineering) contains 135 profiles, the second (Medical School) contains 129, while the last (Humanities and Social Sciences) contains 36 profiles. Table 21 synopsizes the results. The first, and important, observation is that we can see big variations of the number of co-authors, in all schools! This is evident and in the left chart of Figure 5. We can also see that the Medical School has the highest number of average co-authors, followed by the School of Sciences and Engineering. As regards HI/co, we observe that the Medical School has the lowest values. We should also note that the HI/co values of the researchers from Humanities and Social Sciences are relatively high, because only 36 profiles were included in the analysis (many researchers of that school do not have a Google Scholar profile). In any case, this analysis aims at understanding the degree of co-authorship, not to comparatively evaluate the schools of the University of Crete

	Name	wrt Citations	wrt Pubs	wrt H-index	wrt HI/co	wrt Pubs/co	wrt Pubs/(coY)	wrt HI/(coY)
1	Nikos Komodakis	1	16	1	4	14	15	2
2	Ioannis Tsamardinos	2	10	3	12	12	10	4
3	Antonis Argyros	3	3	7	14	11	13	17
4	Ioannis G. Tollis	4	11	4	1	7	11	5
5	Constantine Stephanidis	5	1	5	5	1	1	8
6	Evangelos Markatos	6	5	2	6	8	7	6
7	Yannis Stylianou	7	6	6	2	2	3	3
8	Georgios Tziritas	8	17	10	3	10	18	11
9	Dimitris Plexousakis	9	2	8	11	5	5	9
10	Panagiotis Tsakalides	10	7	11	10	4	4	12
11	Panos Trahanias	11	8	12	9	6	8	14
12	Xenofontas Dimitropoulos	12	19	9	8	15	12	1
13	Manolis G.H. Katevenis	13	15	14	19	21	22	21
14	Haridimos Kondylakis	14	12	13	17	17	9	10
15	Angelos Bilas	15	9	15	23	16	19	20
16	Yannis Tzitzikas	16	4	18	13	3	2	13
17	George Papagiannakis	17	14	16	16	18	17	16
18	Anthony Savidis	18	13	17	7	9	6	7
19	Maria Papadopouli	19	18	19	22	22	21	22
20	Grigorios Tsagkatakis	20	20	20	20	19	14	18
21	Panagiota Fatourou	21	21	21	15	13	16	15
22	Polyvios Pratikakis	22	25	24	21	23	20	19
23	faculty1	23	22	22	18	20	24	24
24	faculty2	24	23	23	24	24	25	25
25	faculty3	25	24	25	25	25	23	23

Table 20 Rankings of the 25 profiles of faculty members of UoC-CSD with respect to various criteria

	avgCo-authors		HI		HI/co	
School	Range	Median	Range	Median	Range	Median
Sciences and Engineering (6 departments)	1.83 to 34.2	5.57	9 to 92	33	1.06 to 19.31	5.79
Medical (1 department)	3.27 to 44.19	8.46	5 to 109	31	0.17 to 13.31	3.75
Humanities and Social Sciences (9 departments)	1.46 to 18.31	3.21	12 to 56	21.5	2.02 to 15.50	6.23
Entire University (all 300 members)	1.46 to 44.19	6.61	5 to 109	30	0.17 to 19.31	4.89

Table 21 Range and medians of average co-authors, HI and HI/co of the Schools of UofCrete

Also the ranking of the researchers in each of these schools is affected significantly if we consider the number of co-authors. This is evident if we look at the top-10 according to the different metrics: In particular, Table 22 shows the top rankings in Sciences and Engineering. Notice that only one person (Eleftherios N. Economou) appears to all the four rankings. Table 23 shows the top rankings in Medical School, here 4 persons

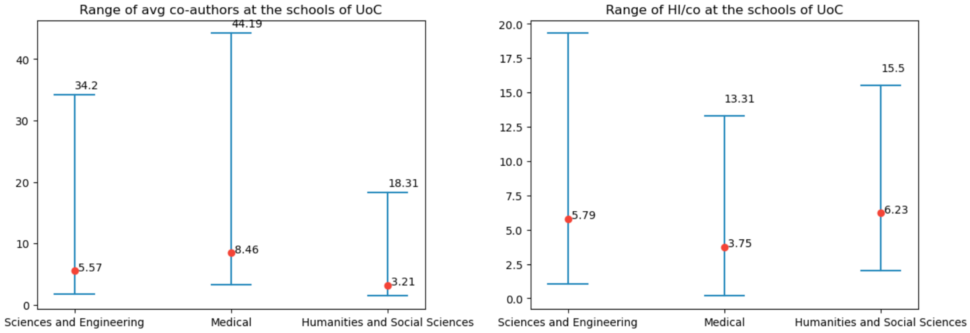

Fig. 5 Range and medians of the average co-authors and HI/co at the Schools of UoC

appear in all rankings. Finally, Table 24 shows the top rankings in Humanities and Social sciences, here two persons appear in all rankings.

