JOURNAL OF
INTERDISCIPLINARY
HISTORY OF IDEAS

2018

Volume 7 Issue 13
Item 11

– Section 3: Notes –

Research Report | Forms, Patterns, Structures

Citation Analysis and the History of Analytic Philosophy

by
Eugenio Petrovich

JIHI 2018

Volume 7 Issue 13

Section 1: Editorials

1. *Editorial* (JIHI)

Section 2: Articles. Special Issue: Contemporary Luther / Luther contemporain

2. *Luther, la Réforme, la Modernité* (M. Albertone, O. Christin)
3. *Les Réformateurs, de l'*ethos *monastique à l'*habitus *académique* (O. Christin)
4. *Modernité catholique, modernité protestante. Batailles historiographiques à l'époque contemporaine* (E. Belligni)
5. *Protestantisme, montagne et environnement: une relation privilégiée?* (L. Tissot)
6. *Luther dans les documents du magistère pontifical du 500e anniversaire de la naissance au 500e anniversaire de la Réforme (1980-2017)* (P. Cozzo)
7. *Luther and his Catholic Readers: the Question of the Nuns* (E. Guillemard)
8. *Is Protestantism the Source of Modern Freedoms?* (V. Zuber)
9. *Elective Affinities and* liaisons dangereuses*: Luther's Heritage and the New Spirit of Capitalism* (D. Spini)
10. *Protestantisme et anarchisme* (P. Adamo)

Section 3: Notes

11. *Research Report | Forms, Patterns, Structures. Citation Analysis and the History of Analytic Philosophy* (E. Petrovich)

Section 4: Reviews

12. *Book Reviews* (J.-L. Bonniol)

Research Report | Forms, Patterns, Structures
Citation Analysis and the History of Analytic Philosophy

Eugenio Petrovich *

In this brief note I discuss two citation analysis-based studies in history of analytic philosophy that I recently published, highlighting some of their methodological features. In the first section, I sum up the two studies, focusing on the three methodologies that were used (citation counting, co-citation analysis, and citation context analysis). In the second section, I advance three remarks on these studies. Firstly, I argue that citation analysis methods produce a formal representation of their object, i.e. they shed light on the form rather than the content of the object. Secondly, I argue that these methods have an ontological counterpart: they frame the object under study at the documental level. I point out that this level should be distinguished both from the intellectual level that is studied by the internalist history of philosophy and from the social level that is studied by the externalist history of philosophy. Thirdly, I point out that citation analysis allows to reach a panoramic point of view on the object under study. Such perspective unveils patterns that are invisible at the micro-scale and that are difficult to study by traditional methodologies. I argue in particular that we need to develop new theories and concepts to better understand the objects and phenomena we observe from this distant point of view. In the third section, I highlight the strengths and weaknesses of citation analysis. The main strengths are epistemological, heuristic, and methodological, whereas the weaknesses relate to the losses caused by the translation of object at the documental level, the risk of being distracted by mathematical properties that lack a clear interpretation, and the problems involved in the validation of the results. I conclude by stressing the need for an interdisciplinary research programme that integrates citation analysis, history of philosophy, and the social studies of science.

* Università di Milano (*eugenio.petrovich @ unimi.it*).

Journal of Interdisciplinary History of Ideas 7(2018), 13, p. 11:1–11:21. Peer-reviewed.

The aim of this brief note is to present some methodological remarks on two citation analysis-based studies in history of analytic philosophy I recently published (Petrovich, 2018; Petrovich & Buonomo, 2018). Both studies applied scientometric methods to the recent history of analytic philosophy. Scientometrics is the discipline investigating the quantitative aspects of science and technology (Nalimov & Mulchenko, 1971). Its main method is the analysis of citations (citation analysis).

Scientific papers are mutually connected by citations, i.e. the cited references that each scientific paper presents. These connections can be studied from different points of view: the number of citations a document or an author collects over time can be counted, developing the so-called scientometric indicators; similarities between papers can be calculated by counting how many times they are cited together (co-citation analysis); the relation between basic science and technology can be assessed by studying how patents cite the scientific literature, and so on (Mingers & Leydesdorff, 2015). The studies I will comment on here were both based on citation analysis and they both addressed the so-called 'late analytic philosophy', i.e. the analytic philosophy of the last thirty to forty years (Tripodi, 2015, Chapter 4).

This note is structured as follows. Firstly, I will briefly sum up the two studies, focusing on the three methodologies employed (citation counting, co-citation analysis, and citation context analysis). Secondly, I will highlight three methodological features shared by both studies: I will argue that they provide a *formal representation* of the object under study, that they address what I will call the *documental level* of analytic philosophy, and that they allow to reach a *panoramic point of view* on scientific and scholarly disciplines. Lastly, I will discuss strengths and weaknesses of using citation analysis for historical purposes, discussing three 'epistemological risks' that may arise in the context of this kind of studies.

