DSI DATA SCIENCE AND INFORMETRICS
VOL.3, NO.3, AUG. 2023

RESEARCH ARTICLE

Towards impact tenacity of academic institutions: A case study in Informetrics

Zhou Fang[a], Win–bin Huang[a,b], Yi Bu[a*]

a. Department of Information Management, Peking University, Beijing, China
b. Publishing Research Institute, Peking University, Beijing, China

ABSTRACT
Academic institutions are important subjects of discipline construction and knowledge production. It is crucial to study and evaluate the impact of institutions, and scientific evaluation is conducive to the development of academic institutions and the improvement of organizational efficiency. Most of the existing bibliometric indicators measure the impact of academic institutions from the perspective of citations (or variations). In this paper, we focus on a different side of their impact, namely tenacity. We conceptualize impact tenacity of academic institutions and define a series of indicators for operationalization. Besides, we implement correlation analysis and principal component analysis to explore whether the impact tenacity indicators and other bibliometric indicators are related and on the same dimension, taking informetrics as the representative discipline. We find that there was a significant negative correlation between the defined impact tenacity and number of papers, number of citations, and number of authors, and involved indicators (i.e., impact tenacity and other involved bibliometric indicators) describe two almost orthogonal dimensions. Moreover, this paper also selects the Max Planck Society and Taiwan Applied Research Laboratories (China) as case studies, and reveals that low-tenacity and high-tenacity institutions have quite different characteristics. Based on these findings, we make some constructive suggestions for research policy makers, such as considering maintaining a high tenacity of institutions by supplementing more academic training.

KEYWORDS
Impact tenacity; Academic institutions; Informetrics

1 Introduction

One of the most striking characteristics of modern science is discipline differentiation and segmentation (Bu et al., 2022; Sugimoto & Weingart, 2015). The scientific revolutions in the 16th, 17th, and 18th centuries were mainly in the domains of astronomy and life science, physics and mathematics, and technology and chemistry. These revolutions led to, especially in natural sciences, the formation of an independent and systematic discipline system. In the 19th century, the rapid development of natural sciences and the continuous improvement of

* Corresponding Author: buyi@pku.edu.cn

societal productivity brought about dramatic changes in the social structure with a more complicated social reality, which has led to the birth of many disciplines in the fields of humanities and social sciences, e.g., sociology, economics, law, and political science that were separated from philosophy (Clark, 1995; Lenoir, 1993).

Indeed, disciplines are the product of modernization. Disciplines consist of basic concepts, categories, principles, methods, expressions, and other elements that form an internal concept and discourse system (Xu et al., 2018, 2019), and an external cultural and educational system composed of universities, colleges, research institutions, societies, and foundations. The maturity of a discipline may best be reflected in its institutionalization (Bowker & Latour, 1987). Among many manifestations of institutionalization, the foundation of academic institutions connotes a quite substantial dimension: Institutions are responsible for "strict training, supervision of behavior and potential condemnation" (Hunt, 1994, p. 2). Academic institutions with degree-awarding department(s) are regarded as a significant milestone of a formally established discipline (Jacobs, 2014). In these operational processes, institutionalization has left a certain mark of authority and form for doctoral training and communication, which is necessary for discipline identity.

Academic institutions are important subjects of discipline construction and knowledge production. It is crucial to study and evaluate the impact of institutions, and scientific evaluation is conducive to the healthy development of academic institutions and the improvement of organizational efficiency. Many governments and funding agencies attach great importance to institutional evaluation and assessment. In 1993, for instance, the U.S.A. Congress enacted the Act (GPRA), which, for the first time, established institutional provisions for performance evaluation of U.S.A. federal departments and agencies, including federal laboratories and projects supported by federal funds. The great significance of GPRA lies in its mandatory performance evaluation provisions for federal laboratories, which institutionalize, standardize, and unify performance evaluation. In the U.K., the British Biotechnology and Biological Sciences Research Council (BBSRC) evaluates the scientific research of the eight research institutes funded under its strategy every five years. This evaluation ensures that the institutes continue to deliver incremental value and meet government policy objectives. For academic institutions, institutional research and evaluation can help them find their niche and make rapid progress. For governments and funding agencies, institutional research and evaluation can, either quantitatively or qualitatively, support strategic decision-making, improve the efficiency of science and technology funding, and identify priority areas for future funding. Therefore, it is crucial and invaluable to study academic institutions and the impact they produce.

However, there are still many limitations of evaluation in institutional impact. Measuring the impact of scientific research institutions is difficult because it is often not possible to establish a direct causal link between research output and outcomes. Evaluation of institutional impact is often limited to metrics such as citation counts and impact factors, which may not provide a complete picture of the quality or significance of the research. Moreover, insufficient attention has been paid to the distribution of impact at the level of institutional scientists. Institutions with the same number of citations may have different characteristics in terms of the distribution of scientists, leading to different performances when evaluating impact.

2 Impact Tenacity of Academic Institutions

Academic institutions, e.g., universities, typically contain multiple layers, for instance, the whole university level (e.g., the University of Oxford, U.K.), the school level (e.g., the School of Information at the University of Washington, U.S.A.), the department level (e.g., the Department of Information Resource Science at the School of Business, Nankai University, China), the discipline level (e.g., Network Science Institute at Indiana University Bloomington, U.S.A.), etc. In this paper, we particularly focus on the discipline level, that is, an institution that focuses on a certain domain of topics. Among the reasons why we focus on the discipline level, the most important one is that we are hoping to make our analyses semantically more focused and, meanwhile, without any "organizational" and/or administrative constraints beyond "academic" factors. In the latter sections of this paper, we will indicate the discipline level when mentioning "academic institutions" for simplicity.

