Infrastructure of Scientometrics: The Big and Network Picture

Jinshan Wu[†]

School of Systems Science, Beijing Normal University, Beijing 100875, China

Jinshan Wu, Professor of School of Systems Science at Beijing Normal University, China. He received his Ph.D. degree from University of British Columbia, Canada. Prof. Wu is a physicist who works in non-equilibrium quantum statistical physics, game theory and experiments, network science and scientometrics. He sees connections beyond boundaries.

Citation: Jinshan Wu (2019). Infrastructure of scientometrics: The big and network picture. *Journal of Data and Information Science*, 4(4), 1–12
DOI: 10.2478/jdis-2019-0017
Received: Nov. 16, 2019
Accepted: Nov. 29, 2019

Abstract

A network is a set of nodes connected via edges, with possibly directions and weights on the edges. Sometimes, in a multi-layer network, the nodes can also be heterogeneous. In this perspective, based on previous studies, we argue that networks can be regarded as the infrastructure of scientometrics in the sense that networks can be used to represent scientometric data. Then the task of answering various scientometric questions related to this data becomes an algorithmic problem in the corresponding network.

Keywords Network science; Scientometrics

1 Introduction

Scientometric analyses are often used in evaluating the scientific performance of journals, universities or institutes, countries, and sometimes even individual papers and authors. We refer to them as research/publication units at various levels. These research units are often ranked according to one or more indicators. In generating such rankings, often simple counting and statistics are used, for example, counting how many citations each paper received and from there to count how many citations other research units received in total or on average. Of course, a final step in such counting and statistics often involves a weighted average to incorporate various counting schemes into a single number. Due to this prevailing paradigm of

[†] Corresponding author: Jinshan Wu (E-mail: jinshanw@bnu.edu.cn).

Perspective

scientometric studies, scientometrics often appears to be a scientific discipline mainly providing tools and results for evaluating science with simple statistical analysis, such as calculating mean, median, mode, distribution, and sometimes also hypothesis testing as its main tools.

In this perspective, we will argue that scientometrics is a scientific discipline that collects and makes use of data related to scientific activities to help the development of the sciences, and for that purpose, we need a framework to represent relational data and to answer the questions that are potentially helpful to the development of other sciences. We claim that this framework is a network.

Of course, we notice that many other scientometric studies have already looked into regularities or laws of various scientific activities aiming to help the development of other sciences and also scientometrics as a science, rather than directly working towards evaluation. See, for example, Waltman (2016) and Mingers & Leydesdorff (2015), Rousseau et al. (2018) and references therein.

Moreover, there exist already quite some studies using network analysis as a tool to answer scientometric questions. For example, Otte and Rousseau point out that many network analysis techniques are applicable to scientometric studies (Otte & Rousseau, 2002, Rousseau et al., 2018). Waltman and van Eck use clustering algorithms to classify papers into clusters of topics in the citation network of papers (Waltman & van Eck, 2012). Chen applies pathfinding methods of network analysis to get a better presentation of the paths over the co-citing or co-cited network of papers or authors (Chen, 2006). West et al. apply PageRank algorithm (Brin & Page, 1998) for networks of web pages as a ranking method in the citation network of journals, authors, or papers (West et al., 2010). In the case of citation networks, the PageRank algorithm assumes that papers cited by more influential papers are themselves more influential. PageRank is derived from the Pinski-Narin approach introduced in the fields of bibliometrics, much earlier (Pinski & Narin, 1976). Amjad et al. extend the PageRank algorithm to multi-layer networks, including both authors and papers (Amjad et al., 2015). In this extension, the PageRank algorithm further assumes that papers cited by more influential papers and written by more influential authors are themselves more influential. Shen et al. (Shen et al., 2016) maps the interrelation among subfields of physics using the citation network of the subfields and a PageRank-like algorithm developed from the Leontief input-output analysis in economics (Leontief, 1941). Shen et al. apply network embedding and clustering method to cluster journals (Shen et al., 2019). We refer the reader to the review by Zeng et al. (Zeng et al., 2017) for more network-related studies in scientometrics.

However, networks can offer much more than just being a technique applicable to scientometric studies. We will show in the following sections how networks can

be used as a language to express relational data and hence answer questions in scientometrics. They can provide at the same time a framework, a language, and a platform to develop algorithms to solve interesting – and maybe even essential – questions within a scientometric framework.