1. Sum up of the two studies

1.1. Citation counting and co-citation analysis

In the first study (Petrovich & Buonomo, 2018), that I conducted with my colleague Valerio Buonomo, we performed two citation analyses on late analytic philosophy production, *citation counting* and *co-citation analysis*, using the VOSviewer software (van Eck & Waltman, 2010). We took as target of our analysis all the articles published in five highly prestigious analytic philosophy journals (*The Philosophical Review, The Journal of Philosophy, Mind, Noûs, Philosophy and Phenomenological Research*) between 1985 and 2014. Thus, we worked on a corpus of 4 966 articles, containing 58 281 references to 17 926 authors. The first analysis (citation counting) is a very basic one, consisting in counting the citations a document or an author receives in a certain set of publications. The outcome of the analysis is classically a *ranking* that allows to gauge the 'impact', i.e. the citation score, of different documents and authors in the selected literature. In our study, we presented rankings of most-cited documents and most-cited authors in the five target journals. These rankings allowed to measure the weight, in terms of citations, of authors belonging to different phases of analytic philosophy (early, middle and late) in contemporary analytic debates (Petrovich & Buonomo, 2018, tabs. 1-4).

The second analysis (co-citation analysis) is a technique developed in the context of science mapping, i.e. the research area that aims at describing the structure of science based on the citation relationships amongst scientific documents (Börner, Theriault, & Boyack, 2015). Co-citation analysis was introduced in the 1970s (Small, 1973). The underlying idea is that documents that are frequently cited together (i.e. frequently co-cited documents) deal with similar topic or belong to the same research area. Mapping co-citation relationships allows then to reconstruct the sub-disciplinary structure of science. Results of co-citation analysis are classically visualized in the form of 'maps', i.e. two- or three-dimensional visualizations that represent visually the co-citation relationships between documents. VOSviewer, in particular, offers a two-dimensional distance-based visualization, where the distance between the nodes representing the documents is inversely proportional to their co-citation frequencies:

the higher the co-citation frequency between two documents is, the closer they will be represented on the map. In our study, we used co-citation analysis to map the structure of the articles published between 1985 and 2014 in the five journals mentioned above. We mapped both the overall production and the articles published in three consecutive decades ([1985-1994], [1995-2004] and [2005-2014]), in order to investigate both the structure and the dynamics of the field in the last thirty years. Regarding the overall map, the main finding was that the map of analytic philosophy presents a clear sub-disciplinary structure, where clusters of documents belonging to different sub-areas of the field (such as metaphysics, epistemology, philosophy of mind, etc.) are easily recognizable (Petrovich & Buonomo, 2018, Fig. 1). Regarding the longitudinal analysis, i.e. the sequence of maps in time, the main result was that the co-citation network becomes, from the first map to the last, more and more structured, changing from a sparse network to a clusterized configuration where distinct clusters of documents are evident (Petrovich & Buonomo, 2018, Figs. 2 a, b, c). We interpreted this pattern as the sign of an increasing *specialization* (i.e. fragmentation into sub-disciplinary areas) of analytic philosophy.

1.2. Citation context analysis

In the second study (Petrovich, 2018) I used a different approach to the study of citations, known as citation context analysis. In citation context analysis, the portion of citing text surrounding the citation is considered. This allows to classify citations according to their *function* (distinguishing for instance positive from critical citations, substantive from perfunctory citations) (Bornmann & Daniel, 2008). The aim of the study was to assess a claim recently made by historians of analytic philosophy, namely that analytic philosophy has approached in the last decades a style of intellectual production close to the Kuhnian normal science (see, amongst others, (Levy, 2003; Marconi, 2014; Putnam, 1997; Richardson, 2008)). This was done by classifying the citations appearing in analytic philosophy articles according to their epistemological function, and assessing how the distribution of the different categories of citations changed over time. Compared to the previous study, this one uses a larger time-window

(1950-2010) in order to trace the evolution from middle to late analytic philosophy. However, the corpus was much smaller, consisting of 60 articles: six sets of 10 papers, drawn from the same five journals of the previous study, with each set drawn from a decade between 1950s to 2000s. The total number of citation analyzed amounted to 1 293. The different size of the corpus depends on the fact that the articles had to be close-read in order to classify the citations.

The main finding of the study was that the rate of *State-of-the art* citations (i.e. the citations that are used to provide an overview of the sub-area to which the paper is meant to contribute) increased steadily from 1960s to 2000s (Petrovich, 2018, fig. 5). At the same time, however, the rate of *Positive* citations (i.e. the citations that are used to support the claims of the citing paper) followed an up-and-down pattern (Petrovich, 2018, fig. 3). I interpreted these results as indicating a fragmentation of the field into several definite sub-areas. Within each of the sub-areas there is a lack of consensus among analytic philosophers, as in traditional philosophical debates. Nevertheless, a consensus emerges at the level of the whole field over the background structure of analytic philosophy, i.e. analytic philosophers seem to agree on the sub-disciplinary divisions of analytic philosophy. Thus, it can be concluded that a sort of *soft paradigm*—concerning the structure of the field—has indeed emerged in the last decades.

I will not comment further the results of these two studies and will now move to commenting some methodological features they shared, discussing also the differences between citation analysis-based methods and the classic methods of history of philosophy.

