Università degli Studi di Milano Bicocca
Dipartimento di Sociologia e Ricerca Sociale

Dottorato di Ricerca in
Sociologia Applicata e Metodologia della Ricerca Sociale
XXV Ciclo

Mapping scientific literature

Structuring scientific communities through Scientometrics

Tutor:
Prof. Antonio Chiesi

Co-Tutor:
Prof. Rickard Danell

Emanuela Riviera
Matr. 732752

Anno Accademico 2011-2012

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

At the basis of this work there is the belief that the collaboration between Sociology and Scientometrics should be intensified: as long as Scientometrics is engaged in the quantitative study of Science it cannot prescind from the fact that its object of study is produced in a social system; it cannot prescind from sociological conceptualisations in interpreting the results of its analyses. With this purpose, a sociologically integrated approach suggesting an interpretation of bibliometric indicators and maps is developed. The programmatic element of this work is the integration of the constructivist perspective relating to the re-production of scientific community with the metric program in Bibliometrics.

The accomplishment of the main purposes introduced above is realised through the following goals, which can be listed according to the level of generality. First of all, this work aims to demonstrate the potentialities of a proposal in the field of citation theory which encompasses different and divergent perspectives as it is based on the integration of structural-functionalist and constructivist approaches. More specifically, this proposal is applied to the field of Scientometrics, where the lack of a theoretical frame for interpreting bibliometric indicators as well as bibliometric maps constitutes a real problem as we are witnessing the affirmation of what can be called the “metric era”: as a matter of fact, like it or not, bibliometric measures and analyses are more and more important for scientists’ career. The purpose of providing a description of the intellectual and cognitive structure of the scientific field of Italian Sociology allows us, at the same time, to test the reliability of those bibliometric techniques employed in the empirical work here presented. In particular, we are testing the effectiveness of Bibliometrics in mapping scientific literature, as well as the effectiveness of the normative approach in describing citers’ behaviour with reference to high citation counts.

According to the scopes mentioned above, the work here introduced is divided into two Sections. The first one deals with Scientometrics at large: starting from its birth, its structural elements, the social and theoretical background that made possible its evolution and, what seem possible to call, its success, passing through a review about citations (dealing with theories and effectiveness matters) and about relational bibliometric studies (thus focusing on co-occurrence techniques as well as clustering and mapping ones), the Section ends with the analysis of scientific communities by means of the autopoietic meta-theory developed in Biology by Maturana and Varela and applied later to social systems by Luhmann. In order to conceive of Science as an autopoietic machine, the use of different theories belonging to different approaches is necessary. Among them we find some classics of the sociological thought, that is to say Giddens, Luhmann and Merton. Citations, conceived of as codes and medium of communication, are supposed to have a triggering role in the re-productive process of the system. Both constructivist and normative theories developed so far concur, in the unified approach developed in this thesis, to describe citers’ behaviour. Furthermore, on the basis of this theoretical apparatus an interpretation of bibliometric indicators as well as bibliometric maps is suggested.

The second Section of the thesis deals with the application of Scientometrics to the case of Italian Sociology. Before proceeding with the empirical work, conceptualisations about the organisation of scientific communities as well as information about the institutionalisation and organisation of Italian Sociology are given. The application of bibliometric techniques to the latter allows us to pursue two goals at the same time: 1) it furnishes a description of the intellectual and cognitive structure of Italian Sociology; 2) it makes possible to test the effectiveness of the normative approach in describing citers' behaviour with reference to high citation counts. These considerations are made on the basis of two bibliometric maps developed by employing the technique called all author co-citation analysis. Finally, data derived from the maps are also analysed trough Network Analysis tools whose compatibility with Relational Bibliometrics lays in the fact that both are immersed in the relational paradigm in Sociology.

SECTION 1
INTRODUCING AND THEORISING SCIENTOMETRICS

1. BIBLIOMETRICS: AN OVERVIEW

The “metrics realm” is a wide one, made of various, and in many cases overlapping, subfields. This situation causes a lot of confusion about the terminology related to this field. Among all the metric terms “Bibliometrics” is the most used and misused one; this misuse causes unclarity about Bibliometrics’ focus and nature. The aim of this Chapter is to make some order furnishing a concise description of the main characteristics of the discipline. For this purpose, the so called “metrics terms” are first introduced, and the main steps characterising the development from Bibliometrics to Webometrics are subsequently presented. The third and forth Sections focus on Bibliometrics’ structure and laws, respectively. In the next one Garfield's Science Citation Index (history, features and consequences of its diffusion) are illustrated; in the final Section an overview of the main bibliometric measures is given.

1.1. ABOUT TERMINOLOGY

Since the early 20th century many terms referring to the study of scientific literature have been introduced; *Bibliometrics*, *Scientometrics* and *Informetrics* are only some of those terms. As Glänzel and Schoepflin claimed at the “Fourth International Conference on Bibliometrics, Informetrics and Scientometrics” (Glänzel & Schoepflin, 1994), there is considerable confusion about the terminology of these three closely related metric fields and, according to them, this situation is partly due to scholars’ use (or abuse) of Bibliometrics as synonym for the other metrics (Hood & Wilson, 2001).

But why is there such confusion? There seem to be two main reasons. First of all, Bibliometrics is divided into various but sometimes not very different specialties, whose boundaries are not always well defined. Therefore, it happens that labels have overlapping, but not identical, meanings. Furthermore, bibliometric techniques are employed as research tools in sundry research fields with different scopes. The following scheme shows what has just been said:

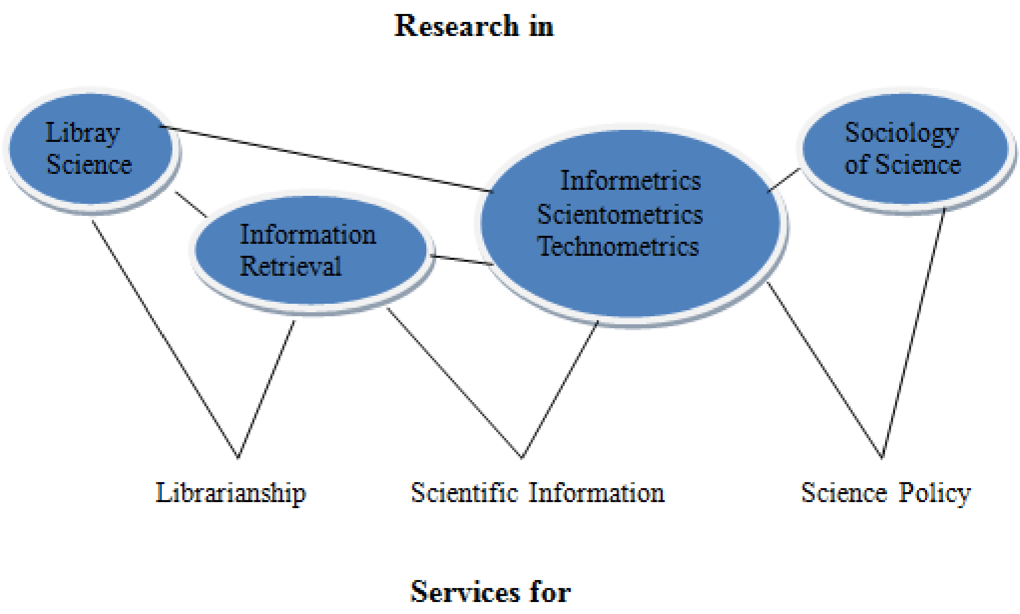


Figure 1: Links of Bibliometrics with related fields and application services
Source: Glänzel, 2003, p. 10.

In 2001 Hood and Wilson conducted a study[1] aiming to demonstrate the variety of terms used in Information Science obtaining the following results:

[1] The study includes only English terms and no attempt was made to remove duplicates (for a definition of duplicates see Footnote 134). The datasets used were 12: ERIC, INSPECT, NTIS, Social SciSearch Dissertation Abstracts Online, Gale Group Magazine DB, LISA, British Education Index, Gale Group Trade & Industry DB, Information Science Abstracts, Education Abstracts, Library Literature (Hood & Wilson, 2001).

Table 1: Number of documents in which each term related to the metric fields in Information Science appeared.

Term	**Frequency**
BIBLIOMETRICS	5097
BIBLIOMETRIC	2653
SCIENTOMETRICS	1326
SCIENTOMETRIC	552
INFORMETRICS	418
TECHNOMETRICS	274
INFORMETRIC	197
BIBLIOMETRY	73
BIBLIOMETRICALLY	40
STATISTICAL BIBLIOGRAPHY	38
BIBLIOMETRICAL	24
TECHNOMETRIC	20
BIBLIOMETRICIANS	17
SCIENTOMETRY	17
LIBRAMETRY	16
SCIENTOMETRICAL	11
SCIENTOMETRICALLY	10
SCIENTOMETRICIANS	9
BIBLIOMETICIANS	7
LIBRAMETRICS	7
SCIENTOMETRICIAN	6
INFORMETRICIANS	5
INFORMETRY	5
LIBRAMETRIC	5
TECHNOMETRICALLY	1

Source: Hood & Wilson, 2001, p. 296.

Table 1 gives us interesting information. First of all, we can observe the presence of terms that have not yet been mentioned in this work, such as *Librametry*, *Statistical Bibliography* or *Technometrics*. Furthermore, it is impossible not to notice the importance of the synonym phenomenon: almost all fields have two different, but almost identical, labels; for example *Bibliometry-Bibliometrics*, *Scientometry-Scientometrics* or *Informetry-Informetrics.*

In the last years other terms have been introduced as the Web became a new field for the application of bibliometric techniques. Thus, besides the three "big metrics" (Bibliometrics, Informetrics and Scientometrics) we also find *Webometrics* and its synonyms *Netometrics* and *Cybermetrics* (as we will see shortly, some scholars distinguish between Webometrics and the latter) (Hood & Wilson, 2001).

At this point it is useful to give some definitions helping us to distinguish among all these terms derived from the fusion of the words "bibliography", "information", "library" and "science" with the word "metrics" (Sengupta, 1992). *Bibliometrics* (or Bibliometry) refers to the measurement, by means of mathematical and statistical tools, of publication patterns as well as the analysis of elements such as citations, authors and semantic items of all forms of written communication regardless of discipline or research field (Potter, 1981a; Sengupta, 1992). *Scientometrics* (or Scientometry) refers to the quantitative study of the scientific and technological progress; it can be defined as the application of quantitative methods to Science conceived of as an information process (Nalimov & Mulchenko, 1969). Where the measurement refers to books, papers, citations and semantic elements, Scientometrics and Bibliometrics cannot be distinguished (Tague-Sutcliffe, 1992). In this regard, I think that the following passage by Glänzel is interesting:

> It is a common misbelief that bibliometrics is nothing else but publication and citation based gauging of scientific performance or compiling of cleaned-up bibliographies on research domains extended by citation data. In fact, scientometrics is a multifaceted endeavour encompassing subareas such as structural, dynamic, evaluative and predictive scientometrics. Structural scientometrics came up with results like the re-mapping of the epistemological structure of science based, for instance, on co-citation, "bibliographic coupling" techniques or co-word techniques. Dynamic scientometrics constructed sophisticated models of scientific growth, obsolescence, citation processes, etc. These models are not only of theoretical interest but can also be usefully applied in evaluation and prediction (Glänzel, 2003, p. 5).

He starts talking about Bibliometrics but proceeds using Scientometrics.

Other definitions, as the one by Van Raan (1998), include also Technology besides Science as object of study of Scientometrics, but *Technometrics* is recognised as a separate area of study and the homonymous journal founded in 1959 has the purpose to develop the employment of statistical methods in the physical, chemical and engineering sciences (Hood & Wilson, 2001). *Informetrics* is a wide term that includes all the metrics as it focuses on the quantitative study of every kind of information in any form (ibid). The subtitle of a book by Egghe and Rousseau gives us a very concise, but at the same time complete, definition of the term: *Informetrics: Quantitative Methods in Library, Documentation and Information Science* (Egghe & Rousseau, 1990). *Webometrics* (or Netometrics) refers to the application of bibliometric techniques and concepts to documents in the web, which means that Webometrics overlaps with both Bibliometric and Informetrics. Sometimes the term *Cybermetrics* is used as synonym for Webometrics but according to the pioneers of this branch, namely Björneborn and Ingwersen, the latter is defined as the study of the quantitative aspects of the construction and use of information resources, structures and technologies by means of informetric approaches, instead the former refers to the whole Internet. So, for example, Cybermetrics includes the statistical study of discussion groups or mailing lists. Below a representation of the relationships among the aforementioned metrics fields (Björneborn & Ingwersen, 2004):

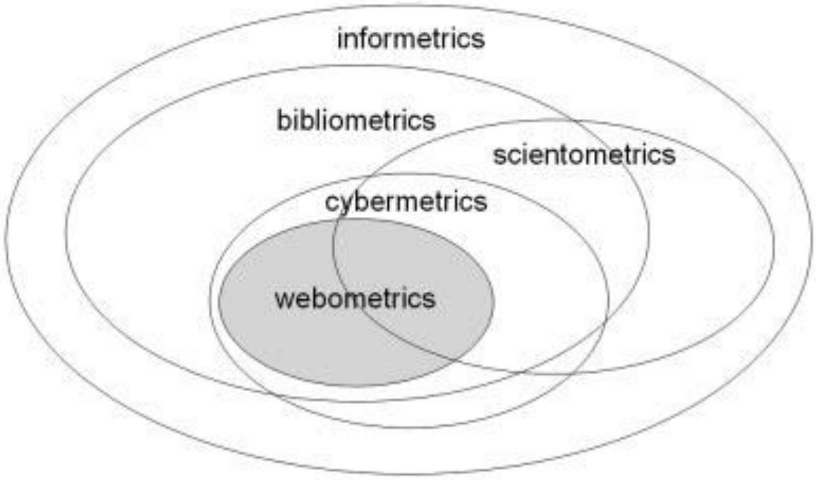


Figure 2: Relationships among the fields of infor-/biblio-/sciento-/cyber-/webo-/metrics. Sizes of the overlapping ellipses are made for sake of clarity only. Source: Björneborn & Ingwersen, 2004, p. 1217.

Finally, *Librametrics* (or Librametrics) refers to the application of mathematical and statistical techniques to library problems. The term was proposed by Ranganathan in the late 1940s (Ranganathan, 1947) and according to Wilson (1999) it can be used to indicate those studies not specifically analysing literature, or at least not specifically linked to the goals of Bibliometrics and Information Retrieval. So, for example, it refers to analyses of books circulation, library acquisitions or shelf allocation. According to the results of Table 1, this is one of the least used among the metric terms.

1.2. FROM BIBLIOMETRICS TO WEBOMETRICS

The term Bibliometrics was introduced in 1969 by Pritchard, who proposed it to indicate that discipline concerning "[...] the application of mathematical and statistical methods to books and other media of communication" (Pritchard, 1969, p. 349). Despite the fact that Bibliometry as autonomous discipline is relatively young,[2] it has age-old roots: it is thought that in the biggest Alexandria library all rolls in it had already been recorded in the third century B.C. (Broadus, 1987). In the second half of 1800 there were the first attempts to measure books usage patterns. So, for example, it turned out that in Italy in 1863 the users of public libraries required 183,528 Mathematics and Natural Sciences volumes and 122,496 books about Literature, History and Philology (Barr, McMullen, & Leach, 1983). The first organised citation study was conducted in 1848 by Jewett with the scope of determining the amount of referenced items that could be obtained in American libraries with reference to the fields of International Law and Chemistry (Jewett, 1849).

In 1923 there was the first attempt to define the bibliometric activity by Hulme, who called it *Statistical Bibliography* (Hertzel, 2003). The librarian, who worked 25 years at the Patent Office of

[2] According to Broadus, in 1987 about 97% of all bibliometrists were probably still alive (Broadus, 1987).

London, suggested the possibility to study relationships among numbers of scientific papers, numbers of patents, amounts of exports and other quantities. But Hulme's idea was harshly criticised at the time, and it was not well received (Broadus, 1987). Another important stage for the development of citational studies was the study by Gross and Gross dated 1927 (Gross & Gross, 1927). The work had the purpose of detecting those periodicals of Chemistry that were worth to be purchased by small college libraries. Furthermore, in 1944 there was the attempt by Rider to quantify the American libraries growth (Rider, 1944).[3]

The potential of bibliometric techniques in measuring Science was recognised only in 1963 when there was the publication of the oeuvre by Derek John de Solla Price *Little Science, Big Science* (Price, 1963), which will decree him the founding father of *Scientometrics*. The term, officially introduced in 1969 and developed in Russia by Nalimov e Mulchenko (1969), refers to the application of those quantitative methods dealing with the analysis of Science viewed as an information process: at the core of Scientometrics there is the study of different aspects of scientific literature (see Section 1.1.). The foundation of the journal *Scientometrics* in 1978 constituted a big contribute to the development of the term (Hood & Wilson, 2001).

The year 1969 was the one in which also the term Bibliometrics, besides the Scientometrics one, was introduced: Pritchard, inspired by Hulme, proposed it as the label for the discipline previously called by the latter *Statistical Bibliography*. According to Pritchard, that term did not properly represent the rising discipline. In the article "Statistical Bibliography or Bibliometrics" (1969) Pritchard claimed: "[...] the term statistical bibliography is clumsy, not very descriptive, and can be confused with statistics itself or bibliographies on statistics" (Pritchard, 1969, p. 348). Later, he will specify that Bibliometrics is the metrology of the information transfer process and that its purpose is to analyse and control it (Pritchard, 1972). A similar definition is given by the *British Standard Glossary of Documentation Terms*, where it is defined as: "[...] the study of the use of documents and patterns of publication in which mathematical and statistical methods have been applied" (Hertzel, 2003, p. 295).

The first official acknowledgement of the paternity of the term to Pritchard was given by a close friend of him, Robert Fairthorne, who employed the word Bibliometrics in an article published in the same year, in the same journal and in the same volume of the friend's one (Hertzel, 2003). Notwithstanding the majority of bibliometrists see in Pritchard the one who coined the term Bibliometrics, there are some who stress its French origins. In 1973 Fonseca, in a criticism about the tendency of English-language authors to ignore works in Romance languages, stressed that the use of the French equivalent of the term, "bibliometrie" was used in the work by Paul Otlet *Traitée de Documentation. Le livre sur le Livre. Théorie et Pratique* published in 1934, where there was a section entitled "Le Livre et la Mesure. Bibliometrie". Despite that, Pritchard is the one who first developed a definition of the discipline (Hood & Wilson, 2001).

[3] The verb "attempt" is used here as author's findings referring to the fact that "[...] doubling every sixteen years has continued now for over three centuries" (Rider, 1944, p. 13) were criticised by Molyneaux (1986), who doubted the reliability of the work (Rider, 1944).

With the passage of time, the boundaries of Scientometrics and Bibliometrics almost disappeared with the result that they are often used as synonyms. As previously mentioned, where Science measurements are based on bibliometric data (e.g. citations or publications counts) the two disciplines are indistinguishable (Tague-Sutcliffe, 1992).

Besides the two above-mentioned metrics, there is another one whose meaning is wider, actually it includes them: it is *Informetrics*. It refers to the study of all quantitative aspects of information, in any form (not only written or bibliometric information) and in any social group (not only scientists). Put differently, it includes all metrics, as each of them analyse a certain type of information. The term has German origins. It derives from the word *Informetrie* suggested by Nacke in 1979 to identify that portion of Information Science committed in measuring information by means of mathematical and statistical tools. In 1984, the Russian agency for the technical-scientific information (Всероссийский Инстиут Научной и Технической Информации)[4] (VINITI)[5] established the Fédération Internationale de la Documentation (FID) under Nacke's chairmanship; in its official documentation Informetrics was used as a generic term for both Bibliometrics and Scientometrics. The official adoption of the term was decreed by its usage in the VINITI monograph by Gorkova (1988) entitled *Informetriya* (the Russian term for Informetrics) (Hood & Wilson, 2001).

In the 1990s there was the introduction of further metric terms as the application of bibliometric techniques to electronic information on the World Wide Web gave birth to a new literature: in 1995 Bossy introduced the term *Netometrics* to describe scientific interactions mediated by Internet; in 1997 Almind and Ingwersen suggested Webometrics for the study of the World Wide Web as well as all network-based communications by informetric methods (Almind & Ingwersen, 1997). In 1997 another term was suggested, that is to say *Cybermetrics*: the homonymous journal covers researches in Scientometrics, Informetrics and Bibliometrics but with special emphasis on their interrelations with the Internet, on the evaluation of electronic journals on the Web, and on the application of informetric techniques to cyberspace communication in general (Wilson, 1999).

1.3. THE STRUTURE OF BIBLIOMETRICS

Bibliometrics can be defined in different ways and both its scopes and structure change depending on scholars' point of view.

Nicholas and Ritchie affirm that the purpose of Bibliometrics is to inform about the structure of knowledge and the way in which it is communicated. They consider the discipline as structured in "descriptive" (describing literature characteristics) and "behavioural" (studying relationships among literature components) bibliometrics (Nicholas & Ritchie, 1978). O'Connor and Voos include in Bibliometrics scopes the study of relationships within literature (e.g. citation studies) and divides

[4] All-Union Institute for Scientific and Technical Information.
[5] Acronym for the transliteration "Vsesojuznyi Institut Naucnoj I Tehnic'eskoi Informac'ii".

Bibliometrics in “descriptive” (description about patterns involving authors, monographs, journals, subject or language) and “relational” (O'Connor & Voos, 1981). Stevens stresses the quantitative aspect of Bibliometrics and splits it into “descriptive” and “evaluative”: the first includes all productivity counts; the second refers to the study of literature usage through reference counts (Stevens, 1953).[6] Finally, Marshakova distinguishes between “plain” (based on the analysis of the dynamics of separate features) and “structural” bibliometrics (focused on the study of correlation between objects, their clustering and classification) (Marshakova, 1996).

Bibliometrics can also be divided according who employs it. Thus, 3 fields of specialisation can be distinguished:

- Bibliometrics for bibliometrists (methodology);
- Bibliometrics for scientific and social-historical disciplines (scientific information): this field can be considered an extension of Scientometrics and its boundaries overlap with those of Information Retrieval;
- Bibliometrics for science policy and management (science policy): it is focused on research evaluation (Glänzel, 2003).

As previously mentioned, Bibliometrics consists of the application of statistical-mathematical tools to books and other means of communication (Pritchard, 1969); however, its applications mainly refer to scientific journals and scientific papers (Glänzel, 2003). But why is it so? In order to answer this question we must go back in time and find out those mechanisms at the basis of the birth and development of scientific papers. Scientific articles were born in 1600 to solve an information crisis, as the amount of scientific books published up to that moment was huge. As Barnaby Rich claimed in 1613:

> […] one of the diseases of this age is the multiplicity of books; they doth so overcharge the world that it is not able to digest the abundance of idle matter that is every day hatched and brought forth into the world (Barnaby cit. in Price, 1963, p. 63).

Therefore, the function of periodicals was to “digest” books.

The first scientific journal was founded in 1665 by the Royal Society of London (founded in 1660) and was called *Philosophical Transactions of the Royal Society*. Thanks to scientific journals “[...] the casual reader might inform himself without the network of personal correspondence, private rumor, and browsing in Europe's bookstores, formerly essential” (Price, 1963, p. 63). However, at the beginning scholars were not so happy about writing short articles instead of books. According to Bernard Barber, this reaction should be considered as a natural one: resistance to innovations is part of an innate

[6] I think it is plausible that Stevens' classification lacks the relational aspect as it precedes the conceptualisation of citations as intellectual links developed after the 1960s, that is to say after the introduction of Garfield's Science Citation Index.

conservatism of scientists that is the counterpart of both the open-minded creativity and the edge of objectivity characterising Science (Barber, 1961). This resistance lasted for a long time and the scientific paper, as we know it today, will assert itself only around 1850, when authors started to explicitly report references to previous works, attitude fundamental to the accumulation of knowledge (Price, 1963). The importance that this practice gained in that period is witnessed by the reaction of Ulrich von Wilamowitz (one of the founding fathers of German classical philology) when Nietzsche published in 1872 *The birth of tragedy* where, even if passages by authors such as Schopenhauer and Wagner were included, there were no bibliographic references. Wilamowitz judged this behaviour as non scientific, violating those rules regulating historical research. He wished about the expulsion of Nietzsche from the academic teaching, which happened seven years later. From then on German, Italian, English and American philologists have paid attention in including in their works accurate references, often unnecessary, to Wilamowitz (De Bellis, 2009).

But why did scientific papers gain more and more importance becoming the basic unit of Bibliometrics? Robert Merton, during his studies on multiple and independent discoveries conducted in the 1950s, found that even if this phenomenon showed a decreasing pattern, it was still an actual one. Therefore, papers did not perfectly fulfil their communicative role meaning that there had to be another function they accomplished very well ensuring their survival; according to Merton this function is the establishment and maintenance of intellectual property (Merton, 1957; Merton, 1961). As Price underlined:

> The prime object of the scientist is not, after all, the publication of scientific papers. Further, the paper is not for him purely and simply a means of communication knowledge. […] If, then, the prototype of the modern scientific paper is a social device rather than a technique for cumulating quanta of information, what strong force called it into being and kept it alive? Beyond a doubt, the motive was the establishment and maintenance of intellectual property. [...] the never-gentle art of establishing priority claims. […] For these reasons scientists have a strong urge to write papers but only a relatively mild one to read them (Price, 1963, pp. 62,65,70).

Scientific papers published in refereed scientific journals demonstrated to be the most suitable unit for bibliometric studies:

> Obviously citation indexes will be effective only to the extent that the bibliographies in published papers are accurate reflections of the earlier literature. In evaluating papers submitted to journals, referees should determine whether all pertinent references have been provided (Garfield, 1964, p. 654).

But the scientific paper is not the unique element on which bibliometric studies focus, as there are also authors, publications and bibliographic references to which basic measures (natural count measures such as number of citations or papers) and bibliometric indicators (statistic functions defined on sets of bibliometric elements and units) refer. The units of analysis can be reviews, subjects, institutions or Countries (Glänzel, 2003).

1.4. THE 3 BIBLIOMETRIC LAWS

In the first half of the 20^{th} century, before Bibliometrics was defined by Pritchard in 1969, 3 bibliometric laws were developed: Lotka's law on scientific productivity (Lotka, 1926); the law by Bradford on the distribution of articles over journals (Bradford, 1934); the Zipf law on words distribution in a text (Zipf, 1949). All of them have been revisited and criticised over time but despite that their centrality in the discipline is still unquestioned (Hertzel, 2003).

Alfred Lotka is considered the founding father of Demography together with Vito Volterra (even if their contributions were developed independently). In 1909 he worked at the Patent office of London where, supposedly, he knew Hulme, who in 1923 made the first attempt to define the bibliometric activity suggesting the name *Statistical Bibliography* (Broadus, 1987). Lotka’s studies led him to develop a law that became a pillar of Bibliometrics, that is to say the homonymous *Lotka's law*. In 1926 Lotka published a pioneering study on frequency distributions of scientific productivity in Chemistry. The work was based on an index referring to publications in both *Chemical Abstract* from 1907 to 1916 and in *Geschichtsafeln der Physik* from the first publication to 1900. Results led him to claim that the number of authors active in a certain research area producing *n* contributions is about $1/n^2$ of those contributing the minimum (n=1) and the proportion of those producing 1 is about 60%:

$$p(n) = K / n^2 \tag{1}$$

where *p* is the number of scholars producing *n* contributions and *K* is a constant that refers to the disciplinary field of reference. So, if 60% of authors in a discipline area produces only one contribute, those contributing two will be 15% ($1/2^2*60$), those contributing three will be 7% ($1/3^2*60$) and so on (Lotka, 1926). This law is also called “the inverse square law of scientific productivity”, as it predicts that the number of authors making two contributions is about one-quarter of those making one, the number of making three is about a ninth and so on (Potter, 1981b).

Lotka's law was ignored for a long time before becoming famous and being appreciated by many scholars in the 1970s. Actually, according to some scholars this Law would not deserve the success it had as not only it would not be valid for many scientific fields but also for the area that constituted the object of Zip’s study. [7] It also seems that the curve used by Lotka to describe observed frequencies does not fit

[7] For a list of publications referring to this argument see De Bellis (2009).

completely the data (Potter, 1981b). Furthermore, in 1982 the MacRoberts examined the database used by Lotka and questioned the validity of the generalisation drawn from it (MacRoberts & MacRoberts, 1982). Nevertheless, this law is still considered a rule of thumb in Bibliometrics (Glänzel, 2003).

Bradford's law deals with the frequency distributions of papers over journals. Samuel Bradford was a librarian at the Science Museum of London. He dreamt about the creation of a universal classification system about technological and scientific productions. Such a project was not a simple one, as indexing and abstracting systems of the time were not able to store the entire scientific production of any disciplinary area (Broadus, 1987). It was the will of understanding this limit that led Bradford to formulate the law according to which:

> [...] if scientific journals are arranged in order of decreasing productivity on a given subject, they may be divided into a nucleus of journals more particularly devoted to the subject and several groups or zones containing the same number of articles as the nucleus when the numbers of periodicals in the nucleus and the succeeding zones will be as $1:b:b^2$ [...] (Bradford, 1934, p. 86).

Thanks to this law, we know that if one wants a complete bibliographic report of a certain disciplinary field or research area she/he must refer to an exponentially growing number of journals than the nucleus ones. Thus, it is necessary to include journals related to contiguous fields to the one considered. That is why specialised libraries cannot gather the complete literature of a certain field (Bradford, 1934).

As Lotka’s law, Bradford's law was criticised but this did not affect its importance for Bibliometrics. The Bradford law found in Garfield a big supporter: he expanded its validity to all scientific fields. In 1979 Garfield developed the *Concentration Law*, generalising the one by Bradford, with the scope of demolishing what he calls the “journal myth”, that is to say a wrong perception about the fact that researches must manage a huge amount of journals in doing their job (Garfield, 1991). Garfield’s law, in fact, affirms that “the tail of the literature of one discipline consists, in a large part, of the cores of the literature of other disciplines” (Garfield, 1979, p. 23). The two laws (Bradford's law and the Concentration law) have always played an important role in the policy of the Thomson Reuters (formerly the Institute for Scientific Information founded in 1960 by Garfield) as it relies on them to legitimate its selection criteria:

> It would appear that, in order to be comprehensive, an index to scientific journal literature might be expected to cover all the scientific journals published. This approach would be not only impractical economically, but as analyses of the scientific literature have shown, unnecessary. It has been demonstrated that a relatively small number of journals publish the bulk of significant scientific results. This principle is often referred to as Bradford's Law (Thomson Reuters).

Garfield (who in 1960 developed the first Citation Index) appeals to these laws to demonstrate that the journals indexed in the *Science Citation Index* and *Social Science Citation Index* cover 50% of all publications and 75% of cited documents (Garfield, 1996a, 1990). Garfield is absolutely sure of that:

> Experience has taught me it is essential that a simple message be repeated regularly: a small number of journals accounts for the bulk of significant scientific results (Garfield. 1996b, p. 13).

George Kingsley Zipf was a German teacher at Harvard University and thanks to his law he is also considered a bibliometrician (Broadus, 1987). According to Zipf, language can be studied as a natural phenomenon, that is to say through the application of Statistics to objective linguistic phenomena (Zipf, 1935). The so called *Zipf's law* states that by multiplying words position in a ranking (where words are classified in descending frequency order) for the appearance frequency we will obtain a constant:

$$r * f = k \tag{2}$$

where r is the word rank in the classification, f the appearance frequency of a word, and k is the constant (Zipf, 1949). Among the reinterpretation and generalisation attempts to which the law has been subject there is the one by Benoit Mandelbrot, who is considered the founding father of Fractal Geometry. By employing Shannon's information theory (Shannon, 1948),[8] he tried to explain the relation between words rank and frequencies (*Zipf-Mandelbrot's law*) (Mandelbrot, 1965).

