

## Article

# VisualRPI: Visualizing Research Productivity and Impact

Chihli Hung <sup>1</sup>  and Wei-Chao Lin <sup>2,3,4,\*</sup> 

<sup>1</sup> Department of Information Management, Chung Yuan Christian University, Taoyuan 320, Taiwan; chihli@cycu.edu.tw

<sup>2</sup> Department of Information Management, Chang Gung University, Taoyuan 333, Taiwan

<sup>3</sup> Center for Institutional Research, Chang Gung University, Taoyuan 333, Taiwan

<sup>4</sup> Department of Thoracic Surgery, Chang Gung Memorial Hospital at Linkou, Taoyuan 333, Taiwan

\* Correspondence: viclin@gap.cgu.edu.tw

**Abstract:** Research productivity and impact (RPI) is commonly measured through citation analysis, such as the h-index. Despite the popularity and objectivity of this type of method, it is still difficult to effectively compare a number of related researchers in terms of various citation-related statistics at the same time, such as average cites per year/paper, the number of papers/citations, h-index, etc. In this work, we develop a method that employs information visualization technology, and examine its applicability for the assessment of researchers' RPI. Specifically, our prototype, a visualizing research productivity and impact (VisualRPI) system, is introduced, which is composed of clustering and visualization components. The clustering component hierarchically clusters similar research statistics into the same groups, and the visualization component is used to display the RPI in a clear manner. A case example using information for 85 information systems researchers is used to demonstrate the usefulness of VisualRPI. The results show that this method easily measures the RPI for various performance indicators, such as cites/paper and h-index.

**Keywords:** h-index; researchers' productivity and impact; information visualization; clustering



**Citation:** Hung, C.; Lin, W.-C.

VisualRPI: Visualizing Research Productivity and Impact.

*Sustainability* **2022**, *14*, 7679.

<https://doi.org/10.3390/su14137679>

Academic Editors: Miltiadis D. Lytras and Andreea Claudia Serban

Received: 7 May 2022

Accepted: 14 June 2022

Published: 23 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In higher education, many universities always keep recruiting suitable researchers to accomplish their teaching and research goals. On the other hand, (young) researchers may want to follow some promising scholars of a specific research field, in order to have a better understanding of this field. To achieve these goals, measuring and/or predicting the research productivity and impact (RPI) of researchers is a very important problem in the domain of academic evaluation.

The number of journal papers published is usually the main indicator used to assess the research performance of a scholar or scientist. The h-index is one of the most commonly used methods for judging the RPI of a researcher's publications [1,2]. It is based on the scientist's most frequently cited papers, and the number of citations that they have received in other publications [3]. The h-index is used to assist in decisions related to hiring, tenure, and promotion for researchers in many subjects, including information science [4], neuroscience [5], and management [6]. In addition to the h-index, several related citation analysis methods were proposed, such as the g-index and e-index, which focuses more on more frequently cited articles [7].

A well-known piece of public software, namely Publish or Perish (<http://www.harzing.com/pop.htm#download> (accessed on 6 May 2022)), was recently provided for researchers to analyze their RPI [8]. It provides many useful statistics, such as total number of papers/citations, average number of citations per paper/author/year, h-index, g-index, and so on.

Using statistics that can be easily collected online allows researchers to understand their current RPI, and compare it with their peers. For hiring, tenure, and promotion

decisions, it is common to select a number of domestic and/or international researchers, who conduct research on similar topics as the applicants, for comparison. However, it may be difficult for decision makers to make an effective assessment of the applicants' RPI over the chosen researchers using only these statistics at the same time.

For example, if there are 100 researchers selected, and each one has 10 different h-index-related statistics (i.e., variables), a  $100 \times 10$  database matrix is formed. In this case, it is relatively easy to rank the researchers using only one specific variable (i.e., dimension), such as the h-index. However, in practice, many of these statistics (variables) could be regarded as being of similar importance in assessing RPI, which makes it hard to rank them in multiple dimensions at the same time. Thus, when different types of statistics are examined simultaneously, one might like to highlight some specific dimension for further examination and comparison. For instance, the g-index is supposed to be more important than the other nine variables, but these other nine variables may still need to be compared for the selected group of researchers at the same time.

To this end, information visualization, which is a technique for producing visual (or graphic) representations of numerical and non-numerical data, can be used [9]. This technique, which allows users to reinforce human cognition, can be applied to visualize the structure of citation networks of scientific documents [10], or can provide an abstract form of engineering text via simple and intuitive graphics [11]. However, differing from the focus in previous studies, in this article, we present a novel prototype, namely visualizing research productivity and impact (VisualRPI), which can be used to visually assess a researcher's productivity and impact. VisualRPI is composed of both clustering and visualization components, where the clustering component is based on applying the growing hierarchical self-organizing map (GHSOM) approach [12,13] to cluster (or group) the selected researchers with similar statistics, and a visualization component (i.e., the interface), which presents not only the general hierarchical clustering results, but also different results of specified statistics that have higher weights than the other statistics in an easily understandable visual form.