2 Core ideas of network science and how they fit into a scientometric framework

Before explaining why networks can serve as the language or infrastructure of scientometrics, let us briefly recall the core ideas behind networks. First, a network is an abstraction of entities and their relations. A set of entities with pairwise (or one-to-one) relations can be represented as a network. Here pairwise (one-to-one) means that the relationship is between two entities, or more accurately from one entity to another. Many relations that might seem to be multi-party can often be seen as a collection of several pairwise relations. For example, the famous three-body question in physics, where three stars are moving under the gravity interaction among them, is, in fact, a collection of three pairwise gravity interactions. When a speaker is talking to a room full of audience, it may seem that the speaker is interacting with all of them together, and there are even interactions among the audience. However, it is also possible to see this as a collection of one-to-one interactions between the speaker and each member of the audience, plus possibly interactions between each pair in the audience. Of course, there might be some inherently multi-party interactions, which current networks cannot describe. Often when using networks, people simplify the interaction even further, for example, by discarding the strength of the interaction and keeping only its existence. Conventionally, the mathematical notation $A_j^i = 1$ means there is an interaction between entity i and entity j (or directed from i to j) and $A_j^i = 0$ otherwise. Thus, this matrix $A = (A_j^i)_{N \times N}$ describes the existence of an interaction between each pair of N entities. Often such an overly simplified description serves as a good platform for many investigations involving these entities. For example, the citation network sets $A_j^i = 1$ when paper i is cited by paper j while we clearly know that not all citations should be counted equally. On one hand, this is indeed a limit, and thus, one might want to consider extensions such as assigning a weight to A_j^i by taking into account the number of citations from j to i and where the citations are in j to a certain degree. On the other hand, however, many analyses can already be done over this limited network. Besides weighted networks, a further extension will be to consider heterogeneous nodes and heterogeneous edges. For example, one might want to present authors and papers all together in a single network. In this case, the relation among papers are citations, the relation among authors are mentor-mentee

Perspective

or other academic or social relations, and the relation between authors and papers are “writing” (authorship). We can see that there are three kinds of relations and two kinds of nodes.

One might wonder why not use co-author as the relation among authors. Doing this would violate a core network idea, namely to consider other relations and phenomena of interest as direct or indirect connections of the simplest possible network. Co-authorship can be seen as an induced relation of the authors-write-papers network. In mathematical term, if we denote the authors-write-papers network as a matrix $W = (W_p^a)_{L \times N}$, then the co-author network can be generated via

$$A_{a_2}^{a_1} = \sum_{p=1}^N W_p^{a_1} W_p^{a_2} = WW^T \quad (1)$$

where a_1, a_2 are two authors and p is a paper. This is one example of “calculating” an induced relation from a more fundamental network. In network analysis, given the fundamental network, which ideally captures all necessary relations, researchers always try to answer other questions via definitions of certain more advanced structures and related algorithms, especially those making use of direct and indirect connections via various orders of the fundamental network just like the WW^T in Eq. (1). Given a simple network A , we call A_j^i the direct (length-1) connections, $(A^2)_j^i$ the length-2 indirect connections, and so on.

Researchers from network science already defined quite some structural quantities over networks, for example, degree, PageRank score, and community. Using again a citation network as an example, we see that the counting of the number of publications of a research unit is the zeroth order quantity of the network, taking no connections in the citation network into consideration at all. A counting of the number of received citations is the first order quantity of the network, considering only the direct connections in the citation network. A PageRank score, say of papers, on the other hand, measures the influence of the papers to other papers along the citation path with all possible lengths,

$$P = \frac{\alpha e}{1 - (1 - \alpha)\mathcal{F}} = \alpha e + \alpha(1 - \alpha)e\mathcal{F} + \alpha(1 - \alpha)^2 e\mathcal{F}^2 + \dots, \quad (2)$$

where

$$\mathcal{F}_j^i = \frac{A_j^i}{\sum_k A_k^i} \quad (3),$$

and αe is the zeroth order influence score of each paper. Sometimes e is taken to be $e = [1, 1, \dots]$; $\alpha(1 - \alpha)e\mathcal{F}$ is the length-1 path influence (the papers directly citing the targeted paper), $\alpha(1 - \alpha)^2 e\mathcal{F}^2$ is the length-2 path influence and so on. There are other structural quantities and algorithms that take various orders of indirect

connections together with direct connections into account, for example the General Input-Output analysis (Shen et al., 2016) and the K-core (Alvarez-Hamelin et al., 2008).