Over time the 3 laws have become more and more linked among each other as their common origins were established in the s*uccess breeds success* (SBS) principle linked to the *cumulative advantage processes* (Price, 1976). The SBS principle applied to Bibliometrics is formulated as follows:

> Items enter the system one by one and […] *are considered* different ways in which a new item can be attributed to a source. When item number t + 1 enters the system, (a) there is a fixed chance, $\alpha \in]0, 1[$ that this item will be produced by a new source, i.e., a source not producing one of the first t items; (b) in case an already active source produces this new item, there is a chance proportional to $nP(t, n)$ that this item is produced by a source that has already n items (here $P(t, n)$ denotes the fraction of sources with n items in a system where the sum of all items is t) (Egghe & Rousseau, 1995, p. 426, italics mine).

[8] According to Meadows (1990), this work, which was based on the three laws discussed in this Section, was fundamental for the rapid development of Bibliometrics in the 1950s.

According to some authors (see, for example, Fairthorne, 1969; Price, 1976; Brookes, 1977), this universal principle can be transposed in Mathematics in a set of functions that are able to describe many and different bibliographic and social phenomena. As Price explains:

> A Cumulative Advantage Distribution is proposed which models statistically the situation in which success breeds success. It is shown that such a stochastic law is governed by the Beta Function, containing only one free parameter, and this is approximated by a skew or hyperbolic distribution of the type that is widespread in bibliometrics and diverse social science phenomena. In particular, this is shown to be an appropriate underlying probabilistic theory for the Bradford Law, the Lotka Law, the Pareto and Zipf Distributions, and for all the empirical results of citation frequency analysis (Price, 1976, p. 292).

1.5. GARFIELD'S SCIENCE CITATION INDEX[9]

Bibliometrics made big step forwards thanks to Eugene Garfield and his Science Citation Index introduced in 1960. Garfield's efforts in classifying and indexing scientific literature go back to 1951, when he participated in the *Johns Hopkins University's Welch Medical Library Indexing Project* promoted by the National Library of Medicine, whose scope was that of establishing if and how computers could be employed to improve biomedical literature indexing and searching. The experience was precious for Garfield, who soon developed the *Current Contents* by the Eugene Garfield Associates (currently the *Current Contents Connect* by the Thomson Reuters), a database of just published (often not yet published) contents, bibliographic information and abstracts from scientific journals (De Bellis, 2009).

Garfield had multiple inspirations. He was inspired by John D. Bernal and H. G Wells. Bernal, in *The social function of Science* (1939),[10] claimed the necessity of a scientific revolution that should have been realised in two steps: abolishment of scientific periodicals; creation of a centralised system to gather, store, coordinate and distribute scientific information (Bernal, 1939). If the first wish would have never been achieved, the second one would have become the commercial mission of the Institute for Scientific Information (ISI). In 1964 Bernal became an ISI consultant (De Bellis, 2009). With reference to Wells, Garfield was inspired by his "world brain" idea (Wells, 1938) about a new, free, synthetic, permanent world encyclopaedia: "[...] a complete Science Citation Index, covering both current and old literature, […] a true encyclopaedia of unified science" (Garfield, 1960b, p. 5).[11]

---

[9] Except when otherwise indicated the term Science Citation Index also stands for the Social Science Citation Index and the Arts and Humanities Citation Index. All of them are published by the Thomson Reuters.

[10] During the Second World War, and also in the post-war period, it was considered the Bible of the so called "red science", the left British and international Science (De Bellis, 2009).

[11] Influences by both Bernal and Wells are also clearly present in the article "Science Citation Index-A new dimension in indexing" (Garfield, 1964).

Furthermore, Garfield was inspired by the project of the Vienna Circle about the creation of an *International Encyclopaedia of Unified Science*. The original and more ambitious project of Garfield was the creation of a *Unified Science Index*:

> The primary purpose of this paper is to discuss a plan for accomplishing what Neurath calls "an encyclopaedic integration of scientific statements", what I call a "Unified Index to Science" (Garfield, 1959, p. 461).

Actually, Garfield's inspiration can be traced further back in time: the inspiration came from the so called *Shepard Citation*. It was an apparatus of legal consultation used by American lawyers since 1875 and developed by Frank Shepard, reporting all the times a case was cited in successive cases and if and which modifications were made to it. As the former vice-president of the company Shepard's Citations Inc. explained:

> The lawyer briefing a case must cite authorities to back up his arguments. So must the court in writing its opinions. This is because of the doctrine of "Stare Decisis" which means that all courts must follow precedents laid down by higher courts and each court generally also follows its own precedents. [...] The lawyer, however, must make sure that his authorities are still good law, that is, that the case has not been overruled, reversed limited or distinguished in someway that makes it no longer useful as a valid authority. Here is where the use of Shepard's Citations comes in. (Adair, 1955, p. 31).

In 1954 Garfield, after browsing through the Shepard's Citation at a public library, wrote the paper "Shepardizing the scientific literature" (Garfield, 1954).

Garfield's first article about the Science Citation Index was published in 1955. In it he claimed:

> This paper considers the possible utility of a citation index that offers a new approach to subject control of the literature of science, […] it tends to bring together material that would never be collated by the usual subject indexing. […] What seems to be needed, then, in addition to better and more comprehensive indexes, alphabetical and classified, are new types of bibliographic tools that can help to span the gap between the subject approach of […] authors and the subject approach of the scientist who seek information. […] The utility of a citation index in any field must also be considered from the point of view of the transmission of ideas. A thorough scientist cannot be satisfied merely with searching the literature through indexes and bibliographies if he is going to establish the history of an idea (Garfield, 1955, pp. 108-110).

Garfield's ideas were completely new, they were about a revolutionary indexing concept, but scientists needed time to really appreciate it. Nevertheless, someone found Garfield's ideas interesting and helped and supported him in developing his project: the geneticists Lederberg and Allen. The latter also thought about possible applications of such indexes and in 1960 he sent a diagram to Garfield showing citation connections among a set of biochemistry articles:

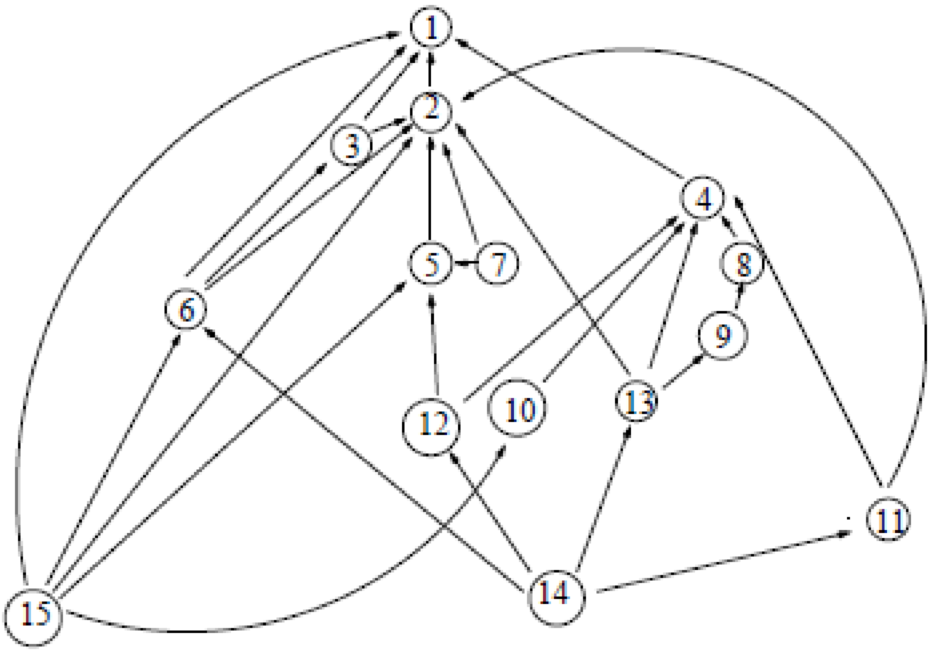


Figure 3: Gordon Allen's citation network as depicted in Garfield (1960a). The circled numbers represent published articles. The arrows indicate citing relations, pointing from the citing to the cited document. Source: Wouters, 1999a, p. 54.

Garfield replied as follows:

> The material you sent was fabulous. Why didn't we think to do this before? I didn't have this in mind when I said I had some examples of the power of the Citation Index. I merely meant specific articles which could be traced through a CI. I once had the idea that some type of network theory could be used with Citation Indexes. I am now convinced more than ever, from your example, that his will be true (Garfield cit. in Wouters, 1999a, p. 51).

The *Glossary of Thomson Scientific terminology* defines the citation index as follows:

> A citation index is a bibliographic tool in print or electronic format that lists all referenced or cited source items published in a given time span. The tool is a useful method for tracking the historical development - backwards and forwards in time - of an idea or given topic within the literature published in a wide selection of journal titles. What distinguishes it from other

> indexes is that it includes all the cited references (footnotes or bibliographies) published with each article it covers (Thomson Reuters).

According to Garfield:

> A citation index is an ordered list of cited articles each of which is accompanied by a list of citing articles. The citing article is identified by a source citation, the cited article by a reference citation. The index is arranged by reference citations. Any source citation may subsequently become a reference citation (Garfield, 1964, p. 652).

Almost immediately after its first publication, the SCI data started to be used in citation analysis; however, it was not welcomed by everyone. Generally speaking, while scientists were in favour of it, librarians were more cautious (Martyn, 1965). The reason of that stands in their different points of view:

> […] the librarian is concerned with information retrieval, whereas the scientist is more interested in information access; regarded as a retrieval tool, the Citation Index is not as efficient as some more conventional approaches to the literature, but as an access tool if functions very well (Martyn, 1965, p. 189).

Furthermore, scientists saw in the Index a tool that could simplify the complex condition of Science: as witnessed by Price's oeuvres *Science since Babylon* (1961/1975) and *Little Science, Big Science* (1963), that period was characterised by a big, uncontrolled growth of Science and scientific information needed to be organised (Leydesdorff, 2005b).

Garfield constructed his SCI according to the Bradford law (1934) and his Concentration law (1979). So, periodicals indexed are those having high citation scores, feature that should have assured their "specificity". It is interesting to note that the SCI is a self-referential system, as the proof of the fact that it encompasses all significant and important journals is given by the same citation indicators used for selection:

> The list of most frequently cited journals shows that the SCI has been remarkably successful in covering all "significant" and "important" journals, insofar as citation counts can be considered a reliable measure of "importance" and "significance" (Garfield, 1970, p. 670).

The citational parameter used in the selection is the Impact Factor, developed by Garfield in the 1960s. It measures the frequency with which the "average article" of a journal is cited, in a given year, by the articles published in the two previous years on a set of journals selected by ISI for indexation (see Section 1.6.2.). The *Journal Citation Reports* (JCR), reporting the IF, is published every year including a Science

Edition (covering over 6,500 journals) and a Social Sciences one (covering over 1,900 periodicals)[12] (Thomson Reuters). The selection of the journals to include is also based on the following parameters: timeliness of publication; full text in English (or bibliographic information in English); cited references in the Roman alphabet; application of the peer review process. In the Web of Science all indexes, except the Arts & Humanities one, which also processes implicit references, refer only to identifiable footnotes and endnotes (ibid).

But what do citations represent according to Garfield? And why should his indexes be useful?[13] His ideas about citations meaning are based on Robert Merton's concept of citation as "intellectual transaction" (Merton, 1979), and Henry Small's conceptualisation of cited documents as "concept-symbols" (Small, 1978).[14] So:

> [...] citations symbolize the conceptual association of scientific ideas [...] By the references they cite in their papers, authors make explicit linkages between their current research and prior work in the archive of scientific literature. […] *These* explicit references imply that an author has found useful a particular published theory, method, or other findings (Garfield, 1994, p. 1, italics mine).

The *Glossary of Thomson Scientific terminology* defines citation counts as:

> [...] formal acknowledgement of intellectual debt to earlier patents and previously-published scientific research papers. They are an important indicator of how new patents are linked to earlier patents and scientific papers. Citation counts are recognized as units of influence or impact on technological development […] (Thomson Reuters).

Even if bibliometricians are aware of the fact that authors cite for many and different reasons (see, for example, Cronin, 1984; Bornmann & Daniel, 2008; MacRoberts & MacRoberts, 1986; Oppenheim & Renn, 1978; Van Raan, 1998; Weinstock, 1971), the concepts expressed in the passages above constitute the basic assumptions of the *citation culture* to which Garfield gave birth, and from which evaluative and relational bibliometrics (establishing relations between units - authors, disciplines, institutions - with the scope of mapping scientific literature) developed.

---

[12] The number of journals increases each year.
[13] The topic related to the meaning of citations is deepened in Chapter 3.
[14] See Sections 3.1.1. and 3.1.3.

1.6. BIBLIOMETRIC MEASURES

As foreseen by Garfield, the SCI demonstrated to be a powerful tool for scientific quantifications; on its basis many and different bibliometric indicators were developed. Performance indicators can be classified as *productivity measures* (relating to the production of papers, or cited papers, by authors or institutions in a given period of time), *impact metrics* (based on the quantification of citation received in a given period of time by different units such as papers, journals or authors) and *hybrid metrics* (aiming to capture at the same time productivity and impact, such as the *h*-index) (Franceschet, 2009). In this Section, with the purpose of illustrating the different evaluative and descriptive scopes for which these indicators can be used, some of the most used measures are introduced: the Price Index; the Journal Impact Factor; the Eigenfactor metrics; the Immediacy Index; the Cited Half-Life; the H-Index.

1.6.1. THE PRICE INDEX

Price (1970) noted that Science was different from other fields with respect to the way in which scientists refer to the literature. More specifically, he found that in hard Science (e.g. Natural Sciences) recent articles were much more likely to be cited than in Social Sciences or Humanities (soft Science). On the basis of these reflections, he developed an indicator informing on citing behaviour. This is given by:

$$PI = (n_1 / n_2) * 100 \tag{3}$$

where, for a given paper, n_1 is the number of cited reference aged less than six years and n_2 is the total number of references (Price, 1970). It is a measure of the recency of the literature cited by a given article or journal; Price called this phenomenon "the immediacy effect".

According to Price, this measure is useful to distinguish among hard science, soft science, technology and non-science:

> Perhaps the most important finding I have to offer is that the hierarchy of Price's index seems to correspond with what we intuit as hard science, soft science, and non-science as we descend the scale. Biochemistry and physics are at the top, with indexes of 60 to 70 percent, the Social Sciences cluster around 42 percent, and the humanities fall in the range 10 to 30 percent (Price, 1970, p. 4)

In 1989 Moed revised the Price Index and proposed an improvement by calculating the average index value across a population of articles excluding the sub-populations with values of 0% or 100%. Thus, while Price's measure is a global one, Moed's measure is a "corrected" one (Moed, 1989).

1.6.2. THE JOURNAL IMPACT FACTOR

The Journal Impact Factor (JIF) is one of the most used bibliometric measures. It was developed by Garfield and Irving Sher (1963) to ease the selection of journals to include in the Science Citation Index (Garfield, 2005). It measures the frequency with which the "average article" published in a periodical is cited in a year or in a given period. It is annually published in the *Journal Citation Reports* (JCR) by Thomson Reuters. The IF of the journal *J* in the year *n* is given by:

$$IF_n(J) = c_n / p_{n-1} + p_{n-2} \tag{4}$$

where c_n is the number of citations received in the year *n* by papers published in the journal *J* in the years *n*-1 and *n*-2 and $p_{n-1} + p_{n-2}$ is the total number of source items published in the journal *J* in these two years (*n*-1 and *n*-2) (Glänzel, 2003).

As Garfield explained in 1955, when for the first time he talked about the impact factor idea, the measure was created with the scope of comparing scientific journals with respect to their probability of being cited:

> Citation frequency is, of course, a function of many variables besides scientific merit. [...] Citation frequency of a journal is thus a function not only of the scientific significance of the material it publishes [...] but also of the amount of material it publishes (Garfield, 1955).

As the Thomson Reuters warns, a careful use of impact data is essential and several caveats must be considered. So, the JIF should be used paying attention to those phenomena influencing citation rates, [15] and reference to informed peer review is advised. Because citation norms are field-dependent, comparison of citation counts across fields should be avoided:

> Thomson Reuters stresses that a journal's impact factor is a meaningful indicator only when considered in the context of similar journals covering a single field of investigation or subject discipline (Thomson Reuters).

In particular, the main flaws of this measure are the following: lack of normalisation for reference practice and traditions in the different fields and disciplines (Pinski & Narin, 1976)[16]; information about the citing journals are not taken into account (Tomer, 1986); journals with long papers (review journals) tend to have higher impact factor (Pinski & Narin, 1976); the fact that citation frequencies are subject to age bias is not taken into account (Asai, 1981; Glänzel & Schoepflin, 1995; Moed et al., 1998; Rousseau,

---

[15] See Section 3.2.

[16] In 2010 the Source Normalized Impact per Paper (SNIP) was introduced by Moed (2011). The idea behind it is that the actual citation rate of a set of target papers in a subject field should be divided by a measure indicating the frequency with which articles in that field cite other documents.

1988); some studies underlined that the two-year period used to calculate the JIF is not suitable for all disciplines (Garfield 1986; Glänzel & Schoepflin 1995; Moed et al. 1998); [17] due to the fact that there is not an adequate operationalisation of the concept of "citable document" for the calculation of the impact factor, JIFs published in ISI's Journal Citation Reports cannot be considered reliable measures with reference to some journals (Moed, 2005; Moed & Van Leeuwen. 1995, 1996; Moed, Van Leeuwen, & Reedijk, 1999). Specifically, the problem refers to the fact that some journals, such as the *Lancet*, contain documents that do not fit into the category "citable document" (usually referring to article, notes and reviews), which is the category used to determine the denominator in the JIF. Thus, while the numerator refers to the total number of citations received by the journal in a period of time, the denominator refers only to some documents (citable documents). In this way there is a certain amount of citations that are "for free" (Moed & Van Leeuwen, 1996). Clearly, as showed by Moed et al. (1999), this leads to distortions in the IF of journals, such as the Lancet, which have a large number of letters and editorials:

Table 2: ISI impact factor is inaccurate: the Lancet 1992.

Type of Article	**Publ.**	**Citat.**	**C/P**
Articles	784	7134	9.1
Notes	144	593	4.1
Reviews	29	232	8.0
SUBTOTAL	957(a)	7959(b)	8.3
Letters	4181	4264	1.0
Editorials	1313	905	0.7
Other	1421	909	0.6
TOTAL	7872	14037(c)	1.8

ISI JCR Impact Factor: (c)/(a) = 14.7

Correct Impact Factor: (b)/(a) = 8.3

Source: Moed et al., 1999, p. 577.

1.6.3. EIGENFACTOR METRICS

In order to solve some of the above-mentioned shortcomings and limits of the JIF, new measures have been developed: the so called *Eigenfactor* Metrics. These measures are based on data retrieved from the JCR (Journal Citation Reports) and, with the purpose of measuring the influence in terms of citations, they not only consider journals citation counts but also the structure of citation networks as a whole. These metrics are available from year 2007.

[17] With the purpose of avoiding such potential distortions, a five-year period IF has been introduced (see Section 1.6.3.).

Among the Eigenfactor Metrics we can find the *Eigenfactor Score* and the *Article Influence Score.* The former differentiates from the JIF with respect to two factors: it is based on a five-year period (the JIF is based on a two-year period); while the Impact Factor weighs each citation to a journal equally, the Eigenfactor Score assigns a greater weight to those citations coming from influential journals, allowing these journals to exert greater influence in the determination of the rank of any journal citing them. The Eigenfactor Score does not count journal self-citations. The sum of Eigenfactor scores for all journals is 100; each journals Eigenfactor Score is a percentage of this total. The *Article Influence Score* measures the relative importance of the journal on a per-article basis. It is the journal Eigenfactor Score divided by the fraction of articles published by the journal. That fraction is normalised, so the total sum of articles from all journals is 1. The mean Article Influence Score is 1.00. A score greater than 1.00 indicates that articles in that journal have an above-average influence; a score less than 1.00 indicates that the articles in the journal have a below average influence (Thomson Reuters).

1.6.4. THE IMMEDIACY INDEX

This measure is strictly linked to the previous one and refers to the amount of citations received by a journal in a given publication year:

$$II_n(J) = c_n / p_n \tag{5}$$

where c_n is the number of citations received by a journal J in a year n and p_n is the number of papers published in that given year n. It can be considered as an estimation of the speed with which the journal "average article" is cited.[18] The Immediacy Factor is included in the annual *Journal Citation Reports* and this calculation is useful in determining those journals publishing in emerging research areas (Thomson Reuters).

1.6.5. THE JOURNAL HALF-LIFE MEASURES

As defined by the *Glossary of Thomson Scientific Terminology*, the *Cited Half-Life* quantifies the age of cited articles by showing the number of years, back from the current one, that account for 50% of the total citations to a journal in the current year. It can be interpreted as the average age of cited articles. Similarly, *the Citing Half-Life* identifies the number of years from the current one that account for 50% of the cited references from articles published by a journal in the current year. Both of them can be considered obsolescence indicators. The Half-Life measures appear only in the JCR by Thomson Reuters (Thomson Reuters).

---

[18] E.g.: the Immediacy Index of a journal *J*, which in 2007 published 50 articles and was cited 10 times, is given by 10/50=0,2.

1.6.6. THE CROWN INDICATOR

The *crown indicator* was developed with the purpose of furnishing a normalised indicator of research performance. It was introduced by De Bruin, Kint, Luwel, and Moed (1993). Specifically, as the average number of citations received by a publication differs according to the field, this indicator corresponds to the number of citations to publications from a specific unit during a given time span, compared to the world average of citations to publications of the same document types, ages and subject areas, seen as a group. The normalisation of citation values is done on the sums of the citations and the field citation scores. Formally:

$$CPP / FCSm^{19} = \frac{\sum_{i=1}^{n} c_i / n}{\sum_{i=1}^{n} e_i / n} = \frac{\sum_{i}^{n} c_i}{\sum_{i=1}^{n} e_i} \quad (6)$$

where c_i stands for the number of citations to publication *i*, and e_i is the expected number of citations to publication *i*. The resulting value shows the relation of the indicator to the world average that equals 1. Thus, for example, a value of 0.9 means that the publications of the unit considered are cited 10% below average and a value of 1.2 that they are cited 20% above average.

1.6.7. THE H-INDEX

The *h*-index is a measure developed by Hirsch in 2005 who defined it as "[...] the number of papers with citation number higher or equal to *h*, is a useful index to characterise the scientific output of a researcher" (Hirsch, 2005, p. 16569, italics mine). Thus, a high *h* index indicates that a scientist has published a considerable body of highly cited work (Hirsch, 2005).[20]

More specifically, a scientist has index *h* if *h* of his/her *Np* papers (number of papers published over *n* years) received at least *h* citations each, and the other $(N_p - h)$ papers have no more than *h* citations each (ibid). In other words, the h-index is the number of citations identified by the intersection point of the 45 degree line and the citation curve:

[19] CPP is the acronym for "citations per publication"; FCSm is the acronym for "mean field citation score".
[20] As Hirsch underlined, this measure must be used with reference to affirmed scholars (Hirsh, 2005).

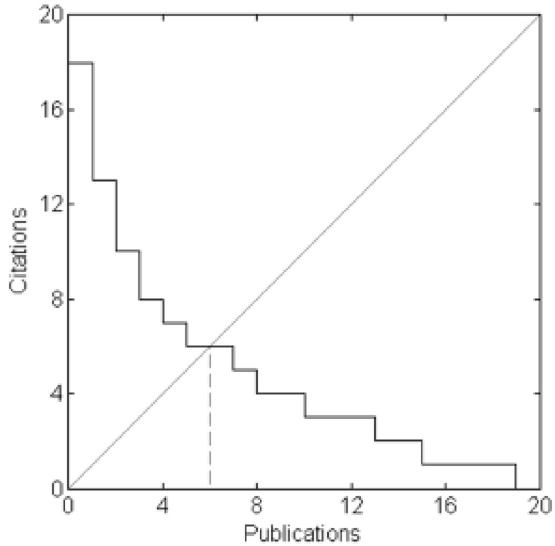


Figure 4: Graphical illustration of the calculation of the h-index, which results from the number of publications corresponding to the intersection point of the 45 degree line and the citation curve. The h-index is 6.
Source: Waltman & Van Eck, 2011, p. 407.

Even if the *h*-index is widely used it has many shortcomings (Bouyssou & Marchant, 2011; Marchant, 2009; Waltman & Van Eck, 2011). Waltman and Van Eck (2011) demonstrated that the *h*-index produces inconsistent results and behaves in a counter-intuitive way. It fails in its purpose of measuring the overall scientific impact of a scientist or a journal mainly because it is a non-normalised size-dependent measure. In case of consistent measures it is sure that if two authors show the same relative or absolute performance improvement their rankings do not change; it is guaranteed that "if scientist X_1 is ranked higher than scientist Y_1 and scientist X_2 is ranked higher than scientist Y_2, then a research group consisting of scientists X_1 and X_2 should be ranked higher than a research group consisting of scientists Y_1 and Y_2" (Waltman & Van Eck, 2011, p. 8); this is not the case of the *h*-index. Furthermore, different bibliometric distributions can have the same h-index (Moed, 2005):

Table 3: Different citation distributions result in the same h-index.

Publications	**Author 1 Citations**	**Author 2 Citations**	**Author 3 Citations**
1	30	30	**100**
2	10	10	**70**
3	8	8	8
4	6	6	6
5	5	5	5
6	1	**4**	1
7	0	**4**	0
8		**4**	
9		**4**	
h-index	**5**	**5**	**5**

notwithstanding the fact that the 3 Authors have very different citations distributions, they have the same h-index. Finally, it is important to stress that all variants and generalisations of the *h*-index suffer from this problem. Here some of these variants: *g*-index (Egghe, 2006); *h*(2)-index (Kosmulski, 2006); tapered *h*-index (Anderson, Hankin, & Killworth, 2008); *w*-index (Wu, 2010); Van Eck and Waltman's generalisation (Van Eck & Waltman, 2008a); Deineko and Woeginger’s generalised Kosmulski-indices (Deineko & Woeginger, 2009).

2. SCIENTOMETRICS: QUANTIFYING SCIENCE

> Scientometrics is the study of the quantitative aspects of science as a discipline or economic activity. It is part of the sociology of science and has application to science policy-making. It involves quantitative studies of scientific activities, including, among others, publication, and so overlaps bibliometrics to some extent (Tague-Sutcliffe, 1992, p. 1).

The birth of Scientometrics is commonly linked to Derek John de Solla Price's contributions and in particular to his oeuvre *Little Science, Big Science* published in 1963. Treating Science as a measurable entity, Price had four objectives: the quantification of the volume of Science (in terms of scientists and publications); tracing the distribution of scientists in a scientific space; conducting an interactionist analysis focused on scientists; detecting Science political and social dimensions (Price, 1963). Results showed that scientific growth was exponential and that had been lasting for 2 or 3 centuries:

> An exponential increase is best characterized by stating the time required for a doubling in size [...]. Now, depending on what one measures and how, the crude size of science in manpower or in publications tends to double within a period of 10 to 15 years. [...] If [...] only scientific work of very high quality is counted, then the doubling period is drawn out so that it approaches about 20 years (Price, 1963, p. 6).

Thus, the growth of Science seems to be very rapid and also constant: historical events, such as World War I, did not affect the growth. According to Price's results, the conflict did not push Science growth; on the contrary, military restrictions caused a mild reduction of the curve trend (Price, 1963). But the growth cannot go on to infinity, "[…] rather, exponential growth eventually reaches some limit, at which the process must slacken and stop before reaching absurdity. This more realistic function is also well known as the logistic curve [...]" (Price, 1963, p. 20). Once reached the ceiling, a period of crisis is expected (lasting a human generation science) followed by three possible situations: a complete reorganisation; a violent fluctuation; the variable death (Price, 1963).

Besides the exponential growth, Price found out another characteristic of Science development, that is to say tendency to crystallisation:

> In the sense that big things growth at the expense of the small ones [...]. Large fields seem to absorb the manpower and subject matter of small ones. Even though new fields, new departments, new institutions, and even new countries arrive on the scientific scene in increasing number, the few previously existing large ones have a natural growth enabling them in general to maintain their lead. It is the exception, rather than the rule, for one of the big blocks to slacken its growth – presumably thorough the existence of some sort of logistic

ceiling that causes it to stagnate – and be overtaken so that it falls in rank (Price, 1963, pp. 56-57).

This implies that the exponential growth of Science should result in the reproduction of the same internal distribution, resulting in a Pareto-like distribution (Price, 1963).

Thanks to data gathered by Garfield in his Science Citation Index (published in 1960), Price demonstrated that in the last two centuries Science recorded the biggest growth ever recorded in any field: "[...] 80 to 90 percent of all the scientists that have ever lived are alive now" (Price, 1963, p. 1). Science has become a monumental, powerful and big enterprise (Price, 1963). But Price's success, and so the birth of Scientometrics, cannot be neither explained nor understood without referring to the social and theoretical context constituting the humus making possible the birth and development of a *Science of Science* (Price, 1963).[21] This second Chapter introduces the social and theoretical backgrounds that paved the way for the birth of Scientometrics. The first Section focuses on those social changes occurring during the 1950s and the 1960s, which help in understanding why the SCI became so popular so quickly. The second one refers to Merton and Kuhn's sociological analyses on Science and scientific knowledge constituting the reference points for subsequent studies on Science.

2.1. SOCIAL BACKGROUND

As previously mentioned, *Little Science, Big Science* (1963) declared the birth of Scientometrics and Derek John de Solla Price its founding father. The oeuvre could not be published in a better period: the time was ripe for a "science of science".[22]

The huge growth of Science and the political awareness of its potentiality required new evaluation tools (Glänzel, 2003). The Sputnik launch in 1957 was a decisive event that determined a big change in Science organisation as it was the demonstration that planning and controlling Science was more successful than a mission orientated research. In Russia in 1952 was founded the All-Union Institute for Scientific and Technical Information (Всероссийский Инстиут Научной и Технической Информации) with the purpose of gathering scientific and technical information from various sources in order to spread it through the Soviet scientific community. Therefore, it is not surprising that Garfield's Science Citation Index (published in 1960) was enthusiastically welcomed by Russian scientists and particularly by Nalimov, who already in the 1950s had developed mathematical models to represent the worldwide scientific growth. Actually, Nalimov proposed the term *Naukometria* (the Russian term for

---

[21] Börner and Scharnhorst (2009) suggested the expression "sciences of sciences". According to them, as Scientometrics involves many and different techniques developed in different disciplinary fields and as its research interests comprise all scholarly activities (Natural Sciences, Social Sciences, Arts and Humanities), the expression "sciences of sciences" would be more appropriate.

[22] The first chapter of the book *Little Science, Big Science* (1963) is entitled "A prologue to a science of science".

Scientometrics) for the rising discipline twice: in 1966 and in 1969. He saw in the SCI a concrete possibility for the realisation of his scientometric project (Leydesdorff, 2005b).