The main contributions of VisualRPI are twofold. First, this is the first attempt to apply a clustering technique (i.e., GHSOM) to visualize RPI. In particular, we show that grouping RPI can be approached using the clustering technique. Since different clustering algorithms and distance functions can produce different grouping results, one future research direction could be examining the results obtained with different clustering algorithms and distance functions. Second, the hierarchical visualization scheme of VisualRPI provides an easy and clear comparison between different researchers in terms of different performance indicators.

The rest of this paper is organized as follows. Section 2 provides an overview of related studies in the literature discussing the use of information visualization techniques for bibliometrics. Section 3 introduces the proposed prototype system, i.e., VisualRPI, and presents a case study for using VisualRPI in the management information systems (MIS) field. Finally, Section 4 concludes the paper.

## 2. Literature Review

### 2.1. Information Visualization

Information visualization is an important technique for data analysis. In general, the visualization approaches apply histograms, scatter plots, surface plots, tree maps, parallel coordinate plots, etc., which help people understand and analyze specific domain datasets [9].

Specifically, the chosen dataset is presented in a visual form, which allows users to obtain insight into the dataset, draw conclusions, and directly interact with the datasets. Visual data mining techniques are useful for exploratory data analysis [14]. For example, let us examine the mining of research trends in computer science [15]. The clustering results present a general concept of the distribution of the identified research areas. They can also be visualized in terms of the relevance between different research areas, and their popularity in computer science.

In the research field of bibliometrics, the information visualization technique was used as a tool to analyze the scientific content of a specific journal during a specific period [16], the characterization and research trends of a specific domain problem [15,17,18], and the most productive institutions in a country in a specific research domain [19].

Recently, Heimerl et al. [10] developed a system, which they call CiteRivers, for the exploration and analysis of the scientific literature. This system is used for the visualization of the structure of citation networks and their exploration. It helps users effectively visualize the content of the cited publications, by interacting with visual abstractions of the dataset to correlate and filter different aspects of the documents.

However, to the best of our knowledge, there are no related studies that focus on visualizing RPI, perhaps because the challenge is multidimensional; that is, there are different indicators used to measure RPI, including the total number of citations, average number of citations per year, h-index, g-index, and so on. This makes application of the current clustering algorithms difficult for the effective visualization of RPI.

## 2.2. SOM and GHSOM Clustering Techniques

One popular information visualization technique is based on the application of clustering algorithms, such as the self-organizing map (SOM), for visual analysis of the chosen datasets [20]. A SOM is a type of artificial neural network, which produces a low-dimensional (usually two-dimensional), discretized representation of the input space of a dataset.

The training stage to construct a SOM is based on competitive learning. That is, when a training example is fed into the SOM, its Euclidean distance to all weight vectors is computed. The neuron with the weight vector, which is the most similar to the input, is called the best matching unit (BMU). The weights of the BMU and neurons close to it in the SOM are adjusted for the input vector. The formula for a neuron with the weight vector  $W_v(t)$  is

$$W_v(t+1) = W_v(t) + \Theta(v,t)\alpha(t)(D(t) - W_v(t)) \quad (1)$$

where  $\alpha(t)$  is a monotonically decreasing learning coefficient, and  $D(t)$  is the input vector. The neighborhood function  $\Theta(v,t)$  depends on the SOM distance between the BMU and neuron  $v$ .

A growing hierarchical self-organizing map (GHSOM) extends the SOM to allow hierarchical decomposition and navigation through sub-parts of the data. In addition, there is no need to pre-define the number of clusters. This technique is based on utilizing a recursive SOM algorithm to automatically provide a SOM hierarchy. The GHSOM starts with  $2 \times 2$  units (or neurons), and grows a whole row or column of units at a time. Two thresholds must be decided before training, which are growing and expansion thresholds. They are used to control the growth of a single map and the expansion of each neuron (i.e., the hierarchical growth) of the GHSOM, respectively. During the training process, the growing process continues until this growing threshold is achieved, or until too many unit-growing procedures are met. On the other hand, the expansion process terminates when a stopping criterion is fulfilled for a desired network size. For each map training cycle, GHSOM defines the topology by using the growing grid [21]. Figure 1 shows the architecture of a trained GHSOM.

Therefore, the clustering results of GHSOM are based on multiple two-dimensional maps, with a hexagonal or rectangular grid. Figure 2 shows an example of a GHSOM, using an animal dataset. There are six clusters on the first level, with large herbivores, medium-sized predators, and large predators on the left, and birds on the right. The second level clusters represent all birds, except the eagle.

Tangripiroj and Samadzadeh [23] shows another example of GHSOM over a collection of reusable components stored in a software repository. The clusters with high data densities are further hierarchically expanded, in which the top layer maps are presented with dark gray, whereas the bottom layer maps are light gray. The first layer map consists of  $3 \times 3$  neurons, which shows the three major clusters of software components: DS, IR, and

AI. Most neurons of the first layer maps are expanded in the second layer maps. Figure 3 shows an example of the AI sub map.