An academic institution typically contains a group of active or inactive scientists that concentrate on one or a set of similar research topics. A "strong" academic institution tends to comprise a number of excellent scientists. Nonetheless, two "strong" academic institutions may have quite different characteristics. For example, institution A may contain a limited number of extremely excellent scientists (e.g., when talking about scientific impact, for example) and a great number of average-level scientists, while institution B may contain many scientists that have balanced excellence in their belonged discipline. As bibliometricians oftentimes adopt the number of citations of publications to operationalize the scientific impact of scientists, we here pay particular attention to their citation impact as a proxy for scientific impact (Waltman, 2016). As a toy example, Figure 1 shows such a difference for institutions A and B. In Figure 1, each character represents a scientist in his/her belonged academic institution, and its size is proportional to the number of citations his/her articles have received. From Figure 1, we see that, in academic institution A, most of the citations may come from the first two authors on the left and that the rest of the authors in the institution only contribute to a limited number of citations compared with the first two scientists. Suppose that these two authors were removed from the institution (e.g., retiring from the institution, moving to another institution by getting a new employment, etc.), the general scientific impact of the original academic institution would fall off a cliff. Yet, academic institution B has quite a different pattern in which the number of citations these authors have received looks quite balanced, despite the fact that the two institutions have roughly received a similar number of citations in total. We name this dimension as the *impact tenacity* of academic institutions.

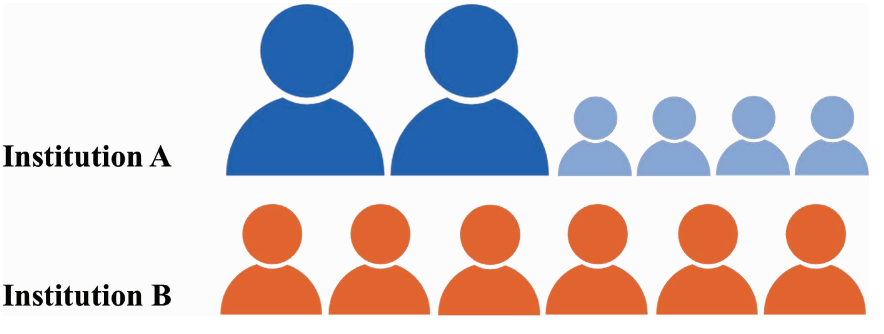

Figure 1 A toy example of academic institutions A and B. Authors in the two institutions are marked as blue and orange characters, respectively. Each character represents a scientist in his/her belonged academic institution, and its size is proportional to the number of citations his/her articles have received. Although the two academic institutions have roughly a similar number of citations in total, they have quite a different distribution.

3 Conceptualization of Tenacity and A Cross–Discipline Examination on Tenacity

The impact tenacity of academic institutions indicates the invulnerability regarding their impact. That being said, an academic institution with high impact tenacity tends to contain a group of scientists that have a similar number of citations (e.g., Institution B in Figure 1); on the other hand, institutions with low impact tenacity usually employ a few high-impact scientists but, simultaneously, some low-impact authors that contribute less to the institutions in terms of impact (e.g., Institution A in Figure 1). Note that impact tenacity highlights balance instead of the absolute level of impact. In other words, two institutions with a similar number of total impacts may have quite different impact tenacity. Tenacity emphasizes what the institution would be like and whether the differences would be revolutionary if it lost some top-ranked scientists from its current situation.

Tenacity highlights a persistent determination and is defined as the quality of not giving up something easily and the quality of being determined per the Oxford Dictionary. The concept and connotation of tenacity have been widely discussed and even quantified/approximated in various disciplines. For instance, in the domain of cybernetics, complex networks and systems, and statistical physics, people tend to measure tenacity with two distinct, yet interrelated, terms, namely percolation and resilience. Here, percolation defines a process where, when nodes or edges are removed from the original network, what the remaining network, as well as its characteristics, looks like (Callaway et al., 2000; Gao et al., 2013; Havlin et al., 2015; Li et al., 2021; Newman & Watts, 1999). In practice, percolation can be divided into two types, namely node percolation and edge percolation. Take a scientometric case as an example: In an international scientific collaboration network where nodes represent countries and edges indicate the number of co-authored publications of two specific countries, the top-ranked countries (nodes) are gradually removed from the original networks. People can then investigate, for example, how the largest connected component, an effective measurement for network connectedness, of the remaining network changes. If observing a sudden change (in a jargon word, phase transition), one would realize a system status transformation at a certain point (Fan et al., 2012). This is a typical case of node percolation.

Another word that is related to percolation is called resilience in network science. Emerging from the study of ecology, resilience is a critical dynamic property of complex systems. Among many fundamental properties of the science system, resilience is one that characterizes a network's ability to maintain its functionality (Holling, 1996; Liu et al., 2022) and quantifies the intensity of change in a network induced by its internal dynamics and exogenous perturbations (Gao et al., 2016). Network resilience provides a useful analytical framework to approach network dynamics and establish a basis for understanding and predicting the structural change of complex networks in science (Hirota et al., 2011; Hughes et al., 2005; Tang & Heinimann, 2018).

The current paper extends these beneficial discussions on the impact of academic institutions. Based on a bibliographic dataset in the domain of informetrics, this paper defines an indicator of the impact tenacity of academic institutions and analyzes the correlation between the tenacity indicators and existing publication indicators. We conclude with a detailed analysis of two typical cases selected from the dataset to further illustrate the specific meaning of institutional tenacity in reality.

4 A Measurement for the Impact Tenacity of Academic Institutions

As aforementioned, one of the purposes of this paper is to define a new measurement for the impact tenacity of academic institutions and to analyze how it relates to existing and frequently adopted bibliometric indicators in practice. Extant studies have explicitly pointed out that there are many distinct, though interrelated, dimensions to characterize impact, e.g., scientific impact, environmental impact, economic impact, societal impact, technological impact, etc.; among these, scientific impact is one of the most straightforward types of impact that sheds light on the science ecosystem itself (Bu, 2020). The number of citations has long been adopted as an important measurement of scientific impact.

Practically, some of these institutions are heavily polarized by their dependence on a few well-known authors for their high citations. The loss of these well-known authors will lead to severe fluctuations in the institution's total number of citations, which is considered low tenacity. The citations of some institutions are evenly distributed, resulting in the fact that the loss of authors will not affect the citations of institutions to a great extent, which is considered high tenacity. The specific calculation method of the tenacity indicator is proposed below.

