Journal of Marketing Analytics (2019) 7:84–93
https://doi.org/10.1057/s41270-019-00052-9

ORIGINAL ARTICLE

# Review of a proposed methodology for bibliometric and visualization analyses for organizations: application to the collaboration economy

**Myriam Ertz**[1] · **Sébastien Leblanc-Proulx**[1]

Revised: 17 April 2019 / Published online: 9 May 2019


**Abstract**
This paper presents the bibliometric and visualization method applied to a dataset of 729 documents published in the collaborative economy research field. Four steps are described in details: (1) the delimitation of the field of study; (2) the selection of databases, keywords, and search criteria; (3) the extraction, cleaning, and formatting; and finally (4) the co-citation analysis and visualization. The method validation section shows the results obtained by applying our methodological procedure to an author network analysis as well as a source title network analysis. This study is unique which presents a co-citation analysis coupled with a network visualization applied to the rapidly growing research area of the collaborative economy as a whole and not only of the collaborative tourism and hospitality research, as has been previously. The originality of this method lies firstly in the fact that the data were extracted from two databases (Scopus and Web of Science) instead of one as is commonly done in analytic studies. Secondly, VOSviewer was our main analytical tool performing the co-citation analysis and the network visualizations.

**Keywords** Bibliometrics · Content analysis · Network analysis · Visualization method · Cluster analysis · Co-citation analysis

## Introduction

Bibliometrics is a quantitatively method that investigates the formal properties of knowledge domains by extracting data from published documents using statistical analysis (Mora et al. 2017; Agarwal et al. 2016). In fact, bibliometrics consists in "mathematical and statistical analyses of patterns that arise in the publication and use of documents" (Diodato 1994, pp. viii–ix). It is considered a discrete, objective, low cost, and reliable approach for analyzing different aspects from publications, journals, scientists, and communities (Zhao and Strotmann 2015). More specifically, the bibliometric procedure enables to describe or evaluate the intellectual structure of a field of study, the diffusion of knowledge, the relationship between academics, and their use of the scientific literature. In addition, bibliometrics allows comparisons over time to assess the history of a specific research domain. Bibliometric analytics are also easily replicable since the data come from search engines widely used by the scientific community. Nowadays, bibliometrics is used by several scholars to quantitatively evaluate research fields. Sometimes, it completely replaces the traditional qualitative peer assessment of performance indicators, such as the h-index or the impact factor of a journal, which creates debates about how we evaluate scientific production (Agarwal et al. 2016; Zhao and Strotmann 2015; De Bellis 2009; Haustein and Larivière 2015).

In a context of exponential growth of research reports and publications, bibliometrics is thus a precious tool for information management and dissemination within an organizational context. The method provides an effective way for mapping the main contributors, key insights, areas of research, important topics and themes, as well as influential works on a given subject (Fahimnia et al. 2015). In addition, bibliometric analytics can be integrated with other methods, such as visualization mapping, for offering new and unique insights. Visualization mapping is part of the science of networks, a multidisciplinary field of research, and is increasingly used with bibliometric analysis (Perianes-Rodriguez

✉ Myriam Ertz
Myriam_Ertz@uqac.ca

Sébastien Leblanc-Proulx
Sebastien_Leblanc-Proulx@uqac.ca

[1] LaboNFC, Université du Québec à Chicoutimi, Boulevard de l'Université 555, Chicoutimi, QC G7H 2B1, Canada

et al. 2016). By using it with co-citation analysis, visualization allows researchers to examine the characteristics, structures, and evolution of a field of research. Several computer programs can be used for this, but our preference went to VOSviewer given its popularity among researchers. Visual depictions of research mappings are particularly useful in organizational contexts for quicker and more efficient grasping of the dynamics pertaining to a specific research topic.

The main difficulty with bibliometrics/visualization, however, is that from its inception, it has been mainly developed, applied, discussed, and improved for academic objectives such as mapping authors in intellectual space (McCain 1990), measuring and mapping the intellectual structure of a field (White and Griffith 1981; García-Lillo et al. 2018), visualizing a discipline (White and McCain 1998), or shedding light on future research directions (García-Lillo et al. 2018; White and McCain 1998), with little application—and applicability—to organizational settings. Besides, the publications used as input for bibliometrics/visualization are often extracted from a single database, typically Web of Science or Scopus, both being the largest scholarly databases (Mongeon and Paul-Hus 2016). Also, the bibliometric/visualization approach needs metadata from a set of publications that are related in a way or another, and especially with their citations (Zhao and Strotmann 2015). Thus, it is necessary to use a search engine that retrieves citations from a document. To date, only Web of Science (Thomson Reuters) and Scopus (Elsevier) offer this possibility (Mongeon and Paul-Hus 2016), and it has been shown that adding Scopus to Web of Science is necessary to accurately map the intellectual impact of studies and more correctly identify the extent to which they influence on the wider researcher community (Meho and Sugimoto 2007). While CiteSeerX also does to some extent, the metadata generated are not as complete as those from Web of Science and Scopus (Mongeon and Paul-Hus 2016). In sum, the ability to extract documents from both databases could be particularly valuable since it enables to have access to a broader range of publication, to render the analysis more comprehensive and refined. It also provide access to complete and reliable metadata to conduct analysis, which is valuable in organizational contexts where automated and systematic processes need to yield maximal output.