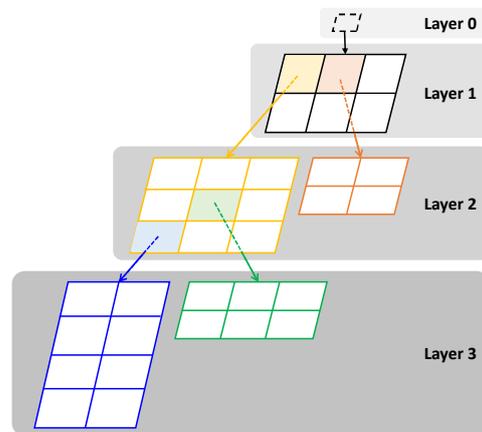


Figure 1. The architecture of a trained GHSOM [21].

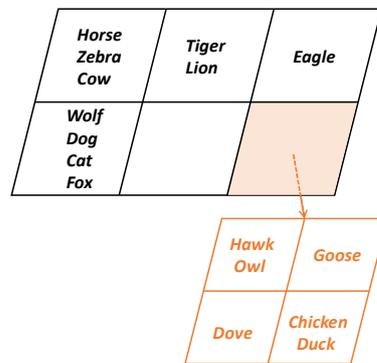


Figure 2. GHSOM trained with an animal dataset [22].

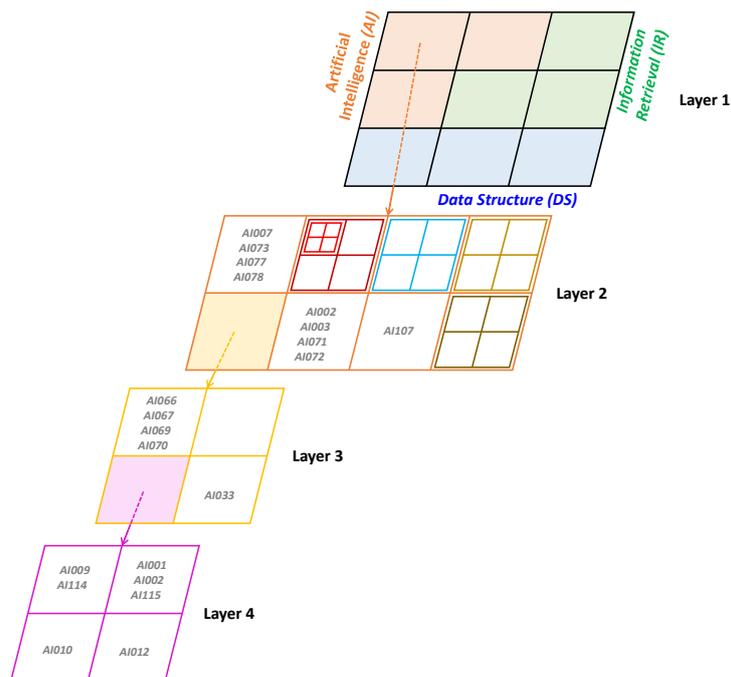


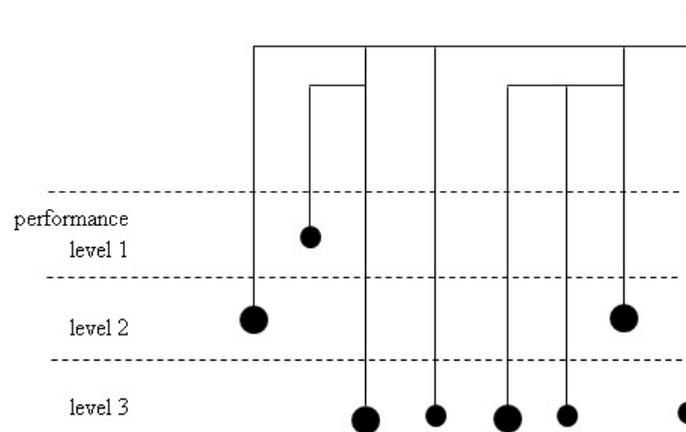
Figure 3. A four-layer GHSOM for AI sub map [23].

### 3. VisualRPI: A Novel Visualization Scheme

The proposed VisualRPI system is composed of both clustering and visualization components. Given a dataset containing the selected number of researchers, which are represented by h-index related statistics, the clustering component is used to cluster similar statistics into a number of groups. The GHSOM technique [13] is applied to handle this task (The GHSOM toolbox was downloaded from: <http://www.ifs.tuwien.ac.at/~andi/ghsom/download.html> (accessed on 6 May 2022)).

However, the hierarchical presentation of GHSOM is sometimes difficult to understand, and the RPI is not clearly displayed. In addition, given that different performance indicators can be used to assess the RPI, the GHSOM results cannot be used to directly represent the different levels of specific performance indicator(s) for each cluster. For example, the average number of citations can be divided into groups: high, middle, and low citations. The hierarchical clustering results of GHSOM for a chosen set of researchers do not allow comparison between different clusters of the three different levels of average numbers of citations.