Decisive for the development of a “science of science(s)” was also the awareness of the importance of scientific knowledge for the social progress. During the 1960s there was the creation of governmental organisations such as NASA (the National Aeronautics and Space Administration), ARPA (Advanced Research Projects Agency) NATO (North Atlantic Treaty Organisation) and OECD (Organisation for Economic Co-operation and Development) (Leydesdorff, 2005b). In 1963 the OECD Countries subscribed the *Frascati Manual* setting the methodology for collecting statistics about research and development. But such measurements were only about the inputs - human and economic sources invested in research – ignoring, thus, the outputs. As a matter of fact, scientific evaluation was still area of expertise of scientists who conducted it by means of the peer review tool: as in 1945 Vannevar Bush affirmed in his report *Science: the Endless Frontier*, addressed to Harry Truman, the institutional autonomy of Science had to be guaranteed, thus governments had to limit their activities in financing research (Bush, 1945/1980). The introduction of other tools for evaluative purposes, besides the classic one (peer review), was supported by scientists themselves: Science was going through a big and uncontrolled growth, therefore there was the need for new instruments able to restore the situation and establish objective standards such as, for example, quantifications related to publications or citations (Leydesdorff, 2005b).

2.2. THEORETICAL BACKGROUND

As all giants also Price stood on other giants' shoulders: he was widely influenced by Merton's Sociology of Science, by Kuhn's analyses on Science, and many of his reflections were possible thanks to Garfield's SCI. Actually, each of these three contributions was fundamental to the realisation of that “science of science(s)” called Scientometrics. I have already talked about Garfield (see Section 1.5.), thus the following sections are dedicated to Merton and Kuhn's Sociology of Science, which constitute the theoretical basis on which Scientometrics developed.

First, however, I think it is worthwhile to dwell on the precursory work by Ludwik Fleck, who mainly inspired Thomas Kuhn and who anticipated Merton in some of his conceptualisations.[23] Fleck was a Polish doctor who in 1935 published the book *Genesis and Development of a Scientific Fact* (Fleck, 1935/1979), where he stressed the collective (and thus social) nature of scientific results. One of the main concepts expressed in the book is the concept of “thought style”; a thought style is shared by what he calls “thought collective”:

[23] It was thanks to Merton that Fleck’s work was translated and then diffused on the international scene. Merton knew Fleck through Kuhn, who seems to be the first one who referred to him in his book *The structure of scientific revolutions* (Trenn cit. in Kuhn, 1970).

> Although the thought collective consists of individuals, it is not simply the aggregate sum of them. The individual within the collective is never, or hardly ever, conscious of the prevailing thought style, which almost always exerts an absolutely compulsive force upon his thinking and with which it is not possible to be at variance (Fleck, 1935/1979, p. 41).

This concept seems very close to Kuhn's notion of paradigm.

Fleck talks as a sociologist of science when he explains that:

> […] when we look at the formal aspect of scientific activities, we cannot fail to recognize their social structure. […] Cognition is the most socially-conditioned activity of man, and knowledge is the paramount social creation. […] *No* themes [...] could have been produced by the isolated thought of any individual. […] But those who consider social dependence a necessary evil and an unfortunate human inadequacy which ought to be overcome fail to realize that without social conditioning no cognition is even possible. Indeed, the very word 'cognition' acquires meaning only in connection with a though collective (Fleck, 1935/1979, pp. 42,43).

It is very interesting to notice that in the passage above we find many concepts that will be at the centre of both Sociology of Science and Sociology of Scientific Knowledge. Perhaps it is more interesting to notice that in his work, Fleck, underlines the necessity to consider both the structural-functionalist paradigm and the constructivist one when analysing Science.

2.2.1. MERTON'S SOCIOLOGY OF SCIENCE

> With the unending flow of achievement […] the instrumental was transformed into the terminal, the means into the end. Thus fortified, the scientist came to regard himself as independent of society and to consider science as a self-validating enterprise which was in society but not of it. A frontal assault on the autonomy of Science was required to convert this sanguine isolation into realistic participation into the revolutionary conflict of cultures (Merton, 1973, p. 268).

Despite these words were stated in the early 1940s[24], in 1963 Price was still wondering: "Why should we not turn the tools of science on science itself? Why not measure and generalize, make hypothesis, and derive conclusions?" (Price, 1963, p. V). Moreover, in *Science since Babylon* he affirmed: "[…] it is perhaps especially perverse of the historian of science to remain purely an historian and fail to bring the powers of science to bear upon problems of its own structure" (Price, 1961/1975, p. 162). Price's words

[24] They were first published in 1942 in an article entitled "The normative structure of science", in a short-lived journal, which became Chapter 13 in *Sociology of Science* (Merton, 1973) (Bourdieu, 2004).

suggest that at the time, that is to say the early 1960s, Science was still an unexplored area. And actually it was. The main reason lies in the predominance of what Merton calls "Pure Science Paradigm":

> One sentiment which is assimilated by the scientists from the very outset of his training pertains to the purity of science. Science must not suffer itself to become the hand maiden of theology or economy or state. The function of this sentiment is to preserve the autonomy of science. […] The exaltation of pure science is thus seen to be a defense against the invasion of norms that limit directions of potential advance and threaten the stability and continuance of scientific research as a valued social activity (Merton, 1973, p. 260).

The toast made at a dinner for scientists in Cambridge can be considered a tacit recognition of this function: "To pure mathematics, and may it never be of any use to anybody" (Merton, 1973, p. 260).

In the mid-20th century Science was still considered an autonomous and context-independent activity producing objective knowledge. In 1945 Vannevar Bush, in the presidential report *Science: the Endless Frontier*, claimed that the institutional autonomy of Science had to be guaranteed and governments' interference in Science could only be linked to research financing reasons (Bush, 1945/1980). In 1963 the *Frascati Manual*, a manual subscribed by the OECD Countries setting the methodology for collecting statistics about research and development, did not include measurements about Science outputs. Scientific evaluation was still area of expertise of scientists who performed it by means of the peer review tool (Leydesdorff, 2005b).

In the 1940s the area of inquiry relating to Science and its institutions was among the most underdeveloped ones in Sociology (Shils, 1972), but it is not surprising as the founding father of the *Sociology of Knowledge*, Karl Mannheim, considered scientific knowledge as a sociological case *sui generis* having a special epistemological status (Mannheim, 1952). As predicted by Merton, sociologists seriously focused on the systematic study of the interaction between Science and Society only when, after the events of World War II (e.g. the explosion of nuclear bombs) and the first genetic engineering experiments, Science started to be considered not only as a social problem but also as a prolific source of social problems (Merton, 1973).

Merton is unanimously considered the founding father of the *Sociology of Science*. He mainly focused on the institutional characteristics of Science, thus his Sociology of Science is also referred to as *Institutional Sociology of Science*. Merton considered Science as a social subsystem, self-regulating, linked to the wider social structure and the other subsystems by a *dynamically interdependent* relationship (Merton, 1949). According to Merton, in order to study the relationship between Science and Society, it was necessary to consider the former as an institution with a normative and organisational structure "[...] thus, we shall consider, not the methods of science, but the mores with which they are hedged about [...]" (Merton, 1973, p. 268). In particular:

> The institutional goal of science is the extension of certified knowledge. The technical methods employed toward this end provide the relevant definition of knowledge: empirically confirmed and logically consistent statements of regularities (which are, in effect, predictions). The institutional imperatives (mores) derive from the goal and the methods. The entire structure of technical and moral norms implements the final objective. The technical norm of empirical evidence, adequate and reliable, is a prerequisite for sustained true prediction; the technical norm of logical consistency, a prerequisite for systematic and valid prediction. The mores of science posses a methodological rationale but they are binding, not only because they are procedurally efficient, but because they are believed right and good. They are moral as well as technical prescriptions (Merton, 1973, p. 270).

Thus, according to Merton the *ethos* of Science is:

> [...] that effectively toned complex of values and norms which is held to be binding on the man of science. The norms are expressed in the form of prescriptions, proscriptions, preferences and permissions. They are legitimized in terms of institutional values. These imperatives, transmitted by percept and example and reinforced by sanctions are in varying degrees internalized by the scientist, thus fashioning his scientific conscience […] (Merton, 1973, pp. 268,269).

Merton identifies four institutional imperatives: *universalism* (scientific results must be subject to preestablished impersonal criteria and are judged independently of authors' personal or social attributes); *communism* (even if scientists must be rewarded for their scientific contributes, these are product of social collaboration and constitute a common heritage); *disinterestedness* (absence of economical or personal motivations in research activities, guaranteed by a pattern of institutional control); *organized skepticism* (it is both a methodological and institutional mandate and refers to the temporary suspension of judgement and the detached scrutiny of beliefs in terms of empirical and logical criteria) (Merton, 1973). Obviously, Merton did not think that all scientists behave according to these norms: the norms refer to Science as an institution and not to scientists' motivation system. Furthermore, sanctions to deviant behaviours would both demonstrate the existence of these norms and reveal their function (ibid).

Bibliometrists found this conceptual frame a suitable one for their citational analyses as it allows them to conceive of scientists' behaviour as rational and predictable.

2.2.2. KUHN'S SOCIOLOGY OF SCIENCE

The structure of scientific revolutions by Thomas Kuhn (1962/1970) is the revolutionary oeuvre which, together with Merton's intuitions, will constitute Scientometrics theoretical framework. The first Dutch scientometricians explicitly placed themselves in it: “Bibliometrics as we use it has to be situated within the Kuhnian sociology of science. The bibliometric study of paradigms has two main themes: 1) scientific growth and 2) the structure of science” (Koefoed, 1976, p. 5).

Thus, Science is no more considered an objective enterprise and its stability would be guaranteed by internal mechanisms of consensus among which the socialisation of scientists is the fundamental one: a “certain science” teaches future scientists how to reproduce itself, making them members of a given scientific community, sharing a certain paradigm, that is to say a complex of theories, values, techniques and in general universally recognised scientific achievements providing, for a certain period, problem-solving models (exemplars) and acceptable solutions. Until big anomalies do not occur the status quo is maintained and the normal proceeding of Science is assured; otherwise, Science will pass through a scientific revolutionary period characterised by the struggle for acceptance among competing paradigms (Kuhn, 1962/1970).

Scientific growth, during normal science periods, depends on previous scientific results (previous knowledge). As Stephen Cole stressed in the 1980s:

> Accumulation of knowledge can occur only during periods of normal science which are characterized by the adherence of the scientific community to a paradigm. It is only when scientists are committed to a paradigm and take it as the starting point for additional research that progress can be made. Without agreement on fundamentals, scientists will not be able to build on the work of others and will spend all their time debating assumptions and first principles (Cole, 1983, p. 26).

According to bibliometricians and scientometricians, citations are signs of Science useful for: depicting scientific communities; intercepting a rising research field; identifying the exemplars on which a community is built (Gilbert, 1977; Small, 1978). In the postscript of 1969 contained in *The structure of scientific revolutions* ([1962/1970) Kuhn claimed:

> How, to take an […] example, would one have isolated the phage group prior to its public acclaim? For this purpose one must have recourse to attendance at special conferences, to the distribution of draft manuscripts or galley proof prior to publication, and above all to formal and informal communication networks including those discovered in correspondence and in the linkage among citations. I think that the job can and will be done […] (Kuhn, 1962/1970, pp. 177,178).

In the postscript Kuhn cites Garfield's *The use of citation data in writing the history of science* (Garfield, 1964b), the work by Kessler (who developed the method called bibliographic coupling)[25] "Comparison of the results of bibliographic coupling and analytic subject indexing" (1925) and that by Price "Networks of scientific papers" (1965).

But Kuhn, as Merton, was a kind of non-intentional protagonist of Scientometrics. He did not totally agree with bibliometrists' derivations of his theory as witnessed by the following passage from Small's "Paradigms, citations, and maps of science" (2003):

> Kuhn's reaction to our initial papers on the cocitation structure of science was one of puzzlement. He wondered why we had focused on highly cited papers and authors, rather than defining the total community of researchers involved with the topic, representing the complete paradigm-sharing community. Any notion that highly cited papers might stand for exemplars or other paradigmatic constructs was not on his radar. Presumably, Kuhn saw bibliometric methods only as a means of performing a social inventory of a specialty, not as way to define the paradigm itself. Therefore, it is unlikely that Kuhn would have accepted any of our bibliometric reinterpretations of his theory (Small, 2003, pp. 395,396).

[25] See Section 4.2.

3. CITATIONS AS INDICATORS OF SCIENTIFIC INFLUENCE

The first Section of this Chapter introduces the main citation theories developed over time: the conceptualisation of citations based on the normative approach (citations as *reward tools* and intellectual links); the constructivist approach (citations as *rhetorical devices*); the multidimensional approaches (Cozzens' *rhetoric-first model*; Small's *citation cube*); the most recent proposals (Wouters' *reflexive indicator theory* and Moed's proposal). The second Section relates to effectiveness matters concerning citation analysis (citations shortcomings in detecting influence as well as databases biases and shortcomings). Section 3 discusses the incessant growth in the use of bibliometric indicators to evaluate and describe Science. The last Section is focused on those effectiveness matters related to the application of Bibliometrics to that wide and heterogeneous field called Social Sciences and Humanities

3.1. CITATION THEORIES: VALIDITY MATTERS

> *The absence of any explicit theory to guide the making and use of indicators may not be good; but the adoption of a single one is likely to be worse.*
>
> Gerald Holton[26]

Scientometrics lacks a citation theory encompassing the following aspects: a theoretical foundation for citation analysis; a clear justification to the use of Science (and Technology) indicators in science policy; an explanation of scientists' citing behaviour (Wouters, 1999b). Notwithstanding the fact that various scholars complained about this lack (Cozzens, 1981; Cronin, 1981; Cronin, 1984; Luukkonen, 1997; Zuckerman, 1987), now, in 2012, what Leydesdorff said in 1987 is still valid: "we still have a theoretically underdeveloped understanding of what these bibliometric data actually mean" (Leydesdorff, 1987a, p. 290). According to Wouters (1999a, b), the solution lies in a different conceptualisation of *citations* and *bibliographic references*, as they are two different entities with different properties. As a matter of fact, often scientometricians (and in general bibliometricians) use them as synonyms. Price (1970) firstly, and then other scholars (Egghe & Rousseau, 1990; Narin 1976), underlined the difference between the two signifiers. As Egghe and Rousseau pointed out:

> If one wishes to be precise, one should distinguish between the notions "reference" and "citation". If paper R contains a bibliographic note using and describing paper C, then R contains a reference to C and C has a citation from R (Price, 1970). Stated otherwise, a reference is the acknowledgement that one document gives to another, while a citation is the acknowledgement that one document receives from another. So, "reference" is a backward-

[26] Holton, 1978.

> looking concept while "citation" is a forward-looking one. Although most authors are not so precise in their usage of both terms, we agree with Price (1970) that using the words "citation" and "reference" interchangeably is a deplorable waste of a good technical term (Egghe & Rousseau, 1990, p. 204).

This differentiation is shared by those that are near to the so called "physical approach" - one of the strongest current supporters of the physical approach is Anthony Van Raan (Moed, 2005). According to the supporters of this approach, citations are objective pieces of information and science indicators are supposed to be interrelated by simple laws. The first supporter of this prospective was Price:

> It seems to me that one may have high hopes of an objective elucidation of the structure of the scientific research front, an automatic mapping of the fields in action, with their breakthroughs and their core researchers all evaluated and automatically signalled by citation analysis (Price, 1965, p. 194).

3.1.1. CITATIONS AS REWARD DEVICES

Merton's studies on Science institutional and organisational mechanisms include analyses on its reward system, in which citations have a specific function.

Science structure and organisation as depicted by Merton attracted bibliometricians and in particular Garfield, who asked him to write the foreword to his book entitled *Citation indexing - Its theory and application* (Garfield, 1979). Scientometricians found in the mertonian norms a theoretical frame useful for understanding the citational behaviour appearing, thus, rational and predictable because regulated by a professional ethics (Small, 2004). In this regard, I think it is meaningful that Eugene Garfield and Derek John de Solla Price (whose importance for Bibliometrics and Scientometrics is illustrated in the first Chapter of this work) participated to the research program developed in the late 1960s by Merton, together with Harriet Zuckerman, Stephen and Jonathan Cole (the so called "first circle") on cumulative advantages both in social stratification in general and in Science.

As Merton underlined in 1942, rewards in the realm of Science are distributed principally through the coin of recognition accorded by fellow scientists. As he explains in the Foreword to Garfield's *Citation Indexing*:

> Since recognition by qualified peers is the basic form of extrinsic reward [...] and since that reward can be accorded only when the work is made known, this historically evolving reward system provides institutionalized incentive for open publication without direct financial reward. Such peer recognition is usually accorded the first published contribution of its kind, later ones presumably being redundant. But since the cognitive structure of science makes for independent

> multiple discoveries - functionally equivalent if not identical discoveries - this social and cognitive complex evokes a concern among scientists to get there first and to establish, through prompt publication, their self-validating claims to priority of discovery (Merton, 1979, p. VI).

As Pliny the Elder said: “I have placed at the beginning of my books the names of my sources. I've done it because I believe that it is a pleasurable endeavour that shows honourable humility. It vouchsafes profound respect to those who have prepared the way to my own achievements” (Pliny cit. in Cardona & Marx, 2008, p. 498).

Intellectual property in Science is anomalous as it is established only when it is given away, that is to say published, and recognition comes by the usage of published scientific results.[27] Usage of scientific literature means to cite:

> Citations and references thus operate within a jointly cognitive and moral framework. In their cognitive aspect, they are designed to provide the historical lineage of knowledge and to guide readers of new work to sources they may want to check or draw upon for themselves. In their moral aspect, they are designed to repay intellectual debts in the only form in which this can be done: through open acknowledgment of them. Such repayment is no minor normative requirement. That is plain from the moral and sometimes legal sanctions visited upon those judged to have violated the norm through the kinds of grand and petty intellectual larceny which we know as plagiarism (Merton, 1979, p. VI).

A practical example of what explained by Merton in the passage above is the following. In 1872 Nietzsche published *The birth of tragedy* where, even if passages by authors such as Schopenhauer and Wagner were included, there were no bibliographic references. Ulrich von Wilamowitz (one of the founding fathers of German classical philology) judged such behaviour as non scientific, violating those rules regulating historical research; seven years after this event Nietzsche was expelled from the academic teaching. And it is not a coincidence if in the next years German, Italian, English and American philologists included in their works accurate and often superfluous references to Wilamowitz. In this way, classical philology community demonstrated approbation and recognition to the scholar (De Bellis, 2009). This event is explicative of what in time became a kind of golden rule of scholarship: in order to receive credits you must give credits (Merton, 1973).

According to this conceptual frame, citation is a device used to acknowledge colleagues' works: it is a reward tool. The symbolic credit flows naturally from the documents to the authors, as it is converted into institutional roles and positions of power, and from the authors to institutions, nations and journals (Merton, 1979). The publication and citation system is successful because it feeds upon itself: it is auto-

[27] Merton underlined that this paradoxical mechanism of Science must be linked to one of its fundamental norms, that is to say communism. This characteristic is reflected in the acknowledgement of scientists of their dependence on a common cultural heritage (Merton, 1973).

referential (Crane, 1976). But, of course, citations have also an important cognitive function: the transmission of knowledge (Merton & Gaston, 1977).

It is important to underlying that, as we will see shortly, Merton was absolutely cautious with reference to the use of citations as means of evaluation, and warned about potential abuses of this device. The scholar who first conceptualised citations according to the "mertonian way" was Norman Kaplan. In 1965 he posited a citation norm, based on the assumption that Science is an institution ruled by norms, according to which authors are expected to acknowledge prior work accurately and respecting original author's intentions (Kaplan, 1965). Put differently, scholars would act in accordance with the *humility norm* (Merton, 1957).

Even if Merton's Sociology of Science played a fundamental role in the development of Scientometrics, aside from the contributions introduced in this Section he did not deepen further the topic and he did not apply himself to citation analysis (Small, 2004). Actually, he warned about citation analysis limits[28] and he also did it in the foreword of the already mentioned Garfield's work *Citation Indexing* (1979):

> It was of course unnecessary for Eugene Garfield to identify this composite communications-intellectual-property-and-reward system in order to arrive at his concept of the citation index. He only needed the sense that the system provided the ingredients for systematically identifying, through citation indexing, links between the work of scientists that could be put to use both for searching the literature and for exploring cognitive and social relationships in science. […] we need to know more than is yet known about what references and citations do and do not represent if citation analysis is to provide further understanding of how science is socially and cognitively organized and practiced (Merton, 1979, pp. VI,VII).

Merton claimed the importance of studying also implicit or tacit citations, besides the explicit ones, as well as all other elements with cognitive importance that can be found in forewords and acknowledgements. Finally, he suggested the necessity to interview authors in order to obtain details on the *creation* of the text. This is because, for example, as Kuhn wrote in the foreword of his work *The structure of scientific revolution* (1962/1970), the influence of authors such as Fleck, Piaget, Sutton, and Quine is not reflected in the (scarce) amount of references included in his book (Merton & Gaston, 1977).

[28] We can find this warning also in Merton and Gaston (1977).

3.1.2. CITATIONS AS RHETORICAL DEVICES

The Mertonian citation conception was the dominant paradigm until the late 1970s, when a new citation theory was suggested. In 1977 Gilbert published the article “Referencing and persuasion” (Gilbert, 1977), where the attention was brought on the rhetorical function of citations. Support arrived from the MacRoberts (MacRoberts & MacRoberts, 1986) and mainly from Bruno Latour (1987), who is one of the major exponents of the constructivists approach in Sociology of Science (see, for example, Collins, 2004; Knorr-Cetina, 1981; Latour & Woolgar, 1979). According to this perspective, scientific knowledge is socially constructed through the manipulation of both financial and political resources, and the use of rhetorical devices should be the predominant feature of scientific communication (Knorr-Cetina, 1981). With reference to this aspect, Gilbert stressed that the citation practice must be interpreted as an author's device used to persuade readers of the validity of his arguments:

> A scientist who has obtained results which he believes to be true and important has to persuade the scientific community (or, more precisely, certain parts of that community) to share his opinions of the value of his work […] Accordingly, authors typically show how the results of their work represent an advance on previous research; they relate their particular findings to the current literature of their field; and they provide evidence and argument to persuade their audience that their work has not been vitiated by error, that appropriate and adequate techniques and theories have been employed, and that alternative, contradictory hypotheses have been examined and rejected (Gilbert, 1977, pp. 115,116).

As Zuckerman (1987) pointed out, actually it is since the time of Bacon that persuasion in Science has been playing an important role in scientific communication, and according to Latour's approach (*actor-netowork theory*) citations have exactly this function: “The presence or the absence of references, quotations and footnotes is so much a sign that a document is serious or not that you can transform a fact into fiction or a fiction into a fact just by adding or subtracting references” (Latour, 1987, p. 33); moreover:

> The effect of references on persuasion is not limited to “prestige” or “bluff”. [...] A paper that does not have references is like a child without an escort walking at night in a big city he does not know: isolated, lost, anything may happen to it. On the contrary, attacking a paper heavy with footnotes means that the dissenter has to weaken each of the other papers, or will at east be threatened with having to do so, whereas attacking a naked paper means that the reader and the author are of the same weight: face to face (Latour, 1987, p. 33).

Thus, according to this approach, even if various motives and rhetorical reasons to cite can be identified, the role of citation is always the same: support claims. So you can cite mechanically (Small talks about *standard symbol*)[29], you can cite to discredit, you can cite changing the meaning conceived in the cited text, you can cite in a haphazard way and the like, but you always cite to persuade. Furthermore, according to this perspective, citations perpetuate patterns of institutional stratification and those factors that influence citation behaviour are mainly linked to the position of authors in the stratification structure of Science.

The authors supporting this approach, such as Edge (1979) or Woolgar (1991), criticise citation analysis: as Edge affirmed, it is an "inappropriately positivist and realist approach" (Edge, 1979, p. 108). By refusing the mertonian concept of citation as reward device, they dispute the utility and legitimacy of Scientometrics.

3.1.3. MULTIDIMENSIONAL APPROACHES: THE RHETORIC-FIRST MODEL AND THE CITATION CUBE

Even if Merton's norms have been playing a fundamental role in Scientometrics, a citation theory satisfying scholars active in this field still lacks. And if some of them complain about it (see, for example, Cozzens, 1981; Cronin, 1981; Cronin, 1984; Luukkonen, 1997; Zuckerman, 1987), there are some who affirm its uselessness (see, for example, Van Raan, 1998). The result is an underdeveloped understanding of what bibliometric data mean (Leydesdorff, 1987a). However, over time, two attempts aiming to unify existing citation theories were made: the one by Cozzens in 1989 and that by Small in 2004.

In 1989 Cozzens suggested her multidimensional model, called *rhetoric-first model*:

> […] this paper argues that we should think of citations first as rhetoric and second as reward. […] citations stand at the intersection between two systems: a rhetorical (conceptual, cognitive) system, through which scientists try to persuade each other of their knowledge claims; and a reward (recognition, reputation) system, through which credit for achievements is allocated. The two systems are analytically distinct […]. But they are concretely indistinguishable; they are both present as impetus and constraint in any given act of citation. […] a third system also needs to be included in the model, the communication system. Among the citation inflators and deflators, journal characteristics, language of publication, and other measures of audience size need to be analyzed as part of this system. Again, the system is only analytically distinct from the reward and rhetorical systems; concretely they coexist (Cozzens, 1989, pp. 440, 444).

The model is called rhetoric-first model as it considers the rhetorical system overlying the communicative and reward one. Thus, according to Cozzens, the latter plays a marginal role with reference to the former: the rhetorical system creates the necessary conditions for citing, and, if we want

[29] See Section 3.1.3.

to model citations statistically, we must measure a new set of concepts, as those referring to the reward system "must be relegated to the role of accounting for the variation" (Cozzens, 1989, p. 445).

In 1978 Henry Small, colleague and collaborator of Garfield and developer of the so called *co-citation analysis*[30], illustrated in the article entitled "Cited documents as concept symbols" (Small, 1978) his citation theory combining the normative approach and the constructivist one.

Unlike the mainstream of the time, which was focused on authors' point of view and motivations, Small's interest was centred on the role played by citations in scientific literature:

> Very little, if any, attention is given in these studies to the scientific content of the citation context. Presumably this is not of interest since it did not shed light on the author's motivation for citing a particular work, or the implicit value judgment rendered by the citing author. Hence these studies have missed the role citations play as symbols of concepts or methods. This cognitive function arises from the formal requirement imposed on the scientist-author of embedding his references to earlier literature in a written text. This leads to the citing of works which embody ideas the author is discussing. The cited documents become, then, in a more general sense, "symbols" for these ideas (Small, 1978, p. 328).

According to Small, if a citation does not derive from a fraud, it must be considered both as a *sign* (with objective character) and as a *symbol* (symbol of the concept contained in the cited work). Put differently: the document cited is considered as symbolic of the idea expressed in the text. Small's symbol concept is the same of that by the anthropologist Edmund Leach, who in the book *Culture and communication: the logic by which symbols are connected* (Leach, 1976) claimed that human communication realises through expressive actions operating as signals, signs and symbols. Small applied Leach's conceptual scheme in the field of bibliographic references: in the "signs realm" there is contiguity, or better a *metonymic* relationship, between the object (citation) and the object meaning (cited document); in the "symbols realm" there is no contiguity between the object (citation) and the object meaning (concept in the cited document), so the relation is defined as *metaphoric*. In case of quoting, and in general when the cited documents contain the ideas which it comes to symbolise, we talk about metonymy (Small, 1978).

Small, drawing again from Leach's terminology, distinguishes between citations as *nonce symbols* and as *standard symbols*: the former is used to convey citer's ideas while the latter produce meanings which are shared by a scientific community or a group of scientist, which therefore are frequently cited (Small, 1978). It is in the concept of standard symbol that we find the basis on which citation and co-citation analyses are based:

[30] See Section 4.3.

> [...] when scientists agree on what constitutes prior relevant literature, including what is significant in that literature, they are in fact defining the structures of their communities (Small, 2004, p. 72).

The above-mentioned process of agreement and construction, as Small stresses, is a process of meaning construction:

> [...] as a document is repeatedly cited, the citers engage in a dialogue on the document's significance. The verdict or consensus which emerges (if one does) from this dialogue is manifested as a uniform terminology in the contexts of citation. Meaning has been conferred through usage and what is regarded and accepted as currently valid theory or procedure has been socially selected and defined (Small, 1978, p. 338).[31]

Small puts in the same logic line the processes of *agreement* and *construction*: “It is here that normative and constructivist approaches can find common ground. […] Here the symbolism of reward and the symbolism of meaning are operating in tandem” (Small, 2004, p. 76).

In 2004 Henry Small developed a three-dimensional model trying to merge Merton's recognition model and constructivist theories. It is a classification system of citations including the dimension of normative compliance, symbolic consensus and disinterestedness (self-citation). It is *the citation cube*. Actually, it consists of the development of Small's idea of citations as *concept-symbols*. This theory is based on the comparison of cited texts with their contexts of citation in the citing texts; once identified the category the citations belong to, they can be placed in the cube.

As shown in Figure 5 below, a possible theoretical framework is to consider the dimensions of “Literalness” versus “Consensus”: “On the vertical axis literalness measures the congruence of the cited work and the citing context from low to high. The horizontal axis labelled consensus indicates the degree of agreement in the citing community. Alternatively, this axis might be labelled revolutionary versus normal science” (Small, 2004, p. 76). Thus, for example, normatively compliant citations, such as paradigmatic or ceremonial citations, concentrate in the high literal-high consensus box, while deviant cases, such as revolutionary negative citations or paradigm breaking reinterpretations, fall mainly into the low literal-low consensus area. “In general, Merton’s recognition model would work best for high literal citations, while the constructivist model would work best for low literal citations” (Small, 2004, p. 77).

---

[31] Small stressed that at the end of this process there is the possibility of transformation of the meaning conveyed by the cited document, in this case the citation will symbolise something different from what the author intended (Small, 1978).

	Consensus Low	High
Literalness Low	CONSTRUCTIVE misattribution, reinterpretation, idiosyncratic negative, revolutionary	CONSTRUCTIVE ritual negative, common misattribution, conventional transformation, obliteration by incorporation
High	NORMATIVE substantive, organic, creative connection, unusual quotation	NORMATIVE perfunctory, ceremonial, common direct quotation, conventional interpretation, paradigmatic

Figure 5: The citation cube. The figure shows two dimensions of the citation cube model, literalness and consensus, each broken into high and low categories. The third dimension of the cube, self citation or disinterestedness is not shown. The dimensions shown highlight the similarities of differences between the original cited text and the citing context, and whether these differences are widely adopted by the community or unique to the citing author.
Source: Small, 2004, p. 77.

3.1.4. A REFLEXIVE INDICATOR THEORY

According to Wouters (1999a, b) the solution to the lack of a citation theory lies in the differentiation between *citations* and *bibliographic references*, as they are two different entities with different properties. In fact, an author *provides* a reference and *receives* a citation and, according to the supporters of this approach, “the distinction is anything but trivial” (Cronin, 2000, p. 440). As Wouters explains:

> If reference R of citing article A points at article B, the corresponding citation C is initially nothing else than a different format of reference R. The citation is the mirror image of the reference. This rather innocent looking inversion has important consequences. By creating a different typographical format of the lists of references – by organizing the references not according to the texts they belong to, but according to the texts they point at – they become attributes of the cited instead of the original, citing texts (Wouters, 1999a, p. 233).