The overarching objective of this paper is therefore to propose a four-step, practical procedure for bibliometric analysis in an organizational context using both Scopus and Web of Science. Our methodological framework for this process is based on Zhao and Strotmann's (2015) book *Analysis and Visualization of citation Network*. Although the general ideas are the same, our process went slightly different to better adapt to an organizational context. More precisely, the process involves a limited number of key steps, which are akin with conventional informational retrieval from information systems. Also, the process focuses on the use of two databases, whereas Zhao and Strotmann (2015) use only one, and the recourse to visualization techniques to map the bibliometrics results more straightforwardly in an organizational context, was mentioned, but not implemented by both authors either.

## Procedure

Three types of citation-based measures are commonly used to evaluate the strength of a relationship between two items: inter-citation analysis, bibliographic coupling frequencies (BCF) analysis, and co-citation analysis (Zhao and Strotmann 2015). Inter-citation is the simplest of them since it counts the frequencies of documents that cite each other's, while BCF counts the frequencies of documents that cite the same documents together. In other words, an inter-citation analysis will count how many times document A cite document B and vice versa, and a BCF analysis will count how many times document A and B cites C and/or D, E, etc. The more complex co-citation analysis examines when two objects appear in a reference list of other documents (e.g., document C, D, and E have each cited A and B).

We chose to perform a co-citation analysis instead of an inter-citation count or a bibliographic coupling Frequencies (BCF) since it enables best to pinpoint the connections between references in the literature. BCF excels in studying the recent research activities of a research field, whereas co-citation analysis examines the past intellectual influences on the field or the knowledge base of the field (Fahimnia et al. 2015). More precisely, BCF maps *citing* publications in the dataset while co-citation analysis maps *cited* publications (Zhao and Strotmann 2015). Using BCF would only identify influential authors and publications that remain limited to our dataset, while co-citation analysis is more encompassing. It includes authors or publications that have been influential in the field of CE although not included in the dataset (Zhao and Strotmann 2015). Thus, choosing between these types of measures for citation count has a dramatic impact on the results afterwards. For example, one of the most iconic books in the collaborative economy, *What's mine is Yours* by Rogers and Botsman (2010) is not in our dataset. Using BCF would ignore the cited references to this document while co-citation would embrace it. The lack of inclusion of a publication on the dataset may be related to the fact that the publication is not stored in the databases used, does not include the defined search terms, or got published outside the timeframe under study. Also, the type of publication has an impact. Books, conference proceedings, and gray literature (reports, working paper, thesis, etc.) are not systematically indexed in databases. With co-citation analysis, all types of publications are considered. The only

prerogative is the need to be cited by some authors in the dataset. Given the superiority of co-citation analysis in mapping extra-sample influential publications and its past-oriented nature, co-citation analysis was favored over BCF for our study. This method discovered by Small (1973) and Marshakova (1973) has been a standard since the seventies and is used for practical purposes among researchers (Boyack and Klavans 2010). The addition of visualization techniques with it generates an interesting tool for studying the structure of a field of study (Small 1999). The example of application of the methodology, namely on the field of CE is derived from Ertz and Leblanc-Proulx (2018).

## Delimitation of the field of study

First, we delimited the research field under study. This field can be broad (e.g., health sciences), focused on a specific science (e.g., medicine), a discipline (e.g., cardiology), a sub-discipline (e.g., pediatric cardiology), and so on (e.g., pediatric cardiology for preterm birth). Our search field was the collaborative economy (CE), a topic of particular interest for managers in many different industries (Sundararajan, 2016; Ertz et al. 2016). The details of the applied study are available in Ertz and Leblanc-Proulx (2018). Within the CE research field, one study, namely Cheng (2016) consisted in a bibliometric analysis but which was circumscribed to the collaborative tourism and hospitality research stream (e.g., Airbnb) within the CE. The current research is unique in that it takes a broader and more encompassing perspective on CE, by including namely car-sharing, goods mutualization, crowdfunding, and other related crowd-based studies that are characteristic of the collaborative economy (Sundararajan 2016; Ertz et al. 2016).

Then, we circumscribed the search field in a timeframe. In our case, the timeframe was between 2010 and November 2017. 2010 was chosen in regard to the media coverage, consultant reports, and conferences, as well as venture capitalists and hedge funds investments into collaborative platforms, which increased tremendously at this time. Then, November 2017 was chosen simply because it was the period when we were doing our study and we wanted to have access to the most recent publications in the domain in order to be up to date.

Finally, we chose which language(s) should be retrieved. We decided to keep documentation in English only, since the majority of the influent publications in the collaborative economy domain are in this language. Optionally, the types of documents may also be chosen at this step (e.g., journal articles, books, etc.). However, we feel it is easier to delimit this matter once the first searches are made. This way, we can see what is possible to obtain and therefore adjust our search criteria in consequence.

## Selection of databases, keywords, and search criteria

After that, we chose which database will provide the information we want. As stated earlier, we decided to include both Web of Science and Scopus as recent studies such as Mongeon and Paul-Hus (2016) demonstrated that the coverage of both databases are not the same. Thus, the results of bibliometric analyses may vary depending on the database used. Following Zhao and Strotmann (2015), a "good approach might be to supplement results retrieved from a citation database with additional publications (which are then indexed by researchers in the same format as the downloaded records) in order to reach the desired level of completeness for the study at hand." (p. 66).

Once the databases are chosen, we defined the search keywords. In our case, we limited our consideration to "sharing economy," "collaborative economy," and "collaborative consumption." Although many synonyms are used in this subject, such as gig economy, platform economy, peer economy, app economy or access-based consumption, our tests with different search queries indicated that they were not relevant for finding new documentation, since for most of them were included in publications which already comprised the initial keywords.