Suppose there is an academic institution $\boldsymbol{I}$ with an article collection $\boldsymbol{C}$. The $\boldsymbol{n}$ authors associated with this institution are sorted by citations: $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_3$, ..., and $\boldsymbol{A}_n$. When all articles written by the most cited author $\boldsymbol{A}_1$ in the institution are removed, a subset $\boldsymbol{C}_1$ of $\boldsymbol{C}$ is obtained, and the proportion of the citations of $\boldsymbol{C}_1$ to the citations of $\boldsymbol{C}$ is denoted as $\boldsymbol{P}_1(=\boldsymbol{C}_1/\boldsymbol{C})$. Next, we continue to sort the authors in the subset by $\boldsymbol{C}_1$ citations, and then remove all articles written by the newly generated author with the greatest number of citations. This process is repeated until all authors have been traversed/removed. The number of removed authors corresponds to $\boldsymbol{n}$ proportional values: $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{P}_3$, ..., and $\boldsymbol{P}_n$. The schematic diagram of the whole process is shown in Figure 2.

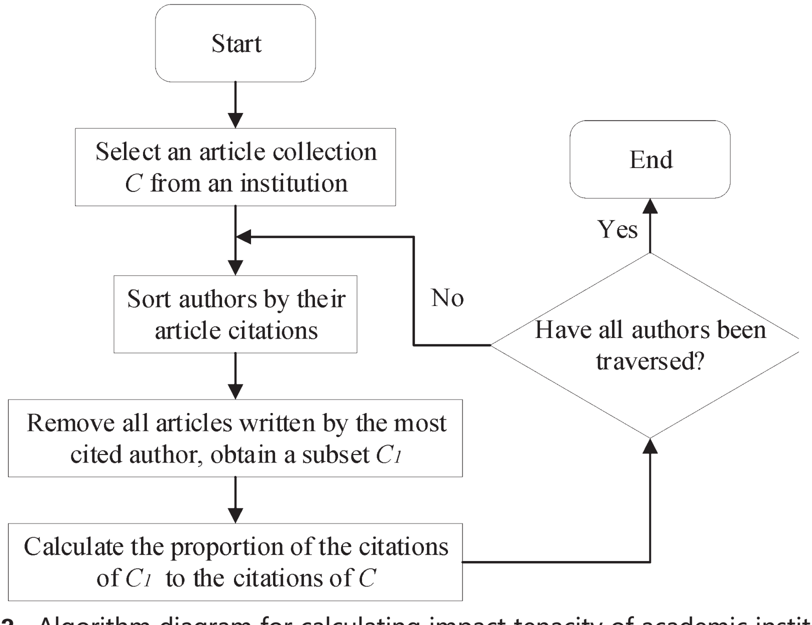

Figure 2 Algorithm diagram for calculating impact tenacity of academic institutions.

Be kindly noted that since each time the article collection of the most cited authors is removed, it affects the article collection of other authors co-authored with A_1, it is necessary to perform a reordering after each removal. This kind of operation also ensures that the calculated proportion of citations is continuously decreasing. In the end, there are even a large number of authors whose articles have been removed before their turn because they co-authored with the previous ones, and the proportion of citations is 0.

Figure 3 presents an example showing the relationship between the number of authors removed and the proportion of citations. S represents the integral area of this discontinuous function. S_1 and S_2 represent the areas of two special triangles, one with a base of the number of all authors in the institution, denoted as n ; the other with a base of the number of authors whose proportion of citations is not 0 according to the above algorithm, denoted as n_0 .

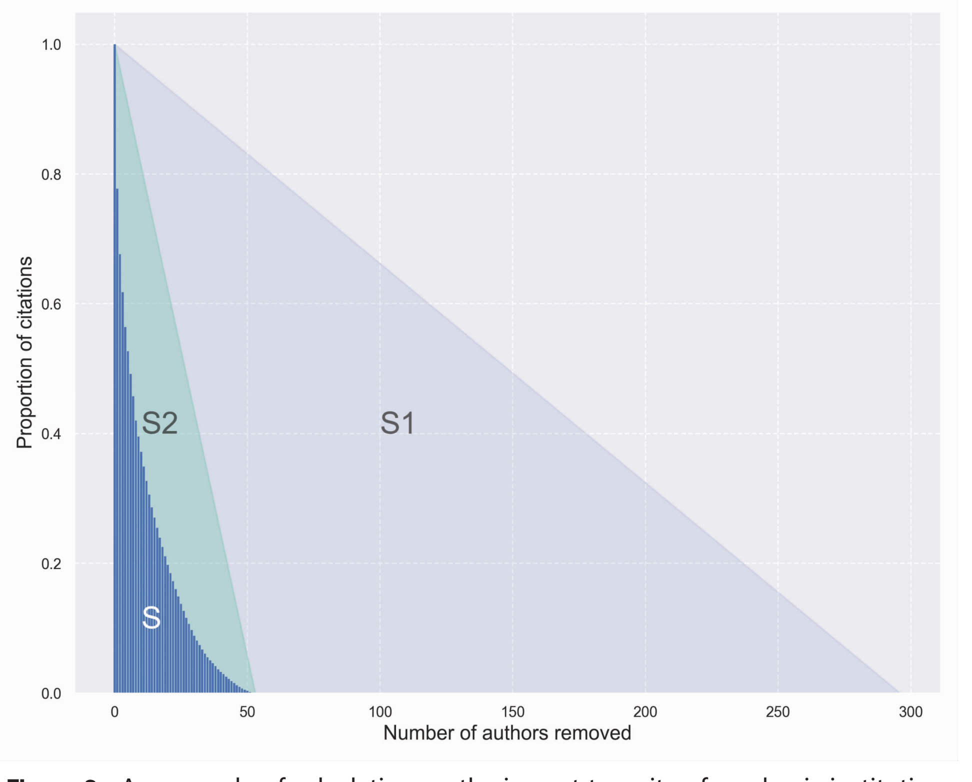

Figure 3 An example of calculation on the impact tenacity of academic institutions.