Thus, citation theory cannot coincide with a theory of citing behaviour. Of course, the act of citing is an important object of study “but not because it gives the citation” (Wouters, 1999a, p. 12). An author cannot give a citation; the citation is the mirror image of the reference, thus they have different semiotic properties:

> […] the citation is the product of the citation indexer, not of the scientist (Wouters, 1999a, p. 4). If the citation is distinct from the reference, it seems natural to construct different citation theories. One set of theories should be focused on the citing behaviour of scientists and scholars […]; the other set of theories should be devoted to the function of the citation [...] (Wouters, 1999b, p. 562).

Furthermore, Wouters distinguishes between two types of Science representations linked to two concepts of information. The first refers to Shannon and Weaver's theory (Shannon & Weaver, 1949), according to which information is a countable entity; this view leads towards a "formalized representation of science"; the second one refers to Bateson's concept of information, which is focused on meaning (*any difference which makes a difference*) (Bateson, 1972) and leads to a "paradigmatic representation of science" (Wouters, 1999a). As Wouters explains:

> Science and technology indicators create a "formalized representation" of science which initially neglects meaning. Of course, to interpret these representations one needs to attribute meaning again. The main point is, however, that this attribution of meaning can be postponed. This is crucial because it enables the manipulation of "meaningless" symbols, such as the citation. The sign citation is an entity like Shannon's information concept and like entropy. Dimensionless, meaningless, countable (Wouters, 1999a, p. 209).

Wouter's *reflexive citation theory* (1999a) can be attractive as it maintains that the quest for a citation theory encompassing citing behaviour and citations is a dead end. It proposes to interpret citation and references as two distinct signs, which should be the object of two analytically independent research problems: on one side there is the study of the patterns of the citing behaviour of scientists, social scientists and scholars in the humanities, and on the other the theoretical foundation of citation analysis. According to Wouters his theory is useful because it:

> […] recognizes the two different domains, to position each indicator theory accordingly, and to establish their interrelations. In this sense, my proposal is also a theory, though a more abstract one: one could call this a proposal for a reflexive indicator theory. First, it is a theory about indicator theories because it explains how they can be related to one another and why the 30 year long quest for a citation theory has not been fruitful. Second, it is a theory about the indicators themselves, starting from the analytical distinction between the reference and the citation (Wouters, 1999b, p. 576, italics mine).

3.1.5. TOWARDS A CITATION THEORY

The latest attempt towards the realisation of a "grand" citation theory, encompassing both the normative and constructivist perspective, is made by Moed (2005). Actually, he did not develop a theoretical proposal but he focused on some reflections that can be considered as a starting point for developing a new and effective citation theory. He tried to show the complementarities between the two approaches, underlying also both their weaknesses and strengths.

First of all, he disagrees with Wouters when he claims that citations and references can be considered theoretically distinguishable: "Reference and citation theories, although analytically distinct, should not be separated from one another. A satisfactory theory of citation should be grounded in a notion of what scientists tend to express in their referencing practices" (Moed, 2005, p. 216). Thus, a citation is not just the product of the citation indexer, as affirmed by Wouters, but also of the scientist (Moed, 2005).

According to constructivists, references are the products of rhetoric strategies, nothing more (see Ection 3.1.2). In agreement with Zuckerman, who in her paper "Citation analysis and the complex problem of intellectual influence" (1987) replied to Gilbert's paper "Referencing as persuasion" (1977), Moed affirms that even if an author uses persuasive strategies, he/she will refer to a paper (or author) having intellectual influence; the reference will reflect cognitive influence:

> Scientists do not merely cite papers because the cited paper or its authors have, in their perception, earned a certain status during the past and can substantiate or add credibility to statements or claims made in a paper. A cited paper can be a strong weapon in persuading colleagues only if it has a certain significance (Moed, 2005, p. 214).

He also agrees with Zuckerman when she underlines that a distinction between motives and consequences of referencing behaviour should be made (Moed, 2005).

With reference to Cozzen's *rhetoric-first model* (Cozzens, 1989), Moed does not agree with her suggestion of separating the rhetoric, the communicative and the reward aspect: even if some rhetoric and communication factors can be separated, there are doubts that this could be done with reference to the reward and the rhetoric systems, as "citations reflect both aspects at the same time" (Moed, 2005, p. 215). The same doubt was also raised by Leydesdorff and Amsterdamska (1990), as they talked about the endogenous multidimensional character of citations. In addition, Moed suggests considering papers not as *concept symbols*, as in Small's view, but as flags or symbols of research programmes. According to him, in this way, it would be possible to account for the phenomenon of "split citation identity" observed by Cozzens (1982).

Furthermore, Moed stresses the importance to not rest on the assumption that errors and violations of norms can be concealed and neutralised by using large datasets. It is fundamental to understand, as claimed by Zuckerman (1987), if such phenomena are "randomly distributed among all subgroups of scientists, or whether they systematically affect certain subgroups" (Moed, 2005, p. 216).

In conclusion, Moed underlines the importance of bearing in mind the fact that reference lists play a fundamental role in the "citation culture", as they are at the basis of the development of bibliometric indicators, and scholars are aware of this. The awareness can result in the implementation of publication and citation tactics by authors; this aspect should be considered in the development of a citation theory (Moed, 2005).

3.2. CITATION ANALYSIS: EFFECTIVENESS MATTERS

It is preposterous to conclude blindly that the most cited author deserves a Nobel Prize. […] the mere ranking by numbers of citations or the numbers of papers published is no way to arrive at objective criteria of importance.

Eugene Garfield[32]

In the 1950s the huge growth of Science required new evaluation tools: Science was affected by a big, uncontrolled growth, therefore new instruments able to establish objective standards, such as number of publications or of citations, were necessary (Leydesdorff, 2005b). Citation counts were welcomed by many and a climate of ecstasy characterised the first publications about the topic. As reported by Merton in the preface of Garfield's oeuvre *Citation indexing - Its theory and application* (1979): "No one reading this book can fail to note its pervading sense of exuberant promise" (Merton, 1979, p. VIII). Garfield's Science Citation Index started to be employed to study and evaluate Science, and even if this will become a growing tendency, criticisms started soon as someone doubted the effectiveness of bibliometric indicators in measuring scientific productivity. Actually, criticisms were of two types: those referred to effectiveness matters, and those referred to validity ones. Previously (see Section 3.1.) I introduced the citation theories that have been developed until now. It should be clear that constructivists doubt the validity of citations as tool for studying and evaluating Science. However, despite all criticisms "citation analysis has conquered the world of science policy analysis" (Leydesdorff & Amsterdamska, 1990, p. 305). In 2002 Elsevier developed a new database, SCOPUS, competing with the indexes by Thomson Reuters (both databases involve strict data-collection and verification protocols); also Google decided to take part in this "bibliometric fashion" by developing its free, but uncontrolled, bibliometric source, namely Google Scholar. Furthermore, over time others databases have been created: disciplinary databases (e.g. MathSciNet, Medline, EconLit); citation databases (e.g. CiteSeerx or Citebase); full-text database (e.g. JSTOR, Science Direct and Wiley Interscience).

The following pages discuss those problems and limits of citation analysis deriving from both citations ineffectiveness in detecting influence and technical problems linked to data storing and data mining process. Attention is paid to Thomson Reuters' databases as they are the most used ones in

[32] Garfield, 1963, pp. 44,45.

research assessment systems, primarily because the articles published in its indexed journals are seen as having reached an internationally recognised standard (Hicks & Wang, 2009).

3.2.1. CITATIONS SHORTCOMINGS IN DETECTING INFLUENCE

Among those scholars harshly critical towards citation analysis we find the MacRoberts, who claimed that citation data are so incomplete and biased that they should not be employed in studies on intellectual influence (MacRoberts & MacRoberts, 1989). They identified various mistakes that authors make in acknowledging colleagues' works. First of all, there are those mistakes resulting in *under-citation* and *over-citation*. The former derive mainly from the *obliteration by incorporation* phenomenon (or briefly OBI): "the obliteration of the sources of ideas, methods, or findings by their being anonymously incorporated in current canonical knowledge" (Merton, 1968, p. 622). Thus, paradigmatic works are not formally acknowledged, consequently intellectual debts are not paid. Under-citation also occurs in case of junior researches collaborating with seniors whose names do not appear in publications. In 1987 the MacRoberts tested the citation analysis assumption according to which "the vast majority of citations are accurate and the vast majority of papers do properly cite their literature" (Garfield, 1980, p. 217). Results did not confirm the assumption: "Most authors simply did not cite the majority of their influences, and none cite all influences" (MacRoberts & MacRoberts, 1987, p. 343). The opposite phenomenon is that of *over-citation* caused by seniors over-citing their students as well as by colleagues "trading" citations in order to be visible in the scientific landscape (MacRoberts & MacRoberts, 1986, 1989).

The MacRoberts also pointed out that all those phenomena detected by Merton and colleagues, such as the Matthew effect (Merton, 1968) or the halo-effect (Crane, 1967), which are cause of distortions, are not considered by scientometricians or are dismissed as unimportant (MacRoberts & MacRoberts, 1989). But, as the Coles stressed some years before, this type of biases affect mainly heavily cited scholars (or documents), thus these phenomena realise in citation counts lower than what expected with respect to authors (or documents) recording high citation counts (Cole & Cole, 1973).

Furthermore, errors derive from inaccuracy in the act of citing, for example, when authors do not cite directly a certain work (or author) but make reference to the intermediary publication serving as "cognitive conduit" (Cole & Cole, 1973; Zuckerman, 1987).

Self-citations are considered a problem by many, but not by all. As Porter underlines:

> The interpretation of self-citations raises alternatives. Self-citations could be used to note the most relevant earlier work, or work with which one is most familiar; they could then reflect real influences and contributions. But there are also potentially self-serving aspects of self-citation, such as propounding one's own work and/or accumulating citations. Co-author citations offer additional self-serving possibilities–better *camouflaged* citation accumulation, and *back-scratching* (Porter, 1977, p. 263).

In this regard, Thomson Reuters is very clear. It affirms that high volumes of self-citations are normal in case of field leaders, mainly because of both the high quality and originality of their works. Thus, this kind of citations is not only quite normal but also expected (Thomson Reuters). According to Glänzel, Thijs, and Schlemmer (2004), indicators based on self-citations are useful measure that should be used both in Informetrics and research evaluation. It is important to underline that besides authors self-citations there are also journals self-citations:

> Journal self-citation occurs if a paper published in a given journal is cited by a paper published in the same journal. A great share of journal self-citations allows the conclusion that the journal in question is highly specialised, a low share indicates in a sense a "lack of originality"; a low share of journal self-citations (for instance, < 10%) is, for example, characteristic for review journals […] (Glänzel, 2003, p. 56).

Of course, distortions in bibliometric measures can be caused by constant high rates of self-citations (McVeigh, 2002).

Finally, one of the main criticisms made to citation analysis refers to the fact that citations are not all of the same type, so, for example, there are affirmative and negative ones. Problems derive from the fact that citation studies make no distinction between them, as SCI databases do not classify citations according to their nature (MacRoberts & MacRoberts, 1989).[33] Another neuralgic issue is the fact that raw citation data should not be used in comparative analysis unless normalised (Garfield, 2005; Van Raan, 2003a) as citation patterns differ according to many factors such as topic, discipline and country. For example, Narin found that:

> Distinct differences appear between fields. Engineering and technology, and mathematics have low referencing and citation/publication counts, in the range of 5 to 6. Psychology and biology form a second group, with 8 to 10 references and citation/publication. The next group contains earth and space science, physics, chemistry, and clinical medicine, all with 12 to 15 references and citation/publication. Finally, the field of biomedical research has substantially higher referencing and citation counts: between 18 and 20 per publication (Narin, 1976, p. 169).

Garfield warns that "all citation analyses should be normalized" (Garfield, 2005, p. 5). Irvine and Martin (1985) showed that Soviet Physics papers contain three-quarters of the references contained in Western Physics papers; Murugesan and Moravcsik (1978) found out that USA journals contain more conceptual and organic citations than Soviet ones; Lange (1985) noticed that the preferred language of cited publications depends on discipline and country; according to Peritz (1983), methodological articles would

---

[33] Different classifications have been developed over time (see, for example, Chubin & Moitra, 1975; Frost, 1979; Hodges, 1972; Moravcsik & Murugesan, 1975; Peritz, 1983).

receive more citation than theoretical ones; Moed, Burger, Frankfort, and Van Raan (1985a) affirmed that studies showed national differences in the life of cited literature (how quickly a paper will be cited, how long a paper will continued to be cited).

3.2.2. DATABASES BIASES AND SHORTCOMINGS

Many criticisms to citation analysis are linked to some structural features of databases. One of the major criticisms which has been made to the SCI since the beginning (which can also be extended to Scopus) refers to its non-objectiveness. The SCI was not created with the purpose of conducting citation analysis (Garfield, 1979), therefore a question naturally arises:

> Is there a logic in the selection process that, although obviously suitable for the intended purpose, is not necessarily suitable for scientometric work? That is, since what is to be included is not a random sample but a selected group, does the selection process reflect the interests and scientific philosophy of the selectors, and in what ways? (MacRoberts & MacRoberts, 1986, p. 346).

Furthermore, there are fundamental literature coverage problems: over-representation of English language journals and western ones (MacRoberts & MacRoberts, 1989; Smith, 1981); over-representation of USA serials (Schoepflin, 1990)[34]; under representation of non-journal literature (Hicks, 2004); coverage of (mainly) international literature poses field-coverage questions as, for example, the Humanities and Social Sciences deal with national, or local, topics, which are published locally as addressed to local audience (Van der Meulen & Leydesdorff, 1991).[35] Moreover, as one basic rule of Thomson Reuters is, according to Bradford's law, to detect (for each discipline or research area) the essential core of the most significant and important journals accounting for the bulk of significant results,[36] concerns about the under-coverage of interdisciplinary literature arise (Glänzel, 1996; Hicks, 2004):

> While science and technology exhibit a high concentration of papers in a select nucleus of special journals, and also in a brief span of time covering a few current years. In contrast, the literatures of the Social Sciences and Humanities exhibit a great dispersion of publications in different forms, on different subjects, and over a comparatively long span of time (Stevens, 1953, p. 12, italics mine).

---

[34] Schoepflin (1990) compared the UNESCO list World List of Social Sciences Periodicals and the SSCI. Results showed that American periodicals constitute 60% of SSCI list and only 17% of that by UNESCO.
[35] For a complete and detailed review see Moed (2005).
[36] http://wokinfo.com/benefits/essays/journalselection

With reference to the geographical coverage, it is interesting to observe the difference between the Web of Science and Scopus:

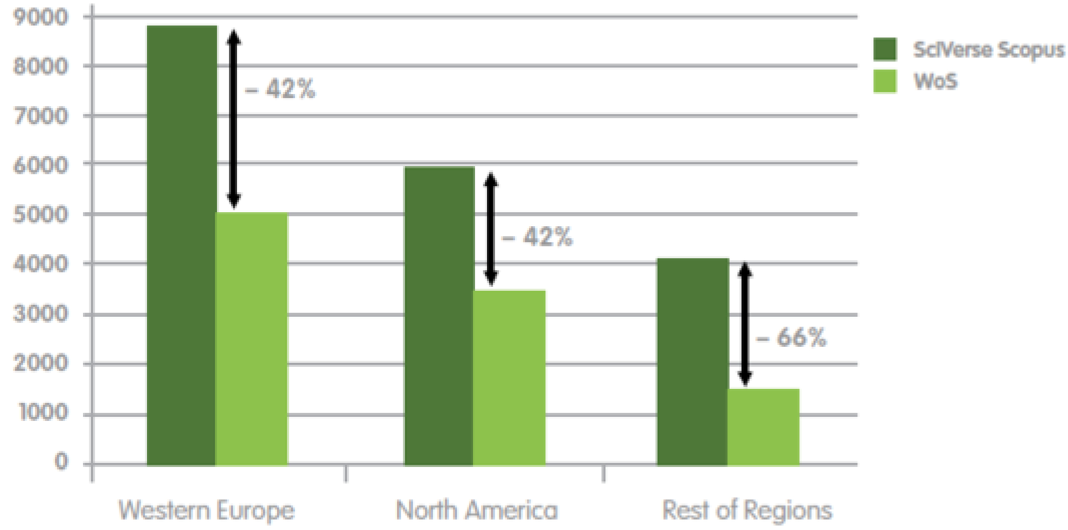


Figure 6: Number of titles in SciVerse Scopus (active) vs. Web of Science (shared titles with SciVerse Scopus) (April 2011).
Source: SciVerse Scopus, 2011.

Figure 7: Number of titles in SciVerse Scopus (active) vs. Web of Science (shared titles with SciVerse Scopus) by geographical region (April 2011).
Source: SciVerse Scopus, 2011.

As Figure 7 shows, the "Scopus Surplus" (Scopus minus WoS) with respect to the coverage of Asiatic journals is 69%. However, both Scopus and WoS over-cover Western Europe and North America journals (see Figure 6).

Another shortcoming of the Web of Science, which, instead, Scopus does not present, is linked to the fact that when downloading references information it is possible to obtain information only about the first authors of co-authored articles (Moed, 2005). This is a problem which is mainly linked to the possibility of making author co-citation analyses (see Section 4.4.).

Finally, we can find those errors which do not derive from the way in which the databases are built, but have different nature. The most recurrent ones derive from: *homonyms* (aggregation of citations which instead refers to different authors); *synonyms* (e.g. "R. K. Merton" and "R. Merton", even if referring to the same person they will be entered in different parts of the index); *clerical errors* (authors' errors in writing bibliographies and mistakes in recording bibliographies) (Cardona, Chamberlin, & Marx, 2007; Marx, 2011; Porter, 1977; Smith, 1981); mistakes in reporting scholars institutional affiliations (wrong affiliation, errors in reporting names); changes in institutional affiliation (Moed, 2005); translation errors (ibid).

As the main experts in Scientometrics repeat when presenting pros and cons of these tools, it is necessary that the ones who are interested in rankings and evaluations based on bibliometric tools are aware of all problems linked to this kind of analyses. The use of different databases at the same time is also suggested. Finally, evaluations and descriptive studies should take in consideration different aspects of the phenomenon studied.

3.3. TOWARDS A CITATION CULTURE

Notwithstanding all the shortcomings previously showed, bibliometric data are more and more used for evaluating Science. They are, at least today, the unique tool which seems able to furnish information on scholars' work to the administrators of the scientific enterprise (Sosteric, 1999). Furthermore, thanks to citation analysis decision makers and funding agencies can detect and monitor the performance of research groups as well as highlight centres of excellence (Moed, Burger, Frankfort, & Van Raan, 1985b).

Despite all its limits, Scientometrics is increasingly used. In France Coutrot talks about "quantophrénie" (Coutrot, 2009),[37] in the UK the decision to tie funding to national research rankings resulted in a strong competition between universities with unprecedented levels (Williams, 1997). Moreover, there are some who talk about the emergence of a cybernetic control culture within academia: Sosteric (1999) considers citation analysis as a form of "Orwellian surveillance net"; Hicks and Potter

[37] In 1990 the organisation OST (Observatoire des Sciences et des Techniques) was founded; it employs only bibliometric tools for evaluation. In 2007 was created the organisation AERES (Agence d'Évaluation de la Recherche et de l'Enseignement Supérieur); it occupies a higher position with respect to the traditional evaluative structures, namely CNRS and CNU (Coutrot, 2009).

(1991) used the Foucauldian perspective to describe the academical situation with reference to what can be called “the citation culture”[38]:

> We have drawn on Foucault [...] because we were interested in the parallels between Foucault’s analysis of the development of institutions such as the penal system and medicine as new technologies of observation and control, and the work of citation analysis (CA) itself, which provides a new way of making scientific practice visible and recordable, and new possibilities for producing hierarchies of difference and categories of normal/abnormal scientific behavior (Hicks & Potter, 1991, p. 475).

As Gingras claimed, it seems that in recent years the European academic realm has been suffering from an evaluation fever as we are witnessing a savage usage of Bibliometrics (Gingras, 2008).

After the publication of the Science Citation Index two types of bibliometric application arose: *evaluative* and *relational bibliometrics*. Evaluative bibliometrics is engaged in evaluating scientific production by means of indicators of research performance, usually to inform research policy and to help direct research funding (Moed, 2005). Relational bibliometrics studies relationships within Science with the scope of detecting the cognitive and intellectual structure of research fields, the emergence of new research fronts as well as national and international co-authorship patterns (Thelwall, 2008). And the “metric” growth increases thanks to the emerging field of Webometrics defined as: “The study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web drawing on bibliometric and informetric approaches” (Björneborn, 2004, p. 1216). As Shadish et al. (1995) underlined, like it or not citation analysis is affecting scientists’ lives more and more. It seems that we cannot escape the so called “citation culture”.

3.3. BIBLIOMETRICS AND THE SOCIAL SCIENCES AND HUMANITIES[39]

While SCI data have been successfully applied to the Natural and Life sciences (see, for example, Van Den Berghe, Houben, de Bruin, Moed, Kint, Luwel, & Spruyt, 1998; Van Raan, 1996), this is not the same for the Social Sciences and Humanities (SSH). The main reason stands in differences between the SSH publication patterns and the Natural and Life Sciences ones (Moed et al., 2002). In many subfields of the Social Sciences and Humanities (SSH) there is lack of standardisation in publication practices, otherwise, for example, the molecular biology subfield seems to be characterised by high consensus on publishing rules (e.g. important works published in English, in international journals, and in a standard

[38] Expression taken from Wouters (1999a).

[39] The SSH (Social Sciences and Humanities) do not constitute neither an homogeneous block nor two separate ones: while some SSH fields share with the Natural and Life Sciences publication patterns and citation behavior, others show Humanities patterns (Nederhof, Zwaan, De Bruin, & Dekker, 1989; Thompson, 2002).

format) and on the most prestigious scientific journals (Van Den Berghe et al., 1998). Furthermore, with reference to the SSH it must be considered the following: many important works are gathered in books; there is not a unique and predominant type of publication; scholars usually publish in national language (Moed, et al. 2002).

Recently, both the Web of Science and Scopus expanded their SSH journals coverage. As regards the Social Sciences Citation Index, this expansion increased by more than 40 per cent during the past 6 years (Moed, Linmas, Nederhof, Zuccala, Illescas, & de Moya, 2009) (currently it covers 2,650 journals across 50 disciplines) and in 2009 it added a set of "regional journals" (currently 1,620) in which periodicals targeting regional rather than international audience are included (Thomson Reuters). Selection criteria for regional journals are fundamentally the same as for international journals: publishing standards, editorial content, international diversity, and citation data are all considered.[40] In 2009 also Elsevier decided to increase Scopus coverage of SSH journals adding 1,450 periodicals (Moed et al., 2009); thus currently it includes a total of 5,900 titles.

The SSH constitute a "special case" and variations are applied to those bibliometric methods developed for monitoring research performance in the Natural and Life Sciences. This is due to differences in publication patterns and citing habits summarised in 5 points by Nederhof (2006):

> 1) *A pronounced national and regional orientation*: the SSH are characterised by a national research frontier[41]. As the great part of their output refers to national or regional topics and addresses a local audience, works in this field are usually locally published (in regional or national monographs or serials) (Broadus, 1971; Hicks, 1999; Nederhof et al., 1989; Nederhof & Zwaan, 1991). Among those fields showing a relatively strong local orientation we find Linguistics, Literature Research, Public Administration, Political Sciences and Sociology (Luwel, Moed, Nederhof, De Samblanx, Verbrugghen, & Van Der Wurff, 1999). Local publications are non-English ones: they are written in the mother tongue (Nederhof, 2006). In a study conducted in the late 1980s, Kyvik (1988) found that in Norway 80% of natural scientists published in a foreign language; conversely, 75% of social scientists and only 35% of natural scientists published in Norwegian.
>
> Authors' citation patterns can also be analysed as indicators of their reading habits. Yitzhaki (1998) found that, with reference to Sociology, American and British authors publishing in US and UK journals cited English language material 99% of the time, although English Sociology publications probably accounts for 70% of world literature. With reference to German authors' citing habits, it was observed that in two German journals the amount of cited German material was 63% and 86%, although German literature probably accounts for 5% of world Sociology.

[40] http://isiwebofknowledge.com/products_tools/multidisciplinary/webofscience/contentexp/

[41] As defined by Stephen Cole "the research frontier, consists of all the work currently being done by all active researchers in a given discipline. The research frontier is where all new knowledge is produced" (Cole, 1983, p. 114).

French authors in two French journals cited French material 65% and 68% of the time, although French literature probably accounts for 8% of world Sociology.

Nevertheless, recent studies showed that the local nature of the SSH is decreasing for a number of reasons, such as the increasing internationalisation of national economies, the growing use of electronic communication, and the growth of documents written and published by institutions from different countries (Hicks, 1999; Katz, 1999);

2) *Less publication in serials and more in books*: the Natural/Life Sciences and the SSH show different publication preferences. As regards the latter, evidences show that books and monographs play an important role in scientific communication (Broadus, 1971; Clemens, Powell, McIlwaine, & Okamoto, 1995; Cronin, Snyder, & Atkins, 1997; Glänzel, 1996; Hicks & Potter, 1991; Small & Crane, 1979). But why is it so? Probably, the most important reason lies in the fact that knowledge obsolescence in the SSH is very slow, therefore the content of books is current for a long period of time. On the contrary, with reference to the hard sciences, the obsolescence rate is much higher, therefore journal articles are the means most widely used by scientists in these fields to disseminate their research results (Hicks, 2004).

In 1999 Glänzel and Schoepflin analysed the references structure of both SCI and SSCI with respect to the year 1993. They found that while 64% of both Psychology and Psychiatry as well as 49% of Economics references concerned journals, the amount of journals cited in Sociology amounted to 40% and only to 35% for History and Philosophy of science. Much higher percentages were registered in solid state Physics (85%) and Analytical Chemistry (84%); Electronic Engineering amounted to 62%. Nonetheless, evidences show that journals are the main communication channel for different fields in the Social Sciences and Humanities, and according to Nederhof et al. this is true mainly for Behavioural Sciences and Economics (Nederhof et al., 1989).

3) *A difference pace of theoretical development*: compared to the majority of the so called hard sciences, many Social Sciences and Humanities seem to be characterised by a slower pace of theoretical development (Cole, 1983; Kuhn, 1962/1970; Storer, 1967; Zuckerman & Merton, 1972). This difference would be reflected in citation patterns of publications: larger cited half-life of publications and higher citation rates of older literature; lower volume of citations than many hard science fields (Glänzel, 1996; Nederhof et al., 1989); reference lists in Social Sciences and Humanities articles often contain publications older than 5, 10, or even 15 years than in the hard sciences (Broadus, 1971; Cole, 1983; Moed, 1989; Thompson, 2002); in the 1970s the Price Index (computing the percentage of references in journals to literature published in the last five years) varied between 60% and 70% in Physics and Biochemistry, whereas for Social Sciences journals it varied between 40% and 50% (Price, 1970). Using a different measure Glänzel and Schoepflin obtained in 1999 very similar results: the mean reference age was lowest in Biomedicine (7–8 years), higher in Sociology (12.5 years), and highest in History

and Philosophy of Science as well as in Social Sciences (39 years) (Glänzel & Schoepflin, 1999).

Furthermore, in 1994 Glänzel and Schoepflin found that hard sciences and the SSH have a different distribution of citations over years: over a 14-year period Psychology articles took more than 8 years to reach 50% of their citations compared to 4,5–6,5 years for Physics articles. Differences in obsolescence are also registered: between 46%–75% of the Physics articles analysed did not receive any citations after 14 years from publication, the amount of the Psychology articles registering the same values, instead, varied between 14%–22% (Glänzel & Schoepflin, 1994).

According to Cole (1983), differences between hard and soft sciences would mainly refer to the "knowledge core" and not to the "research front".

> The core is the starting point, that knowledge which people take as a given from which new knowledge will be produced. The core is characterized by having a relatively small number of theories and substantial consensus on the importance of these theories. The other component of knowledge, the research frontier, consists of all the work currently being done by all active researchers in a given discipline. The research frontier is where all new knowledge is produced (Cole, 1983, p. 114).

Studying the reference pattern of Sociology, Physics and Chemistry undergraduate textbooks of the 1980s (textbooks are considered as indicators of the content of the core knowledge) and conducting a qualitative study on their content, Cole observed that differences between hard and soft sciences can be detected at the core level but not at the research frontier one. He found that in the Natural Sciences only few references (6% in Chemistry and 3% in Physics) were to works published after 1959, while in Sociology 75% of references were to works published after 1959. Furthermore, whereas authors of Natural Sciences textbooks tended to cite approximately 100 articles or books, Sociology ones showed an average of 800 references. Finally, comparing textbooks of the time and 20 years old ones, results showed that the material covered in both was essentially the same in Physics and Chemistry, while in Sociology only a small fraction of the works included 20 years before was cited. Evidences let Cole to claim that there was almost universal agreement on Natural Sciences main theories and exemplars, while just little agreement on Sociology ones, showing a very small core of knowledge but a large research frontier (Cole, 1983).

Even if the claim that the SSH are characterised by a slow pace of theoretical development is true on the whole, something is changing: in some disciplines, such as Archaeology and Anthropology, the use of methods borrowed from the experimental and technical sciences has provided for speedier scientific results. As a result, publications in these disciplines have multiplied and information has become obsolete faster, leading scholars active in these fields to publish their research results in papers rather than books (Iribarren-Maestro, Lascurain-Sánchez, & Sanz-Casado, 2009);

4) S*ingle-scholar approach versus team research*: many SSH (particularly Humanities and, in general, qualitative works) differ from the hard sciences also for a "single scholar" attitude in doing research (Thompson, 2002). For example, Rubio (1992) found that during the period 1986–1988 in Spain only 14% of Social Sciences production was co-authored, and the percentage decreases (3%) if we consider, for example Linguistics. On the contrary, hard sciences show a higher level of collaborations (Moed, 1989; Nederhof, 1989);

5) *A greater share of publications directed at the non-scholarly public*: while hard sciences scholars public mainly for the other scholars, Social Sciences and Humanities ones also address the general public (Nederhof and al., 1989; Van der Meulen & Leydesdorff, 1991). For example, a high percentage of social science studies are directed to policy (Nederhof & Van Wijk, 1997). In some field of the Social Sciences and Humanities more than 75% of publications of researchers address a non-scholarly public (Nederhof et al., 1989; Nederhof & Erlings, 1993). On the contrary, the hard sciences show a different attitude. As witnessed by Willems: "[...] scientists throw up barriers to their work being publicized" (Willems, 2003, p. 470)

It should be clear that differences among fields cannot be ignored if there is the will of using SSCI and AHCI (Art and Humanities Citation Index) data for scientometric analyses: "[...] given the increasing uses of performance indicators based on bibliometrics in Social Sciences and humanities, it is essential to take all due precautions to prevent perverse effects" (Archambault, Vignola-Gagne, Côté, Larivière, & Gingrasb, 2006, p. 330). But, as Van Raan stresses, even if bibliometric assessment in the SSH can be problematic, "we caution against an all too easy acceptance of the persistent characterization of the Social Sciences (and the humanities) as being bibliometrically inaccessible" (Van Raan, 2003a). Of course, there are problems that must be solved, but also the peer review tool employed to assess scientific literature quality has some shortcomings. For example, one problem related to the use of bibliometric tools, particularly when applied to the Social Sciences, is about time dimension, but this is also a dimension pertaining to peer review assessment:

> Citations are given after publication. So, how long must we wait, in other words: what is an acceptable length for the "citation window"? For the Social Sciences this window should be longer than in the Natural Sciences, and around five to six years. This unavoidable time lag (impact is mainly received *after* the work has been published), is often "misused" by critics (even in the Natural Sciences where it is about two to three years) as a general objection against bibliometric analysis. Yet even peers generally need time to see whether research results will "take root"! (Van Raan, 2003a, p. 25).