Another thing to consider here is the search criteria. We use the title, abstract, keywords search for Scopus, and the topic search for Web of Science. In both cases, it was the default way of searching. Instead of having the same process for retrieving information, searching with both controlled (subject), and natural (title, abstract, keywords, etc.) vocabulary offers more variety in the results (Fidel 1991; Savoy 2005).

When the first searches finished, we concluded that many journal articles, conference papers, book, book chapters, editorials, and gray literature[1] were relevant to our field of study. Thus, all of these were considered in the creation of our core set.

## Extraction, cleaning and formatting

When we were satisfied with the results of our keywords and search criteria, we extracted the data. It is normal to use a representative sample of the literature rather than attempting to catch all publications in a field of research. By the very nature of the co-citation analysis, "outsiders" will be taken in consideration if our core set cites them. For example,

[1] Gray literature consists of institutional reports (e.g., European Commission), business magazines, and newspapers (e.g., The Economist), as well as consultant reports (e.g., PwC).

**Table 1** Number of documents per year

| Year | Number of documents |
|---|---|
| 2010 | 1 |
| 2011 | 1 |
| 2012 | 8 |
| 2013 | 16 |
| 2014 | 37 |
| 2015 | 107 |
| 2016 | 257 |
| 2017 | 302 |

*Source* Ertz and Leblanc-Proulx (2018)

**Table 2** Authors with more than 30 citations in the dataset

| Author name | Weight (Citations) |
|---|---|
| Belk, RW | 353 |
| Botsman, R | 208 |
| Schor, JB | 104 |
| Bardhi, F | 100 |
| Hamari, J | 90 |
| Benkler, Y | 82 |
| Zervas, G | 63 |
| John, NA | 61 |
| Lamberton, CP | 59 |
| Shaheen, SA | 53 |
| Tussyadiah, IP | 52 |
| Guttentag, D | 46 |
| Edelman, BG | 44 |
| Mont, OK | 43 |
| Airbnb | 42 |
| Martin, CJ | 42 |
| Pricewaterhousecoopers | 41 |
| Albinsson, PA | 40 |
| Felson, M | 40 |
| Gansky, L | 40 |
| Ozanne, LK | 40 |
| Seyfang, G | 40 |
| Heinrichs, H | 40 |
| Cohen, B | 39 |
| Owyang, J | 37 |
| Weber, Ta | 37 |
| Sundararajan, A | 36 |
| European Commission | 34 |

*Source* Adapted from Ertz and Leblanc-Proulx (2018)

some of the most cited documents in our set were outsiders (e.g., Rachel Botsman, Yochai Benkler, Lisa Gansky).

By using Web of Science and Scopus together, a higher amount of work and precaution is necessary in relation to the extraction, cleaning, and formatting of the data. First, our goal was to obtain a similar format by Web of Science and Scopus in order to be able to analyze them with VOSviewer. At the time of our research, plain text was available for extraction for both of the databases and was readable by VOSviewer. However, although the format (i.e., plain text) was the same, the language between them was different and both VOSviewer and BibExcel were only able to read plain text from Web of Science but not plain text from Scopus. It was thus necessary to translate the plain text sample from Scopus in the Web of Science format. The field tags were easily standardized with regular expression in Notepad++. However, the complexity came with the formatting of authors, affiliations, countries, and citations (about 26,000). VOSviewer needs the same information in order to create a relationship between two information. For example, if the metadata of a document says it was published in England and another says United Kingdom, they will not be linked together when they should have. It was the same for the names of the authors, affiliations, and the titles of documents. All initials, dots, commas, and spaces needed to be in the same order to create a relationship. This task took several hours of work with the help of BibExcel and Notepad++.

All articles were reviewed for relevance when doing the formatting. A dozen of them were retracted for this motive. At the same time, many publications appeared in both databases. They were suppressed from the data extracted from Web of Science since their format offers less information than Scopus (Yong-Hak 2013). After the suppression of non-relevant documentation and duplicates, our dataset went from 1056 to 729 entries. These 729 observations constituted our final sample.

These 729 bibliographic entries, extracted from Scopus and Web of Science, were published between January 2010 and October 2017. Table 1 shows the number of documents related to the year of publication. These data were subsequently used for a co-citation analysis related to the subject of the CE.

## Co-citation analysis and visualization

The 729 entries extracted contained approximatively 26 000 citations, which were subsequently analyzed through a co-citation analysis. Tables 2 and 3 further show the data obtained after the analysis and the *Visualization of Similarities* (VOS) algorithm for clustering used to this end.