2. Methodological remarks

2.1. A formal representation

The two studies summarized above share—even if at different levels—a feature that is typical of scientometrics in general: they provide a *formal representation* of their object (Wouters, 1999b, 1999a). In fact, the three analyses described above (citation counting, co-citation analysis, citation context analysis) all focused on the *relationships* between documents (namely, the citations), not

on the *contents* of the documents. They dealt with the form rather than the content of late analytic philosophy. This is most evident in the case of co-citation analysis: the visual outcome of the analysis is a network map where the crucial point is the *reciprocal distance* between items on the map, not the items in themselves. What a science map shows, is the *structure* of a set of documents, i.e. the formal relations between them, not their content. The same holds for citation counting: the citation score of a document or an author is the result of the sum of the relationships it has with other documents or authors. This is clear if we imagine the cited document as a node in a network: the number of citations it collects is in fact nothing more than the number of *links* it has with other documents. The citation score of a document is always dependent on its being part of a set of other documents, it depends structurally on the fact that it is not isolated. Thus, also the citation score is a formal property of a document, incorporating into a single measure the number of links it has with all the other documents. Lastly, in the case of citation context analysis, the focus is again on the link between the cited and the citing documents. However, in this case, we also take into consideration the *kind of link*, in order to enrich the citation network of the citing document with a further, qualitative dimension (the epistemological function, in our case). Still, this qualitative information denotes the relationship between documents, not directly their content, and is therefore a property of the form.

Thus, all these analyses provide a *formal* representation of the object under study, in the sense that they are mainly concerned with *relations* between items, i.e. with *structures*. More specifically, they are grounded in a *network approach* to their object: they represent their object as a set of nodes connected by a set of links. Now, a crucial aspect of formal representations in general is that they need to be *interpreted* in order to get a *meaning*[1]. Note that the interpretation works differently in the three cases, raising different methodological issues.

In the case of the science maps discussed in (Petrovich & Buonomo, 2018), the formal representation is provided by the VOSviewer algorithm in the form of a two-dimensional visualization of the co-citation relations between documents, the so-called map. As said above, the map presents a clear structure, i.e.

[1] Consider the case of formal languages. They need to be provided with an interpretation in order to get a meaning in the natural language.

documents are not arranged randomly. In particular, they are distributed in different clusters, that the VOSviewer algorithm detects and visualizes in different colors. However, the meaning of the clusters is not directly showed in the map. It needs to be inferred by looking at the contents of the documents of each cluster. For instance, the red cluster showed in (Petrovich & Buonomo, 2018, Fig.1)[1] can be labelled "metaphysics" because it contains several documents (such as Lewis's *On the Plurality of Worlds* and Quine's *World and Object*) the content of which regards this area of analytic philosophy. Now, the interpretation of the map consists precisely in relaying its structure (the clusters) with the intellectual content of the documents it shows. This process is known in scientometric literature as the 'validation' of the map and it is usually done by asking to experts of the field to look at the scientometric map and judge if it matches with their representation of the research area. Therefore, the interpretation always requires a qualitative element (the assessment of field experts) to match the quantitative representation.

In the case of the citation counting, the interpretation of the data consists in attributing a meaning to the citation score. In particular, the question that should be answered is the following: what does a high (or low) number of citations mean? Note that this question opens two very different venues of research: a descriptive and a normative. In the descriptive venue, the question concerns, for instance, the relation between highly cited documents and Kuhnian paradigms (Small, 2003). Can they be considered the same? Do highly cited documents play the role of Kuhnian paradigms? In the normative venue, on the other hand, the question concerns the relation between the citation score of a document and its scientific—or, in the case of analytic philosophy, philosophical—*quality*.

[1] See the final page of this paper.

In fact, the main application of scientometrics is in research performance evaluation: scientometric indicators such as the Journal Impact Factor or the H-Index are increasingly used in evaluation exercises to gauge the scientific quality of journals, universities, laboratories and even individual researchers. The underlying assumption is that citation scores can be used as proxies for research quality. In particular, high citation score would mean a high research quality. Note that the very notion of 'research quality' is intrinsically normative, because it implies the reference to *standards* and *desiderata*. Thus, the interpretation of citation scores, also in the case of analytic philosophy, potentially opens a huge *normative* and *meta-philosophical* problem. Normative because it regards the idea of research quality, and meta-philosophical because it concerns how *philosophical* research should be conducted. I think that the mere use of citation counting cannot answer the question whether citations are proxies for research quality. Nevertheless, *if this equation is accepted*, then citation score would have a clear interpretation and citation counting would become indeed a method to empirically measure philosophical quality.

Lastly, in the case of citation context analysis, the formal representation consists in distributions of categories of citations over time, visualized, for instance, in the form of box plots. The interpretation of these statistical estimates depends on the meaning of the different categories used to classify citations. In turn, the categories are designed based on an *epistemological* theory of the possible roles of citations. If we change the categories or the theory that attributes meaning to them, this affects the interpretation of the trends and patterns visible in the statistics. In other terms, we have always to take into consideration the well-known phenomenon of underdetermination: the evidence constraints but does not determine our interpretation of it (Stanford, 2017). Once again, the interpretation depends on a theoretical element that is not provided by the formal representation alone, but should be supplied by the interpreter.