There are many ways in which experts can be influenced: by subjective elements, narrow-mindedness, or limited cognitive horizons. But there are also positive aspects of a "subjective" view such as the importance of intuitive insights of experts. "We claim however that for a substantial improvement of decision-making [...] bibliometric methods has to be used in parallel to a peer-based evaluation procedure" (Van Raan, 2003a).

Despite all shortcomings and problems related to citation analysis, the "citation culture" (Wouters, 1999a) is becoming more and more a reality. As previously mentioned, the use of Bibliometrics in research performance assessment is increasing, and also the SSH are required to become "bibliometrically measurable". The main effort made in that direction was made in 2009, when the European Science Foundation (ESF), the Agence Nationale de la Recherche (ANR) in France, the Deutsche Forschungsgemeinschaft (DFG) in Germany, the Economic and Social Research Council (ESRC) and the Arts and Humanities Research Council (AHRC) in the UK launched an European Scoping Project aiming to establish the feasibility of developing a robust European bibliometric database for assessing the impact of all types of research output in these domains (Moed et al., 2009). Finally, one should also mention the recent work by Hammarfelt (2012), who applied Bibliometrics to Literary studies analysing the citation patterns in it.

4. RELATIONAL BIBLIOMETRICS: SCIENTOPOGRAPHY

Social life is relational; it's only because [...] blacks and whites occupy particular kinds of patterns in networks in relation to each other that "race" becomes an important variable.

Collins Randall[42]

Relational Bibliometrics studies relationships among scientific publications with the scope of detecting the cognitive and intellectual structure of research fields, the emergence of new research fronts and national and international co-authorship patterns (Thelwall, 2008). This approach encompasses various methods, which differ according to the unit of analysis (authors, papers or semantic elements). Schubert and Braun (1996) proposed for this set of mapping techniques the term "Scientopography", and Small affirmed that Relational Bibliometrics maps can be considered as "[…] a spatial representation of how disciplines, fields, specialties, and individual papers or authors are related to one another as shown by their physical proximity and relative locations, analogous to the way geographic maps show the relationships of political or physical features on the Earth" (Small, 1999, p. 799). The possibility of creating maps based on scientific literature goes back to 1960, before the SCI publication, when the geneticist Allen, who was supporting Garfield in his attempts to create a new, revolutionary bibliographic system for scientific literature, sent to the latter a diagram showing citation connections among a set of biochemistry articles. Since then, three techniques aiming to detect relationships in Science studying papers bibliographic lists have been developed: *Bibliographic Coupling* by Kessler (1963a); *Co-citation Analysis* by Small (1973); *Author Co-citation Analysis* by White and Griffith (1981). All of them share the idea that a bibliographic reference stands for an intellectual link: quoting is an intellectual choice that establishes a relationship between who cites and who is cited. According to Merton's Sociology of Science, a reference is the payment of an intellectual debt and so *the* way used in Science to acknowledge colleagues' work. Moreover, in the early 1980s Callon, Courtial, Turner, and Bauin (1983) developed the so called *Co-word clustering,* and in 1990 Todorov created the technique *Co-heading analysis* (Todorov, 1990), both of them are based on the analysis of the frequency of the co-occurrence of textual elements such as keywords or subject headings. All these techniques aim to map Science structure: "Our task is to depict […] Science relationships in ways that shed light on the structure of Science" (Small & Griffith, 1974, p. 40). But while bibliographic coupling, co-word and co-citation analysis are mainly designed to describe the structure of Science and its evolution at the macro and meso level, co-author clustering and author co-citation analysis aim to reveal structures at the micro and meso level (White & McCain, 1998).

[42] Randall, 1988, p. 413.

4.1. RELATIONAL BIBLIOMETRICS AND THE RELATIONAL PARADIGM IN SOCIOLOGY

Even if the study of relations among the various elements extractable from scientific publications (authors, words, citations, institutions) has been playing a fundamental role in Scientometrics since the beginning, it seems there are no contributions stressing and explicating the epistemological consequences of such approach. I have already stated in the Introduction my conviction about the potentiality and the necessity of a constant collaboration between Sociology and Scientometrics.

We have just seen the general characteristics of Relational Bibliometrics. Its task is to detect and represent relationships among scientific publications in order to shed light on the structure of Science (Small & Griffith, 1974). Relations are at the basis of this approach as it is for Social Network Analysis (SNA). And in fact, as early as 1972 we find studies in which SNA was applied to bibliometric networks (see, for example, Haythornthwaite & Wellan, 1998; Logan & Pao, 1990, 1991; Pritchard, 1984; Shaw, 1985). Recently Otte and Rousseau (2002) stressed the potentialities of the application of SNA in the field of Information Science. In fact, whether a bibliometric network is developed and analysed through SNA techniques or through different ones (see Section 5.1.) we are immersed in a relational dimension. According to Gingras, by combing bibliometric techniques and the Network Analysis approach: “we [...] have the tools to replace a purely metaphoric use of the term ‘‘network’’ by a visible map of the intellectual relations between people [...]” (Gingras, 2010, p. 331).

When Small and Griffith state that through relational bibliometrics it is possible to detect the structure of scientific fields, they are relying on certain assumptions related to a certain conception of social reality. They are assuming that the structures of scientific communities are detectable by studying relations. In Sociology the relational approach states the impossibility to reduce Society to one element (individuals or structures). Simmel conceives of society as an abstract supra-singular structure: “it is not real object. It does not exist outside and in addition to the individuals and the processes among them” (Simmel, 1950, p. 4). According to Luhmann:

> There are no elements without relational connections or relations without elements. [...] Elements are elements only for the system that employs as units and they are such only through this system (Luhmann, 1995, pp. 20,22).

In Chapter 6 it is shown how Scientometrics can be framed in a dynamic and relational sociological approach. Of interest here is the fact that Relational Bibliometrics seems to be able to furnish snapshots of the structure of those social subfields called scientific communities. In claiming this, the developers of this approach locate it in a precise epistemological position. The fact that they did not specify it does not make it less true or of little importance.

4.2. BIBLIOGRAPHIC COUPLING

The first bibliometric method based on references was proposed by Michael M. Kessler in 1963 in the article “Bibliographic coupling between scientific papers” (Kessler, 1963a). Bibliographic coupling (BC) occurs when two works reference a third in their bibliographies; it is based on the assumption that through the analysis of the cited literature it is possible to obtain information about the intellectual environment within which authors operate. Thus, if two papers show similar bibliographies there is an implied relation between them; the coupling strength depends on the number of references shared by them (Kessler, 1963a). As Kessler explained in the early 1960s:

> Most scientific papers contain a bibliography of references. Each reference must be recognized as an acknowledgement of the receipt of information from the past. A symmetry suggests itself. If a paper is published at time $t = t^0$, all its references form a known population of scientists in time $t < t^0$. The readers of the paper comprise an unknown population of scientists in time $t > t^0$. What is the relation between these two populations? Does the bibliographic structure contain a set of properties either numeric or formal that characterize the paper from which it was derived? (Kessler, 1963b, p. 49)

The aim of this method is to determine the similarity of the subject matter between pairs of publications; resulting clusters of coupled works are assumed to represent the same or at least related research topics. BC was the historical antecedents of cocitation analysis but it did not have the same success: Garfield and colleagues at ISI were looking for a technique which, differently from the BC, was able to detect those papers considered important by authors (Small & Griffith, 1974). However, 15 years after its development, Bibliographic Coupling started to be used in information retrieval at ISI (Garfield, 2001).

4.3. CO-CITATION ANALYSIS

Co-citation Analysis (CA) was developed with the scope of mapping the structure of Science identifying clusters relating to specialties or, more generally, research areas. Co-citation maps are used to describe the development of disciplines and specialties and to detect emerging areas of scientific enquiry (Griffith, Small, Stonehill, & Dey, 1974; Small & Crane, 1979; Small & Sweeney, 1985; Small & Griffith, 1974; Narin, 1976; Leydesdorff, 1987b; Todorov & Glänzel, 1988). CA was created simultaneously but independently by Henry Small (1973) and Irina Marshakova (1973). The fundamental premise of cocitation analysis is that the intellectual relation between two documents or authors is given by their co-citation pattern: the more they are cited together, the stronger their intellectual relation (Bellardo, 1980).

In the article “Co-citation in the scientific literature: a new measure of the relationship between two documents” (1973) Small explains:

> Bibliographic citations in scientific papers have been used by a variety of researchers to establish relationships among documents. [...] Unlike bibliographic coupling which links source documents, co-citation links cited documents [...]. Co-citation is the frequency with which two items of earlier literature are cites together. [...] The pattern of linkages among key papers established a structure or map for the specialty which may then be observed to change through time. Through the study of these changing structures, co-citation provides a tool for monitoring the development of scientific fields, and for assessing the degree of interrelationship among specialties (Small, 1973, p. 265,268).

Furthermore, in 1973 Marshakova wrote:

> The proposed method is the logical opposite of the method of bibliographic coupling. From the mathematical point of view citation network is a set of documents with the relation of citing imposed on it. In other words it is a union of a set of citing papers and a set of cited papers. A citation network is a potential base for various classifications of member-papers (Marshakova, 1973, p. 1).

Figure 8 below shows graphically the difference between BC and Co-citation analysis:

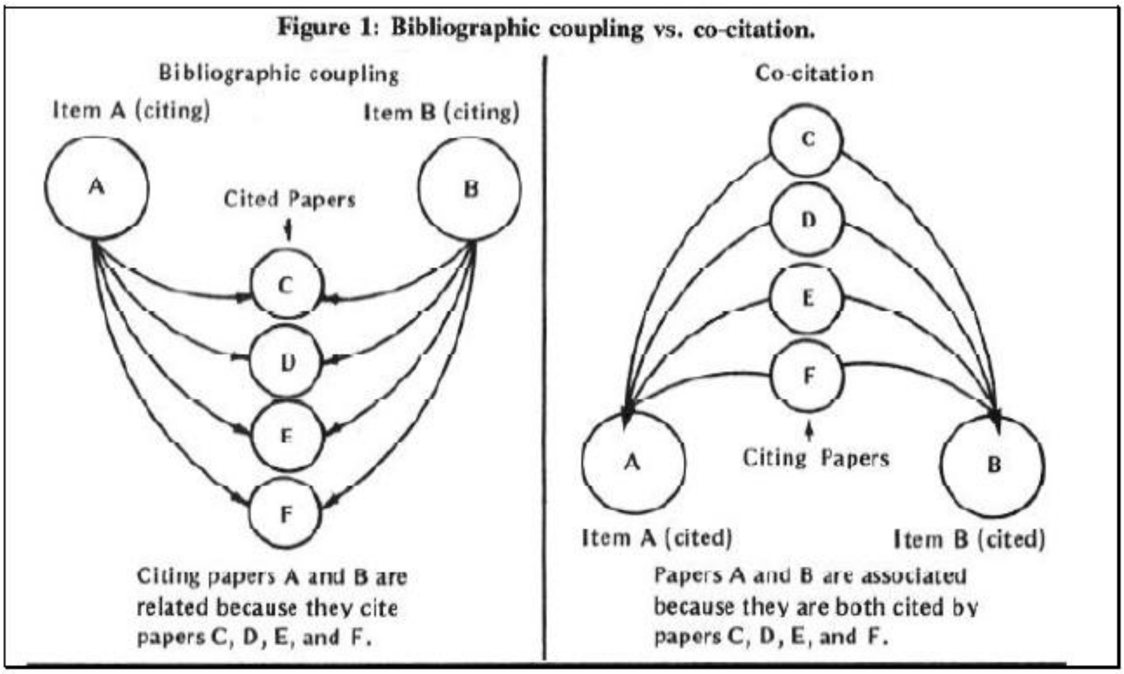


Figure 8: Bibliographic Coupling VS Co-citation analysis.
Source: Garfield, 2001, p. 3.

Small's co-citation model is based on Kuhn's intuitions. In the article "A Co-Citation model of a scientific specialty" (Small, 1977) he explained that the assumption for a citation model of specialty development is that a change in the cognitive and/or social state of a specialty corresponds to a change in its citation patterns. In *The structure of scientific revolutions* Kuhn wrote:

> […] if I am right that each scientific revolution alters the historical perspective of the community that experiences it, then that change of perspective should affect the structure of post revolutionary textbooks and research publications. One such effect - a shift in the distribution of the technical literature cited in the footnotes to research reports - ought to be studied as a possible index to the occurrence of revolutions (Kuhn, 1962/1970, p. XI).

According to Small, co-citation analysis is a tool thanks to which it is possible to explore scientific revolutions or conceptual shifts quantitatively: highly cited and co-cited papers in a cluster would be a concrete representation of the cognitive structure, subject space, or "current paradigm" of the specialty; authors of the highly cited papers would constitute the leading scientists of the specialty (Small, 1977). Thus, co-citation analysis measures the extent to which researchers employ, and thus refer to, the same prior literature. Actually, Kuhn's reaction to the first papers "on the cocitation structure of Science was one of puzzlement" (Small, 2003, p. 395). Kuhn wondered why White and collaborators focused only on highly cited items (papers and authors) instead of considering the whole paradigm-sharing community (i.e. all authors and/or all documents). According to him, highly cited elements were not able to represent paradigmatic constructs such as exemplars (Small, 2003).

Merton's Sociology of Science is fundamental in Small's approach. As he explains in the article "On the shoulders of Robert Merton" (Small, 2004):

> I recall one […] encounter with Merton in the early 1970s when I was struggling with the methodological decision of whether to focus my energies on co-words or co-citations as the structural building blocks for science. The topic of discussion was his memoir "On the Shoulders of Giants" (Merton, 1965) and it occurred to me that the use of the plural of giant suggested that a scientist could in effect be standing on the shoulders of many giants, and that the aggregate pattern of giant straddling was equivalent to the co-citation idea I had been toying with. The further implication was that the resulting groups of giants rubbing shoulders, as it were, would be a structure of the co-allocation of rewards […] Thus [...] the structure of science is generated by patterns of co-recognition. I am sure that this insight strengthened my resolve to pursue the co-citation rather than the co-word route, and led me to the work of Gene Garfield and the field of "citationology". Later on, reading Merton on originality and priority in science, I realized how these norms could be expressed in the choices scientists make when citing the literature. When scientists agree on what constitutes prior relevant literature, including what is

significant in that literature, they are in fact defining the structures of their communities (Small, 2004, pp. 71,72,73).

Also Small's collaborators share this "passion" for Merton:

> Belver Griffith, my late collaborator on mapping science, was fond of saying that this or that sociologist of science, mainly of the constructivist persuasion, was in need of a strong dose of Robert Merton. His prescription was: take two Mertons and see me in the morning (Small, 2004, p. 72).

The potentially of co-citation studies in reconstructing the history of scientific fields was already illustrated in 1964 by Garfield, Sher, and Torpie when, using citation data to study the historical development of works on the genetic code, they obtain what Garfield called an "historical map" of the field (Garfield, Sher, & Torpie, 1964). Nowadays CA is one of the most used techniques to map the structures and dynamics of scientific development. However, its validity and usefulness for science policy purposes has become subject of recent studies and debates (Hicks, 1987; Franklin, 1988; Hicks, 1988). King (1987) summed up a number of problems related to co-citation analysis, such as loss of relevant papers, inclusion of non-relevant papers, over-representation of theoretical articles, time lag (between emergence of new specialties and the detection of them in a co-citation map) and subjectivity with reference to threshold levels strongly affecting both size and content of clusters. Furthermore, interpretation of results is considered problematic mainly with reference to the following question: to what extent can we claim that what represented is the cognitive structure of specialties instead that the social structure of research? (Callon, Courtial, Turner, & Bauin., 1983; Rip & Courtial, 1984; Rip, 1988). Moreover, Sullivan, White, and Barboni (1977) asked if entire specialties or just part of them are represented and Rip (1988) doubts that, as claimed by Small (1973), authors in clusters constitute the relevant scholars in a specialty, while, according to him, just subgroups with shared legitimatory tactics are represented. As Braam, Moed, and Van Raan (1991) claimed:

> Of course, research specialties lacking "focused" referencing can never be identified using cocitation analysis. Thus, in as far as co-citation clusters identify research specialties, it is clear that they do so only partially, and probably only cover the part where consensus exists about important or useful earlier contributions (Braam et al., 1991a, p. 249).

Further limits and problems are linked to the fact that co-citation analyses are mainly conducted by using SCI data, so results can be affected by SCI shortcomings[43] (Glänzel, 2003). In order to overtake

[43] See Section 3.2.2.

some of the above-mentioned problems, Braam et al. from 1988 to 1990 conducted a number of studies[44] from which they derived that:

> [...] in order to generate significant results in the field of "Mapping of science," it will be necessary to analyze different structural aspects of publications in combination in a quantitative fashion instead of the exclusive use of single aspects (Braam et al., 1991a, p. 250).

4.4. AUTHOR CO-CITATION ANALYSIS

Author Cocitation Analysis (ACA) was introduced by White and Griffith in 1981 with the article "Author co-citation: A literature measure of intellectual structure" (White & Griffith, 1981). As explained by White and McCain:

> Author co-citation (ACA) is the subcategory of co-citation analysis that maps oeuvres, and by implication, the people who produce them. The raw data are counts of the times that selected author pairs are cited together in articles, regardless of which of their works are cited (White & McCain, 1997, p. 327).

In ACA co-citation frequencies inform about the strength of the intellectual link among authors; the more two scholars are co-cited, the stronger their intellectual relationship. Therefore, maps based on author cocitation data are considered a valid representation of the intellectual structure in Science domains (McCain, 1986).

ACA was described in technical details by White (1986) in terms of cocited authors retrieval and by McCain (1990) in terms of cocitation mapping. Over time different authors contributed to the improvement of this method (see, for example, Leydesdorff, 2005a; Leydesdorff, 2008; Leydesdorff & Vaughan, 2006; Persson, 2001; White, 2003) and a debate on the normalisation of data took place in the early 2000s (Ahlgren, Jarneving & Rousseau, 2003, 2004; Bensman, 2004; White, 2003)

Even if, as we have just seen, there is no agreement on some technical aspect, it seems that all scholars active in ACA agree about the interpretation of co-citation maps. In this regard, White in 1990 compiled a list about the basic elements of a co-cited author interpretation:

- Author maps reveal the "cognitive" or "intellectual structure" of a field by showing the consensus of citers to important contributors and works;
- Author maps show who is central and who is peripheral to a field;

[44] Braam, Moed, & Van Raan, 1988; Braam, Moed, & Van Raan, 1989; Braam et al., 1991a,b.

- The maps show who is central and who is peripheral within clusters representing specialties or schools of thought;
- The maps show broad dimensions on which clusters are arranged. Usually one of these can be interpreted as a subject dimension and "style of work" dimension. A qualitative-quantitative polarity often appears;
- The knowledgeable interpreter of a map may see much to explicate in the fine structure of author points: for example common nationality, temporal conjunctions, teacher-student relationships, collegial and co-author relationships, or common philosophical orientations (White, 1990, p. 103).

ACA is usually performed by using data retrieved from databases, thus results can be affected by databases shortcomings. For example, one of the main problems relates to a rule established by Thomson Reuters is the following: if a reference list include co-authored publications, it is possible to retrieve only the information relating to the first authors. This led to the development of the technique called *first author co-citation* (White & Griffith, 1982), which, therefore, develops maps including only a part of the authors of the co-cited publications considered in the analyses. In order to overtake this limit, Olle Persson in 2001 developed the *all author* form:

> [...] all author citation counts should be preferred when visualizing the structure of research fields. *First author citation* studies distort the picture […] Supporters of *(first) author co-citation* studies may object to such a criticism by saying that the aim is not to rank authors but rather to identify research themes, and as representing such themes first cited authors might suffice. Well, that may very well be the case, but then one should also test if the structure of author co-citation maps is dependent on the counting method used (Persson, 2001, pp. 339, 343).

It is worth to underline that, unlike the first-author method, the all-author version also includes the strong intellectual link among co-authors. This latter is an important factor that allows researchers to develop clearer maps characterised by a high cohesion level of the specialties represented in them (Zhao, 2006).

In 2004 Rousseau and Zuccala developed a classification of ACA variants (Rousseau & Zuccala, 2004). Specifically, we can distinguish among: *pure first-author cocitations* (it occurs when the reference list of a publication includes at least one publication with A as the first or sole author and at least one publication with B as the first or sole author; all articles co-authored by A and B are excluded in the computation of co-citation frequency); *pure cocitations* (in this case at least one publication with A as a co-author - regardless of rank order - and one publication with B as a co-author - regardless of order - must co-occur in the reference list of an article; articles co-authored by A and B as co-authors are not taken into account.); *general co-citations* (at least one publication with A as a co-author, and one – additional- publication with B as a co-author co-occur in the same reference list. If A and B only occur in a reference list as co-authors their co-authored articles is not included in the computation); *co-authors/cocitations* (in addition to the first three co-citation forms, a publication with A and B as co-authors can be included in the co-citation computation, in this case the intellectual link between co-publishing authors is recognised) (Rousseau & Zuccala, 2004). Figure 9 below furnishes a graphical representation of the classification of ACA forms:

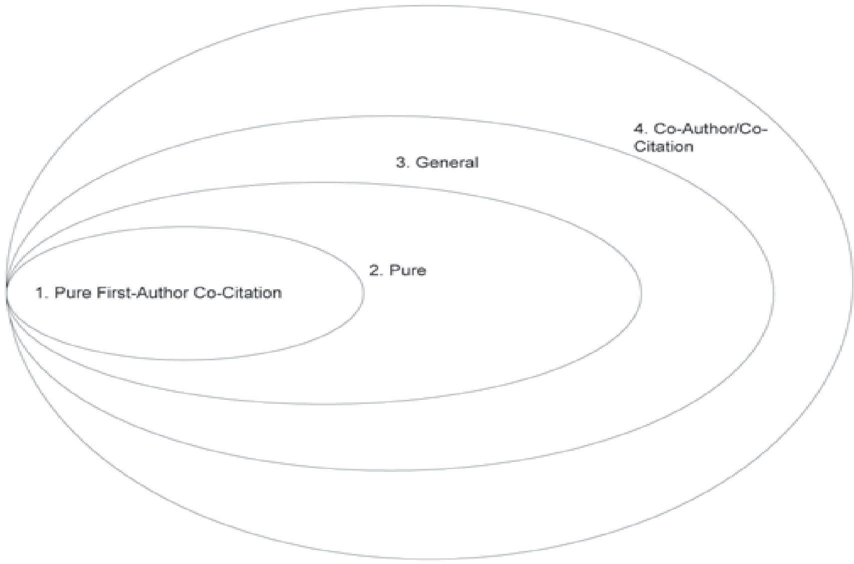


Figure 9: Classification hierarchy of author co-citation forms.
Source: Rousseau & Zuccala, 2004, p. 516.

4.5. TEXTUAL CO-OCCURENCES ANALYSES

Textual co-occurrences analyses encompass mainly two techniques: *Co-word* and *Co-heading analysis*. The major purpose of these techniques is that of detecting those key-words representative of the cognitive structure of documents (with reference to its content or the title), and thus of disciplines or subfields.

Co-word analysis is a content analysis technique that is effective in mapping the strength of association between keywords in textual data. This method has been developed in France in the 1980s by Callon, Courtial, Turner, and Bauin (Callon et al., 1983). It is based on frequency analyses of co-occurrence of keywords extracted from titles, abstracts or text, in general. Co-word clustering is a standard technique at ISI and it is often used in association with other techniques (see, for example, Braam et al., 1989) (Glänzel, 2003).

Co-heading analysis was developed by Todorov in 1990. It is based on the co-occurrence of subject headings used in specified bibliographic databases. As Todorov warned, this technique has some limitations and it should be used together with other methods, such as co-citation and co-word analysis (Todorov, 1990).

4.6. MAPPING SCIENTIFIC LITERATURE

In 1985 Small and Garfield wrote: "The purpose of *our* analysis is to create what we call maps of science which show the topography of science at various levels of aggregation" (Small & Garfield, 1985, p. 147, italics mine). The idea that Science could be mapped was first introduced by Price in 1965 in the article "Networks of scientific papers":

> The total research front of science has never, however, been a single row of knitting. It is, instead, divided by dropped stitches into quite small segments and strips. From a study of the citations of journals by journals I come to the conclusion that most of these strips correspond to the work of, at most, a few hundred men at any one time. Such strips represent objectively defined subjects whose description may vary materially from year to year but which remain otherwise an intellectual whole. If one would work out the nature of such strips, it might lead to a method for delineating the topography of current scientific literature. With such a topography established, one could perhaps indicate the overlap and relative importance of journals and, indeed, of countries, authors, or individual papers by the place they occupied within the map, and by their degree of strategic centralness within a given strip. Journal citations provide the most readily available data for a test of such methods (Price, 1965, p. 515).

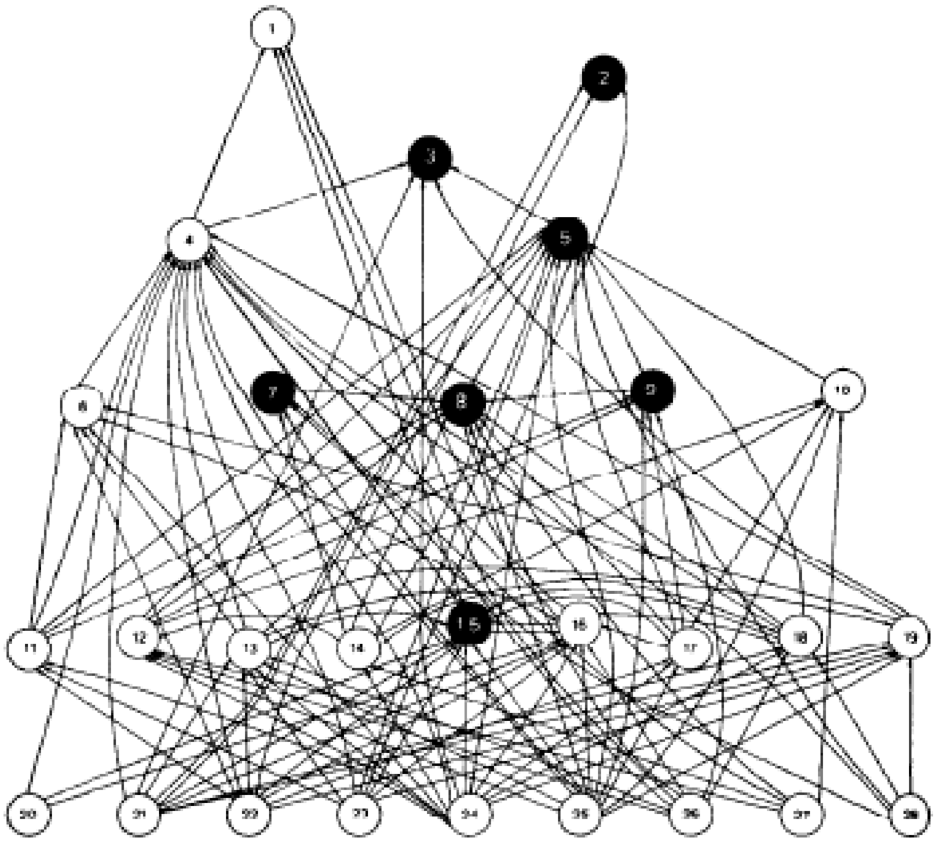
Actually, already in the 1940s we can find precursors; for example, in 1948 Bradford wrote:

> If *x* represents any class, as men, its individuals will have many different qualities. We may separate those individuals, which are distinguishable from one another as having different, in sub-class. The symbols of all sub-classes will have only the possible numerical values 0 and 1. So let us draw straight lines, of unit length, from a convenient point, to represent these symbols. The lines will be distinguished by direction, and, if we like to use a space of three dimensions, the unit lines will terminate in point upon the surface of a unit sphere. The aggregate of points will represent the class of beings men. Similarity let more lines be drawn to represent all we wish to talk about. The aggregate of pints, where all these lines end on the surface of the sphere, represents the universe of discourse [...] And so we get a picture of the universe of discourse as a globe, on which are scattered in promiscuous confusion, the mutually related, separate things we see or think about (Bradford 1948, p. 137).

In the article "Visualizing science by citation mapping" Small (1999) explains why the mapping of scientific literature is useful: it can enhance information usability and retrieval; it can facilitate the understanding of conceptual relationships and developments; it can provide insight into the state of knowledge; it can help in developing a good history of science, or as Small explains "a map of bibliographic data is a useful heuristic device by providing a visible organizing structure to information" (Small, 1999, p. 812). According to him, mapping scientific literature is not only useful but also natural:

> A map of science is a spatial representation of how disciplines, fields, specialties, and individual papers or authors are related to one another as shown by their physical proximity and relative locations, analogous to the way geographic maps show the relationships of political or physical features on the Earth. […] Nevertheless, we find arranging information in space a natural and useful heuristic tool, perhaps because spatial relations play such an important role in everyday experience (Small, 1999, p. 799).

One of the earliest attempts to develop a graph from citation data is that by Garfield in 1979. It consists in the historiography of the major advances in Genetics between 1958 and 1967:



Nodes: 1 Sheehan, 1958; 2 Bray, 1960; 3 Nirenberg, 1961; 4 Marcker, 1964; 5 Nirenberg, 1964; 6 Marcker, 1965;7 Brenner, 1965; 8 Khorana, 1965; 9 Nirenberg, 1965; 10 Khorana, 1965; 11 Marcker, 1966; 12 Khorana, 1966; 13 Marcker, 1966; 14 Khorana, 1966; 15 Adams, 1966; 16 Webster, 1966; 17 Nirenberg, 1966; 18 Ochoa, 1966; 19 Nakamoto, 1966; 20 Berberich 1967; 21 Lucas-Leonard, 1967; 22 Caskey, 1967; 23 Ochoa, 1967; 24 Khorana, 1967; 25 Nirenberg, 1967; 26 Ochoa, 1967; 27 Khorana, 1967; 28 Ochoa, 1967.

Figure 10: Historiography of the major advances in genetics between 1958 and 1967, based on a citation analysis of a review of the 1967 literature. Each circle represents a paper cited five or more times by the papers listed in the bibliography of the review. The papers represented by solid black circles were cited 15 times or more in the 1967 SCI. Source: Garfield, 1979, p. 94.

The historiographical approach allows us to depict papers links forward and backward in time, thus it is possible to trace the lineage of ideas. The introduction of co-citation thanks to Small and Marshakova's intuitions (Marshakova, 1973; Small, 1973) allowed the creation of maps of Science whose building blocks are highly cited papers and their co-citation frequencies. Co-citation maps are used for showing microstructures and macrostructures of scientific domains. Since the beginning Multidimensional Scaling (MDS) (Kruskal & Wish, 1978) has been used for mapping. The first computer visualisation of such networks was by Yermish (1975), who employed a one-dimensional multidimensional scaling to determine documents and publications position along horizontal and vertical axis. Two-dimensional maps were used to show Science structure; co-citations provided a coefficient of similarity among documents, and a measure to determine distances among objects. Furthermore, thanks to the application of Cluster Analysis (Tryon, 1939) high inter-citation regions could be collapsed obtaining a network of super-nodes (Griffith et al., 1974; Small & Griffith, 1974). In 2001 Chen and Paul presented a three-dimensional spatial model for visualising a knowledge domain intellectual structure (Chen & Paul, 2001).[45]

Cocitation maps are usually characterised by areas with different density levels. According to Small and Griffith (1974), high density areas correspond to subject areas or specialties. Results furnished by ACA are considered a proof of the fact that Science is a mosaic of specialities as well as a proof of the reliability of the co-citation technique. Outcomes are also in line with claims about the centrality of few key scientists and few research centres in specialty organisation (Crawford, 1971).