By calculating these variables, we can get:

$$S = \int_0^n P_i$$

$$S_1 = \frac{n}{2}$$

$$S_2 = \frac{n_0}{2}$$

We then define two tenacity indicators:

$$R_1 = \frac{S}{S_1}$$

$$R_2 = \frac{S}{S_2}$$

If the proportion of citations decreases at a completely uniform rate with the number of authors removed, then the value of the indicators is equal to 1, which means that the distribution of authors is the most balanced and that the institutional tenacity reaches its greatest. While the faster the rate of decline, the smaller S, which leads to the fact that smaller area ratio and smaller of the indicators. This symbolizes the uneven distribution of authors and the small tenacity of the institution. Therefore, the area ratio can represent the indicators of impact tenacity. Note that the value range of these two indicators is between 0 and 1, and the greater the value of the indicator, the greater the impact tenacity of the corresponding academic institution. Here, R_1 represents the institutional tenacity in the entire author group, and R_2 represents the institutional tenacity in the core author group. Obviously, we have $R_1 < R_2$.

5 Empirical Studies

5.1 Data

As a pilot study, this study does not aim to present a cross-disciplinary analysis on impact tenacity of academic institutions. Instead, we are hoping to showcase some descriptive analyses on one single domain with which the authors are familiar. This inspires us to select our domain, Informetrics, as the field of dataset. Informetrics is an interdisciplinary discipline that integrates bibliography, library and information science, mathematics, computer science, and statistics, which plays an important role in the information science system. *Journal of Informetrics* and *Scientometrics* are among the most authoritative and influential journals in the field of international informetrics. Therefore, this paper uses informetrics as a case discipline by taking the literature published in *Journal of Informetrics* and *Scientometrics* in the past ten years as representative data.

The article data comes from the Web of Science (WoS) database, with "*Journal of Informetrics*" OR "*Scientometrics*" as the search term in the publication name field. The time frame is limited to 2012-2021. A total of 4583 documents were retrieved, and their bibliographic data were downloaded as the initial data set for article analysis.

We do admit that the selection of these two journals cannot fully reflect all related publications in Informetrics. For example, there are many relevant works on *Journal of the Association for Information Science and Technology, Research Policy, Journal of Information Science, Journal of Documentation, Online Information Review*, and even multi-disciplinary journals (e.g., *Nature Human Behavior, Nature Communications*, etc.). Yet, as there are also publications in this journal that focus on other topics (e.g., information retrieval and behavior, social media, and user studies) and it is challenging to adopt any automatic strategy to exclude these contributions on large-scale datasets, we have to select only these two journals. As this is not a bibliometric analysis on this domain or not a descriptive analysis that shows the nuanced structure in this domain, we believe that such a strategy would not heavily affect the patterns behind the impact tenacity of academic institutions.

5.2 Overview

Table 1 shows the descriptive statistics of some indicators related to academic institutions, specifically the number of papers published by the academic institution, the total number of citations, the number of authors, and two indicators of tenacity R_1 and R_2. To ensure that R_1 and R_2 are greater than 0, the filtering condition for academic institutions is that the number of citations is greater than 0. Finally, a sample of 1,723 academic institutions is obtained.

Table 1 Descriptive statistics of academic institutions.

Indicator	N	Minimum	Maximum	Mean	Std. Dev.
Number of papers	1723	1	155	5.33	11.825
Number of citations	1723	1	4422	97.92	287.886
Number of authors	1723	1	307	12.23	21.541
R_1	1723	0.0097	0.6415	0.1642	0.0961
R_2	1723	0.1327	1.0000	0.6762	0.1715

Figure 4 is a histogram of the frequency distribution of R_1 and R_2 in the sample set, none of which completely satisfy the normal distribution. As is shown in the figure, the values of R_1 are mainly concentrated in the range of 0.1 to 0.2, and the values of R_2 are mainly concentrated in the range of 0.6 to 0.8. The majority of institutions have a more dispersed distribution of tenacity in the entire author group (R_1) with small values, suggesting that they are more dependent on a few highly cited authors. However, the distribution of tenacity in the core author group (R_2) is relatively concentrated, and the values are quite large. It is easy to understand that there is no strong dependence on the core author group.

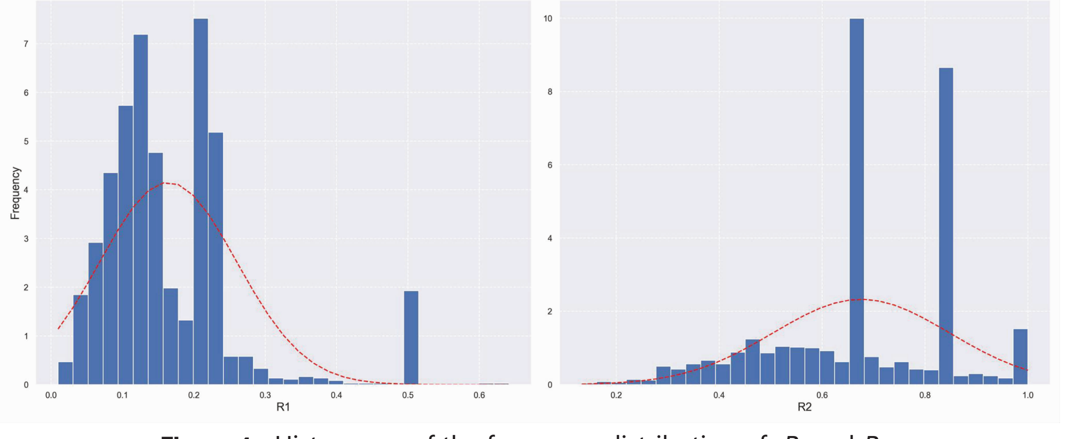

Figure 4 Histograms of the frequency distribution of R_1 and R_2 .

5.3 Correlation Analysis of Impact Tenacity and Other Bibliographic Indicators

It is necessary to analyze the correlation between the impact tenacity indicators and other bibliometric indicators. Table 2 shows the results of the correlation analysis, in which we can see the Spearman correlation coefficient and the significance among these indicators. The reason why a Spearman correlation coefficient instead of a Pearson's correlation coefficient is utilized is that most of these focused indicators do not follow a normal distribution.