In sum, the first methodological remark I would like to make on these two studies is that scientometrics and citation analysis provide a formal representation of their object. This form is a pure mathematical or statistical structure that has no meaning *per se*. The meaning needs to be introduced by the human element, and this is the role of the historian. Therefore, the use of quantitative methods does not cancel out the historian from the equation, but, on the contrary, highlights his/her role as an interpreter and, even, a theory-builder.

2.2. The documental level

The different approaches to the history of philosophy can be classified in a spectrum ranging from internalist to externalist approaches (Rorty, Schneewind, & Skinner, 1984). Internalist approaches focus on *intellectual contents*, such as philosophical theories, arguments, notions, concepts, etc., whereas externalist approaches focus on the historical, social and political *contexts* of philosophical ideas. In the label history of philosophy, internalists put the emphasis on philosophy, rather than on history: the object of historical reconstruction is the 'context of justification' of philosophical doctrines, not their historical and sociological genesis (the 'context of discovery'), which is the object of externalist history of philosophy (Reichenbach, 1938).

The most internalist kind of history of philosophy is the so-called "rational reconstruction" (Rorty, 1984). In this kind of history, all historical factors are ignored, and the doctrines of the past philosophers are discussed and assessed as they were contemporary colleagues. Past philosophical theories are discussed from a sort of a-temporal point of view, in the pure "space of reasons". On the other end of the spectrum, we find full-fledged externalist approaches to the history of philosophy. Sociology of philosophy is the main representative of this kind of approach (Kusch, 1995, 2000). Following the tradition of Mannheim sociology of knowledge as well as the Strong Programme in sociology of science, it aims at showing how the ideas are shaped (and even determined) by the social context in which they are produced. In between these two extremes there are intermediate approaches that mix internalist and externalist considerations, attributing from time to time different weights to intellectual and social factors.

What is the position of the scientometric approach in the internalist-externalist spectrum? I think that in order to answer this question, we need to frame the internalist-externalist debate in *ontological* terms, i.e. we need to focus on the kind of object the two approaches focus on. As we said above, internalist approaches consider philosophy as a set of 'ideas', i.e. as a set of intellectual 'items' that have at least a certain degree of independency from their creators. In other terms, ideas can be discussed *per se*: they can be put in relation with other ideas, if they are arguments they can be rejected, if they are notions they can be clarified, etc. In internalist history of philosophy, ideas are considered as inhabi-

tants of what Popper calls the Third World, the world of "objective knowledge", that is distinct both from the First World (the physical world) and the Second World (the mental world) (Popper, 1979). On the other hand, in externalist approaches, ideas are considered as essentially connected with their context. In fact, they are considered as the *product* of a certain disposition of social features. They cannot be considered as independent entities. In fact, the true objects of externalist histories of philosophy are not ideas, but *agents*, i.e. the producers of ideas. Agents use ideas as resources to advance their own interests, and the evolution of ideas in time can be explained only by taking in consideration their strategic use (Collins, 1998; Kusch, 1995)[1]. Ideas are social objects that are not individuated by a set of intellectual commitments, but by the *community of actors* advancing them.

The scientometric approach is, I think, *intermediate* between the internalist and externalist approach. This intermediate status is a direct consequence of the methodology that is typical of scientometrics. Scientometrics does not address directly scientific activity, but only its public outcome, i.e. articles published in scientific journals. From these articles, citations are extracted, and the formal relations between pieces of knowledge are studied via citation analysis. Therefore, publications and citations are the true *object* of scientometrics, they are the *ontological counterpart* of the scientometric method. Every application of scientometric methods to an object of study, such as analytic philosophy, will involve a translation of that object into a relevant set of documents and citations. This translation is a form of *operationalization*, i.e. the transformation of a concept into something to which a certain form of research method can be applied successfully (Chang, 2009)[2]. In both the studies I am commenting, late

[1] This approach has never been fully developed in history of philosophy. In Science and Technology Studies, however, it is extensively promoted by the Strong Programme and, from a certain point of view, by the Actor-Network-Theory of Bruno Latour (2003).

[2] Providing "operational definitions" of elusive concepts is a fundamental part of social science methodology (Calhoun, 2002).

analytic philosophy was operationalized into a set of documents, defining it at what we may call the *documental level*.

Thus, in scientometrics, the object of study is neither defined as a set of *ideas* (internalist approach) nor as a set of *agents* (externalist approach), but as a set of *published documents* (publications). Publications can be considered as the *interface* between the abstract realm of ideas and the social field of actors, since they share features of both "Worlds", to use Popper's phrase[1]. They share with ideas the property of existing independently of their creators, since they are public objects that can circulate even beyond the intentions of their authors. At the same time, they share with the social level the property of being concrete and spatio-temporally determined. Publications have definite coordinates and boundaries, they are not abstract objects in the intellectual space. Therefore, scientometrics allows to capture an object of study at an intermediate level that does not coincide neither with the strictly intellectual point of view, typical of internalist history of philosophy, nor with the strictly social point of view, typical of externalist history of philosophy. Scientometrics is therefore a sort of *ontological middle point* between internal and external history[2].