To solve this limitation, we introduce a novel visualization scheme to better represent the clustering results by GHSOM. Figure 4 shows an example of this visualization scheme. This visualization interface displays the level of difference between the groups, based on specific statistics (i.e., performance indicators).



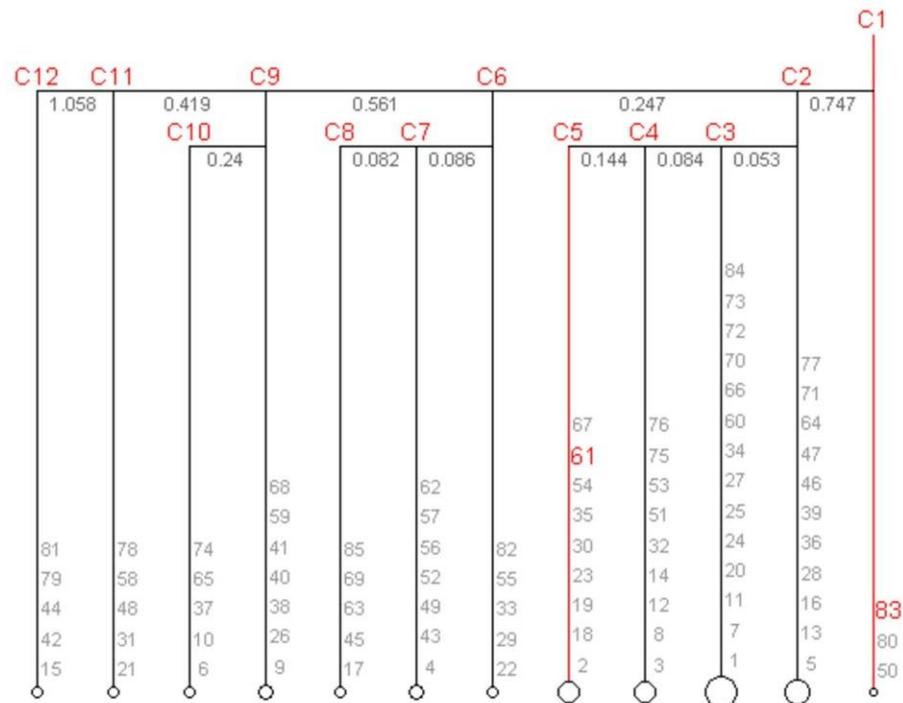
**Figure 4.** The visualization scheme of VisualRPI.

In this hierarchical scheme, each cluster is represented by a circle, with a larger circle meaning that there is a larger number of researchers in the group. It is also easy to identify which clusters are parent and child nodes. Moreover, the dotted lines separate the clusters, based on their performance levels. Clusters assigned to a higher performance level mean that the researchers in these clusters produce better research performances. For example, there is only one cluster with the highest average number of citations per paper.

### 4. A Case Study

A case study was conducted to show the performance of VisualRPI (VisualRPI is implemented by the GitHub packages of Python, and the computing environment is based on Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz with NVIDIA GeForce RTX 2080 Ti GPU, 64.0 GB memory, and the Window 10 operation system). In this case, 85 leading information systems (IS) researchers are considered in the data sample [24]. The Publish or Perish software is used to collect statistics on these researchers for a 5 year period, between 2016 and 2020. Particularly, 12 related types of statistics (i.e., features) are collected for each researcher (i.e., data sample), which are the number of papers, number of citations, years, cites/paper, papers/author, cites/year, cites/author/year, h-index (annual), h-index, g-index, hc-index, and h-index (norm). Note that these twelve features can be replaced by different statistics, if necessary.

Figure 5 shows the clustering results based on VisualRPI. Note that the 85 data samples are indicated only by ID numbers, to ensure personal privacy. In addition, two data samples (i.e., IDs 61 and 83, highlighted in red) are selected for comparison in this case study.



**Figure 5.** VisualRPI clustering results.

VisualRPI presents a general view of the clustering of these 85 researchers into several groups. In total, there are twelve groups identified in a three-level hierarchy, where cluster C1 is the root, and the values under the lines between two different clusters indicate their (Euclidean) distances in the 12-dimensional feature space. It indicates that researchers in the same groups have similar RPIs. On the other hand, the researchers in different groups, e.g., C2, C6, C9, C11, and C12, have different RPIs. In particular, the root node (i.e., cluster) may be divided into a number of sub-clusters, or child nodes, based on the distances between the members in the cluster. For example, C2 in the second level of the hierarchy. The members in C2 can be split into three sub-clusters, i.e., C3, C4, and C5. The distances between C2 vs. C3, C2 vs. C4, and C2 vs. C5 are 0.053, 0.084, and 0.144, respectively. In addition, the level of similarity between these sub-clusters (i.e., the distance between items) is larger than that of the clusters in the second level of the hierarchy (i.e., C2, C6, C9, C11, and C12).

It should be noted that clustering algorithms start by randomly choosing a data sample as the cluster center, around which similar data samples are grouped, which usually affects the grouping results of the members in the root node (i.e., C1). Thus, members in C1, obtained by performing the clustering algorithm several times over the same dataset, are not necessarily the same. However, related statistics for the members in the same cluster should be similar.