The co-citation approach works in groups of two. If two objects are cited inside the same document, they each earn a mention and become linked. For example, if A quotes B and C, B and C become coupled. This is the most common approach when analyzing a citation network. However, it

**Table 3** Source titles with more than 20 citations in the dataset and their associated cluster

| Source title | Cluster | Weight (citations) |
|---|---|---|
| Belk R, sharing versus pseudo-sharing in web 2.0 (2014a) the anthropologist, 18 (1) | 1 | 29 |
| Benkler Y, sharing nicely: on shareable goods and the emergence of sharing as a modality of economic production (2004) the yale law journal, 114 (2) | 1 | 47 |
| Benkler Y, the wealth of networks: how social production transforms markets and freedom (2006), Yale University Press New Haven | 1 | 41 |
| Botsman R, Rogers r, what's mine is yours: the rise of collaborative consumption (2010b), Harper Collins New York | 1 | 175 |
| Gansky L, the mesh: why the future of business is sharing (2010), Portfolio Penguin New York | 1 | 65 |
| John N, sharing and web 2.0: the emergence of a keyword (2012) new media and society, 15 (2) | 1 | 24 |
| John N, the social logics of sharing (2013) commun. Rev., 16 (3), pp. 113–131 | 1 | 32 |
| Lessig L, (2008) remix: making art and commerce thrive in the hybrid economy, New York: Penguin Books | 1 | 22 |
| Ostrom E, governing the commons (1990), Cambridge University Press: Cambridge, UK | 1 | 23 |
| Putnam R, (2000) bowling alone: the collapse and revival of american community, New York: Simon and Schuster | 1 | 23 |
| Rifkin J, the zero marginal cost society: the internet of things, the collaborative commons, and the eclipse of capitalism (2014), Palgrave Macmillan New York, NY | 1 | 32 |
| Albinsson P, Perera B, alternative marketplaces in the 21st century: building community through sharing events (2012) journal of consumer behaviour, 11 (4) | 2 | 35 |
| Bardhi F, Eckhardt G, access-based consumption: the case of car sharing (2012) journal of consumer research, 39 (4) | 2 | 81 |
| Belk R, sharing (2010) journal of consumer research, 36 (5) | 2 | 80 |
| Belk R, why not share rather than own (2007) the annals of the american academy of political and social science, 611 (1) | 2 | 66 |
| Felson M, Spaeth J, community structure and collaborative consumption: a routine activity approach (1978) american behavioral scientist, 21 (4) | 2 | 39 |
| Lamberton C, Rose R, when is ours better than mine? A framework for understanding and altering participation in commercial sharing systems (2012) journal of marketing, 76 (4) | 2 | 47 |
| Leismann K, Schmitt M, Rohn H, Baedeker C, collaborative consumption: towards a resource-saving consumption culture (2013) resources, 2 (3) | 2 | 24 |
| Ozanne L, Ballantine P, sharing as a form of anti-consumption? An examination of toy library users (2010) journal of consumer behaviour, 9 (6) | 2 | 29 |
| Belk R, you are what you can access: sharing and collaborative consumption online (2014b) journal of business research, 67 | 3 | 132 |
| Botsman R, Rogers R, beyond zipcar: collaborative consumption (2010a) harvard business review, 80 (10) | 3 | 21 |
| Cohen B, Kietzmann J, ride on! Mobility business models for the sharing economy (2014) organization and environment, 27 (3) | 3 | 31 |
| Ert E, Fleischer A, Magen N, trust and reputation in the sharing economy: the role of personal photos in airbnb (2016) tourism management, 55 | 3 | 31 |
| Guttentag D, airbnb: disruptive innovation and the rise of an informal tourism accommodation sector (2015) current issues in tourism, 18 (12) | 3 | 45 |
| Hamari J, Sjöklint M, Ukkonen A, the sharing economy: why people participate in collaborative consumption (2016) journal of the association for information science and technology, 67 (9) | 3 | 106 |
| Möhlmann M, collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again (2015) journal of consumer behaviour, 14 (3) | 3 | 45 |
| Heinrichs H, sharing economy: a potential new pathway to sustainability (2013) gaia ecol. Perspecties sci. Soc., 22 (4) | 4 | 30 |
| Martin C, the sharing economy: a pathway to sustainability or a nightmarish form of neoliberal capitalism? (2016) ecological economics, 121 | 4 | 29 |
| Sundararajan A, the sharing economy: the end of employment and the rise of crowd-based capitalism (2016), MIT Press Cambridge | 4 | 25 |

*Source* Ertz and Leblanc-Proulx (2018)

is quite complex to set up and requires the use of computer programs to automate the process. For example, if an article has more than one author, each author must receive its mention when pairing with another document. On a scale of a few hundred documents, authors' couplings become a real puzzle. This is why we imported the standardized data into VOSviewer. This program allows us to do both the co-citation analysis and the visualization at the same time.

Furthermore, network analysis through co-citation analysis requires using either distance-based or graph-based data mapping techniques (Perianes-Rodriguez et al. 2016). In distance-based maps, smaller distance between two items reflects the strength of the relation between the items (Van Eck et al. 2010). In graph-based maps, the difference between two items need not reflect the strength of the association between the two items (Pampel 2004). Items distribute in a uniform way and the most connected nodes move to the center of the network while the more isolated nodes move to the borders (Fahimnia et al. 2015). Both approaches have pros and cons. However, Van Eck et al. (2010) suggest that with graph-based mapping, it is more difficult to see the strength of the relation between two items, and clusters of related items may be difficult to detect. Given that we seek to find clusters of publications to identify research themes within the CE, we favor distance-based maps.

Past research showed that the VOS mapping technique shows superior performance than other distance-based algorithms (e.g., multidimensional scaling, VsOrd, Kopcsa-Schiebel) (Van Eck et al. 2010). The VOS mapping technique is fully integrated in the VOSviewer software. Thus, no additional computer program (e.g., Pajek) is needed for constructing VOS maps (Van Eck et al. 2010).