2.3. Towards a massive and panoramic history

Traditional history of philosophy basically relies on one key method: close reading of texts. This poses serious limitations to the amount of sources that can be handled by the historian of philosophy. One of the main motivations for turning to 'distant reading' methodologies is in fact the intention to overcome the cognitive limits of the individual scholar, using computers and algorithms to enlarge enormously the amount of data that can be used in the analysis. This paves the road to a massive, data-driven history. Scientometrics shares with

[1] In Science and Technology Studies, the role of documents in science is gaining considerable attention in the last years (see Felt, Fouché, Miller, & Smith-Doerr, 2017, Chapter 2).

[2] The question whether the documental level may be considered as a sort of Hegelian synthesis between the intellectual and the social is an interesting one, that I think should be addressed in future research.

distant reading (and with digital humanities more generally) this 'big-data' approach. Citation analyses are conducted on thousands, even millions of documents. Furthermore, it has been argued that scientometric indicators are meaningful *only if* they are performed on huge amounts of data, adopting a sort of 'thermodynamic' approach that focuses on the aggregate rather than the individual behavior (Price, 1986; van Raan, 1998). As we saw above, the map of late analytic philosophy we presented in the first study was the result of processing almost 5 000 articles: it would have been simply impossible to read all of them in the traditional way. The same holds for the citation counting analysis: recording manually the citations to authors and documents would have been a tremendously time-consuming task.

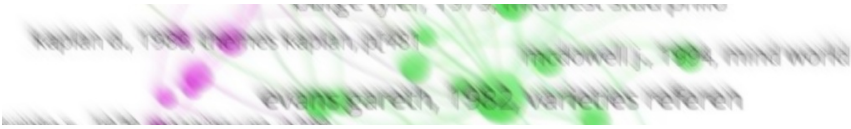
The shift from small to large sets of texts has two important consequences: it shifts the scale from the micro to the macro and the focus from the individual to the collective. This has the effect of moving towards what we may call a *panoramic* point of view. The panoramic point of view allows to observe phenomena that cannot be perceived at the micro level, or that can be perceived only qualitatively as anecdotal evidence. This is the case of the phenomenon of specialization. The trend towards an increasing specialization is perceived by analytic philosophers but it can hardly be investigated at the micro-level and with close reading methods, because it is a trend involving the *whole* field, not a specific author or philosophical theory. Hence, it is difficult to test by traditional methods whether or not the perception of analytic philosophers is true. The panoramic view of the science mapping, on the other hand, allows to observe the dynamics of the whole field, and therefore to check the perception of the field members against quantitative patterns. Indeed, as we saw above, the use of longitudinal co-citation analysis (i.e. the mapping of analytic philosophy production by decades) allowed to observe a clusterization pattern that can be interpreted as the effect of the fragmentation of analytic philosophy into definite sub-specialty areas.

I said that the panoramic point of view changes the scale of the analysis, allowing to detect macro-phenomena such as specialization, but also the focus, highlighting the collective dimension of a field. Again, this is evident for all the kinds of citation analysis. Citation analysis as such involve the use of citations, and citations are a sign of the embedding of the documents into a fabric of documents, i.e. a network, that is constitutively a collective entity. The individual

authors, when they cite, contribute to the weaving of the fabric of documents, and their collective efforts results in the overall structure of the network.

Now, I suggest that we should consider the idea that the direction of causation goes not only from the individual to the structure, but also on the other way, from the structure to the individual. In other terms, I suggest the existence of *feedback mechanism* from the structure to the individual. This means that the individual action is constrained by the structure, i.e. the structure reduces the degrees of freedom of the individual. If this suggestion is correct, the structure and the trend that are revealed by citation analysis can be interpreted as the *collective constraints* that the individual actors face when they interact with the field. The physical metaphor of the 'field' is particularly suitable for describing this interaction. A gravitational field is revealed by the effect it has on the masses interacting with it—and this effect is manifested as gravitational attraction on the masses. In the same way, I suggest interpreting the clusters in the maps as different centers of gravity that exert a force on the individual actors, restraining their possible actions. Furthermore, the trend revealed in the longitudinal analysis can be interpreted by another physical notion: the notion of *inertia*. An individual actor that wants to invert the trend (for example, address general philosophical themes instead of specialized and delimited philosophical puzzles) must deal with the inertia of the whole field, that pushes all the actors towards specialization.

In sum, the panoramic point of view that results from operating with huge amounts of data does not only modify the scale of the studies but allows to unveil new types of dynamics. I think that an important task for historians of philosophy that want to use these methods consists in developing a good theory of these dynamics, forging new concepts for phenomena that are visible only from the panoramic perspective[1].



[1] This point was also highlighted by Franco Moretti in his lecture "Patterns and interpretation", delivered at the "Distant Reading and Data-Driven Research in the History of Philosophy" Conference, Università di Torino, 16-18 January 2017.