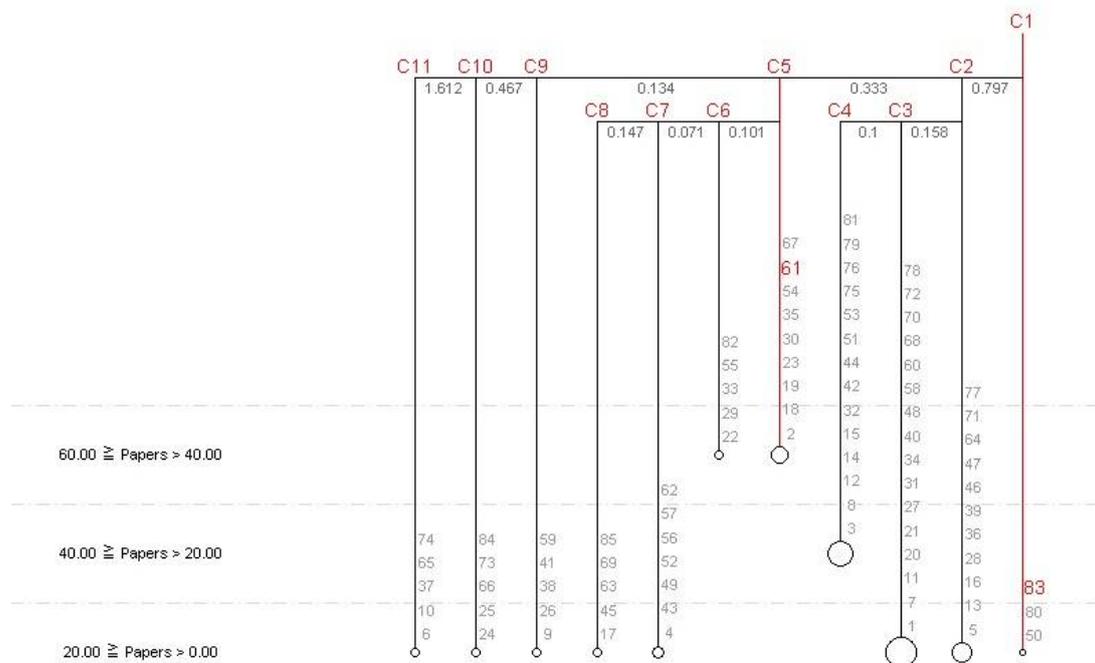
The cluster size (i.e., the circle size) reflects the number of researchers in the same group. Data samples in the same group are more similar to each other than to those in other groups. This generally allows us to understand that samples 61 and 83 certainly have different research productivity and impacts. Table 1 lists their related statistics.

**Table 1.** Related statistics for data samples 61 and 83.

ID	Papers	Citations	Years	Cites/ Paper	Papers/ Author	Cites/ Year	Cites/Author/ Year	h-Index (Annual)	h-Index	g-Index	hc-Index	h-Index (Norm)
61	240	3578	5	14.91	72.11	596.33	144.23	2	29	56	28	12
83	3	440	5	146.67	1.25	73.33	35.45	0.5	3	3	3	3

In some circumstances, some features may be given more weight during consideration. For example, the number of papers can simply represent a researcher’s productivity, and cites/paper represents his/her research impact. Therefore, a weight (e.g., higher than 1) can be assigned to each specific feature if it is regarded as more important than the other features. In other words, 1 out of the 12 input features can have a weighting value assigned, which is higher than 1, and the other 11 features’ weighting values are equal, i.e., 1. In this case, different weights ranging from 1 to 1.5, increasing by 0.1 intervals were examined. We find that assigning a weight of 1.2 makes the clustering results more distinguishable. That is, it is easier to compare differences in performance between different researchers. It should be noted that there are other weighting strategies for this purpose, e.g., the sum of weights for all features is 1. In addition, the optimal weighting value varies between different datasets.

Figures 6–8 show five different clustering results obtained by assigning a weight of 1.2 to each of the five different features. In addition, each higher weighted feature is divided into three scales. In terms of the number of papers and citations (i.e., Figures 6 and 8), researcher 61 (and the others in groups C5 and C6) has the highest level of research productivity. Clearly, without doubt, researcher 61 has significantly higher productivity than researcher 83. One possible reason may be that there may be a relatively larger number of publications in the research domain related to technically oriented IS than managerial-related research. This also directly affects the total number of citations. However, the aim here is not to discuss this issue.