VOSviewer is based on Van Eck et al. (2005) visualization of similarities (VOS). It is a clustering technique that provides a low-dimensional visualization in which objects are located in such a way that the distance between any pair of objects reflects their similarity as accurately as possible (Van Eck and Waltman 2007). The weighted sum of the squared Euclidean distances with all pairs of objects is minimized with VOS and the similarity between two objects will affect positively the weight for their squared distance. In the lines of Van Eck et al. (2005, p. 2), if there are $n$ objects (i.e., citations), denoted by 1,…, $n$ and an $n \times n$ similarity matrix S, then element $s_{ij}$ of S denotes the similarity between the objects $i$ and $j$. Then if there is an $n \times m$ matrix X, where m denotes the number of dimensions of the space that is used, contains the coordinates of the objects 1,…$n$, the vector $x_i\left(x_{i1}, \ldots, x_{im}\right) \in \mathbb{R}^m$ denotes the $i$th row of X and contains the coordinates of object $i$. The objective function to be minimized in VOS is expressed as follows:

$$E(X;S) = \sum_{i<j} s_{ij} \left|\left|x_i - x_j\right|\right|^2, \tag{1}$$

where $\|\cdot\|$ denotes the Euclidean norm. The objective function is minimized in accordance to the constraint shown in Eq. (2), in which the distances $\left\|x_i - x_j\right\|$ in the constraint are not squared.

$$\sum_{i<j} \left|\left|x_i - x_j\right|\right| = 1. \tag{2}$$

Another important point when doing a visualization network is the threshold for citations. In this case, the researcher has control of the thresholds above which documents will be retrieved (McCain 1990; Shaw 1985). When the threshold is too small, a "giant component" may appear where most of the items are related to each other (Small 2009). This result may be interesting since it demonstrates the unicity of a group. However, it is impossible to observe distinctiveness or particular associations. This was also an issue that we faced when we applied thresholds that were too small. However, applying a very high threshold may break meaningful relationships and alter the results. According to Shaw (Shaw 1985), a threshold between 3 and 35–40 may be statistically significant depending on the situation. However, it is ambiguous to use a specific threshold based on similar studies because of the singularity of our own dataset (e.g., number of documents, associations between them, subject studied, etc.). After some experiments with different thresholds, we concluded that a cut-off value set at 30 citations for the author visualization, and at 20 citations for the literature classification, were the best for visualizing meaningful clusters.

Table 2 shows the authors with more than 30 citations in the dataset. All of these were given a distance ($x$, $y$) by VOSviewer with the algorithm described earlier.

The results of the authors visualization analysis are shown in Fig. 1. We observe meaningful information for the researchers interested in the field of CE. For example, Fig. 1 shows at a glance, which is connected to whom and can demonstrate allegiances, leaders, or excluded authors. Authors in the same colors are thus more strongly related to one another.

Table 3 displays the source titles with more than 20 citations in the dataset as well as their associated cluster. These were also given a distance ($x$, $y$) by VOSviewer. The results of the source title visualization analysis are visible in Fig. 2. This figure shows which documents are the most popular in the field of CE and those who share similar approaches or methodologies.

Cluster 1 "Managerial and conceptual elaboration": red; cluster 2 "Goods redistribution and mutualisation practices": green; cluster 3 "technology-mediated mutualisation systems": blue; cluster 4 "Sustainability in the collbaorative economy": yellow.

## Discussion

The objective of this research was to propose a generic framework for bibliometric/visualization analysis applicable in an organizational research context. The example of application of the methodology, namely on the field of the CE derived from Ertz and Leblanc-Proulx (2018), revealed