	Wrt HI	Wrt HI/co	Wrt Pubs	Wrt Pubs/co
1	Nikos Mihalopoulos	**Eleftherios N. Economou**	Nikos Mihalopoulos	**Eleftherios N. Economou**
2	Costas Stoumpos	Eleftherios Zouros	A. Zezas	Giorgos Tsironis
3	Elias Kiritsis	Nikos Papanicolaou	Vassilis Charmandaris	JOHN DAMILAKIS
4	Vassilis Charmandaris	Elias Kiritsis	**Eleftherios N. Economou**	Yannis Stylianou
5	**Eleftherios N. Economou**	Vasilis Niarchos	JOHN DAMILAKIS	Yannis Tzitzikas
6	Maria Kanakidou	Ioannis G. Tollis	Nikos Tzanakis	Panagiotis Tsakalides
7	Euripides G. Stephanou	Yannis Stylianou	Maria Kanakidou	A. Zezas
8	Eleftherios Zouros	Nikos Frantzikinakis	Alexandros I. Georgakilas	Panos Trahanias
9	Nikos Tzanakis	Georgios Tziritas	Giorgos Tsironis	Ioannis G. Tollis
10	A. Zezas	Giorgos Tsironis	Costas Stoumpos	Evangelos Markatos

Table 22 The top-10 profiles of UoC/Sciences and Engineering with respect to various criteria

To conclude, we have seen big variations of the number co-authors in all schools of the University of Crete, which affects significantly the ranking of researchers.

6.7 Famous Scientists of the Past

Just for curiosity we also extracted and computed these metrics of some great scientists, that have a Google Scholar profile. The results are show in Table 25.

Of course, the eras (17th to 21st century) are not comparable in terms of the rate of production of scientific articles and the size of the scientific community. We should also consider that here we have scientists who are no longer alive, and their Google Scholar profile has many modern reprints of their old work. Consequently, the number

	Wrt HI	Wrt HI/co	Wrt Pubs	Wrt Pubs/co
1	panos vardas	Nektarios Tavernarakis	panos vardas	Aristidis Tsatsakis
2	Nektarios Tavernarakis	Aristidis Tsatsakis	Aristidis Tsatsakis	Ioannis G. Pallikaris
3	Aristidis Tsatsakis	Ioannis G. Pallikaris	MAVROUDIS DIMITRIOS	panos vardas
4	MAVROUDIS DIMITRIOS	A. Moschovakis	Christos Lionis	Apostolos Karantanas
5	George Samonis	Apostolos Karantanas	Nektarios Tavernarakis	Christos Lionis
6	Dimitris Mavroudis	George Samonis	Ioannis G. Pallikaris	Nektarios Tavernarakis
7	Christos Lionis	Filippatos Theodosios	Apostolos Karantanas	George Samonis
8	Ioannis Koutroubakis	panos vardas	George Samonis	Elias Castanas
9	Panagiotis Simos, Ph.D	Tzanakakis George	Dimitris Mavroudis	MAVROUDIS DIMITRIOS
10	Ioannis G. Pallikaris	Panagiotis Simos, Ph.D	Ioannis Koutroubakis	Achille Gravanis

Table 23 The top-10 profiles of UoC/Medical with respect to various criteria

of publications does not correspond to what they themselves wrote when they were alive. However, our data shows the low degree of co-authorship at that time, given the difficulty that existed then for remote collaboration. Although the eras are not comparable, while the HI of these famous scientists of the past is small (range 43-293) relative to the top university researchers we looked at (who have range 48-332, as shown in Section 6.3.6), their HI/co is higher: the range of HI/co of the famous scientists of the past is 20.02 - 159.06 while the range of the HI/co of the top researchers of the universities that we studied is 2.14-93.68. This is shown more evidently in Figure 6.

Obviously, evaluating past scientists through bibliographic indices requires a different approach (therefore the ranking presented in Table 26 is just given for reasons of completeness), but we have seen that HI/co can distinguish them, from contemporaries, better than HI.