Besides mapping specialties by focusing on strong ties (highly co-cited documents), in 1974 Griffith, Small, Stonehill, and Dey decided to focus on weak links in order to relate the clusters in an "overall picture of scientific literature" (Griffith et al., 1974, p. 341). Results led them to affirm that:

> We [...] confirmed the notion that science and its literature can be conceived as a network of specialties, each specialty being the centre of a highly interactive, intense, communication system (Griffith et al., 1974, p. 341).

[45] A review of mapping and clustering techniques is given in Chapter 5.

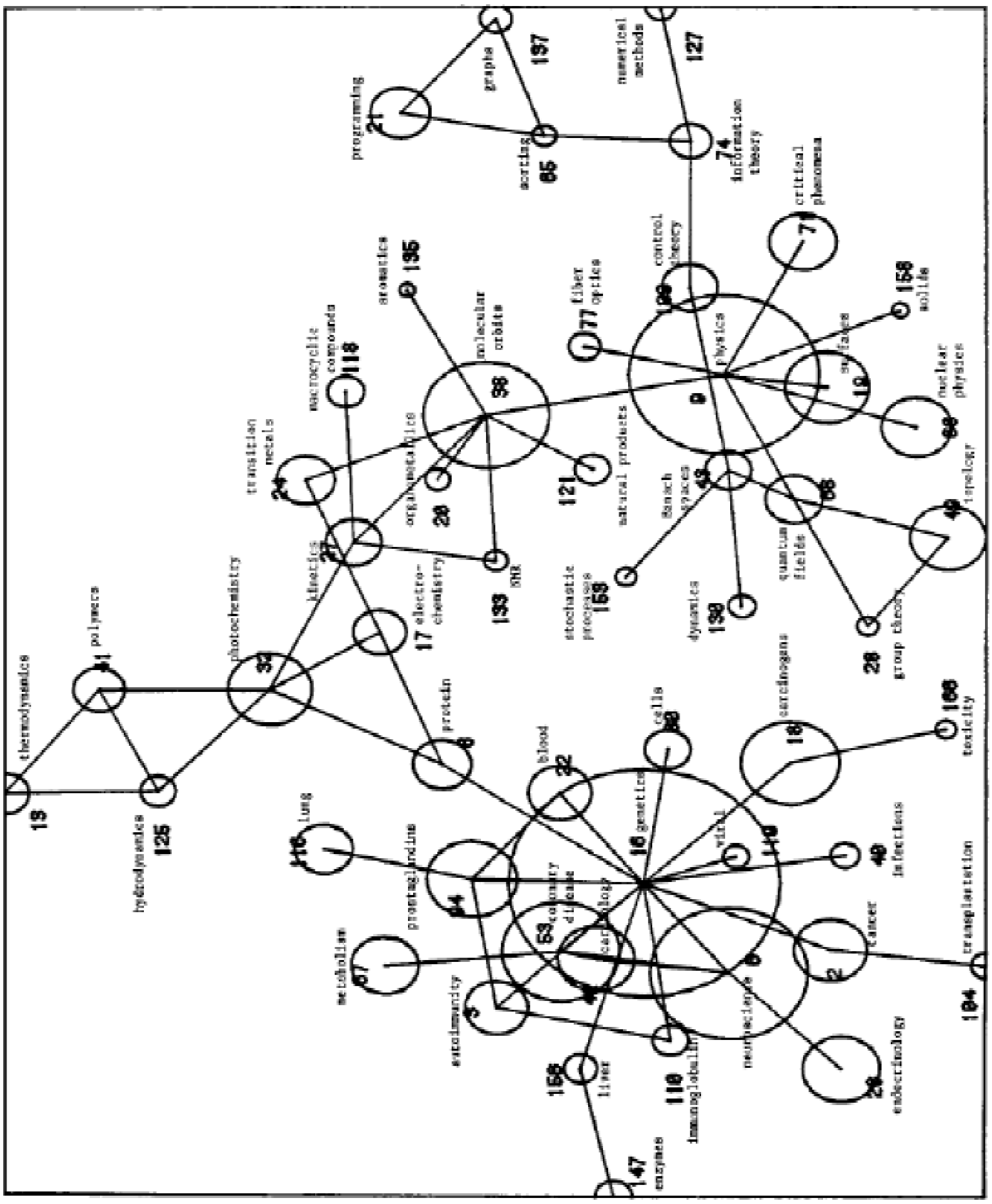


Figure 11: An early co-citation map showing the major disciplines of the Natural Sciences: Biology, chemistry and physics. Data are from a five-level co-citation analysis. Multidimensional scaling was used to position macro-clusters.
Source: Small & Garfield, 1985, p. 154.

5. CLUSTERING AND MAPPING BIBLIOMETRIC NETWORKS

Relations became a sociological object of study with Georg Simmel. According to him, relations constitute Society and individuals are sociological subjects as they are placed in a complex relational context: a society exists where a group of individuals are in mutual action. Simmel conceives of Society as nothing but the result of the experiences of individuals mutually involved; it does not have a proper consistence as it is a *process* (Simmel, 1950). Relational bibliometrics is based on the assumption that relations constitute the changing structures of Society: by studying relationships patterns in scientific literature is possible to detect the intellectual structure of Science (White & Griffith, 1981).

As previously mentioned, the idea of analysing citation networks arose almost together with the SCI. According to Small, maps of Science are as natural and useful as geographical ones (Small, 1999).[46] The possibility to represent Science spatially attracted the second generation of information scientists whose efforts were directed towards the visualisation of scientific disciplines. Debates and numerous improving attempts have been characterising the field of mapping scientific literature in the 21th century: in recent years a debate concerning the appropriate similarity measures to evaluate the proximity of items arose (Ahlgren et al., 2003, 2004; Bensman, 2004; Leydesdorff, 2008; Leydesdorff & Vaughan, 2006; White, 2003a); efforts have been made in order to improve both the unfolding community methodology (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008; Boyack, Börner, & Klavans, 2009; Chen, 2004; Klavans & Boyack, 2006; Schneider & Borlund, 2007a/b; Small, 2006) and information visualisation techniques (Chen, 1997; Chen, 1999a; Chen, 1999b).

With the purpose of furnishing a review of the mapping and clustering techniques used in the science mapping domain (methods are, thus, discussed with reference to bibliographic networks), this Chapter is organised as follows. In the first Section the most important methods for uncovering clusters in co-citation networks are introduced. Next, an overview of the most used dimensionality-reduction techniques is given. Finally, a unified approach for mapping and clustering bibliometric networks is illustrated.

5.1. GRAPH CLUSTERING: UNFOLDING COMMUNITIES

In literature there are different graph clustering techniques as scholars both disagree on what a cluster (or a module) is and technique suitability depends on graphs complexity and research targets. Therefore, questions such as “under which conditions do we have a cluster? What is the level of connectivity required? What is the most suitable cut size? How to measure the clustering quality? Are overlaps conceived?” are attracting more and more scholars and in particular physicists, who have been showing an increasing interest in unfolding communities in complex networks.

However, scholars agree on a loose definition of cluster. A good cluster is defined by high inter-cluster density (within the subgraphs) and low intra-cluster density (among the subgraphs) (Girvan & Newman, 2002; Newman, 2004b; Palla, Derényi, Farkas, & Vicsek., 2005). In particular:

[46] See page 74.

Globally speaking, the internal density[47] of a good clustering should be notably higher than the density of the graph (*G*) and the intercluster density of the clustering should be lower than the graph density (Schaeffer, 2007, p. 33).

Figure 12 below represents graphically what has been said in the above passage:

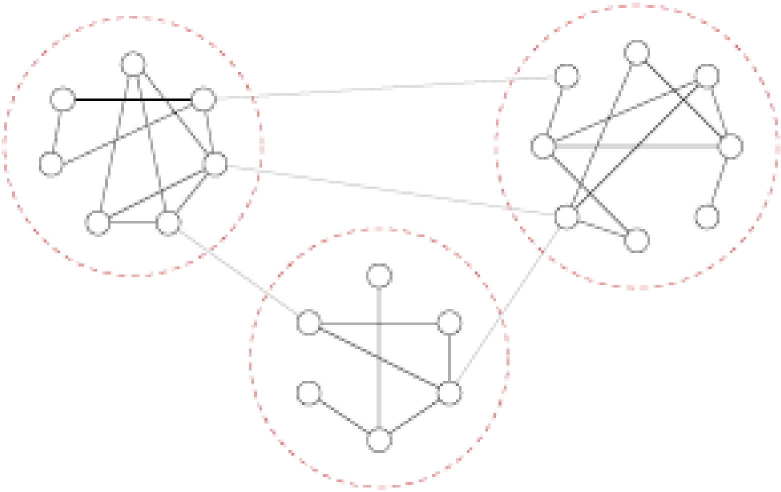


Figure 12: A small network made of three communities denoted by the dashed circles, which have dense internal links but between which there are only a lower density of external links.
Source: Newman, 2004a, p. 321.

In this Section the main methods for unfolding communities are introduced, namely Cluster Analysis, Concor and Blockbusting, Modularity, Clique Percolation, Map Equation, and the Louvain method. The aim is that of illustrating the improvements made over time with respect to this research field, therefore these methods are introduced in chronological order.

5.1.1. CLUSTER ANALYSIS

The term "Cluster Analysis" (CA) was first used by Tryon (1939) in the late 1930s. It encompasses a number of different algorithms aiming to discover natural division of networks into groups (clusters), on the basis of similarity matrices or the connection strength between vertices.

The purpose of this method is to maximise intra-cluster similarity and minimise inter-cluster similarity. Thus, clusters are supposed to be highly internally homogeneous but highly externally heterogeneous. In 2003 Börner, Chen, and Boyack (2003) tried to solve the problem relating to the

[47] The density of a subgraph is defined as the number of lines in the subgraph divided by the number of lines that could be present in the subgraph (Wasserman & Faust, 1994, p. 102).

selection in case of alternative partitions. They introduced the concept of *utility measure* "that contrasts the sum of within cluster similarities *(wSim)* with the sum of between-cluster similarities *(bSim)*" (Börner et al., 2003, p. 206):

$$Utility = wSim / (wSim + bSim) \tag{6}$$

the partition showing the highest utility is selected.

Cluster Analysis encompasses more than 150 techniques that can be divided in *agglomerative* and *divisive* algorithms (actually, according to the criterion chosen, other classifications are possible)[48] depending on whether they work adding or removing edges to or from the network. In the *agglomerative method* similarities are calculated between vertex pairs. For each pair of vertices a *weight* representing their connection strength is calculated, and, starting from a network of only vertices (*n* vertices and no edges), edges are iteratively added starting with the vertex pairs with the highest similarity. The procedure can be halted at any point, and the resulting components in the network are taken to be the communities. The *divisive method*, instead, starts with the whole graph and iteratively cuts weak edges (it starts from the least similar connected vertex pairs), thus dividing the network progressively into smaller and smaller disconnected sub-networks. Also in this case it is possible to stop the procedure at any time (Girvan & Newman, 2002; Newman & Girvan, 2004). Both techniques have shortcomings. The main limit of the agglomerative method is the tendency to find only the cores of communities and leave out the periphery. The core nodes in a community often have strong similarity, hence they are connected early in the agglomerative process, but peripheral nodes, which have no strong similarity to others, tend to be neglected. The crucial point in a divisive algorithm is the selection of the edges to be cut: obviously, they must be those connecting communities and not those within them (ibid).

In 2002 Girvan and Newman (2002) developed a divisive algorithm called *Girvan-Newman* (GN). This method differs from the previous ones because "instead of trying to construct a measure that tells us which edges are most central to communities, we focus instead on those edges that are least central, the edges that are most between communities" (Girvan & Newman, 2002, p. 7822). The GN algorithm is based on the *edge betweenness*[49] concept derived from the generalising of Freeman's betweenness centrality (Freeman, 1977):

> The betweenness centrality of a vertex *i* is defined as the number of shortest paths between pairs of other vertices that run through *i*. It is a measure of the influence of a node over the flow of information between other nodes […]. To find which edges in a network are most between other pairs of vertices, we generalize Freeman's betweenness centrality to edges and define the

[48] See Börner et al., 2003.

[49] The "edge betweenness" concept was first developed by Anthonisse in an unpublished work dated 1971 where it was called "rush" (Anthonisse, 1971).

> edge betweenness of an edge as the number of shortest paths between pairs of vertices that run along it. If there is more than one shortest path between a pair of vertices, each path is given equal weight such that the total weight of all of the paths is unity. If a network contains communities or groups that are only loosely connected by a few intergroup edges, then all shortest paths between different communities must go along one of these few edges. Thus, the edges connecting communities will have high edge betweenness. By removing these edges, we separate groups from one another and so reveal the underlying community structure of the graph (Girvan & Newman, 2002, p. 7822).

Girvan and Newman tested the method on both real and artificial networks whose structure was both already known and unknown; it was successful in extracting structures in all cases considered (Girvan & Newman, 2002). However, the algorithm had two main limits: 1) it was computationally costly as it run in time $O(n^3)$ in sparse graphs (where n is the number of vertices), therefore also working with moderately large networks was unfeasible (Girvan & Newman, 2002); 2) without a previous knowledge on the nature of the network it was difficult to state if the identification of a community was reliable (Newman & Girvan, 2004).[50] Various efforts have been made in time to overcome the shortcomings of the GN algorithm, such as those by Newman (2004b, 2006), Newman and Girvan (2004), Radicchi, Castellano, Cecconi, Loreto, and Parisi (2004), and Wilkinson and Huberman (2004).

5.1.2. CONCOR AND BLOCKMODELING

In the 1970s an approach based on the concept of equivalence classes (positions) for decomposing networks into meaningful subgroups was introduced (Breiger, Boorman, & Arabie, 1975; Lorrain & White, 1971).

Even if there is not agreement among social scientists on the concepts of social position, social status and social role, in Network Analysis the term position is unanimously used to refer to a collection of nodes which are similarly embedded in relational networks (Wasserman & Faust, 1994). The concept of class equivalence, or position, was discussed by Lorrain and White (1971) in terms of *structural equivalence*: by conceiving of nodes in a graph as having equivalent positions relating in a similar way to other positions, a network can be transformed into positions and the relations between nodes are, thus, considered as relations between positions. Thus, two nodes are structurally equivalent if they have identical ties to and from all other nodes in a network (Wasserman & Faust, 1994).

There are different methods for partitioning nodes into subsets so that actors within each subset are closer to being equivalent than actors in different subsets. Among the most used methods there are Hierarchical Clustering, CONCOR and Blockmodeling.[51]

[50] As stressed by Radicchi et al. (2004) there are networks which have not a community structure.
[51] For a discussion on partitioning methods see Wasserman and Faust (1994).

CONCOR is one of the earliest approaches to partitioning (or clustering) into positions based on the structural equivalence principle; the term CONCOR stands for "convergence if iterated correlations". The procedure is applied to sociomatrices transformed in correlation ones (by calculating correlations among rows and/or columns) and correlations are used as structural equivalence measures. The correlation matrix becomes the basis for a further correlations calculation process producing a third correlation matrix which will be used to calculate again correlations, and so on. The procedure goes on until to obtain a correlation matrix with +1s and -1s occurring in a pattern such that correlated items can be partitioned into two subsets where all correlations between items assigned to the same subset are equal to +1 and all correlations between items in different subsets will be equal to -1. One of the main problems with CONCOR is the decision about stopping the partitioning. Furthermore, as Hierarchical Clustering, it is a discrete method resulting in mutually exclusive and exhaustive subsets (Wasserman & Faust, 1994).

Blockmodeling encompasses various partitioning techniques based on the so called "blockmodels" introduced by White, Boorman, and Breiger (1976).[52] A blockmodel is a model or a hypothesis about a multirelational network. It has two components: the mapping describing the assignment of positions (there are a number of ways to do it, including CONCOR and HC); the image matrix specifying the absence or presence of links between and within positions on each relation. Each node can be assigned only to one position and the assignment remains invariant across relations. A *block* is each of the entries in the image matrix; a block containing the value 1 (presence of link from the row position to the column position) is called *oneblock*, while the block containing the value 0 (indicating the absence of the link) is called *zeroblock* (Wasserman & Faust, 1994). In the early 2000s Doreian, Batagelj, and Ferligoj (2005) developed a new approach for identifying equivalent positions: *Generalized Blockmodeling* (GBM). It considers ideal blockmodels and employs optimisation methods to fit them to empirical data. Its advantage is that it allows the use of context information for hypotheses formulation and gives a criterion function (i.e. inconsistencies) which measures the fit of a specified blockmodel to the actual data (Hsieh & Magee, 2008). Generally speaking, problems related to Blockmodeling mainly refer to the partitioning technique employed and to the method used in defining whether or not there are links between and within positions. Finally, subjectivity plays a fundamental role in selecting the way in which a blockmodel is interpreted (validation using node attributes; description of node position; description of the overall blockmodel) (Wasserman & Faust, 1994).

[52] For deepening the topic, see Wasserman and Faust (1994).

5.1.3. MODULARITY

Newman and Girvan (2004) developed a measure to quantify the strength of communities' structure, namely *modularity:*

> Consider a particular division of a network into k communities. Let us define a k X k symmetric matrix **e** whose element e_{ij} is the fraction of all edges in the network that link vertices in community i to vertices in community j. […] The trace of this matrix Tr $\mathbf{e} = \sum_i e_{ii}$ gives the fraction of edges in the network that connect vertices in the same community, and clearly a good division into communities should have a high value of this trace. The trace on its own, however, is not a good indicator of the quality of the division since, for example, placing all vertices in a single community would give the maximal value of Tr $\mathbf{e} = 1$ while giving no information about community structure at all. So we further define the row (or column) sums $a_i = \sum_i e_{ii}$ which represent the fraction of edges that connect to vertices in community i. In a network in which edges fall between vertices without regard for the communities they belong to, we would have $e_{ij} = a_i a_j$. Thus we can define a modularity measure by:
>
> $$Q = \sum_I (e_{ii} - a_i^2) = \text{Tr } \mathbf{e} - \|\mathbf{e}^2\|$$
>
> where $\|\mathbf{e}\|$ indicates the sum of the elements of the matrix x (Newman & Girvan, 2004a, p. 7).

According to Girvan and Newman, the "minimum cut approach", which is based on the minimisation of the number of edges among clusters as criterion to detect communities, cannot be accepted anymore mainly because in many cases there is not previous knowledge of the community structure (Newman & Girvan, 2004; Newman, 2006). "A good division of a network into communities is not merely one in which there are few edges between communities; it is one in which there are fewer than expected edges between communities" (Newman, 2006, p. 8578).

The modularity can be either positive or negative (between -1 and +1); positive values indicate the presence of a community structure. According to Newman, on the basis of various successful applications of the algorithm (see, for example, Guimerà & Amaral, 2005; Danon, Díaz-Guilera, Duch, & Arenas, 2005), maximisation of the modularity is the best method of community detection (Newman, 2006). In 2004 Newman expanded the technique to weighted graphs (Newman, 2004c):

$$Q = \frac{1}{2m}\sum_{ij}(A_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j) \qquad (7)$$

where A_{ij} represents the weight of the edge between i and j, $k_i = \sum_j A_{ij}$ is the sum of the weights of the edges attached to vertex i, c_i is the community to which vertex i is assigned, the δ function $\delta(u, v)$ is 1 if $u = v$ and 0 otherwise and $m = ½ \sum_{ij} A_{ij}$ (Newman, 2004c).

5.1.4. CLIQUE PERCOLATION

The Clique Percolation Method (CPM) was introduced in 2005 by Palla, Derényi, Farkas, and Vicsek with the article "Uncovering the overlapping community structure of complex networks in nature and society" (2005). The original version of the algorithm had the target to detect the k-clique communities of unweighted and undirected networks, but in 2007 it was expanded to directed and weighted networks as well (CPMw[53]) (Farkas, Ábel, Palla, & Vicsek, 2007).

CPM was developed as an alternative to all deterministic methods of the time producing separated communities when partitioning: real networks are made of highly overlapping cohesive groups of nodes. As the developers of this method explained:

> In general, each node *i* of a network can be characterized by a membership number m_i which is the number of communities that the node belongs to. In turn, any two communities α and β can share $s^{ov}_{\alpha,\beta}$ nodes, which we define as the overlap size between these communities. Naturally, the communities also constitute a network, with the overlaps being their links. The number of such links of community a can be called its community degree, d^{com}_{α}. Finally, the size s^{com}_{α} of any community a can most naturally be defined as the number of its nodes. Thus, we define a community, or more precisely a *k*-clique community, as a union of all *k*-cliques (complete subgraphs of size *k*) that can be reached from each other through a series of adjacent *k*-cliques (where adjacency means sharing *k* - 1 nodes) (Palla et al., 2005, p. 814).

In Figure 13 below the concept of k-clique community is shown:

[53] "w" stands for weighted.

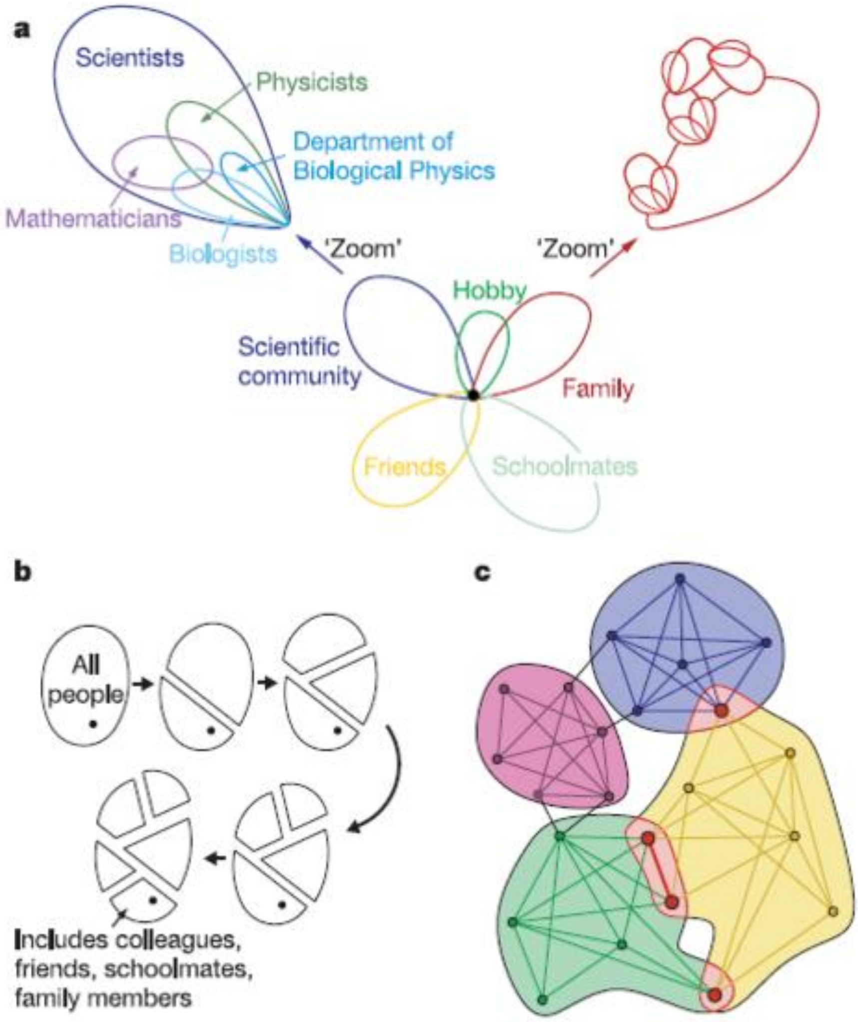


Figure 13: Illustration of the concept of overlapping communities. (a) The black dot in the middle represents either of the authors of a paper *i*, with several of his communities around. Zooming in on the scientific community demonstrates the nested and overlapping structure of the communities, and depicting the cascades of communities starting from some members exemplifies the interwoven structure of the network of communities. (b) Divisive and agglomerative methods grossly fail to identify the communities when overlaps are significant. (c) An example of overlapping k-clique communities at k=4. The yellow community overlaps the blue one in a single node, whereas it shares two nodes and a link with the green one. Those overlapping regions are emphasized in red. Notice that any k-clique (complete subgraph of size k can be reached only from the k-cliques of the same community through a series of adjacent k cliques. Two k-cliques are adjacent if they share k – 1 nodes.
Source: Palla et al., 2005, p. 2.

First, all cliques (maximal complete subgraphs) in the network are identified, subsequently a standard *component analysis* of the clique–clique overlap matrix is performed and communities are detected (Palla et al., 2005). With reference to weighted networks, the correlation between link weights can be studied: two links are adjacent if they have one node in common, and link weights are *assortative* (or *disassortative*) in a network if the weights of neighbouring links are correlated (or anti-correlated). Focusing on triangles, *assortativity* of link weights in triangles (with nodes *i*, *j* and *k*) can be assessed by measuring the weight of a link, $w_{i,j}$, as a function of the geometric mean of the other two links' weights $w_{j,k}$ and $w_{i,k}$:

$$w_{i,j} = F\ (\ [w_{i,k}\ w_{j,k}]^{1/2}\ ) \tag{8}$$

F is an increasing (or decreasing) function if the link weights in a triangle are similar (or very different). This definition is closely related to the *intensity*, *I(g)*, of a subgraph, *g*, defined as the geometric mean of its link weights; the intensity of a subgraph is equal to the geometric mean of its link weights. A *k*-clique is included into a module only if it has intensity larger than a fixed threshold value, *I* (ibid).

5.1.5. MAP EQUATION

The method called Map Equation was developed by Rosvall and Bergstrom (2007) with the purpose of detecting and describing "[...] the dynamics across the links and nodes in directed, weighted networks that represent the local interactions among the subunits of a system" (Rosvall & Bergstrom, 2007, p. 1118). Rosvall and Bergstrom had in mind the concept of "information flow", thus they considered links among nodes as channels of information flows:

> These local interactions induce a system-wide flow of information that characterizes the behavior of the full system. Consequently, if we want to understand how network structure relates to system behavior, we need to understand the flow of information on the network (Rosvall & Bergstrom, 2007, p. 1118).

On the basis of these presumptions, Rosvall and Bergstrom developed a way to identify the modules (clusters) of a network by finding an efficient description of how information flows in the network. By referring to Shannon's information theory (1948) and Huffman coding (1952), the algorithm identifies groups of nodes among which information flows quickly, thus they are aggregated in a module that is weakly connected with the others (or, more generally, with the rest of the network) by links representing the avenues of information flow between them (Rosvall & Bergstrom, 2007).

The information flows are compressed in codes by following compression/coding rules which exploit regularities in the process generating the stream of information (Shannon, 1948). A "random walk" is used as a proxy for the information flow. As a random walk uses all of the information in the network representation and nothing more:

> Taking this approach, we develop an efficient code to describe a random walk on a network. We thereby show that finding community structure in networks is equivalent to solving a coding problem (Rosvall & Bergstrom, 2007, p. 118)

Map Equation employs a variant of Shannon's code assignment, which ensures, in addition to compression, that codewords given to important structures in the networks are unique. In order to give names to nodes a Huffman code is used: "Huffman codes save space by assigning short codewords to common events or objects and long codewords to rare ones, much as common words are short in spoken languages" (Rosvall & Bergstrom, 2007, p. 1118). A second step is to create a map; in order to do this the network is divided into two levels of description:

> We retain unique names for large-scale objects, the clusters or modules to be identified within our network, but we reuse the names associated with fine-grain details, the individual nodes within each module. This is a familiar approach for assigning names to objects on maps: most U.S. cities have unique names, but street names are reused from one city to the next, such that each city has a Main Street and a Broadway and a Washington Avenue and so forth. The reuse of street names rarely causes confusion, because most routes remain within the bounds of a single city. A two-level description allows us to describe the path in fewer bits than we could do with a one-level description. We capitalize on the network's structure and, in particular, on the fact that a random walker is statistically likely to spend long periods of time within certain clusters of nodes (Rosvall & Bergstrom, 2007, p 1119)

The procedure is illustrated in Figure 14:

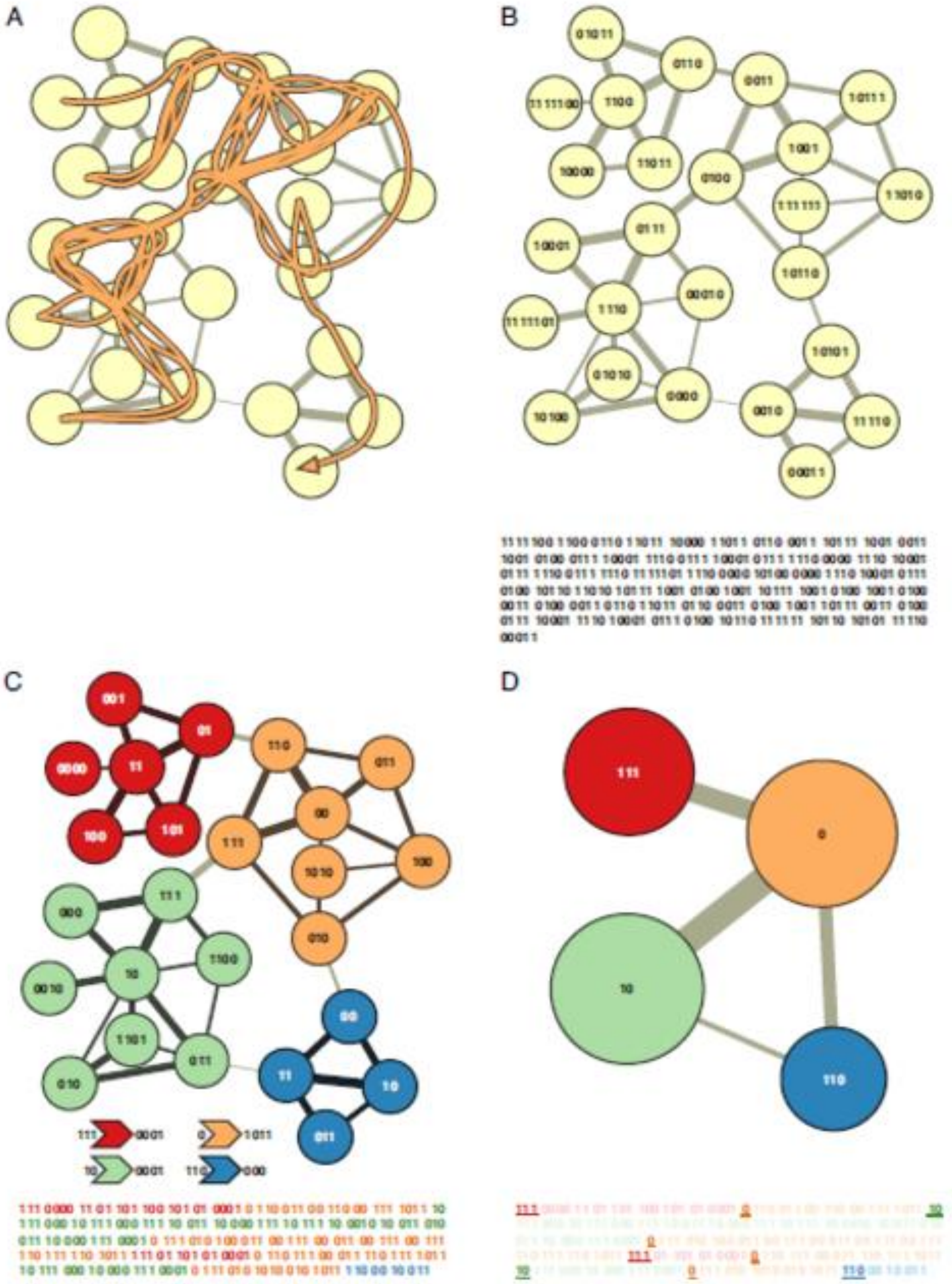


Figure 14: Detecting communities by compressing the description of information flows on networks. (A) We want to describe the trajectory of a random walk on the network such that important structures have unique names. The orange line shows one sample trajectory. (B) A basic approach is to give a unique name to every node in the network. The Huffman code illustrated here is an efficient way to do so. The 314 bits shown under the network describe the sample trajectory in A, starting with 1111100 for the first node on the walk in the upper left corner, 1100 for the second node etc., and ending with 00011 for the last node on the walk in the lower right corner. (C) A two-level description of the random walk, in which clusters receive unique names, but the names of the nodes within clusters are reused, yields on average a 32% shorter description for this network, The codes naming the modules and the codes used to indicate an exit from each module are shown to the left and the right of the arrows under the network, respectively. Using this code, we can describe the walk in A by the 243 bits shown under the network in C. The first

three bits 111 indicate that the walk begins in the red module, the code 0000 specifies the first node on the walk, etc. (D) Reporting only the module names, and not the locations within the modules, provides an efficient coarse graining of the network.
Source: Rosvall & Bergstrom, 2007, p. 1119.