**Figure 6.** The clustering result by number of papers.

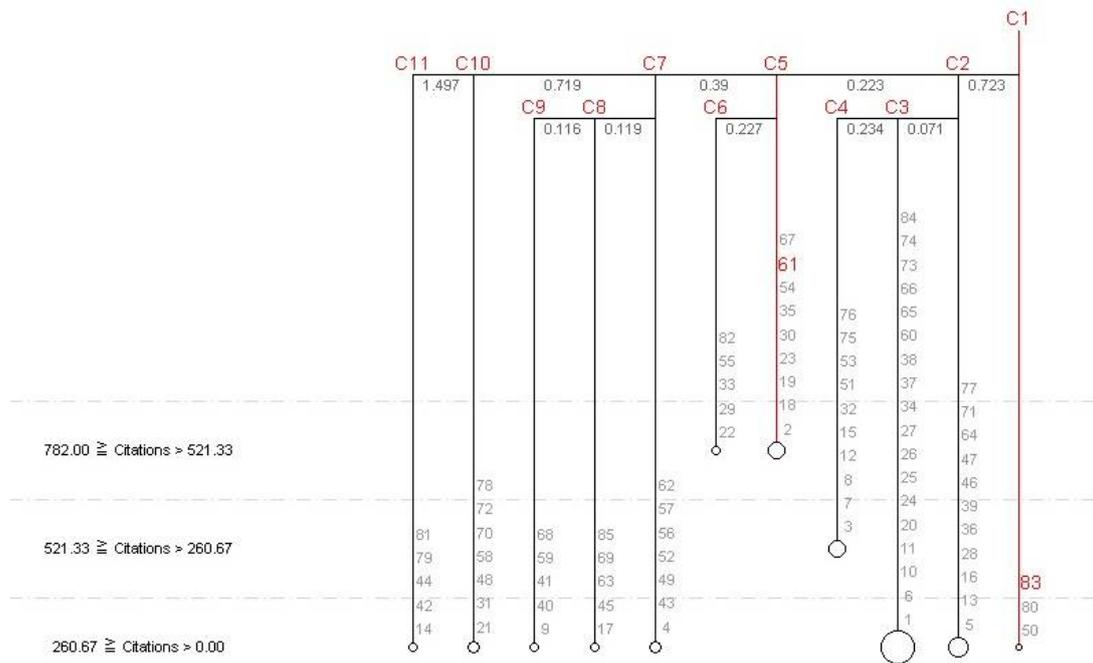


Figure 7. The clustering result by number of citations.

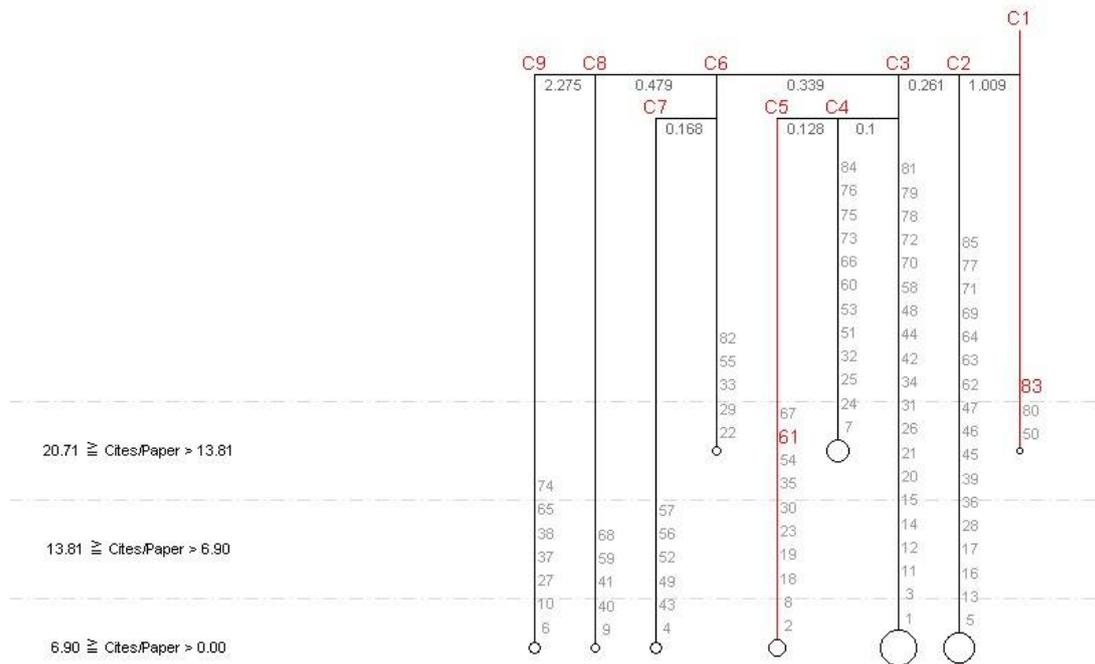


Figure 8. The clustering result by cites/paper.

On the other hand, when we examine the average number of citations per paper, i.e., Figure 8, researcher 83 has the highest level of research impact, significantly higher than researcher 61. This implies that researcher 83 focuses on higher impact research.

Sometimes, a researcher would like to see if his/her research performance is similar to other researchers. Suppose researcher ID 6 would like to understand his or her own RPI in comparison to the chosen IS researchers. Take cluster 10 (i.e., C10) as an example (c.f. Figure 5), Table 2 shows that these five researchers, IDs 6, 10, 37, 65, and 74, have similar research performances in terms of the 12 chosen features (i.e., performance indicators).