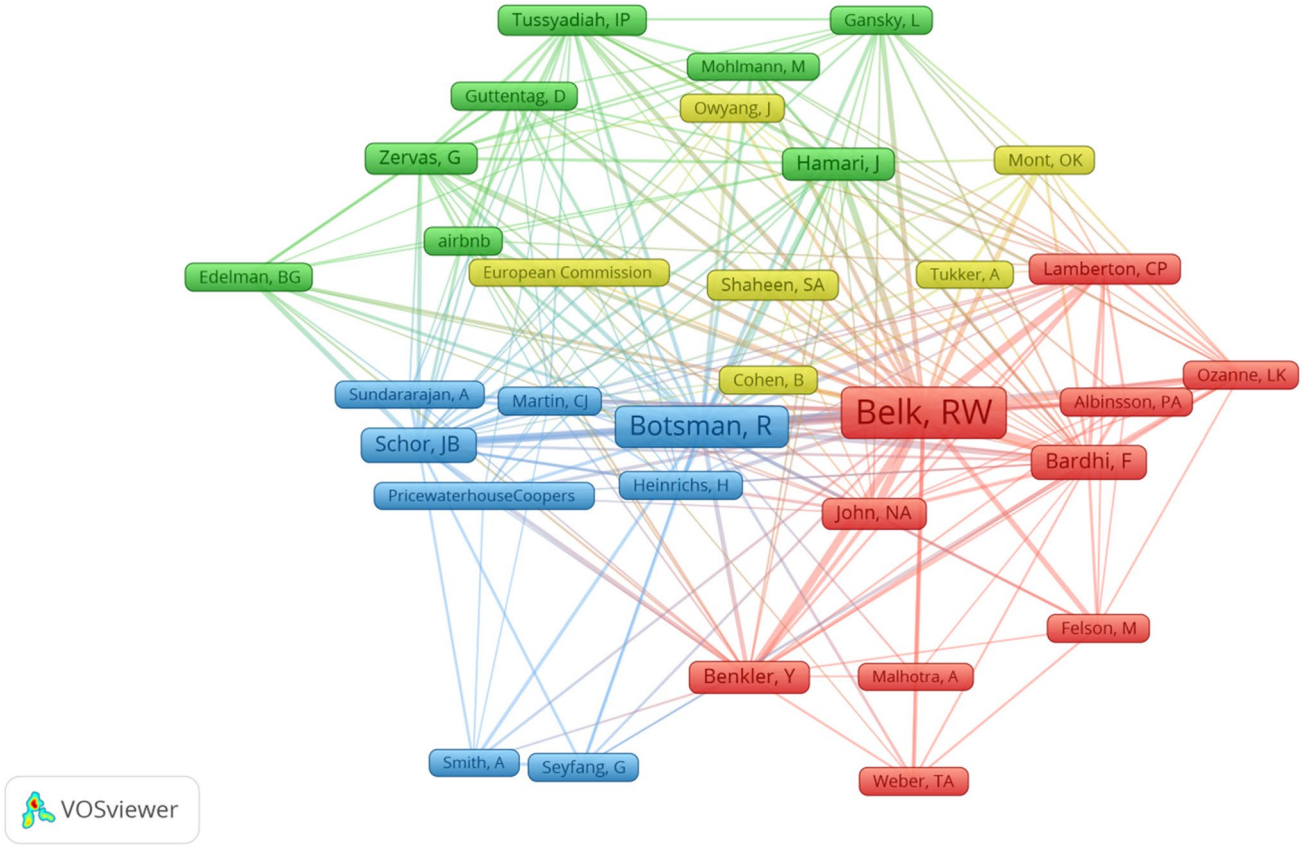


**Fig. 1** Authors visualization. *Source* Ertz and Leblanc-Proulx (2018)

valuable insights and perspectives on the CE. Four clusters are distinguishable, which refer to four major research streams within the CE field that follows also a certain chronological order. Cluster 1 entitled "managerial and conceptual elaboration" encompasses mainly conceptual elaboration on the CE concept as well as critics, debates and avenues for future research. These works were produced by both scholars and practitioners altogether, hence the "managerial" and "conceptual" denominations used in its title. These works are also older and more exploratory in nature, acting as foundational works to the research field. Cluster 2 entitled "Goods redistribution and mutualisation practices" is composed of hypothetico-deductive research, drawing on theoretical framework to test empirically conceptual frameworks. The research is mainly focused on the study of systems enabling the mutualization (e.g., rentals, leasing, sharing) and redistribution (e.g., donations, swapping, secondhand marketplaces) of tangible goods including cars or bicycles. The third cluster entitled "technology-mediated collaborative systems" consists of studies analyzing mutualisation or sharing systems involving all types of resources not only goods. These studies also take a predominantly technological perspective focusing on electronic platforms and discussing topics akin to information technology/information system preoccupations such as building trust or optimizing reputation systems, for example. Finally, the fourth cluster, entitled "sustainability in the collaborative economy," covers the sustainability theme in the CE. Authors draw on a great variety of foundational theories in order to adopt a critical stance of the CE with regard to its potential to reach sustainability.

The bibliometric classification of publications in four major clusters as well as its visualization both offer powerful means for rapidly assessing the state of the art of a given research field or topic, identify key contributors as well as key research themes and perspectives. The method enables therefore a quick and meaningful understanding of the state of the art of a specific research domain or topic in order to aid managerial research processes.

## Conclusion

The conventional bibliometric and visualization methods have mostly been developed and implemented in scholarly research context with typically lower implementation in organizations. However, the exponential growth in publications and research increases the difficulties related