We hope that the above examples sufficiently illustrate why networks can serve as a language or infrastructure of scientometrics. Networks describe relations among entities, and often network analysis starts from some fundamental network and makes use of both direct and indirect connections in the fundamental network via proper mathematics/algorithms to answer more advanced questions. We refer to this collection of connected entities as a system. From the detailed connections, we observe the micro-level structures of the system, and from the more advanced network analysis, which takes both direct and indirect connections into consideration, we can see the macro-level structures of the system. Therefore, the network framework and network analysis serve as a bridge between micro and macro-level structures.

Equipped with these two core ideas of network science, which, from now on, we refer to respectively as “relation” and “propagation”, we now try to use networks to represent, at least many if not all, scientometric data, research questions, and analysis.

3 The three-layer fundamental network of scientometrics

As we mentioned earlier, scientometrics studies all kinds of scientific activities to help the development of the sciences. What are the major actors – to become nodes in the fundamental network of scientometrics – of those scientific activities? Since the core scientific activities are researchers performing and publishing their researches, researchers, research questions (also methods, instruments, and materials), and papers should be considered as the major actors. We call them respectively, the authors, concepts, and papers, and represent them as a three-layer network in Fig. 1. What are the relations between them? Within each layer, among authors, there are academic or social relations, such as the mentor-mentee relation; among the papers, of course, there are citations; among the concepts, there are logic relations from disciplinary knowledge. Between the layers, authors “write” papers; papers “work on” or “use” certain concepts.

Let us denote the network as a matrix W , where $W_{a_2}^{a_1} = 1$ means author a_1 supervises author a_2 ; $W_{p_2}^{p_1} = 1$ means paper p_1 is cited by paper p_2 ; $W_{c_2}^{c_1} = 1$ means concept c_1 provides logically the basis of concept c_2 , while W_p^a means author a write paper p ; W_c^p means paper p works on concept c .

Then other common relations studied in scientometrics can be defined according to this fundamental network. For example, co-authorship becomes an induced

Figure 1. A three-layer network of scientometric relational data: Authors, papers and concepts on the one hand; and inventors, patents, and technology concepts on the other.

relation from the fundamental network W , $A_{a_2}^{a_1} = \sum_p W_p^{a_1} W_p^{a_2} = \sum_p W_p^{a_1} (W^T)^p_{a_2}$.

Co-cited relation between papers becomes $C_{p_2}^{p_1} = \sum_p W_p^{p_1} W_p^{p_2} = \sum_p W_p^{p_1} (W^T)^p_{p_2}$,

while co-citing relation becomes $B_{p_2}^{p_1} = \sum_p W_p^{p_1} W_p^{p_2} = \sum_p (W^T)^{p_1}_p W_p^{p_2}$. Co-occurrence

(or co-studying) of concepts becomes $Co_{c_2}^{c_1} = \sum_p W_p^{c_1} W_p^{c_2} = \sum_p (W^T)^{c_1}_p W_p^{c_2}$. Author a 's expertise, which are the concepts that the author has been working on can be find from $E_c^a = \sum_p W_p^a W_c^p$.

Other properties of the entities can also be added to the fundamental network. For example, research groups are sets of authors; countries, and also journals are sets of papers; and disciplines are clusters of concepts. All those sets can be either

pre-defined or generated from the fundamental network via clustering algorithms. For example, a research group is in principle a set of authors who often publish the same papers and similarly disciplines or fields are sets of concepts, often studied by the same papers. In network science, clustering algorithms are designed to find communities of nodes, where there are more intra-community links than inter-community links. Therefore, it is quite possible that those communities can be found via clustering algorithms on the network, instead of defining them ahead of time using certain heuristic rules or conventions.