**Table 2.** Related statistics for cluster 10.

ID	Papers	Citations	Years	Cites/ Paper	Papers/ Author	Cites/ Year	Cites/Author/ Year	h-Index (Annual)	h-Index	g-Index	hc-Index	h-Index (Norm)
6	52	721	6	13.87	16.78	120.17	46.37	1.33	14	26	15	8
10	52	599	6	11.52	17.61	99.83	33.02	1.33	11	23	12	8
37	50	715	6	14.3	16.39	119.17	40.71	1.33	13	26	13	8
65	48	758	6	15.79	21.28	126.33	55.9	1.33	11	27	12	8
74	57	737	6	12.93	16.91	122.83	40.87	1.33	12	27	12	8
20	58	691	6	11.91	27.49	115.17	46.83	1.5	13	25	12	9

More specifically, when some performance indicators are considered to be more important than others, there are several researchers who, except for the researcher IDs 10, 37, 65, and 74, perform similarly to researcher ID 6, which are grouped into the same cluster. For example, for the number of citations in C3 (Figure 6), researcher ID 20 is grouped in the same cluster as researcher ID 6. In the general clustering results shown in Figure 7, the papers/author, h-index (annual), and h-index (norm) of researcher ID 20 are somewhat different from the other five researchers, therefore, the research performance of researcher ID 20 is regarded as dissimilar for the other researchers in C10.

However, if it is assumed that the h-index is a more critical performance indicator for the assessment of RPI (and the other performance indicators have the same level of importance, but are lower than the h-index), one can assign a higher weighted value to the h-index. In this case, the six researchers (as can be seen in Table 2) have a similar RPI.

## 5. Discussions

Based on the results of this case example, VisualRPI demonstrates its usability for visualizing the productivity and impact of researchers from various viewpoints. In particular, visualizing several different important statistics (i.e., higher weighted features) allows us to better understand the researchers' overall performances. Moreover, researchers whose performances are similar are clustered in one group. On the contrary, researchers with different performances are grouped separately. More specifically, the 'distance' between the clusters represents their level of similarity.

For individual researchers, the clustering results of VisualRPI easily identify how similar different researchers' performances are according to specific statistics, i.e., performance metrics. He or she can understand which statistics are insufficient, superior, or similar to his or her peers. On the other hand, for related academic and research institutions, the clustering results of VisualRPI are very helpful for making hiring, tenure, and promotion decisions. That is, by collecting related statistics of target and comparative researchers, such as tenured professors, within different periods of time, the clustering results based on different time periods recognize the progress of the target researchers when compared with the comparative ones.

However, it should be noted that for practical usage and implementation of VisualRPI, the initial dataset should be carefully collected in the first stage, depending on the research topic, the target researchers, comparison periods, citation statistics, etc.

## 6. Conclusions

Measuring researcher productivity and impact (RPI) is always a critical issue for hiring, tenure, and promotion decisions. In this article, we present a prototype system, called visualizing research productivity and impact (VisualRPI), which is designed to clearly visualize researchers' RPI over a range of various statistics, such as the average number of papers/citations, the average cites per paper/year, etc.

In our case study, by examining multiple statistics for researchers simultaneously, such as the number of papers, number of citations, cites/year, and cites/paper, the visualization results produced by VisualRPI demonstrate that researchers who have similar research

productivities are grouped in the same (sub-)clusters. In other words, the clustering results allow us to understand similarities in the research performance of researchers in the same groups, and vice versa. Moreover, when a specific statistic is regarded as more important than the others, it can be assigned a weighting value, which is higher than 1. Our results show that assigning a weighting value of 1.2 to some different statistics produces more explainable clustering results. That is, this can help us to obtain a more comprehensive understanding of a researcher's performance, in terms of research productivity and impact, which can effectively lead to more objective hiring, tenure, and promotion decisions.

Despite its usefulness, in the future, VisualRPI could be further improved by providing some other flexible functionality for users, such as manually defining the level of the presentation hierarchy, and the feature weight. This is because the assessment of RPI varies between different subjects. Moreover, some user studies or human judgments to evaluate VisualRPI should be conducted, in order to fully understand its performance. Statistical learning or optimization techniques can also be employed to obtain the weight for each feature automatically when needed. Last but not least, since the visualization scheme described in this paper is simply designed to coordinate the hierarchical clustering results, user studies should be conducted to discover the optimal presentation style of the clustering results because not all users will be satisfied and/or familiar with the results of this type of visualization.