Below the algorithm developed to detect a module partition M of *n* nodes into *m* modules minimising the expected description length of a random walk is introduced; the average description length of a single step is given by:

$$L(M) = q_{\curvearrowright} H(2) + \sum_{i=1}^{m} p_{\circlearrowright}^{i} H(P^{i}) \qquad (9)$$

Considering that the term "entropy" in Shannon's information theory relates to the amount of information contained in a message in contrast to the portion of the message that is determined or predictable (Shannon, 1948), the equation consists of two terms: the first one is the entropy of the movement between modules; the second is the entropy of movements within modules (exiting the module also is considered a movement). Both are weighted by the frequency with which they occur in the particular partitioning. In particular, $q_{\curvearrowright}$ is the probability that the random walk switches modules on any given step; $H(Q)$ is the entropy of the module codewords; $H(P^i)$ is the entropy of the within-module movements, including the exit code for module *i*. The weight $p_{\circlearrowright}^{i}$ is the fraction of within-module movements that occur in module *i*, plus the probability of exiting module *i* such that $\Sigma_{i=1}^{m} p_{\circlearrowright}^{i} = 1 + q_{\curvearrowright}$ (Rosvall & Bergstrom, 2007).

5.1.6. THE LOUVAIN METHOD

In 2008 Blondel, Guillaume, Lambiotte, and Lefebvre introduced a new method for detecting specialties which does not include any free parameters neither pre or post data processing (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008). This algorithm has many important advantages compared to the classical methods employed to detect clusters: it avoids the use of similarity measures; it relies entirely on the topology of the weighted network; it can be applied to relatively large networks; it does not require subjective interpretations of the bibliometric data analysed or of the communities found. As Blondel, Guillaume, Lambiotte and Lefebvre affirmed:

> We propose a simple method to extract the community structure of large networks. Our method is a heuristic method that is based on modularity optimisation. It is shown to outperform all other known community detection methods in terms of computation time. Moreover, the quality of the communities detected is very good, as measured by the so-called modularity (Blondel et al., 2008, p. 1).

The method is *iterative* and *unsupervised* (no information concerning the correct clusters is provided to the network during its formation) and it is has the aim to optimise the value of the modularity function (Blondel et al., 2008). Specifically, the algorithm is divided into two phases, which are repeated iteratively. Starting from a weighted network of *N* nodes, the first step is to assign a different community to each node of the network, obtaining as many communities as the nodes number. Thus, in this initial partition there are as many communities as there are nodes. Then, for each node *i* the neighbours *j* are considered and the gain of modularity resulting by removing *i* from its community and by placing it in the community of *j* is calculated. The node *i* is placed in the community for which the gain is maximum and positive (in case of ties is used a breaking rule) otherwise, namely in case of negative gain, no movements are made. The algorithm showing the gain in modularity ΔQ obtained by moving an isolated node *i* into a community *C* is formulated as follows:

$$\Delta Q = \left[\frac{\Sigma_{in} + 2k_{i,in}}{2m} - \left(\frac{\Sigma_{tot} + k_i}{2m} \right)^2 \right] - \left[\frac{\Sigma_{in}}{2m} - \left(\frac{\Sigma_{tot}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right] \tag{10}$$

where $\sum_{in}$ is the sum of the weights of the links inside *C*, $\sum_{tot}$ is the sum of the weights of the links incident to nodes in *C*, k_i is the sum of the weight of the links incident to node *i*, $k_{i,in}$ is the sum of the weights of the links from *i* to nodes in *C* and *m* is the sum of the weights of all the links in the network. A similar expression is employed to evaluate the change in modularity when *i* is removed from its community and moved into a neighbouring community (Blondel et al., 2008).

The first phase of the process is completed only when all nodes have been processed and no further improvement in modularity can be achieved by any node movement (a node often is considered several times). At this point, the second phase of the algorithm can start and a new network, whose nodes correspond to the communities found during the first phase, is built. The weights of the links between the new nodes are given by the sum of the weights of the links between nodes within the corresponding communities. Links between nodes of the same community lead to self-loops for this community in the new network. After completing this passage, the first phase of the algorithm is reapplied iteratively to the resulting weighted network; the iteration goes on until the maximum value of modularity is obtained (ibid) (see Figure 15).

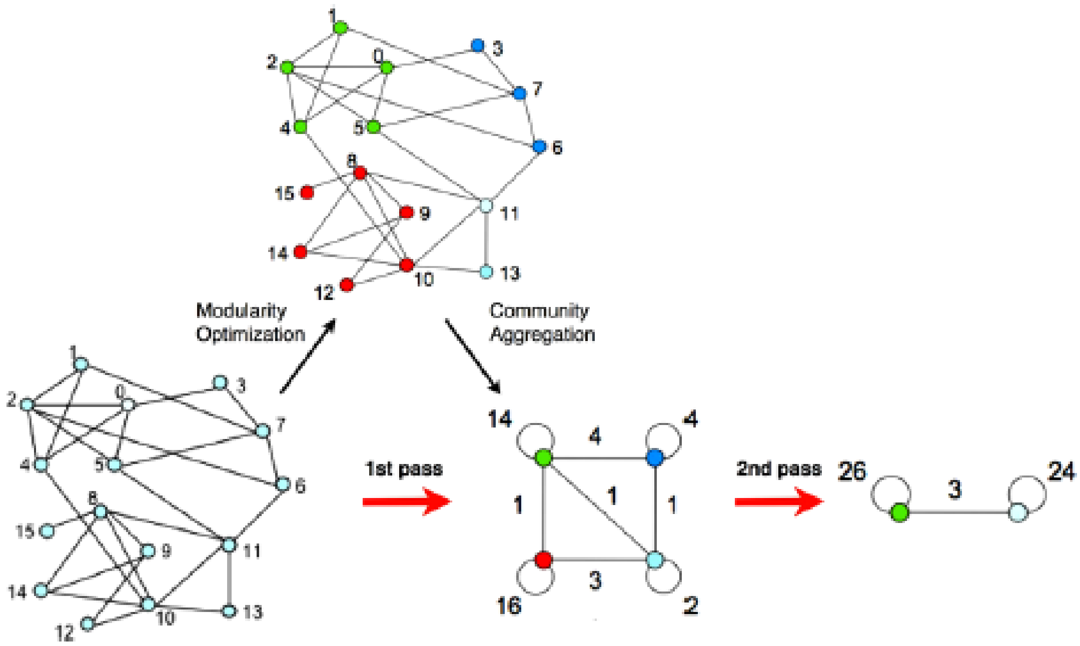


Figure 15: Visualisation of the steps of the modularity optimisation algorithm. Each pass is made of two passes: one where modularity is optimized by allowing only local changes of communities; one where the communities found are aggregated iteratively until no increase of the modularity is possible.
Source: Blondel et al., 2008, p. 5

Even if the Louvain method constitutes a great step forward in detecting modules (clusters) in complex networks, it has some drawbacks (obviously the importance of such limits depends on research targets). First of all, the algorithm produces separated modules; thus it do not allow for overlapping communities, although overlaps are generally assumed to be crucial features of communities. Secondly, it suffers from the so called "resolution limit":

> We find that modularity optimisation may fail to identify modules smaller than a scale which depends on the total size of the network and on the degree of interconnectedness of the modules, even in cases where modules are unambiguously defined (Fortunato & Barthélemy, 2007, p. 36).

The resolution-limit[54] problem is usually demonstrated by means of a ring network of cliques (the maximum complete sub-graph). Modularity merges the cliques depending on the size of the network:

[54] In 2010, Waltman, Van Eck, and Noyons developed a weighted and parameterised variant of modularity-based clustering (Waltman, Van Eck, & Noyons, 2010); furthermore a proposal for a resolution free algorithm was developed by Traag and Van Dooren (2011).

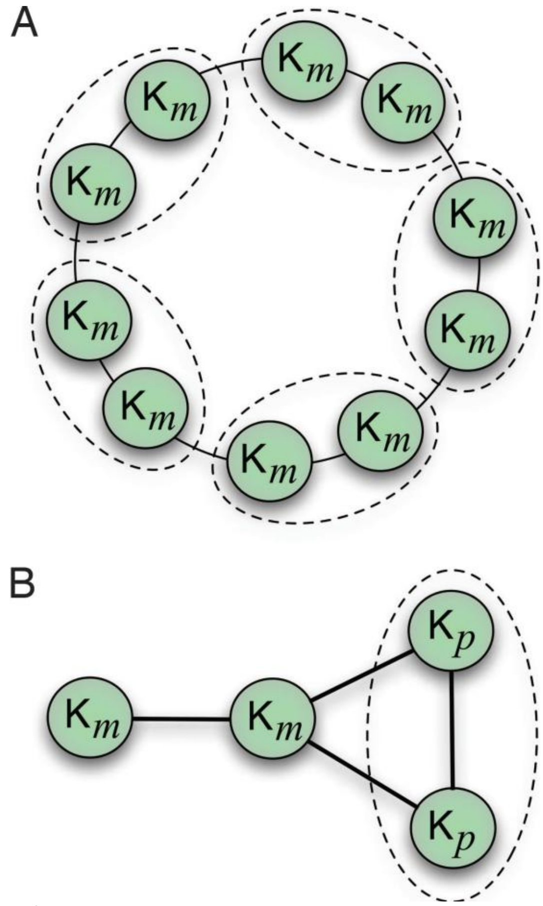


Figure 16: Schematic example of clique merging in modularity optimisation method. (A) A network made out of identical cliques (which are here complete graphs with m nodes) connected than about √L (L is the total number of links in the network), modularity optimisation would lead to a partition where the cliques are combined into groups of two or more (represented by dotted lines). (B) A network with four pairwise identical cliques (complete graphs with m and p < m nodes, respectively); if m is large enough with respect to p (e.g., m 20, p 5), modularity optimisation merges the two smallest modules into one (shown with a dotted-line).
Source: Fortunato & Barthélemy, 2007, p. 39

5.1.7. SUMMARISING

In the previous sections I introduced different, but sometimes similar, techniques which have been developed in time with the scope of unfolding communities (or clusters) in complex networks. Currently, the Louvain method, Map Equation (ME) and Clique Percolation (CP) are the most widely used. Even if each of them is successful in detecting clusters, they do not lead to the same results; two main differences can be highlighted. The first big difference is the one between CP and the other two, since the former allows identifying overlaps while the other two force the division in modules. The second one is the difference between the Louvain method and Map Equation, which is illustrated in Figure 17. Briefly, while the first method focuses on links number, the second refers also to weights; thus if ME is better in detecting flows information paths, the Louvain method is more successful in detecting structures.

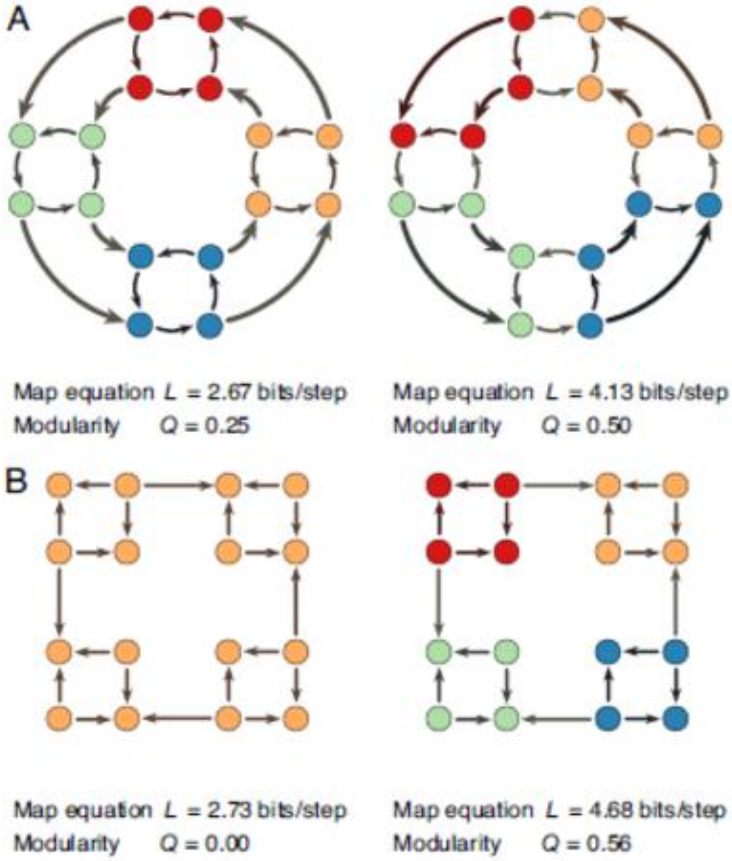


Figure 17: Mapping flow highlights different aspects of structure than does optimising modularity in directed and weighted networks. The colouring of nodes illustrates alternative partitions of two sample networks. (Left) Partitions show the modular structure as optimised by the map equation (minimum L). (Right) Partitions show the structure as optimized by modularity (maximum Q). In the network shown in A, the left-hand partition minimizes the map equation because the persistence times in the modules are long; with the weight of the bold links set to twice the weight of other links, a random walker without teleportation takes on average three steps in a module before exiting. The right-hand clustering gives a longer description length because a random walker takes on average only 12/5 steps in a module before exiting. The right-hand clustering maximizes the modularity because modularity counts weights of links, the in-degree, and out-degree in the modules; the right-hand partitioning places the heavily weighted links inside f the modules. In B, for the same reason, the right-hand partition again maximizes modularity, but not so the map equation. Because every node is either a sink or a source in this network, the links do not induce any long-range flow, and the one-step walks are best described as in left-hand partition, with all nodes in the same cluster.
Source: Rosvall and Bergstrom, 2007, p. 1120

5.2. DIMENSIONALITY-REDUCTION TECHNIQUES

This Section introduces some of the most used techniques for analysing and visualising knowledge domains. These tools are grouped under the label "dimensionality-reduction techniques", as they are used to represent multi-dimensional data through a small number of salient dimensions. The techniques here discussed are *Factor Analysis* (FA), *Multidimensional Scaling* (MDS), *Pathfinder network scaling* (PF) and *Self-Organising Maps* (SOMs). As specified in the Introduction of this thesis, Section 2 consists in a case study focused on Italian Sociology, which will be analysed through bibliometric techniques with the aim of obtaining a bibliometric map (derived from an author co-citation analysis) representing the intellectual and cognitive structure of the Italian sociological community. For this reason, the dimensionality-reduction techniques above mentioned are discussed with reference to their application to co-citation data, and more specifically to author co-citation data.

5.2.1. FACTOR ANALYSIS

Factor analysis (FA)[55] is a multivariate technique usually employed to reduce the number of variables, to detect variables relationships structure, and for classification purposes (Börner et al., 2003). When used in Network Analysis it takes the name *Factor Network Analysis* and, similarly to FA, it is designed for analysing large amounts of highly interrelated processes and quality data by means of principal component analysis and partial least squares analysis. It was widely used in the early history of network analysis to reveal aspect of network structure (both by analysing directly sociomatrices and by factor analysis of a correlation or covariance matrix derived from rows or columns) (Wasserman & Faust, 1994). FA can be used in a complementary way with MDS and clustering display (McCain, 1990).

As McCain explains, in Author Co-citation Analysis (ACA):

> [...] a factor is interpreted as a subset of authors *loading* on it – i.e. making substantial contributions to its construction. Essentially it reveals their underlying subject matter, as perceived by citers. In ACA, every author loads on (contribute to) every factor, and interpretation or definition of each new factor is based on those authors with high loadings (McCain, 1990, p. 440).

Principal Component Analysis (PCA) is a key method in FA and it is often used in ACA (see, for example, Chen & Carr, 1999; McCain, 1990; White & McCain, 1998). It works by an orthogonal rotation of extracted factors producing factors that are uncorrelated, with most authors having high loadings on only one. However, authors may contribute to more than one factor, so FA can demonstrate the breadth of contributions by authors loading substantially on more than one factor (McCain, 1990).

---

[55] The term "Factor Analysis" was introduced by Thurstone in 1931 (Thurstone, 1931).

Factor Analysis has the big advantage of classifying items into multiple factors while traditional clustering techniques forced them into one cluster (Börner et al., 2003).

5.2.2. MULTIDIMENSIONAL SCALING

The term Multidimensional Scaling (MDS) refers to a family of techniques for the analysis of similarities or dissimilarities on a set of objects with the scope of revealing the hidden structure underlying the data.[56] MDS is the most used approach for constructing bibliometric maps; so, for example, in collaboration studies, those collaborating authors who exhibit a high co-occurrence in the multidimensional space are placed close to one another in a low dimensional one (Katz, 1994). By placing objects as points in a low dimensional space, the observed complexity in the original data matrix can often be reduced while preserving the essential information in the data. Thanks to a representation of the pattern of proximities, this multivariate technique is used to identify the dimensions that best explain similarities and differences between network nodes (McCain, 1990). MDS is not usually applied directly to co-occurrence frequencies, as they do not properly reflect similarities among the elements analysed,[57] therefore usually co-occurrence frequencies are transformed employing direct or indirect similarity measures[58] (Van Eck & Waltman, 2010).

MDS provides important insights on the relationships between the elements of a graph (it is often used together with agglomerative hierarchical clustering or Factor Analysis to resort to decomposition in order to determine the "optimal" partitioning scheme). With reference to ACA:

> The major output of MDS is a display of points, usually mapped in two or three dimensions. Points are placed on the map according to their proximity in the original matrix (where high values may reflect high similarities or high dissimilarities). If a correlation matrix is used, the points in the map represent the individual authors placed according to their inter-author similarities. Points representing authors with high similarities will be placed close together in "intellectual space," while points representing authors with high dissimilarities will be placed farther apart (McCain, 1990, p. 438).

MDS major purpose is to capture as much of the original data as possible in only two or three dimensions, but ACA research focuses mainly on two-dimensional outputs. As affirmed by McCain, two dimensions generally capture a high proportion of the variance (85% or more) in the proximities matrix thus providing rich matter for interpretation:

---

[56] For a detailed discussion see Borg & Groenen, 2005; Cox & Cox, 2001.

[57] For a detailed discussion see Van Eck & Waltman, 2010.

[58] For a discussion on direct measures see Van Eck & Waltman, 2009. For a discussion on indirect measures see Van Eck & Waltman, 2008b.

> Interpretation of MDS cocited author maps is based on author (point) placements and author cluster orientations along the horizontal and vertical axes. A three-dimensional solution is more complex but generally adds little explanatory power (McCain, 1990, p. 439).

Of course, the projection is seldom perfect and often important distortions occur due to the information loss in the process of dimension reduction (Katz, 1994).

Alternatives to MDS have been developed and the most used ones are the following: Kamada and Kawai's technique (Kamada & Kawai, 1989); the VxOrd mapping technique (Boyack, Klavans, & Börner, 2005); VOSviewer mapping method[59] (Van Eck & Waltman, 2007b).

5.2.3. PATHFINDER NETWORK SCALING

Pathfinder Network Scaling is a structural and modelling technique extracting underlying patterns in proximity data (based on pairwise similarity). Pathfinder algorithms estimate the proximity measures between pairs of items and preserve only the most important links. The output is a spatial representation of the latter in a class of networks called Pathfinder networks (PFnets) (Schvaneveldt, 1990).

PFnet consists of items as nodes and a set of links connecting pairs of nodes. The set of links is determined by patterns of proximities in the data and parameters of Pathfinder algorithms; similarities can be obtained by either a subjective estimation or a numerical computation (ibid).

Pathfinder network scaling is employed in Generalized Similarity Analysis (GSA), a generic framework for structuring and visualising distributed information resources (Chen, 1997a, 1998a,b, 1999a) and was initially conceived to handle a number of intrinsic interrelationships in hypertext documents (Börner et al., 2003). Chen (1999a, 1999b) incorporated Pathfinder network scaling into the methodology for author cocitation analysis producing author cocitation maps of the hypertext community (Chen & Kuljis, 2003). A further application of Pathfinder to ACA data is by White (2003b).

5.2.4. SELF-ORGANISING MAPS

One of the most important contributions made by the field of artificial neural networks (ANNs) to the information visualisation one consists of the paradigm of the so called *self-organising maps* (SOMs) developed by Kohonen in the 1980s (Kohonen, 1985).

> The main applications of the SOM are in the visualisation of complex data in a two dimensional display, and creation of abstractions like in many clustering techniques (Kohonen, 1995, p. 69).

---

[59] See Section 5.3.

The principle at the basis of the SOM approach is that our knowledge organisation at higher levels is created during the process of learning by algorithms that promote self-organisation in a spatial order; the SOM is then a neural network that pays attention to spatial order (Polanco, François, & Lamirel, 2001). The SOMs belong to the category of *unsupervised learning networks* because no information concerning the correct clusters is provided to the network during its formation. Thanks to this feature this approach can be employed to detect clusters in the input data, and to identify an unknown data vector with one of the clusters. The SOMs also belong to the category of *competitive learning networks* (Hinton, 1989; Kohonen, 1995) for two reasons: there is a set of nodes that compete with one another to become active; a number of nodes compare the same input data with their internal parameters, and the node with the best match then tunes itself to that input, in addition the best matching node activates its topographical neighbours in the network to take part in tuning to the same input. The more a node is distant from the winning node the weaker is the learning (Polanco et al., 2001). SOM maps constitute a combined representation of what Campanario calls "relations maps and domain maps" (Campanario, 1995) (see Figure 18 below).

If from one side the SOMs appear to be among the most promising algorithms for organising large volumes of information, they have some significant deficiencies such as the absence of a cost function, and the lack of a theoretical basis for choosing learning rate parameter schedules and neighbourhood parameters to ensure topographic ordering (Kohonen, 1995). In 2001 Polanco et al. provided some improvements to the method: an automatic way of naming the clusters; the division of the map into logical areas; the introduction of a map generalisation mechanism (Polanco et al., 2001). Furthermore, they introduced multi-map extension of SOMs: Multi-SOMs. These latter are communicating Self-organising Maps which introduce in the analysis the use of "viewpoints": each different viewpoint is achieved in the form of map and each map is a spatial order in which the information is represented into nodes (clusters) and spatial areas (group of clusters). The multi-maps enable a user to highlight semantic relationships between different themes belonging to different viewpoints (ibid).

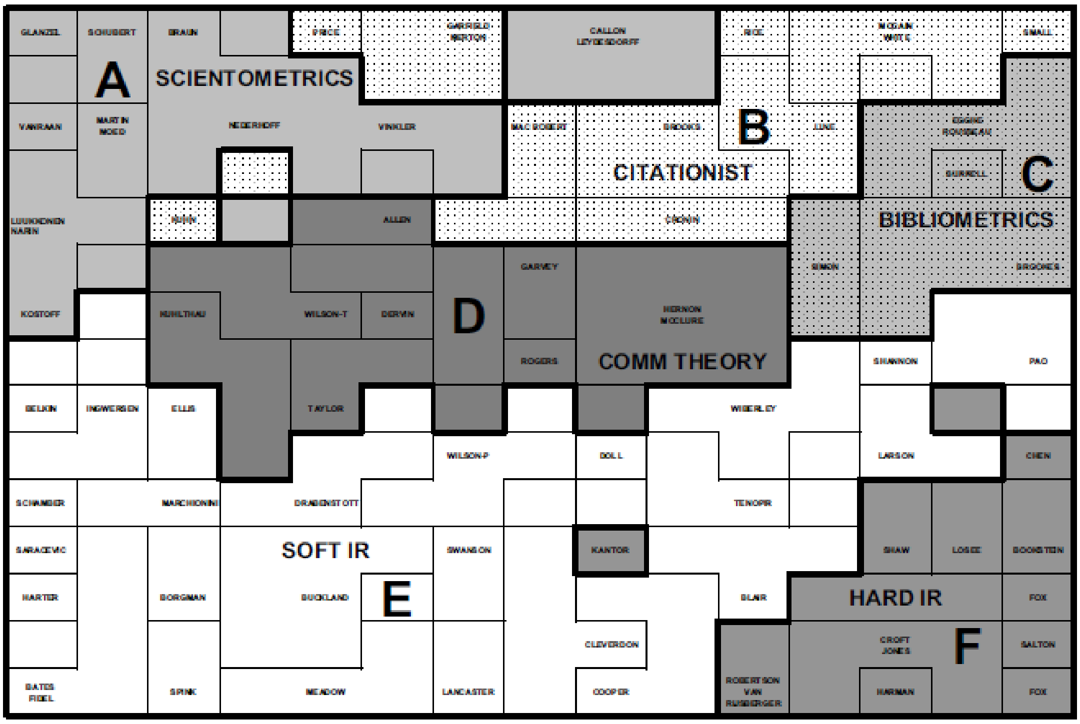


Figure 18: SOM of an ACA study of Library and Information Science. Capital letters identify LIS specialties (A: scientometrics; B: citationist; C: bibliometrics; D: communication theory; E: soft information retrieval; F: hard information retrieval).
Source: Moya-Anegòn, Herrero-Solana, & Jiménez-Contreras, 2006, p. 67

5.3. VOS VIEWER: A UNIFIED APPROACH TO CLUSTERING AND MAPPING

In 2010 Waltman, Van Eck, and Noyons introduced a unified approach to both clustering and mapping bibliometric networks, as an alternative to the combined use of clustering and mapping techniques relying on different assumptions (Waltman, Van Eck, & Noyons, 2010). Their proposal is based on a weighted and parameterised variant of the Louvain Method (introduced in Section 5.1.6), which also includes a parameter solving the resolution limit problem (the parameter γ) (Waltman et al., 2010) and a variant of the well known Multidimensional Scaling technique (Van Eck & Waltman, 2007b). This new approach is implemented in the software VOSviewer[60] (Van Eck & Waltman, 2007b), which employs a direct probabilistic similarity measure to mapping bibliometric data. For these properties, which will be discussed shortly, this approach is the one used for developing the empirical work introduced in Section 2.

Similarity measures employed in mapping techniques can be divided into direct and indirect ones.[61] Indirect similarity measures estimate the similarity between two objects *i* and *j* by comparing the i^{th} and the j^{th} row (or column) of the co-occurrence matrix C. Intuitively, the more similar the co-occurrence profiles, the higher the similarity between *i* and *j*. Direct similarity measures estimates the similarity

---

[60] http://www.vosviewer.com/

[61] For a detailed analysis of the differences between the two measures see Van Eck & Waltman, 2009.

between two objects i and j by adjusting the number of co-occurrences for the total number of occurrences or co-occurrences of the objects analysed. The most popular direct similarity measures are the Cosine, the Jaccard index, and the inclusion index. As explained by Van Eck and Waltman (2009), direct measures can be divided in set-theoretic similarity measures and probabilistic similarity measures:

> [...] a natural approach to determine the similarity between two objects i and j seems to be to determine the similarity between, on the one hand, the set of all documents in which i occurs and, on the other hand, the set of all documents in which j occurs. We refer to direct similarity measures that take this approach as set-theoretic similarity measures. In other words, set-theoretic similarity measures are direct similarity measures that are based on the notion of similarity between sets (Van Eck & Waltman, 2009, p. 639).

The cosine, the Jaccard index, and the inclusion index belong to the category of set-theoretic similarity measure, while the so called association strength belongs to the second. The association strength (Van Eck, Waltman, Van den Berg, & Kaymak, 2006; Van Eck & Waltman, 2007a) (also referred to as "the proximity index"[62], and "the probabilistic affinity -or activity- index"[63]) is a proportional measure. It is proportional with respect to the ratio between the number of co-occurrences of objects i and j, and the expected number of co-occurrences of the two objects under the assumption of statistically independence. Thus, *the association strength* of nodes i and j is given by:

$$S_{ij} = 2mC_{ij} / C_i C_j \tag{11}$$

where C_i denotes the total number of links of node i and m denotes the total number of links in the network.

> The higher the association strength of two nodes, the stronger the attractive force between the nodes. Since the strength of the repulsive force between two nodes does not depend on the association strength of the nodes, the overall effect of the two forces is that nodes with a high association strength are pulled towards each other while nodes with a low association strength are pushed away from each other (Waltman et al., 2010, p. 631).

The unified approach for mapping and clustering is, in the context of mapping, based on the notion of finding for each node i a vector $x_i \in R^p$ that indicates the location of node i in a p-dimensional map (in chapter 9 only 2 dimensional maps will be considered). In the context of clustering we instead seek to

[62] See for example Peters & Van Raan, 1993; Rip & Courtial, 1984.
[63] Leydesdorff, 2008; Zitt, Bassecoulard & Okubo, 2000.

find for every node *i* a positive integer x_i that indicates the cluster to which node *i* belongs. Given this setup, the approach aims to minimise:

$$V(x_1,...,x_n) = \sum_{i<j} s_{ij} d_{ij}^2 - \sum_{i<j} d_{ij} \tag{12}$$

with respect to $x_1...,x_n$. Where d_{ij} represents the distance between nodes *i* and *j* and in case of mapping is given by:

$$d_{ij} = \| \mathrm{x}_i - \mathrm{x}_j \| = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2} \tag{13}$$

while in case of clustering by:

$$d_{ij} = \begin{cases} 0 & if \quad x_i = x_j \\ 1/\gamma & if \quad x_i = x_j \end{cases} \tag{14}$$

where γ represents the resolution parameter hinted above; the larger the value of this parameter, the larger the number of clusters that we obtain. The parameter was developed in order to solve the resolution limit problem[64] (Fortunato & Barthélemy, 2007), which creates problems in detecting small clusters. Thanks to the parameter γ small clusters can always be identified (Waltman et al., 2010). It should be observed that if the resolution parameter is set to 1, the clustering approach correspond to the (weighted) modularity function presented in equation (7) and the minimisation of (12) in the context of mapping is a version of multidimensional scaling described in the previous Section.

[64] See Section 5.1.6.