Besides unifying various relations, the fundamental network can also provide new insights concerning some scientometric questions. For example, we might be able to measure the creativity of a paper by considering the concepts on which the paper works. If a paper p proposes a new concept c or a new connection such as a theorem between existing concepts c_i and c_j , denoted as a new concept c_{ij} . Then, we should see in the fundamental network that $W_c^p = 1$, $W_{c_{ij}}^p = 1$, $W_c^q = 0$ and $W_{c_{ij}}^q = 0$ for all $q \neq p$. Therefore, we may define a creativity metric based on those quantities, for instance by comparing W_c^p and $\sum_{q \neq p} W_c^q$.

Consider next the task of recommending the most relevant papers to researchers as another example. For such a system, we first need to rank all papers according to their quality in a certain way, and secondly, we also need to measure the relevance of all fields (concepts) for each given author. We can then choose to recommend the top papers in each most relevant field to the researcher. For the first task, i.e., the quality measure we can, roughly speaking, extend the PageRank-like algorithm to the three-layer network, implying that papers working on more influential concepts are themselves more influential, and the concepts worked on by more influential papers are themselves more influential. The second task, i.e., relevance measure, might be done simply by covering all the fields that author a has been working on, i.e., the c s such that $E_c^a \neq 0$, or ranking the c s according to the value of E_c^a . A more complex but potentially valuable algorithm would be to propagate E_c^a on the concept layer, meaning that not only the fields c s that have been directly worked on by author a but also the fields that are logically close to these should be taken into consideration. For the task of relevance measure, we can even try to solve it by turning it into a task of more general similarity measure between papers. Once we have such a similarity measure between papers, we can choose to recommend those papers which are most similar to author a 's papers and also of high quality. Working on the same (or closely related) concepts, citing the same (or closely related) papers and being cited by the same (or closely related) papers, are all attributes that make papers more similar. Therefore, for such a similarity measure, it is natural to consider a propagation algorithm over the whole three-layer

fundamental network. From this example, although we have not implemented and explained all the details, we can already see that the task of recommending papers becomes an algorithmic problem in the fundamental network and we can also see how the two key ideas of networks – “relation” and “propagation” – can potentially help in designing a better recommendation system of papers to researchers. This example shows how the fundamental network can be used to represent scientometric data, rephrase scientometric questions, and to establish concepts and algorithms to answer the questions with the data.

The fundamental network can be expanded further if we need to look into other scientific activities by including new entities and relations into the network. For example, a similar three-layer network can be established for patents: inventors, patents, and technology concepts. We can even link the two three-layer networks by connecting the papers and patents via their citations. Furthermore, if we have data on how each patent and paper are used in producing products, we can even add another layer, the product layer, to the fundamental network. Papers and patents are connected to products via the “used-in” relation while the products themselves are connected by the “made-from” relation. In this way, we might be able to measure the contribution of papers and/or patents in final products. Moreover, if we have a citation network of textbooks and corresponding citations between textbooks and papers (patents), we can add another layer of textbooks so that we can even measure the contribution of papers and patents to the accumulated knowledge of humanity.

We hope that the above examples illustrated the value of the fundamental network of scientometrics. However, in reality, it is quite challenging to establish such a network. Besides the citation network, we will need to disambiguate authors and find their academic relations, build a concept network for human knowledge, and annotate each paper with their concepts. With the accumulated data, like the open data from *academic tree*^① and *dblp*^②, and with the fast development of natural language processing, we are getting closer to the dream of really establishing such a network. One possible proxy of the fundamental network can be the network of researchers, papers, and reactants/reactions in chemistry. In the concept layer of this proxy, chemical reactants are connected via chemical reactions, and it is also much easier to link papers to specific chemical reactants and reactions than linking papers to more general concepts. In fact, there are already commercial products such as *Reaxys*^③ and *SciFinder*^④ that realize at least part of these ideas. Creating an open data version is, in principle, plausible. It will be interesting to create a three-layer fundamental network for the discipline of scientometrics itself.

^① <http://www.academic tree.org>

^② <https://dblp.uni-trier.de/>

^③ <https://www.reaxys.com/>

^④ <https://scifinder.cas.org>

4 The closed-system and open-system approach

Often in our investigations, we try to collect and represent all necessary data, and then we search for concepts and algorithms to answer the research questions. We want to make sure that all entities relevant to the research questions are covered by the framework and data. This is called a closed-system approach. Using the network language, this means that our network is complete such that all entities and their relations relevant to the research questions become nodes and edges in the network. We do not need any other information beyond the network, i.e., the system, to answer the research questions.