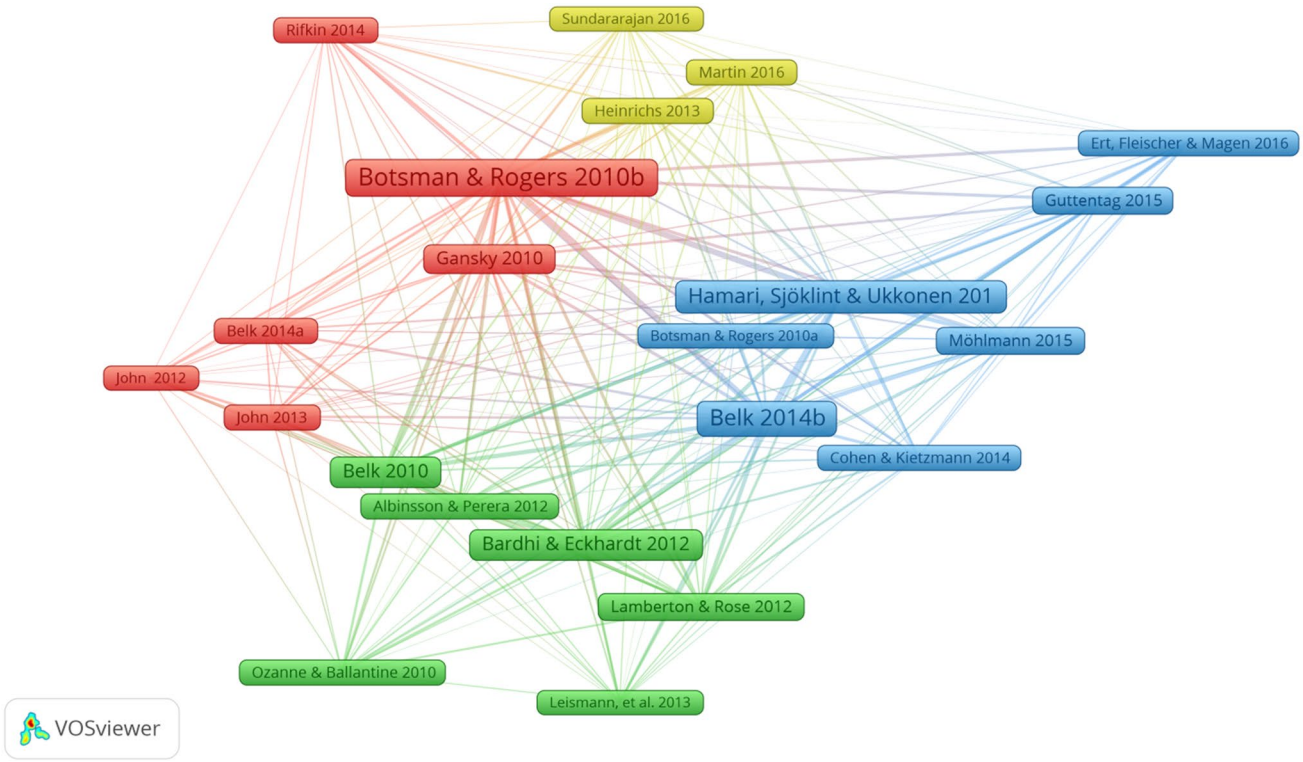


**Fig. 2** Source titles visualization. *Source* adapted from Ertz and Leblanc-Proulx (2018). (Color figure online)

to information search and processing, while constituting opportunities for exporting the bibliometric analyses to the managerial sphere. However, the existing bibliometric/visualization methods have been mainly designed for academic purposes and are not adapted to organizational contexts. Besides, they privilege the recourse to one database, typically Web of Science, while other databases such as Scopus have been shown to add meaningful insights to the results when integrated in the analysis (Meho and Sugimoto 2007).

This paper proposes a four-step approach to bibliometric/visualization analysis, and applies the proposed methodology to assess the field of the collaborative economy (Ertz and Leblanc-Proulx 2018) across both Scopus and Web of Science. The process relies mainly on specific software and codes to enable the kind of automation typically required in organizational settings for extracting meaningful intelligence in an efficient manner. The parsimonious nature of the four generic phases also allows for the kind of adaptation that are typically required in varying organizational contexts. In fact, the four stages act as meta-guidelines that remain highly specific but generic enough to allow for modulations to better suit organizational processes and specificities.

The first step aims at delimitating the field of the study, namely the search field, the timeframe as well as the language. The second phase deals with the selection of the databases, in our case Scopus and Web of Science, but other could be added in the future. The second phase also entails the choice of the keywords used to collect relevant publications, and the search criteria specify the conditions under which publications are acceptable to be retained in the final sample. The third stage involves the extraction, cleaning, and formatting of the citations across databases. Finally, the last stage consists in the co-citation analysis and application of the visualization method in order to extract meaningful information on the studied research theme or field.

The databases used in this framework are the two basic ones: Web of Science and Scopus. Yet, other databases are evolving rapidly and could potentially reach the same coverage and metadata quality as the two previously cited. Examples include CiteSeerX or Google Scholar. The framework allows for the addition of databases but does not explain how to extract information from those. Future research could determine to what extent extraction could extend to such databases as well. The paper further draws on the literature to assume that the threshold of the number of citations is a rule of thumb that can be fixed arbitrarily. In fact, we used 30 citations for the author visualization, and at 20 citations for the literature classification. However, this choice may be debatable and future research might investigate to what extent the use of varying thresholds could significantly produce changes in the results.

**Acknowledgements** We thank Vincent Morin, head of the department of economics and social sciences, at Université du Québec à Chicoutimi (UQÀC), for his valuable support and unaltered trust throughout the conduct of this study. We would also like to express our gratitude to Érick Chamberland, professor at the department of economics and social sciences, at Université du Québec à Chicoutimi (UQÀC) for introducing us to the field of bibliometric analysis and his initiation of the project.

## References

Agarwal, A., D. Durairajanayagam, S. Tatagari, S.C. Esteves, A. Harlev, R. Henkel, S. Roychoudhury, S. Homa, N.G. Puchalt, R. Ramasamy, A. Majzoub, K.D. Ly, E. Tvrda, M. Assidi, K. Kesari, R. Sharma, S. Banihani, E. Ko, M. Abu-Elmagd, J. Gosalvez, and A. Bashiri. 2016. Bibliometrics: Tracking research impact by selecting the appropriate metrics. *Asian Journal of Andrology*. https://doi.org/10.4103/1008-682x.171582.

Boyack, K.W., and R. Klavans. 2010. Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology* 61 (12): 2389–2404.

Cheng, M. 2016. Sharing economy: A review and agenda for future research. *International Journal of Hospitality Management* 57: 60–70.

De Bellis, N. 2009. *Bibliometrics and citation analysis: From the Science citation index to cybermetrics*. Toronto: Scarecrow Press.

Diodato, V. 1994. *Dictionary of bibliometrics*. New York: Haworth Press.

Ertz, M., F. Durif, and M. Arcand. 2016. Collaborative consumption: Conceptual snapshot at a buzzword. *Journal of Entrepreneurship Education* 19 (2): 1–23.

Ertz, M., and S. Leblanc-Proulx. 2018. Sustainability in the collaborative economy: A bibliometric analysis reveals emerging interest. *Journal of Cleaner Production* 196: 1073–1085.