6.8 Suggestions for the Community and the Bibliographic Services

Difficulties. Currently, the task of fetching and obtaining complete and accurate information about publications, authors, and citations, is not an easy task. We had to develop a scrapper for performing the analysis that we presented. Below we list the main difficulties and how we could overcome them, for making sush analyses more easy.

Access. Bibliographic sources do not enable in a straightforward manner to fetch all publications of a researcher in a structured manner. This is important for reasons of transparency.

	Wrt HI	Wrt HI/co	Wrt Pubs	Wrt Pubs/co
1	**Stamatios Papadakis**	**Stamatios Papadakis**	**Elena Anagnostopoulou**	Vassilios Makrakis
2	**Elena Anagnostopoulou**	Vassilios Makrakis	Emmanuel Petrakis	Kostas Vlassopoulos
3	S Giakoumaki	Kostas Vlassopoulos	Evangelos Karademas	Emmanuel Petrakis
4	Leonidas A. Zampetakis	Emmanuel Petrakis	Vassilios Makrakis	Vangelis Tzouvelekas
5	Emmanuel Petrakis	Leonidas A. Zampetakis	**Stamatios Papadakis**	**Elena Anagnostopoulou**
6	Gaganis Chrysovalantis	Dimitrios Stylidis	Ioannis N. Mammas	Panagiotes Anastasiades
7	George Panagis	Leonidas A. Zampetakis	Vangelis Tzouvelekas	Vangelis Tzouvelekas
8	Leonidas A. Zampetakis	**Elena Anagnostopoulou**	Maria Kousis	Maria Kousis
9	Vassilios Makrakis	Gaganis Chrysovalantis	Panagiotes Anastasiades	**Stamatios Papadakis**
10	Evangelos Karademas	Nicholas Zaranis	Vangelis Tzouvelekas	Stergios Chatzikyriakidis

Table 24 The top-10 profiles of UoC/Humanities and Social sciences with respect to various criteria

	Name	Citations	Publications	H-index	avgCo-A	HI/co
1	Sigmund Freud (1856-1939)	688855	3000	293	2.05	142.86
2	Friedrich Nietzsche (1844-1900)	356187	3000	230	1.45	159.06
3	Claude E Shannon (1916-2001)	222264	358	64	3.11	20.57
4	Charles Robert Darwin (1809-1882)	207469	2062	96	3.06	31.39
5	Albert Einstein (1879-1955)	164905	1044	125	1.61	77.54
6	Stephen Hawking (1942-2018)	149491	1062	133	2.29	58.08
7	John von Neumann (1903-1957)	147778	870	98	2.39	41.07
8	Richard Feynman (1918-1988)	117436	157	61	1.87	32.57
9	Paul Erdős (1913-1996)	101122	1608	130	2.53	51.48
10	Edsger Wybe Dijkstra (1930,2002)	77538	487	49	2.45	20.02
11	Alan Turing (1912-1954)	63571	299	44	1.77	24.82
12	Isaac Newton (1643-1727)	31730	945	70	2.36	29.68
13	Nikola Tesla (1856-1943)	9819	276	43	1.59	26.97

Table 25 Ranking of the famous scientists of the past with respect to Citations

Incompleteness. Some bibliographic sources contain, or display, *incomplete* information, e.g. in the bibliographic entry of a paper sometimes dots are used in case of many authors, reducing in this way the effectiveness of web scraping techniques, i.e. it makes scrapping more difficult.

	Name	wrt Citations	wrt Pubs	wrt H-index	wrt HI/co	wrt Pubc/co
1	Sigmund Freud	1	1	1	2	2
2	Friedrich Nietzsche	2	2	2	1	1
3	Claude E Shannon	3	**10**	9	**12**	12
4	Charles Robert Darwin	4	3	7	8	3
5	Albert Einstein	5	6	5	3	4
6	Stephen Hawking	6	5	3	4	6
7	John von Neumann	7	8	6	6	8
8	Richard Feynman	8	13	10	7	13
9	Paul Erdős	9	**4**	**4**	5	5
10	Edsger Wybe Dijkstra	10	9	11	13	9
11	Alan Turing	11	11	12	11	11
12	Isaac Newton	12	7	8	9	7
13	Nikola Tesla	13	12	13	10	10

Table 26 Ranking of the famous scientists of the past with respect to various criteria