3. Strengths and weaknesses of citation analysis

Scientometric methods based on citations present both strengths and weaknesses when they are used for historical purposes. I already mentioned some of them in the previous paragraphs. I want now to systematize them, starting from the strengths. I think that they can be divided into epistemological, heuristic, and methodological strengths:

1. From an epistemological point of view, the main strength of scientometrics is that it allows to *operationalize* elusive notions (such as 'specialization') into *definite* and *measurable* features. This happens because the translation of the object of study into a documental object (i.e. into a set of documents) plays a crucial role in the very methodology of scientometrics. The operationalization allows to test *by quantitative means* historical claims put forward by historians of philosophy. I think this is an important step towards a more robust and empirically oriented history of philosophy.

2. The panoramic point of view that is reached when a massive amount of documents is analyzed via citation analysis allows to unveil trends and patterns that are invisible at the micro scale and from a close-reading standpoint. Thus, this perspective can generate new ways of reading the historical material, providing a significant *heuristic gain* to the historian of philosophy.

3. Regarding the methodology, it must be stressed that scientometric and citation analysis are not incompatible with close reading or traditional methodologies. In fact, quantitative and qualitative methods can be *coupled*, for instance in the citation context analysis. This allows to enhance both kinds of analysis and, most of all, to interpret more clearly the results of quantitative analyses. As we saw above, data do not speak by themselves: they need always to be interpreted within a theoretical framework. The formal representation has to be filled with a meaning. Quali-quantitative methods can indeed enable the interpretation of results.

At the same time, however, there are several weaknesses of citation analysis as an historical tool that cannot be ignored. There are technical limitations as well as what we may call 'epistemological risks'.

Concerning the technical limitations, it must be noticed that the available citation databases (Clarivate's Web of Science, Elsevier's Scopus and Google's Google Scholar) have a limited coverage and historical depth. Books and monographs are not indexed in the database[1] and there are biases in the coverage of journals. In particular, English-language journals are disproportionally represented in the databases. Moreover, citation data before 1980s are only partially reliable. In general, the coverage of Humanities and Social Sciences is unsatisfactory, compared to the natural sciences and bio-medical areas. These technical limitations are widely discussed in the scientometric literature and the reader can refer to this literature for more details on this point[2]. For the purposes of this note, I think it is more interesting to discuss some 'epistemological risks' that are involved in using citation analysis for historical purposes. They concern the losses in the operationalization process, the independent life of mathematical objects, and the problem of validation of results.

3.1. Losses in the operationalization process

When the intellectual level (the level of ideas) is translated to the documental level (the level of publications), this does not happen without losses. Furthermore, if we consider only the *publications* indexed in a database, this imposes a further limitation to the object we study. By focusing only on the published product of philosophical research, we lose the entire *process of production* of the philosophical contents. We lose what we may call 'philosophy in action', echoing Latour's notion of "science in action". The losses involved in the operationalization should always be kept in mind when we interpret the results. It is important to be clear about what aspect of our object we are describing and

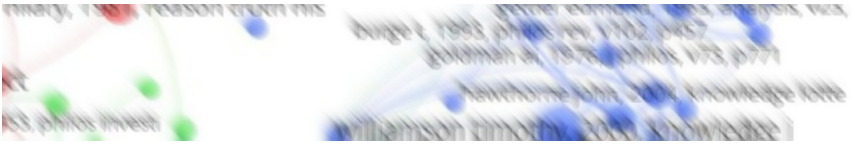
[1] Web of Science has recently launched a Book Citation Index, but, at the moment, it is still too limited in coverage to be useful.

[2] See (Nederhof, 2006) and (Hellqvist, 2009) for an overview of the main technical limitations in applying citation analysis to the humanities.

about the fact that the description is constrained by the method we use. We should always be wary of excessive generalizations.

3.2. The independent life of mathematical objects

Quantitative methods in general produce quantitative images of their object. The quantitative features can hence be manipulated by mathematical means (for instance, by statistical analysis). In the case of scientometrics, we saw above that the documental level is represented as a *network*. A network can be described mathematically as a mathematical object—a graph—so that the documental level can be related to the mathematical theory of graph. This theory offers a lot of interesting tools for studying the properties of networks (for example, all the different measures of 'centrality', such as degree centrality, betweenness, Eigenvalues, etc.). However, a concrete risk arises of being lost into the mathematical properties of the network, losing sight of the meaning that such properties have for our object of study. An example would help to clarify this point. Imagine we analyze the science map of late analytic philosophy with a cluster algorithm. This algorithm "recognizes" the different sub-communities of the network, grouping similar items into the same cluster. Then, it is possible to ask whether there is a relation between the number of clusters and the number of items contained in each cluster. Is there any mathematical relation holding between these two quantities? Is it dependent from the cluster algorithm we choose? Questions like these are interesting, but it is not straightforward that their answers would shed light on the *object* that is represented with the science map. They seem to concern more the *representation* than the object represented, or, to use the terminology introduced above, they seem to regard only the *formal* properties of the object. Therefore, quantitative methods raise technical and mathematical problems that have 'a life of their own': the risk is to pursue these problems losing sight of the target.