Table 2 Correlation analysis results.

	Number of papers	Number of citations	Number of authors	R_1	R_2
Number of papers	1.000	0.790**a	0.792**	−0.434**	−0.557**
Number of citations	0.790**	1.000	0.669**	−0.399**	−0.522**
Number of authors	0.792**	0.669**	1.000	−0.748**	−0.572**
R_1	−0.434**	−0.399**	−0.748**	1.000	0.542**
R_2	−0.557**	−0.522**	−0.572**	0.542**	1.000

a. **. Correlation is significant at the 0.01 level (2-tailed).

There is a significant positive correlation between the number of papers, the number of citations, and the number of authors, and the correlation coefficient is relatively large, which is in line with people's common sense. However, there is a significant negative correlation between R_1 , R_2 and other bibliometric indicators. This indicates that the greater the number of papers, citations, and authors an institution has, the lower its impact tenacity, and the more likely it is to rely on well-known authors with high citations. One possible explanation is that well-known authors tend to work in larger institutions (with more students, workforce, and/or computational and social resources, etc.), leading to an extremely unbalanced distribution of citations, which, in turn, results in lower impact tenacity of institutions; when the size of the institution is small, the citation volume itself is limited and, there, is not much variability among authors, resulting in a high calculated tenacity of institutions. The details of the relationship among these indicators will be shown in the following sections.

5.4 Principal Component Analysis of the Impact Tenacity Indicators and Other Bibliographic Indicators

To implement a principal component analysis, the very first step is to analyze whether the research data are suitable for principal component analysis. It can be seen from Table 3 that the Kaiser-Meyer-Olkin (KMO) test value is 0.704, greater than 0.6, which meets the prerequisite requirements of principal component analysis. And the data passed Bartlett's Test of Sphericity ($p < 0.05$), indicating that the research data are suitable for principal component analysis.

Table 3 KMO and Bartlett's Test.

Kaiser–Meyer–Olkin Measure of Sampling Adequacy.		0.704
Bartlett´s Test of Sphericity	Approx. Chi–Square	6252.599
	df	10
	Sig.	0.000

As shown in Table 4, a total of two principal components are extracted from the principal component analysis. The eigenvalues of these two principal components are greater than 1, and the rotation variances explained by these two principal components are 52.577% and 33.049%, respectively, with a cumulative variance explained of 85.626%. Their corresponding weighted variances explained are: 52.577 / 85.626 = 61.40%; 33.049 / 85.626 = 38.60%.

Table 4 Total Variance Explained.

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.196	63.922	63.922	2.629	52.577	52.577
2	1.085	21.704	85.626	1.652	33.049	85.626
3	0.371	7.414	93.041			
4	0.273	5.457	98.498			
5	0.075	1.502	100.000			

Figure 6 shows the coefficients of the four indicators on the two rotated components. As shown in the figure, the three bibliographic indicators citations: the number of papers, citations, and authors fall on Component 1 with a large projection, but on Component 2 with a smaller projection. However, the impact tenacity indicators exhibit the opposite pattern. Therefore, these two sets of indicators actually represent two orthogonal dimensions. One dimension represents the size of the academic institution, such as the number of articles published by the institution and the number of authors belonging to the institution; the other dimension represents the impact tenacity of the institution, which specifically refers to the sustainability of an institution's development and whether it relies on a small number of authors to produce most of its content. The concepts discussed so far are relatively abstract, and the following section will give concrete examples to help illustrate them.

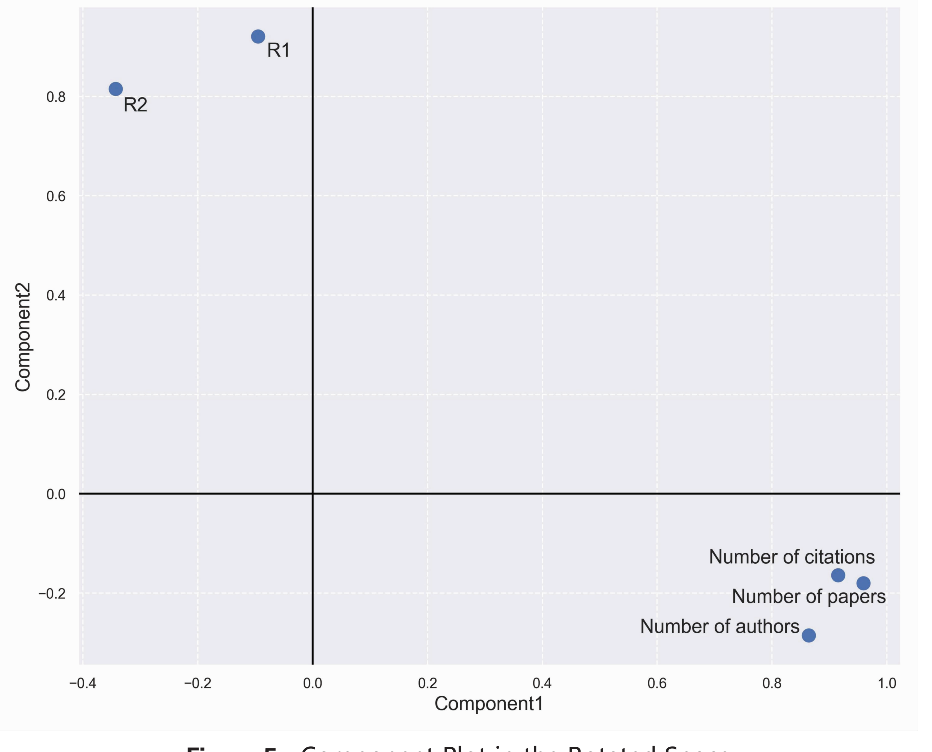

Figure 5 Component Plot in the Rotated Space.

5.5 Case Studies on Two Academic Institutions

We will describe two cases of academic institutions, one with low impact tenacity and one with high impact tenacity, so as to explain the specific meaning of tenacity indicators in reality.