**Author Contributions:** Conceptualization, C.H.; methodology, C.H. and W.-C.L.; software, C.H.; validation, C.H. and W.-C.L.; data curation, W.-C.L.; writing—original draft preparation, C.H. and W.-C.L.; writing—review and editing, C.H. and W.-C.L.; funding acquisition, W.-C.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported in part by the Center for Institutional Research, Chang Gung University from the Framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan, under grant UZRPD3L0041; and in part by Chang Gung Memorial Hospital at Linkou, under grants BMRPH13 and CMRPG3J0732.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Furnham, A. What I have learned from my Google scholar and h index. *Scientometrics* **2020**, *122*, 1249–1254. [[CrossRef](#)]
2. Ivanovic, D.; Ho, Y.-S. Highly cited articles in the information science and library science category in social science citation index: A bibliometric analysis. *J. Librariansh. Inf. Sci.* **2016**, *48*, 36–46. [[CrossRef](#)]
3. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [[CrossRef](#)] [[PubMed](#)]
4. Cronin, B.; Meho, L. Using the h-index to rank influential information scientists. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 1275–1278. [[CrossRef](#)]
5. Ameer, M.; Afzal, M.T. Evaluation of h-index and its qualitative and quantitative variants in neuroscience. *Scientometrics* **2019**, *121*, 653–673. [[CrossRef](#)]
6. Mingers, J. Measuring the research contribution of management academics using the Hirsch-index. *J. Oper. Res. Soc.* **2009**, *60*, 1143–1153. [[CrossRef](#)] [[PubMed](#)]
7. Ali, M.J. Understanding the 'g-index' and the 'e-index'. *Semin. Ophthalmol.* **2021**, *36*, 351–353. [[CrossRef](#)] [[PubMed](#)]
8. Harzing, A.W. *The Publish or Perish Book*; Tarma Software Research Pty Ltd.: Melbourne, Australia, 2010.
9. Ware, C. *Information Visualization: Perception for Design*, 4th ed.; Morgan Kaufmann: Burlington, MA, USA, 2019.
10. Heimerl, F.; Han, Q.; Koch, S.; Ertl, T. CiteRivers: Visual analytics of citation patterns. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 190–199. [[CrossRef](#)] [[PubMed](#)]
11. Sun, J.; Lei, K.; Cao, L.; Zhong, B.; Wei, Y.; Li, J.; Yang, Z. Text visualization for construction document information management. *Autom. Constr.* **2020**, *111*, 103048. [[CrossRef](#)]
12. Malondkar, A.; Corizo, R.; Kiringa, I.; Ceci, M.; Japkowicz, N. Spark-GHSOM: Growing hierarchical self-organizing map for large scale mixed attribute datasets. *Inf. Sci.* **2019**, *496*, 572–591. [[CrossRef](#)]

13. Rauber, A.; Merkl, D.; Dittenbach, M. The growing hierarchial self-organizing map: Exploratory analysis of high-dimensional data. *IEEE Trans. Neural Netw.* **2002**, *13*, 1331–1341. [[CrossRef](#)] [[PubMed](#)]
14. Cao, J.; Zhang, X. A survey on visual data mining techniques and applications. In Proceedings of the 2021 7th International Conference on Systems and Informatics (ICSAI), Chongqing, China, 13–15 November 2021; pp. 1–6.
15. Hoonlor, A.; Szymanski, B.K.; Zaki, M.J. Trends in computer science research. *Commun. ACM* **2013**, *56*, 74–83. [[CrossRef](#)]
16. Cobo, M.J.; Martinez, M.A.; Gutierrez-Salcedo, M.; Fujita, H.; Herrera-Viedma, E. 25 years at Knowledge-Based Systems: A bibliometric analysis. *Knowl. Based Syst.* **2015**, *80*, 3–13. [[CrossRef](#)]
17. Cheng, Q.; Lu, X.; Liu, Z.; Huang, J. Mining research trends with anomaly detection models: The case of social computing research. *Scientometrics* **2015**, *103*, 453–469. [[CrossRef](#)]
18. Garechana, G.; Rlo-Belver, R.; Cllleruelo, E.; Sarasola, J.L. Clusterization and mapping of waste recycling science: Evolution of research from 2002 to 2012. *J. Assoc. Inf. Sci. Technol.* **2015**, *66*, 1431–1446. [[CrossRef](#)]
19. Singh, V.K.; Uddin, A.; Pinto, D. Computer science research: The top 100 institutions in India and in the world. *Scientometrics* **2015**, *104*, 529–553. [[CrossRef](#)]
20. Kohonen, T. *Self-Organizing Maps*, 3rd ed.; Springer: Berlin/Heidelberg, Germany, 2001.
21. Fritzke, B. Growing grid-a self-organizing network with constant neighborhood range and adaptation strength. *Neural Process. Lett.* **1995**, *2*, 9–13. [[CrossRef](#)]
22. Pampalk, E.; Widmer, G.; Chan, A. A new approach to hierarachical clustering and structuring of data with self-organizing maps. *Intell. Data Anal.* **2004**, *8*, 131–149. [[CrossRef](#)]
23. Tangsripairoj, S.; Samadzadeh, M.H. Organizing and visualizing software repositories using the growing hierarchical self-organizing map. *J. Inf. Sci. Eng.* **2006**, *22*, 283–295.
24. Hu, P.J.; Chen, H. Analyzing information systems researchers' productivity and impacts: A perspective on the h index. *ACM Trans. Manag. Inf. Syst.* **2011**, *2*, 7. [[CrossRef](#)]