3.3. Lack of validation vs. reinventing the wheel

In discussing the strengthens of scientometrics as a tool, I highlighted the fact that it can unveil patterns that cannot be observed with classical methods. However, this also raises the issue of the *validation* of the results obtained by scientometric methods. If they unveil patterns that are invisible at the qualitative level, how can they be assessed? Imagine that a science map cannot be easily interpreted by the experts of the field: should we reject the results of the map or the assessment of the experts? There is also a mirror risk: if the experts recognize easily the map as a faithful representation of their field, what is the advantage that is gained by using science maps instead of the classic reviews of the literature? The use of quantitative images of our object of study presents therefore a double risk: if they match the already-known representation, then they add nothing to our knowledge; if they contradict the judgement of the experts, then we lack an independent source of validation for the results. In sum, scientometric and citation analysis are in between the Scylla of reinventing the wheel and the Charybdis of meaningless results.

4. Conclusions

In this brief note, I discussed three main methodological aspects of the two studies I recently published on the history of recent analytic philosophy. I focused on the methodological consequences of the use of scientometrics and citation analysis methods as research tools. Firstly, I argued that these methods produce a formal representation of their object, i.e. they shed light on the form rather than the content of the object. The role of the historian is to attribute a meaning to the formal mathematical structure by providing an interpretation of the data. Secondly, I argued that these methods have an ontological counterpart: they frame the object under study at the documental level, i.e. as a set of (published) documents. This level of analysis should be distinguished both from the intellectual level (the ideas) and the social level (the social actors) and considered as the interface between them. Thirdly, I pointed out that scientometrics allows to reach a panoramic point of view on the object under study. Such perspective allows to unveil patterns that are invisible at the micro-scale

and that are difficult to study by traditional methodologies (i.e. close-reading). I argued that we need to develop new theories and concepts to better understand the objects and phenomena we observe from this distant point of view. Lastly, in the third section I highlighted the strengths and weaknesses of scientometrics and citation analysis. The main strengths are epistemological (enhancing the empirical test of historical claims), heuristic (generating new hypotheses) and methodological (coupling with close-reading method). On the other hand, the weaknesses regard the losses caused by the translation of object at the documental level, the risk of being distracted by mathematical properties that lack an interpretation, and the problems involved in the validation of the results.

I think that the scientometric study of philosophy is just at the beginning. It should be extended both in the scope, considering other areas of philosophy, and in the theoretical machinery: as I stressed many times in this note, we need to develop *new concepts* to interpret insightfully the results of quantitative methods. Therefore, I believe that we need a closer interaction with other areas of philosophy (namely, philosophy of science, historical and social epistemology) and with the social sciences (sociology of science, Science and Technology Studies), promoting a true inter- and trans-disciplinary research programme.

References

Börner, K., Theriault, T. N., & Boyack, K. W. (2015). Mapping Science Introduction: Past, Present and Future. *Bulletin of the Association for Information Science and Technology*, *41*(2), 12-16. DOI:10.1002/bult.2015.1720410205

Bornmann, L., & Daniel, H. (2008). What Do Citation Counts Measure? A Review of Studies on Citing Behavior. *Journal of Documentation*, *64*(1), 45-80. DOI:10.1108/00220410810844150

Calhoun, C. (2002). *Dictionary of the Social Sciences*. Oxford: Oxford U.P.

Chang, H. (2009). Operationalism. In *The Stanford Encyclopedia of Philosophy* (Fall 2009 Edition). Retrieved from `https://plato.stanford.edu/archives/fall2009/entries/operationalism`

Collins, R. (1998). *The Sociology of Philosophies: a Global Theory of Intellectual Change*. Cambridge, MA: Belknap Press of Harvard U.P.

Felt, U., Fouché, R., Miller, C., & Smith-Doerr, L. (Eds.). (2017). *The Handbook of Science and Technology Studies* (Fourth edition). Cambridge, MA: The MIT Press.