5.5.1 Case Study 1: Max Planck Society

The Max Planck Society is the largest non-university scientific research and academic organization in Germany. Since its establishment in 1948, it has produced at least 18 Nobel laureates. The Society currently has 84 research institutes and laboratories, and has established 17 international research centers in cooperation with top scientific research and academic institutions in various countries. A total of more than 15,000 academic papers have been published on the website of the Max Planck Society, many of which are highly cited papers in related fields, reflecting the international influence of German scientific research. In this case study, we mainly examine its impact tenacity in the domain of informetrics (see details in the Section "Data").

Table 5 shows the value of each indicator of Max Planck Society. It is worth noting that Max Planck Society has a significant scale in terms of the number of papers and authors. It is the second most cited institution in the sample set, with 3,729 citations. However, its tenacity indicators are very low, among which R_1 is the last one, and R_2 is the penultimate one.

Table 5 Indicators of Max Planck Society (Case Study 1).

Institution	Number of papers	Number of citations	Number of authors	R_1	R_2
Max Planck Society	142	3729	74	0.0098	0.1463

Figure 6 shows a schematic representation of the calculation on tenacity indicators for Max Planck Society, and Table 6 lists the specific data of the authors removed one by one during the calculation process. The institution's tenacity is rather low, with only 5 core authors (i.e.,

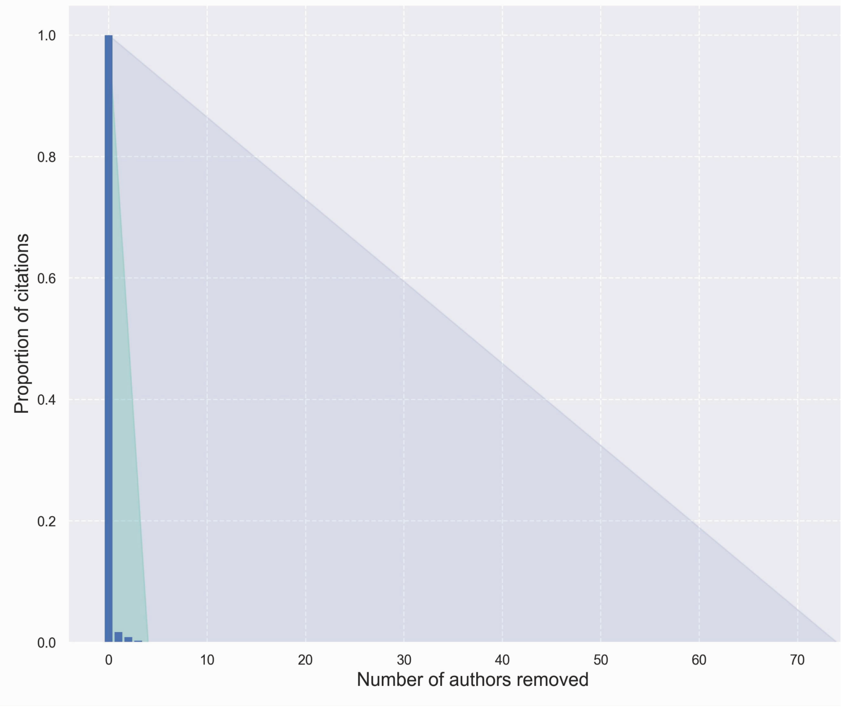

Figure 6 Calculation on the tenacity of Max Planck Society (Case Study 1).

removed authors with a non-zero proportion of citations) out of 74 authors. Obviously, Lutz Bornmann is the most "core" author in this institution, and the citations related to him account for 98.28%. This highlights the high degree of imbalance in the institution, which means that the institution's achievements in informetrics cannot be sustained once this author leaves the institution.

Table 6 Author distribution in calculation of Max Planck Society (Case Study 1).

Ranking	Author name	Removed citations	Remaining citations	Removed proportion	Remaining proportion
1	Bornmann, Lutz	3665	64	98.28%	1.72%
2	Haunschild, Robin	32	32	0.86%	0.86%
3	Korom, Philipp	22	10	0.59%	0.27%
4	Subbotin, Alexander	9	1	0.24%	0.03%
5	Howey, Riaz	1	0	0.03%	0.00%

5.5.2 Case Study 2: Taiwan Applied Research Laboratories (China)

Taiwan Applied Research Laboratories (China) is a regional institute under the Ministry of Science and Technology of Taiwan, China and entrusted with four core missions: to develop Taiwan Province's science and technology infrastructure, to support academic research of excellence, to enable frontier science and technology exploration, and to nurture high-tech human capital in Taiwan Province.

Many highly resilient institutions in the sample set are unrepresentative in terms of citation distribution due to the small number of publications and authors. Under the condition that the number of papers is greater than or equal to 10 and the number of authors is greater than or equal to 20, we screen out Taiwan Applied Research Laboratories, the institution with the largest R_1 value, as a representative of highly resilient institutions. Table 7 shows the value of each indicator of Taiwan Applied Research Laboratories. Taiwan Applied Research Laboratories is an academic institution of a certain size with a high degree of tenacity.

Table 7 Indicators of Taiwan Applied Research Laboratories (China) (Case Study 2).

Institution	Number of papers	Number of citations	Number of authors	R_1	R_2
Taiwan Applied Research Laboratories (China)	11	147	23	0.2043	0.7005

Figure 7 shows a schematic representation of the calculation on tenacity indicators for Taiwan Applied Research Laboratories, and Table 8 lists the specific data of the authors removed one by one during the calculation process. The impact tenacity of this institution is high, and the proportion of citations decreases uniformly as the number of removed authors increases. It has six core authors out of 23 authors, accounting for more than 25%, indicating that most of the authors have contributed to the institution's citation achievements in informetrics, and that the distribution of citations among authors is balanced. Even if individual authors leave the institution, this will not have a significant impact on the academic institution.