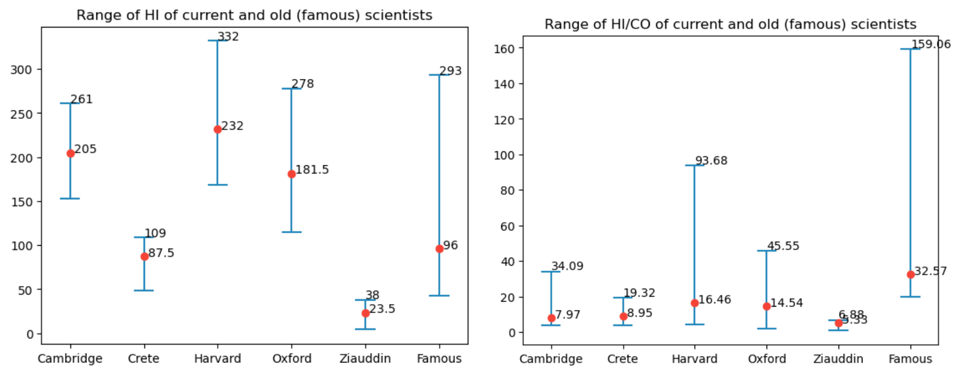

Fig. 6 Range and medians of HI and HI/co of the top-10 profiles of 5 universities and the group of 13 famous scientists of the past

Suggestions. It should be straightforward to fetch, in a structured manner, all publications of a researcher, and for each publication to get the complete set of authors, enabling in this way the computation of the average number of co-authors. Two suggestions follow:

Structured CVs. Just like each researcher maintains a CV in pdf, it would be beneficial to maintain (and have published) a file that contains in a structured manner all publications and complete information about each publication. That would enable the computation of the number of publications, average authors and years, easily, without having to rely to bibliographic sources, and without having to perform web scrapping.

Bibliographic Sources. Based on our analysis in this paper, we believe that the systems that compute citations and provide related access services (like Google Scholar, ResearchGate, and others), should not provide ranking by citations and year. We suggest as default method for ranking the number of citations divided by the average co-authors, or HI/co. In general, such systems should offer various options for sorting (not just by citations and date).

Moreover, for transparency and for fostering the development of new metrics, it would be beneficial if such sources offer an API through which one can get all citations of one particular paper, without having to perform web scrapping. An even better service of such systems (like Google Scholar, ResearchGate, etc) would be to allow the user to define the formula (or code) to be used for computing the desired metric and get the induced ranking.

6.9 Suggestions for Comparing Researchers

In brief, one fair measure to measure the *productivity* of a researcher is Pubs/co-authors, while to measure the *impact* of her research is Citations/co-authors. If we want to use a single metric, then we suggest using *HI/co*. Finally, if we want to compare two or more researchers of different academic age, with a single metric, we suggest HI/(coY) and Pubs/(coY).

If we want to compare several researchers, using more than one metrics, one step is to compute the *efficient set* (else called *Pareto front*, *maximal vector*, or *skyline*) according to Pubs/co and HI/co, i.e. to exclude those candidates for which there is at least a candidate with higher values on both Pubs/co and HI/co. This can reduce a lot the number of candidates. For instance, the Pareto front of the 25 faculty members of Section 6.5, comprises the following three members: Ioannis G. Tollis, Constantine Stephanidis, and Yannis Stylianou. In Figure 7 we can see the plot of the 25 members, the members in the Pareto front are in red.