Fahimnia, B., J. Sarkis, and H. Davarzani. 2015. Green supply chain management: A review and bibliometric analysis. *International Journal of Production Economics* 162: 101–114. https://doi.org/10.1016/j.ijpe.2015.01.003.

Fidel, R. 1991. Controlled vocabulary or free-text searching. *Journal of the American Society for Information Science* 42 (7): 501–514.

García-Lillo, F., E. Claver-Cortés, M. Úbeda-García, B. Marco-Lajara, and P.C. Zaragoza-Sáez. 2018. Mapping the "intellectual structure" of research on human resources in the "tourism and hospitality management scientific domain" Reviewing the field and shedding light on future directions. *International Journal of Contemporary Hospitality Management* 30 (3): 1741–1768.

Haustein, S., and A. Larivière. 2015. The use of bibliometrics for assessing research: Possibilities, limitations and adverse effects. In *Incentives and performance*, ed. I. Welpe, J. Wollersheim, S. Ringelhan, and M. Osterloh, 121–139. Cham: Springer.

Marshakova, I. V. (1973). A system of document connections based on references. *Scientific and Technical Information Serial of VINITI, 6*, 3–8. 5, 33.

McCain, K.W. 1990. Mapping authors in intellectual space: A technical overview. *Journal of the American Society for Information Science* 41 (6): 433–443.

Meho, L.I., and C.R. Sugimoto. 2007. Mapping the intellectual impact of library and information science research through citations: A tale of two databases–Scopus and Web of Science. *Proceedings of the American Society for Information Science and Technology* 44 (1): 1–7.

Mongeon, P., and A. Paul-Hus. 2016. The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics* 106: 213–228. https://doi.org/10.1007/s11192-015-1765-5.

Mora, L., R. Bolici, and M. Deakin. 2017. The first two decades of smart-city research: A bibliometric analysis. *Journal of Urban Technology* 24: 3–27. https://doi.org/10.1080/10630732.2017.1285123.

Pampel, F.C. 2004. *Exploratory data analysis, encyclopedia of social science research methods*. London: SAGE Publications Inc.

Perianes-Rodriguez, A., L. Waltman, and N.J. van Eck. 2016. Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics* 10: 1178–1195. https://doi.org/10.1016/j.joi.2016.10.006.

Rogers, R., and R. Botsman. 2010. *What's mine is yours: the rise of collaborative consumption*. New York: HarperBusiness.

Savoy, J. 2005. Bibliographic database access using free-text and controlled vocabulary: An evaluation. *Information Processing and Management* 41: 873–890. https://doi.org/10.1016/j.ipm.2004.01.004.

Shaw, W.M. 1985. Critical thresholds in co-citation graphs. *Journal of the American Society for Information Science* 36 (1): 38–43.

Small, H. 1973. Cocitation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science* 24: 265–269. https://doi.org/10.1002/asi.4630240406.

Small, H. 1999. Visualizing science by citation mapping. *Journal of the American Society for Information Science* 50 (9): 799–813.

Small, H. 2009. Critical thresholds for co-citation clusters and emergence of the giant component. *Journal of Informetrics* 3: 332–340. https://doi.org/10.1016/j.joi.2009.05.001.

Sundararajan, A. 2016. *The sharing economy: The end of employment and the rise of crowd-based capitalism*. Cambridge: MIT Press.

Van Eck, N.J., and L. Waltman. 2007. VOS: A new method for visualizing similarities between objects. *Advances in Data Analysis*, 299–306. Cham: Springer.

Van Eck, N.J., L. Waltman, R. Dekker, and J. Van Den Berg. 2010. A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology* 61: 2405–2416. https://doi.org/10.1002/asi.21421.

Van Eck NJ, Waltman L, Van den Berg J. (2005) A novel algorithm for visualizing concept associations. In *6th International Workshop on Database and Expert Systems Applications (DEXA'05)* (pp. 405–409). IEEE

White, H.D., and B.C. Griffith. 1981. Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science* 32 (3): 163–171.

White, H.D., and K.W. McCain. 1998. Visualizing a discipline: An author co-citation analysis of information science: 1972-1995. *Journal of the American Society for Information Science* 49 (4): 327–355.

Yong-Hak, J. 2013. *Web of science*. Philadelphia: Thomson Reuters.

Zhao, D., and A. Strotmann. 2015. *Analysis and visualization of citation networks*. San Rafael: Morgan & Claypool.



**Myriam Ertz, PhD** is an Assistant Professor of Marketing, at the Department of Economics and Administrative Sciences of the Université du Québec à Chicoutimi (UQAC), and founder and head of the LaboNFC (Laboratory of research on New Forms of Consumption). Her research

interests include modeling of consumer choice, branding, and business ethics. She published articles on these topics in reputable journals (e.g., *Journal of Cleaner Production, Journal of Environmental Management, Journal of Business Research, Environment & Behavior, Technological Forecasting & Social Change, International Marketing Review*), and serves on the editorial review board of the *Journal of Consumer Marketing*.

**Sébastien Leblanc-Proulx, MSI** is a postgraduate Master's student in organization management. He holds a master's degree in information sciences and his research interests include the study of the collaborative economy and product lifetime extension practices. He published on this subject in the *Journal of Cleaner Production*, wrote a book chapter on the reconfigurations of commercial exchanges, and presented an article on this topic at the annual conference of the *Association Francophone pour le Savoir* (ACFAS).