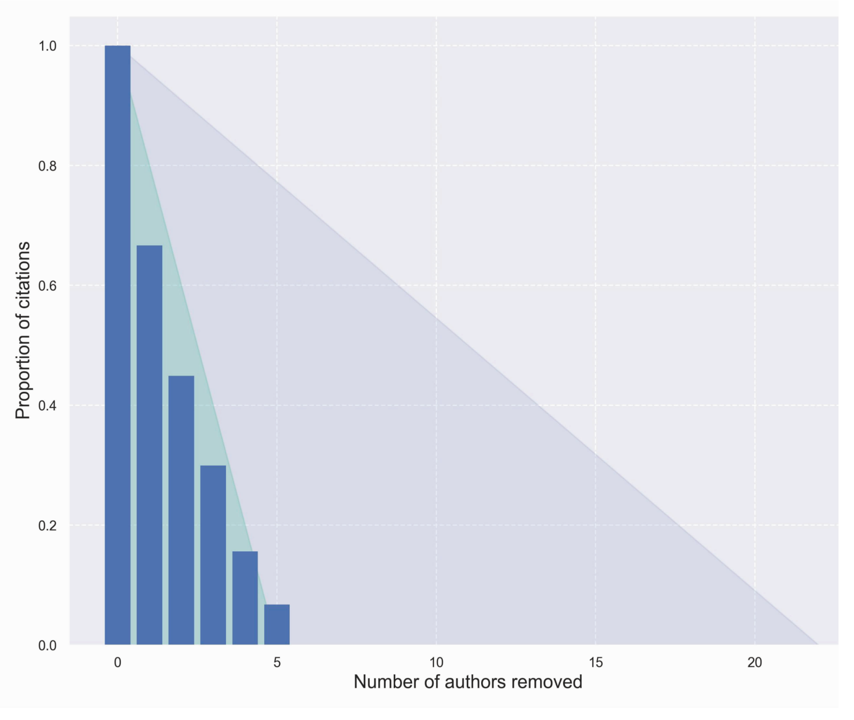

Figure 7 Calculation on the tenacity of Taiwan Applied Research Laboratories (China) (Case Study 2).

Table 8 Author distribution in calculation of Taiwan Applied Research Laboratories (China) (Case Study 2).

Ranking	Author name	Removed citations	Remaining citations	Removed proportion	Remaining proportion
1	Su, Hsin–Ning	49	98	33.33%	66.67%
2	Hung, Wen–Chi	32	66	21.77%	44.90%
3	Liu, Hsuan–I	22	44	14.97%	29.93%
4	Liaw, Yi–Ching	21	23	14.29%	15.65%
5	Yeh, Hsi–Yin	13	10	8.84%	6.80%
6	Chang, Shu–Hao	10	0	6.80%	0.00%

6 Conclusions

This paper proposes a new perspective for analyzing the performance of academic institutions and defines a new indicator: impact tenacity of academic institutions. The paper finds that there is a significant negative correlation between the defined impact tenacity and an academic institution's bibliometric indicators, the number of papers, citations and authors. Moreover, we implement a principal component analysis and notice that involved indicators describe two almost orthogonal dimensions. One dimension represents the size of the institution, while the other dimension indicates the impact tenacity of the institution. On this basis, the paper selects Max Planck Society and Taiwan Applied Research Laboratories (China) as representatives of low tenacity institutions and high tenacity academic institutions, and conducts a detailed analysis of the calculation process of their tenacity indicators and the distribution of authors' citations.

One thing that is worth noticing is that the definition and operationalization of impact tenacity of academic institutions is discipline-oriented. That being said, the value of impact tenacity of a certain discipline might be quite different from that of another discipline, even if one is considering the same academic institution. Thus, one important implication of the proposal of impact tenacity of academic institutions is that research policy makers may think of maintaining a high tenacity of institutions by supplementing more academic training. For instance, newly entered assistant professors in a university may attend some discipline-level mentoring programs that involve senior professors to supervise junior faculty members. Moreover, establishing a more collaborative academic environment might also be beneficial. In the context of the impact tenacity of academic institutions, we pay particular attention to these intra-institution collaborative works' citations; yet, if more publications in an institution are collaborative, a loss of senior, highly cited authors may not be a huge issue because of the sturdily growing junior members.

Our study still has many limitations. First of all, the applicable scenarios of the tenacity indicators are limited. The tenacity indicators are more meaningful in highly cited institutions. Otherwise, if an institution has almost no impact or has quite low impact, it does not matter how high its impact tenacity is. Secondly, the indicators we currently propose are only an approximate measure of the impact tenacity. They are more about using the balance of the scientists in an institution to understand the impact tenacity of academic institutions. This leads to the fact that it is not a problem for institutions to have top researchers and less excellent young scholars in the short term if there is a better mechanism for the intergenerational development of scientists. In terms of indicator improvement, consideration can be given to defining the extent to which how members of the institution differentiate from each other when measuring the impact tenacity in the future.

As aforementioned, the impact tenacity indicators constructed in this paper are used to evaluate academic institutions and are more appropriate for highly cited institutions. These indicators and the citation-based indicators are not in the same evaluation dimension, so the two may complement each other. They estimate tenacity mainly by measuring the balance of scientists in an institution and are, therefore, subject to some bias. In future research, the tenacity indicators of institutions can be further improved, and its relationship with other existing indicators of institutions is to be explored. The tenacity indicators constructed in this paper turns out to show a high level when the number of papers published by the institution is extremely small, but, in this case, it does not fully reflect the tenacity of the institution and can be adjusted as a breakthrough. Furthermore, in this paper, the indicator we define in this paper to measure the impact tenacity of academic institutions can be extended across disciplines and journals in the future as well.

Acknowledgements

This work was supported by the Social Science Foundation of China under Grant #20BTQ054. The authors are grateful for the fruitful discussions with all members of the Knowledge Discovery Lab at Peking University.

Declaration of Interest Statement

The authors declare no conflict of interest.

Disclosure Statement

The authors have nothing to disclose.

Biographical Note

Zhou Fang (ORCID: 0000-0002-8600-8331) is a master student majoring in Information Science at the Department of Information Management, Peking University. She holds a bachelor degree in Information Management and System from the Department of Information Management, Peking University. Zhou is doing research in bibliometrics, scientometrics, and scholarly communication.