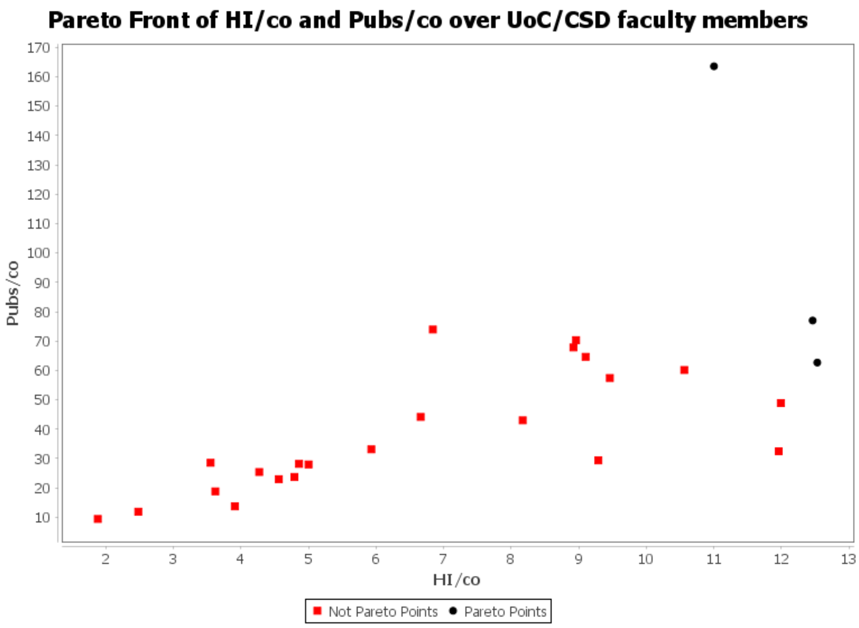

Fig. 7 Pareto front of the 25 faculty members of the UoC/CSD

In case the number of compared persons is low, it can be convenient to visualize the above metrics as a radar chart, for being able to show more than 2 metrics. An example, of a normalized radar chart, that shows the values of Pubs/co, HI/(coY) and HI/co for 3 professors from Section 6.5 is shown in Figure 8.

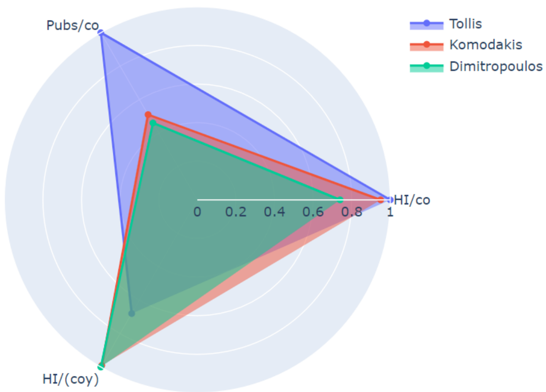

Fig. 8 Radar chart for comparing three researchers according to HI/co, HI/(coy) and Pubs/co

The above metrics usually are computed over all publications of an author. However, one one might decide to consider only the publications of an author in top-tier conferences and journals, and compute the H-index by considering only this restricted dataset. Again, the computation of the average number of co-authors (over that restricted dataset) will lead to more fair evaluation.

Other implications. As we stated in the introductory section, collaboration is good, not only for the involved individuals, but for the research community in general for various reasons: complementarity of expertise, resource sharing, improved quality, more impact, etc. Papers with many authors are not necessarily written to hack bibliographic metrics. Our proposal is not for discouraging collaboration, but for avoiding cases of long lists of non contributing authors. However we should be try to avoid as much as possible unfair evaluation.

7 Concluding Remarks

We need good measures not only to evaluate scientific output fairly, but also because they affect the goals and the activities of the scientific community. Obviously, the collaboration of researchers is not bad, quite the opposite, and papers with many authors are not necessarily written to hack bibliographic metrics. As we mentioned earlier, collaboration is beneficial, not only for the involved individuals, but for the research community in general: complementarity of expertise, resource sharing, improved quality, more impact, and others. However unfair evaluation is another thing, and we should try to use metrics that are as fair as possible, since they are used for hiring, promotion, funding, and recognitions. Therefore, it would be good for the community to discourage the misuse of the concept of author. Towards this objective, the key findings of our work, are the following:

- Without dividing the number of publications and the number of citations by the number of paper authors, each author gets the full credit of a joint work, something that is not fair. Through simulation scenarios we have showcased the impact of the factor F, i.e. the number of "friend" co-authors. For instance, in 10 years time, with the same effort a "lonely" researcher can get HI equal to 12, while a group of 5 researchers will each get a HI equal to 49.
- To tackle the weakness of HI, we proposed metrics, i.e HI/co, and HI/(coy). The results of simulations indicated that they can tackle these problems, i.e. equally strong researchers that have dedicated the same amount of effort, obtain the same values, independently of how many friend researchers they have.
- The measurements performed over real data of researchers from five universities, top ones as well as weak ones, revealed big variations of the number of co-authors. In total, we analyzed 526 authors, having in total more than 127 thousands publications, and 16.7 million citations. The range of the average co-authors of the top-10 researchers (according to citations) from these 5 universities, is from 1.94 to 114.45. We have also seen big variations in the number of co-authors, both at department level, school level and university level. Consequently, the consideration of the number of co-authors (through Pubs/co and HI/co), affects significantly the ranking of researchers. Indeed, the normalized Kendall's tau distance of these rankings ranged from 0.28 to 0.46, which is quite high.
- We have also seen that the metrics that consider the number of co-authors, are capable to distinguish the famous scientists of the past, from the current ones.
- One fair way to measure the *productivity* of a researcher is publications divided by the average co-authors, while to measure research *impact* we can use the number of citations divided by the average co-authors. If we want to use a single metric, then we suggest using HI/co. Finally, if we want to compare two or more researchers of different academic age, with a single metric, we suggest HI/(coy).