Hellqvist, B. (2009). Referencing in the Humanities and its Implications for Citation Analysis. *Journal of the American Society for Information Science and Technology*, *61*(2),310-318. DOI:10.1002/asi.21256
Kusch, M. (1995). *Psychologism: a Case Study in the Sociology of Philosophical Knowledge*. London; New York: Routledge.
Kusch, M. (Ed.). (2000). *The Sociology of Philosophical Knowledge*. Dordrecht; Boston: Kluwer Academic Publishers.
Latour, B. (2003). *Science in Action: how to Follow Scientists and Engineers through Society* (11th print). Cambridge, MA: Harvard U.P.
Levy, N. (2003). Analytic and Continental Philosophy: Explaining the Differences. *Metaphilosophy*, *34*(3), 284-304. DOI:10.1111/1467-9973.00274
Marconi, D. (2014). *Il mestiere di pensare: la filosofia nell'epoca del professionismo*. Torino: Einaudi.
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
Nalimov, V., & Mulchenko, B. M. (1971). *Measurement of Science: Study of the Development of Science as an Information Process*. Washington, DC: Foreign Technology Division.
Nederhof, A. J. (2006). Bibliometric Monitoring of Research Performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, *66*(1), 81-100. DOI:10.1007/s11192-006-0007-2
Petrovich, E. (2018). Accumulation of Knowledge in Para-Scientific Areas: the Case of Analytic Philosophy. *Scientometrics*. DOI:10.1007/s11192-018-2796-5
Petrovich, E., & Buonomo, V. (2018). Reconstructing Late Analytic Philosophy. A Quantitative Approach. *Philosophical Inquiries*, *VI*(1), 149-180.
Popper, K. R. (1979). *Objective Knowledge: an Evolutionary Approach* (Rev. ed). Oxford; New York: Clarendon Press; Oxford U.P.
Price, D. J. de S. (1986). *Little Science, Big Science… and Beyond*. New York: Columbia U.P.
Putnam, H. (1997). A Half Century of Philosophy, Viewed from Within. *Daedalus*, *126*(1), 175-208.
Reichenbach, H. (1938). *Experience and Prediction: an Analysis of the Foundations and the Structure of Knowledge*. Chicago: University of Chicago Press.
Richardson, A. W. (2008). Scientific Philosophy as a Topic for History of Science. *Isis*, *99*(1), 88-96. DOI:10.1086/587534
Rorty, R. (1984). The Historiography of Philosophy: Four Genres. In *Philosophy in History: Essays on the Historiography of Philosophy* (pp. 49-75). Cambridge; New York: Cambridge U.P.
Rorty, R., Schneewind, J. B., & Skinner, Q. (Eds.). (1984). *Philosophy in History: Essays on the Historiography of Philosophy*. Cambridge; New York: Cambridge U.P.

Small, H. (1973). Co-Citation in the Scientific Literature: A New Measure of the Relationship between Two Documents. *Journal of the American Society for Information Science*, *24*(4), 265-269. DOI:10.1002/asi.4630240406

Small, H. (2003). Paradigms, Citations, and Maps of Science: A Personal History. *Journal of the American Society for Information Science and Technology*, *54*(5), 394-399. DOI:10.1002/asi.10225

Stanford, K. (2017). Underdetermination of Scientific Theory. In *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition). Retrieved from `https://plato.stanford.edu/archives/win2017/entries/scientific-underdetermination/`

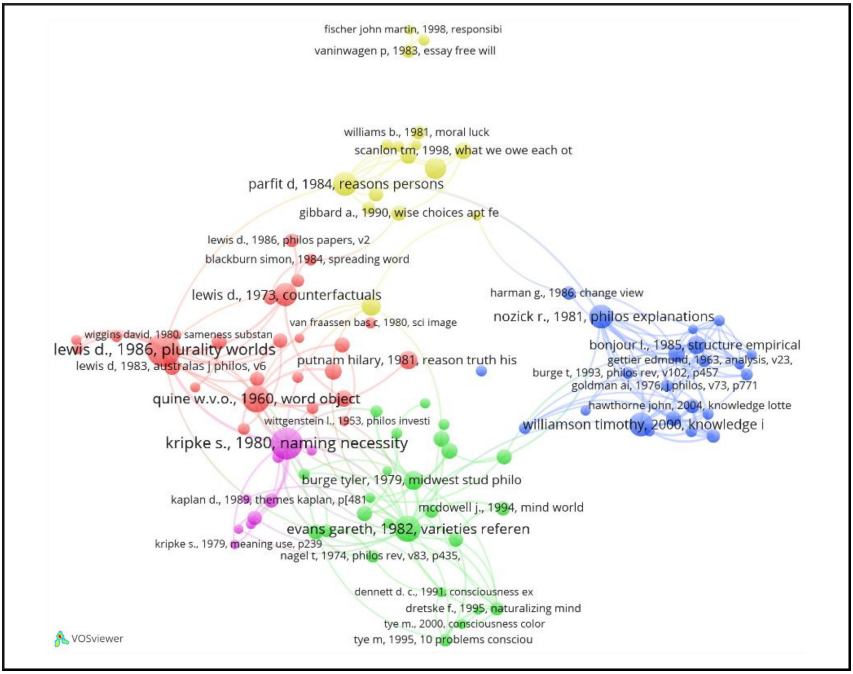
Tripodi, P. (2015). *Storia della filosofia analitica: dalle origini ai giorni nostri*. Roma: Carocci.

van Eck, N. J., & Waltman, L. (2010). Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping. *Scientometrics*, *84*(2), 523-538. DOI:10.1007/s11192-009-0146-3

van Raan, A. F. J. (1998). In Matters of Quantitative Studies of Science the Fault of Theorists is Offering Too Little and Asking Too Much. *Scientometrics*, *43*(1), 129-139. DOI:10.1007/BF02458401

Wouters, P. (1999a). Beyond the Holy Grail: From Citation Theory to Indicator Theories. *Scientometrics*, *44*(3), 561-580. DOI:10.1007/BF02458496

Wouters, P. (1999b). *The Citation Culture*. Unpublished PhD thesis, University of Amsterdam.



Petrovich & Buonomo, 2018, Fig. 1 (modif.).