On the other hand, there might be situations where some of the relevant entities or some relations are not covered by the framework and data, but we still need to answer those research questions. In that case we call this an open-system approach. Sometimes, we have to take the open-system approach due to the limitation of available data. Alternatively, we might do so on purpose. When we work with open systems, often, we need to take some exogenous information/quantities of the missing entities into account via, for example, propagation algorithms. Rephrased in another way, for a presumed closed system, if for some reason we think that the closed system is missing some very relevant entities and we cannot easily get full data about those entities, it is then a good idea to treat the presumed closed system as an open system and make use of some available data about the missing entities. We will illustrate this via examples.

First, let us show that while by definition determining if a framework is a closed-system or that an open-systems approach is called for, is in principle a clear-cut question, in practice, almost all closed systems are just presumed closed systems, which are more likely open systems. For instance, consider the task of measuring the influence of papers. For simplicity, we decide to do so using citation counts of papers. For that, we need the whole citation network of all publications. We then first need to specify which types of publications to consider. For example, we might consider only research papers but not patents, books, reports, policy papers, and other documents. In doing so, we assume that there is a “boundary” between papers and other documents. In principle, we can call it a closed-system approach only when there are indeed no citations at all between papers and other documents. However, often, we still regard it as a closed-system approach if the citations among the papers are much denser than the citations between papers and other documents. Even considering only the citation network of research papers, paper coverage is always an issue. Again, often, we regard a citation network based on certain data set to be a closed-system approach if most of the citations are already covered by the data set. We then count the received citations of each paper in the presumed “closed” citation network. Therefore, on the one hand, we have to isolate the central

system of interest from the rest of the world to focus on the central “closed” system, while on the other, we should always be aware that the so-called “closed” systems are very likely open systems.

Second, we want to show that an open-system approach, which makes use of partial data when the full data of the missing entities are not available, often helps us to go beyond the presumed closed system. Let us continue with the above example of the citation network. Now we want to measure the reliability or trustworthiness of papers. Taking the closed-system approach implies that the citation network has captured all the relevant entities and relations regarding the reliability of papers. However, this assumption very likely does not hold. There might be some correlation between citation counts and reliability, but these notions are not the same, and the citation network only represents citations among papers but not reliability. Therefore, the citation network is not enough for the task of finding or determining a reliability measure. One way out of this would be to define a new network that captures the essence of reliability, about which we do not really know much. An open-system approach will be making use of the citation network together with some exogenous information. Let us assume that we have a small set of papers whose reliability scores have been evaluated by human experts. Then if we assume that papers citing more reliable papers are themselves more reliable or papers cited by more reliable papers are themselves more reliable, we can propagate the exogenous reliability scores of a selected set of papers to all papers via the citation network. This is exactly the idea of the TrustRank algorithm (Gyöngyi et al., 2004) in computer science, which has been used to rank most and least trustworthy web pages to either recommend the highly trustworthy papers or filter out the least trustworthy ones in search engines.

Besides measuring reliability, we might be able to extend the idea of propagation and open systems to measure the contribution of research papers towards human knowledge. For example, we can still use the citation network of papers as the open system, and we take the number of citations from textbooks to papers as the exogenous contribution scores. We then propagate this exogenous contribution scores to all papers via the citation network. Of course, if we also have the full data on the citations between textbooks, we turn back again to the closed-system approach, and it will even be better. However, the open-system approach only requires citation counts from textbooks to papers and the citation network of papers, but not the full citation network of textbooks and papers.

What is the relation between networks on the one hand, and open and closed systems on the other? With networks, it is explicit that the algorithms that we are using are closed-system ones or open-system ones depending on whether or not all the relevant entities and relations are covered by the network. Networks also inspire

us to make use of partial data when some relevant entities and some or all of their data are missing, but partial data is available. Furthermore, networks also provide a natural platform to develop algorithms to take partial data into account via, for example, propagation.