The are several directions for future research. One is to investigate diagrams and plots that facilitate the comparative evaluation of researchers. Another one, is to refine the notion of co-authorship and consider also the order of authors. Finally, another interesting direction is to elaborate on how additional criteria like open datasets, open source code, and others (e.g see [18] and the San Francisco Declaration on Research Assessment[11]), could be considered as well.

Acknowledgements. Many thanks to Evangelos Markatos, Xenofontas Dimitropoulos and Yannis Marketakis, for valuable comments and suggestions.

[11] https://sfdora.org/read/

Statements and Declarations

• Funding: None
• Conflict of interest/Competing interests: None
• Ethics approval: Not Applicable
• Consent for publication: Yes
• Availability of data and materials: The extracted datasets over which the metrics were computed are publicly accessible in JSON format at https://drive.google.com/drive/folders/1zB6zgJl4gP_vnMfl_9Oe9OyBuVejhy4n?usp=sharing
• Code availability: Upon request to the authors.
• Authors' contributions: All authors contributed to The study conception, design and writing of this work was done by Yannis Tzitzikas. The implementation of the system was done by Giorgos Dovas. All authors read and provided comments for improving the manuscript

References

[1] Hirsch, J.E.: An index to quantify an individual's scientific research output. Proceedings of the National academy of Sciences **102**(46), 16569–16572 (2005)

[2] Malesios, C., Psarakis, S.: Comparison of the h-index for different fields of research using bootstrap methodology. Quality & Quantity **48**, 521–545 (2014)

[3] Guns, R., Rousseau, R.: Simulating growth of the h-index. Journal of the American Society for Information Science and Technology **60**(2), 410–417 (2009)

[4] Engqvist, L., Frommen, J.G.: The h-index and self-citations. Trends in ecology & evolution **23**(5), 250–252 (2008)

[5] Bartneck, C., Kokkelmans, S.: Detecting h-index manipulation through self-citation analysis. Scientometrics **87**(1), 85–98 (2011)

[6] Zaorsky, N.G., O'Brien, E., Mardini, J., Lehrer, E.J., Holliday, E., Weisman, C.S.: Publication productivity and academic rank in medicine: a systematic review and meta-analysis. Academic Medicine **95**(8), 1274–1282 (2020)

[7] Shanmugasundaram, S., Huy, B., Shihora, D., Lamparello, N., Kumar, A., Shukla, P.: Evaluation of h-index in academic interventional radiology. Academic Radiology **30**(7), 1426–1432 (2023)

[8] Bornmann, L., Daniel, H.-D.: What do we know about the h index? Journal of the American Society for Information Science and technology **58**(9), 1381–1385 (2007)

[9] Egghe, L., *et al.*: An improvement of the h-index: The g-index. ISSI newsletter **2**(1), 8–9 (2006)

[10] Meštrović, R., Dragovic, B.: Extensions of Egghe's G-Index: Improvements of Hirsch H-Index. Available at SSRN 4408038 (2023)

[11] Bi, H.H.: Four problems of the h-index for assessing the research productivity and impact of individual authors. Scientometrics, 1–15 (2022)

[12] Egghe, L.: Mathematical theory of the h-and g-index in case of fractional counting of authorship. Journal of the American Society for Information Science and Technology **59**(10), 1608–1616 (2008)

[13] Hirsch, J.E.: An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. Scientometrics **85**(3), 741–754 (2010)

[14] Bihari, A., Tripathi, S., Deepak, A.: A review on h-index and its alternative indices. Journal of Information Science **49**(3), 624–665 (2023)

[15] Fire, M., Guestrin, C.: Over-optimization of academic publishing metrics: observing goodhart's law in action. GigaScience **8**(6), 053 (2019)

[16] Ioannidis, J.P., Maniadis, Z.: In defense of quantitative metrics in researcher assessments. Plos Biology **21**(12), 3002408 (2023)

[17] Kumar, R., Vassilvitskii, S.: Generalized distances between rankings. In: Proceedings of the 19th International Conference on World Wide Web, pp. 571–580 (2010)

[18] Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., Rafols, I.: Bibliometrics: the leiden manifesto for research metrics. Nature **520**(7548), 429–431 (2015)