Win-bin Huang (ORCID: 0000-0002-9174-5467) is a tenured Associate Professor at the Department of Information Management and at the Academy for Advanced Interdisciplinary Studies, Peking University, and the assistant to the dean at the Department of Information Management. Win-bin is doing research in data science, machine learning, learning science, and science of science.

Yi Bu (ORCID: 0000-0003-2549-4580) is an Assistant Professor at the Department of Information Management, Peking University. He holds a bachelor in Information Management and System from Peking University, a master in Data Science and a Ph.D. in Informatics from Indiana University. Yi is doing research in science of science and scholarly data analytics.

Data availability Statement

The authors will release all data records adopted in our analyses after publication upon request.

References

Bowker, G., & Latour, B. (1987). A booming discipline short of discipline: (Social) Studies of science in France. *Social Studies of Science, 17* (4), 715–748.

Bu, Y. (2020). *Understanding the citation–based impact of scientific publications through ego–centered citation networks* [Doctoral dissertation, Indiana University]. Indiana University.

Bu, Y., Chen, H., Xu, J., & Wang, Y. (2022). Connotations of interdisciplinarity from the perspective of paradigms: Towards "paradigms" in information science research and practices. *Information Studies: Theories and Applications, 45* (2), 1–10. (in Chinese).

Callaway, D. S., Newman, M. E., Strogatz, S. H., & Watts, D. J. (2000). Network robustness and fragility: Percolation on random graphs. *Physical Review Letters, 85* (25), 5468.

Clark, C. (1995). Competence and discipline in professional formation. *The British Journal of Social Work, 25* (5), 563–580.

Fan, J., Liu, M., Li, L., & Chen, X. (2012). Continuous percolation phase transitions of random networks under a generalized achlioptas process. *Physical Review E, 85* (6), 061110.

Gao, J., Barzel, B., & Barabási, A.–L. (2016). Universal resilience patterns in complex networks. *Nature, 530* (7590), 307–312.

Gao, J., Buldyrev, S. V., Stanley, H. E., Xu, X., & Havlin, S. (2013) . Percolation of a general network of networks. *Physical Review E, 88* (6) , 062816.

Havlin, S., Stanley, H. E., Bashan, A., Gao, J., & Kenett, D. Y. (2015) . Percolation of interdependent network of networks. *Chaos, Solitons & Fractals, 72*, 4–19.

Hirota, M., Holmgren, M., Van Nes, E. H., & Scheffer, M. (2011) . Global resilience of tropical forest and savanna to critical transitions. *Science, 334* (6053) , 232–235.

Holling, C. S. (1996) . Engineering resilience versus ecological resilience. *Engineering within Ecological Constraints, 31* (1996) , 32.

Hughes, T. P., Bellwood, D. R., Folke, C., Steneck, R. S., & Wilson, J. (2005) . New paradigms for supporting the resilience of marine ecosystems. *Trends in Ecology & Evolution, 20* (7) , 380–386.

Hunt, L. (1994) . The virtues of disciplinarity. *Eighteenth–Century Studies, 28* (1) , 1–7.

Jacobs, J. A. (2014) . In defense of disciplines. In *In Defense of Disciplines*. University of Chicago Press.

Lenoir, T. (1993) . The discipline of nature and the nature of disciplines. *Knowledges: Historical and Critical Studies in Disciplinarity*, 70–102.

Li, M., Liu, R.–R., Lü, L., Hu, M.–B., Xu, S., & Zhang, Y.–C. (2021) . Percolation on complex networks: Theory and application. *Physics Reports, 907*, 1–68.

Liu, X., Li, D., Ma, M., Szymanski, B. K., Stanley, H. E., & Gao, J. (2022) . Network resilience. *Physics Reports, 971*, 1–108.

Newman, M. E., & Watts, D. J. (1999) . Scaling and percolation in the small–world network model. *Physical Review E, 60* (6) , 7332.

Tang, J., & Heinimann, H. R. (2018) . A resilience–oriented approach for quantitatively assessing recurrent spatial–temporal congestion on urban roads. *PLoS One, 13* (1) , e0190616.

Waltman, L. (2016) . A review of the literature on citation impact indicators. *Journal of Informetrics, 10* (2) , 365– 391.

Xu, J., Bu, Y., Ding, Y., Yang, S., Zhang, H., Yu, C., & Sun, L. (2018) . Understanding the formation of interdisciplinary research from the perspective of keyword evolution: A case study on joint attention. *Scientometrics, 117* (2) , 973–995.

Xu, J., Ding, Y., Bu, Y., Deng, S., Yu, C., Zou, Y., & Madden, A. (2019) . Interdisciplinary scholarly communication: An exploratory study for the field of joint attention. *Scientometrics, 119* (3) , 1597– 1619.

Appendix

The appendix contains Tables A1 and A2.

Table A1 Top 10 institutions by their total number of citations received.

Ranking	Institution	Number of citations
1	Leiden University	4422
2	Max Planck Society	3729
3	University of Wolverhampton	3217
4	Consejo Superior de Investigaciones Cientificas (CSIC)	2679
5	University of Amsterdam	2595
6	Universite de Montreal	2412
7	KU Leuven	2347
8	Wuhan University	2265
9	Consiglio Nazionale delle Ricerche (CNR)	2201
10	Chinese Academy of Sciences	2194

Table A2 Top 10 institutions by their average number of citations per paper (screening threshold: published papers ≥ 5).

Ranking	Institution	Numbers of papers	Number of citations	Average citations per paper
1	University of Edinburgh	5	707	141
2	Middlesex University	7	934	133
3	University of Melbourne	12	1138	95
4	Pohang University of Science & Technology (POSTECH)	7	432	62
5	University of Southern Denmark	13	773	59
6	Universite de Montreal	41	2412	59
7	Southern Cross University	6	337	56
8	Deutsche Zentralbibliothek fur Wirtschaftswissenschaften (ZBW)	5	248	50