5 Working towards a unified framework of scientometrics

We hope that the above discussion has already shown that the three-layer network of scientometrics can be a fundamental network and lead to a unified framework of scientometrics. Firstly, the fundamental network is capable of representing most, if not all, relevant entities and their relations regarding various scientific activities of scientometrics in a broad sense, which is to discover regularities for all kinds of scientific activities to help the development of science. Secondly, with the fundamental network already in place, scientometric questions can be rephrased as questions waiting to be solved via concepts and algorithms over the network. Finally, with the two key ideas, i.e., relation and propagation, the fundamental network also provides a good platform to develop concepts and algorithms to solve these scientometric questions. Furthermore, with the concept layer and the links between the papers and the concepts of this fundamental network, we can now dive into a contents-based scientometric analysis rather than only making use of metadata of publications.

However, as we mentioned earlier, building up such a unified framework for scientometrics is a very demanding task. Mapping a concept network of human knowledge is challenging. Connecting each paper to the concept network is also not easy. Even author name disambiguation and getting correct academic relations are not simple tasks. To illustrate the power and beauty of this unified framework of scientometrics, we may start from a small scale example or a proxy of it. That is the ultimate purpose of this article: Let us implement a small scale example and work together towards a unified framework of a general scientometric theory (some may refer to it as informetrics, or as a science of science) that goes beyond the evaluation of various research units.

References

- Alvarez-Hamelin, J.I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2008). K-core decomposition of Internet graphs: Hierarchies, selfsimilarity and measurement biases. *Networks and Heterogeneous Media*, 3(2), 371–393. doi:10.3934/nhm.2008.3.371. Annual European Conference on Complex Systems, Dresden, GERMANY, OCT 01-06, 2007.
- Amjad, T., Ding, Y., Daud, A., Xu, J., & Malic, V. (2015). Topic-based heterogeneous rank. *Scientometrics*, 104(1), 313–334. doi:10.1007/s11192-015-1601-y.

Perspective

- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30, 107–117. In proceedings of the Seventh International World Wide Web Conference. doi:10.1016/S0169-7552(98)00110-X.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359–377. doi:10.1002/asi.20317.
- Gyöngyi, Z., Garcia-Molina, H., & Pedersen, J. (2004). Combating web spam with trustrank. In proceedings of the Thirtieth international conference on Very large data bases, 30, 576–587.
- Leontief, W. (1941). *The Structure of American Economy, 1919-1929*. Cambridge: Harvard University Press.
- Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, 246(1), 1–19. doi:10.1016/j.ejor.2015.04.002.
- Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science*, 28(6), 441–453. doi:10.1177/016555102762202123.
- Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing & Management*, 12(5), 297–312. doi:10.1016/0306-4573(76)90048-0.
- Rousseau, R., Egghe, L., & Guns, R. (2018). *Becoming Metric-Wise*. Chandos Information Professional Series. Chandos Publishing. doi:10.1016/B978-0-08-102474-4.00010-8.
- Shen, Z., Chen, F., Yang, L., & Wu, J. (2019). Node2vec representation for clustering journals and as a possible measure of diversity. *Journal of Data and Information Science*, 4(2), 79–92. doi:10.2478/jdis-2019-0010.
- Shen, Z., Yang, L., Pei, J., Li, M., Wu, C., Bao, J., Wei, T., Di, Z., Rousseau, R., & Wu, J. (2016). Interrelations among scientific fields and their relative influences revealed by an input-output analysis. *Journal of Informetrics*, 10(1), 82–97. doi:10.1016/j.joi.2015.11.002.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391. doi:10.1016/j.joi.2016.02.007.
- Waltman, L., & van Eck, N.J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12), 2378–2392. doi:10.1002/asi.22748.
- West, J.D., Bergstrom, T.C., & Bergstrom, C.T. (2010). The Eigenfactor Metrics (TM): A network approach to assessing scholarly journals. *College & Research Libraries*, 71(3), 236–244. doi:10.5860/0710236.
- Zeng, A., Shen, Z., Zhou, J., Wu, J., Fan, Y., Wang, Y., & Stanley, H.E. (2017). The science of science: From the perspective of complex systems. *Physics Reports – Review Section of Physics Letters*, 714–715, 1–73. doi:10.1016/j.physrep.2017.10.001.

This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).