6. FRAMING CITATIONS: A SOCIOLOGICALLY INTEGRATED APPROACH IN CITATION THEORY[65]

That Scientometrics lacks a citation theory encompassing both a theoretical foundation for citation analysis and a description of scientists' citing behaviour is an established fact. Even if over time there have been various complaints about this situation (Cronin, 1981, 1984, 1998; Cozzens, 1981; Leydesdorff, 1998; Luukkonen, 1997; Zuckerman, 1987), what Leydesdorff said in 1987 is still valid: "[...] we still have a theoretically underdeveloped understanding of what these bibliometric data actually mean" (Leydesdorff, 1987, p. 290). Moreover, if we consider that the use of bibliometric indicators in research evaluation (at micro, meso and macro level) is increasing, and that, like it or not, citation analysis affects scientists' life (Shadish, Tolliver, Gray, & Sen Gupta, 1995), one cannot avoid sharing Zuckerman's statement: "by now, it may seem redundant to say that a theory of citation is badly needed" (Zuckerman, 1987, p. 336).

The quest for a grand theory of citations has been stressed from different perspectives; for example, in the early 1980s Cozzens (1981) and Cronin (1981) raised independently the problem from the perspective of Sociology and Information Science, respectively. The incapacity and/or unwillingness of sociologists to overcome the old antagonism between normative and constructivist approaches resulted in the availability of interesting but middle-range theories[66] (Merton, 1949), each of them illuminating a specific aspect of the citation phenomenon. Thus, from the normative side citations are conceived of as reward tools, intellectual links, devices for the payment of intellectual debts (here the focus is on the content of contributions) (Kaplan, 1965; Merton, 1973, 1979), while on the constructivist side the stress is on the rhetorical function and the impossibility to standardise citers' behaviour, as it is linked to mental states and emotions of scientists as human beings (here the focus is on both the influence and the role of the one cited) (Gilbert, 1977; Latour, 1987). With the scope of demonstrating the existence of a cognitive relation between citing and cited works, and thus trying to corroborate the validity of the normative approach, Peters and Van Raan in 1995 published a study in which word-profile similarities between citing and cited publications were measured. Results show that in case of publications with a citation relationship, the similarity in terms of content is significantly higher than in case of no relation (Peters & Van Raan, 1995). More generally, as Van Raan stated: "all our findings are contrary to the opinions of these constructivism-dominated circles of sociologists of science" (Van Raan, 1998, p. 134). Furthermore, in 1998 Baldi conducted a study in which the normative argument and the constructivist one are compared.

[65] This Chapter is based on the article "Scientific communities as autopoietic systems: the re-productive function of citations" (Riviera, in press).

[66] According to Merton: "Our major task today is to develop special theories applicable to limited conceptual ranges – theories for example, of deviant behavior, the unanticipated consequences of purposive action, social perception, reference groups, social control, the interdependence of social institutions – rather than to seek the total conceptual structure that is adequate to derive these and other theories of the 'middle range'. […] Sociological theory, if it is to advance significantly, must proceed on these interconnected planes: 1. by developing special theories from which to derive hypotheses that can be empirically investigated and 2. by evolving a progressively more general conceptual scheme that is adequate to consolidate groups of special theories" (Merton, 1949, p. 51).

The study is based on a network-analytic model whose aim is to asses if and to what extent the respective features of potentially citing and potentially cited papers play a role in the likelihood that a citation is given or not. The relationship between potentially citing and potentially cited papers is assumed to be dyadic (Baldi, 1998). Results are "consistent with a normative interpretation of the allocation of citations in which scientists cite the best, most relevant regardless of who wrote it" (Baldi, 1998, p. 843). The overall pattern seems to indicate that authors tend to cite papers which are intellectually relevant (with respect to their work) in terms of subject, recency of knowledge and theoretical orientation, showing that authors' characteristics of the cited papers have little influence.[67] Further evidences show no support to the constructivist approach, bolstering in this way the effectiveness of bibliometric impact measures (Nederhof & Van Raan, 1987, 1988; Nederhof 1998; Van Raan, 1998).

The results mentioned above suggest a model in which the normative approach plays a fundamental role in citing. However, it is clearly not sufficient to describe the citation phenomenon in its totality: the normative approach is useful with reference to high citation counts, while it does not furnish any hint with reference to low counts, which, in my opinion, are effectively described by the constructivist approach. Citation studies play an important role in different fields, such as Library Science, Information Science, all the "metrics" fields and Sociology of Science. Therefore, it seems necessary to work towards the realisation of an approach encompassing both normative and constructivist theories on citations. Actually, it is possible to identify two attempts in this direction, namely that by Henry Small (1978, 2004) and Susan Cozzens (1981), but such attempts have not been very successful. The reason could be linked to the fact that in the citation process many dimensions are at work, such as the cognitive, communicative, historical, social and textual. Therefore, schematisations about different types of citations seem to be too reductive (see, for example, Baldi, 1998; Cronin, 1994). Exhortations in considering the multidimensionality of citations have been made by Leydesdorff and Amsterdamska (1990) and Leydesdorff (1998). In particular, Leydesdorff (1998) underlines the fact that Science should be considered an order emerging from networks of reflexive communication:

> […] the interaction between a second-order layer of reflexive communications among (first-order) authors provides the condition for the generation of a more abstract cognitive structure of concept-symbols. The dual-layered network system of (1) social relations among scholars and (2) relations among communications is expected to resonate dynamically on the basis of interactions in some directions, while not in others. The structural dimensions of this complex network are partially correlated to individual authors as carriers of the communication, as they are also correlated to the textual dimension of the knowledge content. The interactively emerging dimension itself remains latent while it develops, and thus a strategic vector is

[67] The field inspected is Celestial Masers, a research area in Astrophysics. Only one of the variables used to test the constructivist hypothesis showed a significant negative effect on citation occurrences, that is to say, the percentage of women cited, suggesting a devaluation of their work in the field studied (Baldi, 1998).

> induced that can be recognized (Abernathy & Clark, 1985). The recognition recursively assumes and refines cognition within the observing systems (Maturana & Varela, 1980) (Leydesdorff, 1998, pp. 11,12).

Leydesdorff's contribution constitutes a step forward in the direction of creating a theoretical frame for understanding and describing the phenomenon of citations. A further step forward could be a theoretical formulation that also refers to the social system in which citations are produced, providing in this way a conceptual frame for describing citers' behaviour. In order to do this, in line with Leydesdorff, in the following pages Science is conceived of as an autopoietic cognitive domain, and communicative events as the elements of this system, with the consequence that citations, as codes and medium of communications, can be conceptualised as devices through which the structuration process in scientific communities is accomplished.

Autopoiesis theory has already been applied to social systems and to scientific communities (Maturana, 1990; Maturana & Varela, 1980). In sociology the most noteworthy attempt has been made by Luhmann (1995), whose theory is developed around the concept of autopoiesis: components of social systems are communicative events, which are recursively produced and reproduced by a communicative network. Even if Luhmann’s work encompasses many and important aspects of autopoiesis, his theory lacks some technical details on the functioning of social systems, which instead are conceptualised in Giddens' structuration theory. A complementarity between the two theories on social systems has already been identified by Mingers (2004) and operationalised by Leydesdorff (2010) who recombined them in a theory of the structuration of expectations. By integrating the two theoretical approaches with Merton and Small's conceptualisations (in the first case related to both the institutional features of Science and citations, while in the second related only to citations), the effectiveness of bibliometric indicators in furnishing information about the influence of authors and ideas and of bibliometric maps in furnishing representations of the cognitive structure of scientific domains will be demonstrated.

Citations are assumed to play a fundamental role in the re-production of scientific communities viewed as self-organising and self-referential systems. In self-organisation theories and in communication theories (see, for example, Luhmann, 1986; Leydesdorff, 1998; Maturana, 1975b) the network metaphor has often been employed: networks of relations, networks of processes, communication networks and so on. What is suggested here is that, starting from the consideration that Science is a cognitive domain (Maturana, 1990), and that citations can be considered both as codes and medium of communication (Leydesdorff, 1995, 2001, 2007), maps created through relational bibliometric techniques[68] enable us to render such a metaphor concrete. Thus, a bibliometric map can be conceived of in the way here stated:

[68] Such as author co-citation (White & Griffith, 1981), bibliographic coupling (Kessler, 1963), co-citation analyses (Small, 1973), co-word clustering (Callon et al., 1983) and co-heading analysis (Todorov, 1990).

> […] a spatial representation of how disciplines, fields, specialties, and individual papers or authors are related to one another as shown by their physical proximity and relative locations, analogous to the way geographic maps show the relationships of political or physical features on the Earth (Small, 1999, p. 799).

At the basis of the theory developed in the following pages there is the general belief that efforts should be made towards the formulation of theories offering the possibility to build a bridge between structural-functionalist and constructivist approaches. With the scope of describing the self-reproductive nature of Science, and the consequences of such a conceptualisation for Scientometrics, the remainder of this Chapter is organised as follows. First the concepts of "autopoiesis" and "self-organisation" are introduced, and the application of Luhmann's autopoietic theory to social systems is subsequently illustrated. The third Section focuses on the development of the integrated theoretical proposal based on the application of the autopoietic theory to scientific communities, aiming to demonstrate the re-productive function of citations and introducing a model for citers' behaviour. Implications for Scientometrics deriving from the conceptualisations made in the third Section are showed in the next one. Finally, the main elements of the theory proposed are summarised, and conclusions are given.

6.1. AUTOPOIETIC AND SELF-ORGANISING SYSTEMS

Even if strictly interwoven between them, the concepts of "autopoiesis" and "self-organisation" imply different things. The concept of autopoiesis goes back to Kant: in "Critique of judgement", referring to the purposiveness of nature, he wrote:

> In such a natural product as this every part is thought as owing its presence to the agency of all the remaining parts, and also as existing for the sake of the others and of the whole, that is as an instrument, or organ. […] the part must be an organ producing the other parts – each, consequently, reciprocally producing the others. […] Only under these conditions and upon these terms can such a product be an organized and self-organized being, and, as such, be called a physical end (Kant, 1790/1980, p. 65).

The term "autopoiesis" has been developed in Biology by Humberto Maturana and Francesco Varela in 1970s (Maturana, 1975b; Maturana & Varela, 1980; Maturana & Varela, 1987), with the aim of differentiating between living and non-living entities. According to the concept of autopoiesis, a system can be defined as a living one (an autopoietic machine) if it is able to fulfil the functions of self-production, self-reproduction and self-organisation. Components of an autopoietic machine are not physical entities but processes (such as the metabolic networks of a cell). Moreover, living systems are structure-determined systems, that means that external intrusions are not admitted, they can just trigger a structural change but not determine it, as it can be determined only by internal dynamics: what happens in an autopoietic machine can be caused only by the machine itself (Maturana, 1990). Put differently, by definition an autopoietic system transforms itself into itself. Autopoietic systems are *organisationally closed* but *structurally open*: the structure (conceptualised as the components and their relations) may change over time, but the *organisation* maintains its relations of self-production. An autopoietic system has definite boundaries. Formally:

> […] a composite unity whose organisation can be described as a closed network of productions of components that through their interactions constitute the network of productions that produce them, and specify its extension by constituting its boundaries in their domain of existence, is an autopoietic system (Maturana, 1987, p. 349).

The concept of self-organisation is not the same as autopoiesis: while autopoiesis includes self-organisation, self-organisation does not necessarily entail autopoiesis:

> Self-organising process are such physico-chemical processes that reach a specific ordered state […] under a (more or less extensive) domain of initial conditions and constrains […]. Arriving at a given state of order is not (or at least not essentially) imposed on the process from outside, but is the result of the specific characteristics of the components involved in the process. The state of order is achieved spontaneously (Roth, 1986, p. 154).

There are various definitions of self-organisation, as this concept has been applied in many and different disciplinary fields. However, it is possible to isolate some recurring features: self-regulated differentiation of autonomised systems into complex subsystems which are dynamic networks open to reorganisation; dynamic stability (system's re-production contemplates structural changes but requires organisational stability); circular causality (outputs are at the same time also inputs in a given system) (Krohn, Küppers, & Nowotny, 1990).

The main difference between an autopoietic system and a self-organised one is the degree of neutralisation of influences external to the system. Autopoietic systems accept elements produced only by themselves: the system is made of these elements. External influences can have consequences in triggering or stimulating the production of new elements (ibid).

6.2. SOCIAL SYSTEMS AS AUTOPOIETIC MACHINES: THE COGNITIVE PROGRAM OF CONSTRUCTIVISM

> *It is perhaps not the least important function of constructivist epistemology to make society irritatingly aware of the fact that it produces science.*
>
> Niklas Luhmann[69]

The first attempt to apply autopoiesis theory to social reality was made by Niklas Luhmann (Luhmann, 1981/1982, 1982/1986a, 1986b, 1984/1995). His theory entails different intellectual traditions, that is to say: Parsons' systems-theoretical approach; the cybernetic interpretation of the relationship between the system and its environment; a phenomenological understanding of meanings and of experience; the application of the autopoietic theory to social systems (Luhmann, 1986/1989).

Luhmann found in the self-organising paradigm a set of epistemological considerations constituting a valid alternative to classic responses to the question: "how is knowledge possible?" In Luhmann's perspective this is possible by substituting the Kantian polarity *transcendental/empiric* to the constructivist one *environment/system*:

> What we call "environment" today had to be conceived of as the state of being contained and carried (periechon); and what we call "system" had to be thought of as order according to a principle. Both of these were already objects of knowledge. In order to answer the question of how knowledge is possible without falling into a self-referring circle the distinction transcendental/empirical was developed. Hardly anyone accepts this distinction today [...]. But if one drops this distinction how does one then avoid the circle of the self-founding of knowledge? The serving as medium foundation for dealing with these questions offers up the distinction system/environment and, in this context, a worked-out systems theory (Luhmann, 1990, p. 66).

Cognition is the means through which human beings come into contact with the external environment; a system is able to know only by making distinctions based on different unities of distinctions (Maturana, 1980; Maturana, 1990; Luhmann, 1990; Luhmann, 1984/1995). *Recursivity* is at the basis of these systems, as they use their own outputs as inputs:

> [...] recursivity requires a continuous testing of consistency and it has been shown by investigations in perception and memory that this necessitates a binary schematization, [...], which holds in readiness the possibility of acceptance and rejection. The states of the system that have been produced by its own operation serve then as criteria for the acceptance or

[69] Luhmann, 1990, p. 81.

> rejection of further operations [...] Decisive, however, is the continuous self-evaluation of the system – which always operates in a state of irritation or agitation by means of a code that permits acceptance and rejection with regard to the adoption of further operations. The brain functions in this way. And the same will be true for physic and social systems (Luhmann, 1990, p. 72).

According to the *cognitive program of constructivism*, knowledge results from the construction of a relation to the environment, therefore only non-knowing systems can know. Furthermore, as distinctions are internal processes, the information deriving from this process is nothing more than an internal achievement, that is to say the product of the recursivity process:

> Apparently it is fundamental for the functioning of the brain that selected information is enclosed and not that it is let through. As if it were already information (or data) before it motivates the brain to form a representation. Such knowledge as this was not made use of by theoretical epistemologist and it is only a formulation in terms of systems theory that leads to an insight which must seem surprising to epistemologists: only closed systems can know. [...]. Whoever [...] maintains *this* [...] is forced [...] *to face* a paradox: it is only non-knowing systems that can know; or, one can only see because one cannot see. [...] The effect of the intervention of systems theory can be described as a "de-ontologization" of reality. This does not mean that reality is denied, for then there would be nothing that operated [...]. It is only the epistemological relevance of an ontological representation of reality that is being called into question. If a knowing system has no entry to its external world it can be denied that such an external world exists. But we can just as well – and more believably – claim that the external world is as it is. Neither claims can be proved; there is no way of deciding between them (Luhmann, 1990, p. 67, italics mine)

But which are the epistemological consequences of these claims for Sociology? Put differently, under which conditions is the study of social phenomena by sociologists possible? It is possible by making distinctions and observing the observation of the first observer. "In terms of Sociology one could say that observation is directed now to the observed observer's latent structures and functions" (Luhmann, 1990, p. 73). The cognitive program of cognition, entailing the recursivity of all observations, including the observations of latency, allows us to abandon the classic Sociology of knowledge claims about the fact that "latent structures give a false picture of the world [...]. The assumption [...] that latent structures, functions and interests lead to distortions of knowledge, if not blatant errors can and must be abandoned" (Luhmann, 1990, p. 73). Of course, as previously mentioned, it is impossible to distinguish the distinction used for making distinctions, as it is an inescapable precondition of cognition, thus questions about motivations behind a certain rationale arise only at the second order of observation. "The claim of ideological distortion can then be observed in the person making the claim (for which he has to be observed as observer, that is, in relation to what he is observing)" (Luhmann, 1990, p. 73).

The application of the autopoietic concept to social systems required the specification of additional properties other than those listed in the previous section. As Luhmann specified:

> The theory of self-producing, autopoietic systems can be transferred to the domain of action systems only if one begins with the fact that the elements composing the system can have no duration, and thus must be constantly reproduced by the system these elements comprise (Luhmann, 1984/1995, p. 11).

Furthermore, the elements of a social system are events localisable at a precise point in time. According to Luhmann, such elements are communicative events: communication is used by social systems for autopoietic reproduction and, at the same time, they are recursively re-produced by internal networks of communications. Social systems are conceptualised as expectations embodied in the reflexive systems of communication: discourses produce standards working as codes of communication, which in turn is a discursive construct. Finally, as we will see shortly in detail, it must be specified that Luhmann, differently from Maturana, considers organisation as an integrating retention mechanism at the social level given systems differentiation (Luhmann, 1984/1995).

Luhmann's theory implies a paradigm shift in Sociology inasmuch as society is considered to be made up of communicative processes, and not of human beings: "we assume that social systems are not composed of psychic systems, let alone of bodily human beings" (Luhmann, 1984/1995, p. 25). According to Luhmann, communication involves three elements: information (message content), utterance (form of production of the information plus sender's intentions) and understanding (generated meaning from receiver's point of view) (Luhmann, 1984/1995). He also refers to Bateson's concept of information as "meaningful information", that is, a difference that makes a difference (Bateson, 1972). Meaning is understood as *a certain strategy amongst alternative possibilities*. Meaning can be generated only in human interactions, but human beings, with their emotions and feelings, are excluded from social systems – they are part of the psychological system (Luhmann, 1984/1995). Obviously, without human beings there is no communication, thus Luhmann employs Parsons' concept of "interpenetration" (Parsons, 1951), which allows us to consider an element as part of two systems at the same time. As stressed by Leydesdorff (2010) these two systems can interpenetrate because they process meaning reflexively. Even if Luhmann did not explicit it, interpenetration adds an operational coupling to the structural one: "in *the* second contingency – the layer of expectations as different from the contingency among variables – the meaning of communication at the individual level and at the social level are operationally coupled in terms of possible reflections in additional to the structural coupling" (Leydesdorff, 2010, p. 2140, italics mine).[70]

[70] According to Leydesdorff, this additional coupling makes the communication system of Luhmann quasi-autopoietic: "cogitantes are not only the carriers of cogitate, but reflexively they also have access to

In summary, social systems are an ongoing network of interacting and self-referring communications of different types. Furthermore, autopoietic systems are operationally closed but structurally open, which means that while the structure (conceptualised as the components and their relations) may change over time, the organisation maintains its relations of self-production. Communication would be able to define subsystems. Referring to the Parsonian concept of functional differentiation of society[71] (Parsons, 1951), and Husserl's (1929/1973) conceptualisation of "horizon of meanings"[72], Luhmann interprets communication as a device through which inclusion and exclusion processes operate in defining domain-specific selections (Luhmann, 1981/1982, 1986/1989).

Luhmann's theory is based on the concept of *differentiation* (Luhmann, 1981/1982). First of all, a system differentiates from its environment (which entails also the other systems): "systems are orientated by their environment not just occasionally and adaptively, but structurally, and they cannot exist without an environment" (Luhmann, 1984/1995, p. 16). A social system's environment is constituted by consciousness. Systems and subsystems are characterised by a high level of autonomy and self-referentiality, distinguishing themselves from their environments; they are operationally closed (Luhmann, 1981/1982). As suggested before, communication, in Luhmann's view, performs exactly this function. Communicating something involves a selection among different possibilities. If differentiation is at the basis of such a theory it means that boundaries have a fundamental role. More specifically, their function is twofold: they separate and connect at the same time systems, subsystems and environments. Boundaries separate events but not relations (Luhmann, 1984/1995). Therefore, they constitute a precious object of study for those who, for example, are interested in innovation processes (see, for example, Etzkowitz & Leydesdorff, 2000; Leydesdorff & Etzkowitz, 1996, 1998). Notwithstanding the fundamental role of boundaries, Luhmann does not describe if and in which way it is possible to detect them.

Luhmann demonstrated the applicability of his theory in works such as "Love as passion" (1982/1986) and "Ecological communication" (1986/1989), but for our purposes we need to integrate it with information about the structural properties of such systems, in order to understand the mechanisms at the basis of the re-productive process we need Gidden's structuration theory with its strictly related concepts of "duality of structure" and "structuration" (Giddens, 1984).[73] Finally, in Luhmann's theory there is a big

---

their substantive content" (Leydesdorff, 2010, p. 2141).

[71] According to which concept, societies move from a simple to a complex state by increasing specialisation of different subsystems and institutions within the society (Parsons, 1951).

[72] "[...] no meaning-constituting system can escape the meaningfulness of all its own processes. But meaning refers to further meaning. The circular closure of these references appears in its unity as the ultimate horizon of all meaning: as the world. Consequently, the world has the same inevitability and unnegatability as meaning. […] Husserl outlined this situation in the metaphor 'horizon' […]" (Luhmann, 1984/1995, p. 69).

[73] As Luhmann wrote: "[…] systems are not merely relations among elements. The connection among relations must also somehow be regulated. This regulation employs the basic form of conditioning. […] relations among elements can condition themselves reciprocally; one occurs only when the other also occurs. […] Successful conditionings, which are achieved by the emergence of what they enable, work as constraints" (Luhmann, 1984/1995, p. 23).

hole; the lack of descriptions about the creation of self-referential social systems. How does it happen that a self-organising and self-referential system gains its autonomy?

6.3. SCIENTIFIC COMMUNITIES AND THE RE-PRODUCTIVE FUNCTION OF CITATIONS

Maturana and Varela tried to extend the domain of their theory developed in Biology to social systems (Maturana & Varela, 1980, 1987). In order to do this, they had to apply the concept to non-biological and non-physical systems; thus, an autopoietic system can also be defined as consisting of concepts, descriptions, rules or communications (Maturana, 1975a, Maturana & Varela, 1980). We have just considered the attempt made by Luhmann to describe social systems as autopoietic machines, and we have also underlined the limits of this attempt. In the following pages an integrated approach consisting of different sociological theories is suggested. By demonstrating the feasibility of applying the autopoietic theory to scientific communities and the re-productive role of citations in such systems, a theoretical frame for interpreting bibliometric measures and maps will be provided.

As the concept of autopoiesis has a central role in this work, I think it is useful to recall its main features. A system is defined autopoietic if it is *operationally closed* (it is characterised by self-organisation and self-production), if it is *structurally open* (it has interactions with its environment and the structure may change over time) and if it is also *structure-determined* (external inputs cannot directly provoke changes in the system, they can just trigger them). Autopoietic systems are thus *autonomous* and *self-referential*. In the following pages I will try to reply to the following questions: what are the components constituting the social system? What are the re-productive mechanisms which take place in the process of production and re-production? Are a system's boundaries detectable? If they are, in which way is it possible to detect them?

First, however, a question should be posed: assuming that Science is an autopoietic system, how did it happen? In the article "Self-organisation and autopoiesis in the development of modern Science" (1990), Rudolf Stichweh analyses the developmental steps of the social system of Science between the 16th and the 20th century. According to him, such a process consists of two phases: self-organisation and autopoiesis. Before a system becomes self-constructing and self-referential, the organisation of the system as autonomous entity against a surrounding disorder is required. Thus, this stage is characterised by the duality order/disorder. As Stichweh underlines, European modern Science (the period extending from the 16th century to the second half of the 18th century) is mainly an enterprise dealing with received knowledge; it means that "there are huge quantities of comparatively new knowledge not produced by early modern Science itself. This received knowledge comes from a plurality of sources […] From the perspective of the institutionalized learning […] this knowledge, in most of its types, is fortuitous knowledge, since learning has neither controlled its production nor the accidents of its advent" (Stichweh, 1990, p. 196). With the expression "received knowledge" we refer to the corpus of ancient scientific, rhetorical and literary texts and papyri; information from extra-European scientific traditions, for example, from the Islamic world or China; knowledge produced in different contexts such as the mechanical arts or travellers' descriptions (Stichweh, 1990). As stated above, early modern Science is characterised by the organisation of received knowledge; conservation and organisation constitute its two main activities. Conservation and organisation are achieved through the following values (or ideals): imitation (as the social ideal for learned knowledge); preservation (temporal ideal); order (the material value for knowledge). As ordering is achieved through the imperatives of imitation and preservation, it leads to the discovery of structures and the perception of connections hidden in the received traditions. As Stichweh affirms: “to me, this achievement seems to be the most important self-organisation property of early modern Science” (Stichweh, 1990, p. 198). Imitation entails the translation of every bit of knowledge into Latin, which was considered the language of learning; this process also implies the adaptation of knowledge to the conceptual structures available in Latin. Furthermore, received knowledge is now available in structured text instead of unstructured papyri: “editing these texts, and commenting upon texts, it may be argued that these activities never have been simple preservations of received stocks of knowledge only, but have always included the construction of a new order of the text”; libraries constitute the final result of the ordering activity.

The passage from self-organisation to autopoiesis is a gradual one, and it comprises the passage from a system accepting any external elements and focused on the integrative function, to a system focused on the selection of external elements to be internalised (placing the refused ones at the boundaries of the system).[74] “The establishment of a critical and selective awareness in contradistinction to the earlier preference for structure formation and ordering can be observed […]” (Stichweh, 1990, p. 200). Such a shift is part of a broader one involving the whole society. It took place between the 17th and the 18th century and was caused by two factors: scepticism and critical spirit. These are two of the major imperatives of the Enlightenment, which occurred in Europe in the 18th century. The cultural and philosophical movement was characterised by the effort to separate Science and intellectual traditions such as Alchemy and Astrology, and found its highest expression in the Encyclopédie of Diderot e D'Alembert. Moreover, in this period a new imperative crystallised: usefulness of knowledge. As Schölzer claimed at the University of Göttingen: “being very learned and working for the common weal are pure synonyma” (Stichweh, 1990, p. 201). It is interesting to find in this description trace of those features of Science that Merton calls “major norms” or “institutional imperatives” (Merton, 1973).[75]

---

[74] Obviously, such a reconstruction is an oversimplification, as many elements characterising modern Science appeared before the 17th century. Among these elements we can mention the need to credit prior authors and the symbolic function of references (De Bellis, 2009).

[75] Universalism (scientific results are to be subject to preestablished impersonal criteria and are judged independently of author's personal or social attributes); communism (even if scientists have to be rewarded for their scientific contributes, these are product of social collaboration and constitute a common heritage); disinterestedness (absence of economical or personal motivations in research activities, guaranteed by a pattern of institutional control); organised scepticism (it is both a methodological and institutional mandate and refers to the temporary suspension of judgment and the detached scrutiny of beliefs in terms of empirical and logical criteria) (Merton, 1973).

To sum up, the reaching of the autopoietic status has been possible thanks to three elements: order, systematisation of knowledge and the reflection upon knowledge. The result of these processes is that Science has a new essence: the self-production of new elements. This entails a shift from the imperative of usefulness of scientific knowledge to the production of truths and addition of novelty. It is clear that the novelty, in such a system, derives from the past knowledge reproduced by the system itself. This is exactly the mechanism which makes modern Science what it is; it consents to the accumulation of scientific knowledge:

> The idea that you have a collection of masterpieces from the past which remain in the chest of cultural treasures [...] started in the nineteenth century (Feyerabend, 2011, p. 26).

The accomplishment of the autopoietic era for scientific communities is determined by the exclusive reliance on knowledge that is self-produced by the system itself**.** When the systematic unity becomes the distinguishing criterion of Science:

> The acceptance of empirical and experimental knowledge means at the same time the acceptance of the self-production of the elements of knowledge and a confrontation with self-reference. […] Synthetic substances, the institutionalization of laboratories, and the dependence on self-constructed instruments confronts Science with the reality of self-referential circularity (Stichweh, 1990, p. 204).

The disciplinary differentiation of Science ensures an indefinite internal self-stimulation: differentiation and constitution become two complementary and fundamental mechanisms of Science as autopoietic system. Finally, another important fact must be underlined: publications become the universal medium of scientific communication (Stichweh, 1990).

Now that we have sketched the historical path that made Science what it is today, we can try to identify the re-productive mechanisms taking place in this autopoietic environment. As we discussed previously, Luhmann furnishes a general frame for applying autopoietic theory to social systems, but its conceptualisation lacks specifications about the mechanisms involved in the process of self-re/production. Furthermore, he did not specify if and how it is possible to identify those boundaries constructed and maintained by the system, which in turn allow the system to exist. It seems plausible to identify the origin of this lack of specification in the impossibility to furnish a description of all dynamics active in a system as well as in the fuzziness of boundaries characterised by translation and overflow mechanisms (Callon, 1986, 1998; Latour, 1987, 1988). However, in the following pages a model specifying the above mentioned elements is proposed. The basic idea is to show how Merton and Small's conceptualisations – in the first case related to both the institutional features of Science and citations, while in the second related to citations – can be inserted in Leydesdorff's reformulation of Luhmann and Giddens' theories, with the purpose of proposing a description of the way in which scientific communities re-produce themselves.

Luhmann's idea that communication is constitutive of social systems is here shared. Communication is the engine of the autopoietic machine. Actually, in society there are various communicative networks, which refer to specific communities. As previously mentioned, communication works as an inclusion and exclusion device. The type of communication considered in these pages is *scientific communication*, and "the difference between true and false is what matters for Science's code" (Luhmann, 1986/1989, p. 76). Scientific communication conveys information about representations of the world, which are derived from the cognitive process of knowledge of this latter. When extending autopoietic properties also to social systems, Maturana and Varela (1980) focused on the cognitive mechanisms involved in the process of knowing the surrounding reality.

In Maturana's words, the cognitive domain of Science is conceived of as follows:

> The use of criterion of validation of scientific explanations defines and constitutes scientific explanations. The use of scientific explanations to validate a statement makes that statement a scientific statement. The use of scientific explanations by the members of a community of standard observers[76] to directly or indirectly validate all their statements defines and constitutes science as a cognitive domain that defines as a scientific community the community of those observers that use it. Therefore, ontologically, in its manner of constitution as a cognitive domain, science is no different from other cognitive domains because it is defined and constituted as all cognitive domains are, namely, as a domain of actions defined by a criterion of validation or acceptability used by an observer or by the members of a community of observers to accept those actions as valid in a domain of science defined by that very same criterion of acceptability (Maturana, 1990 p. 24).

Thus, the usage of scientific explanations by scientists establishes their validity. In that way, the intellectual structure of a community is dynamically re-created; the users are defined as members of a community. The key element in this mechanism seems to be "the usage". But what does "the usage of scientific explanation" consist of in a scientific community? Communication in Science is realised through publications. Thus, scientific explanations, and in general scientific knowledge, are contained in written documents constituting scientific literature. An interesting description of how publications spread in a community is given by Bruno Latour:

> A statement is [...] always in jeopardy, much like the ball in a game of rugby. If no player takes it up, it simply sits on the grass. To have it move again you need an action, for someone to seize

[76] As Maturana states: "We scientists do Science as observers explaining what they observe. As observers we are human beings. We human beings already find ourselves in the situation of observers observing when we begin to observe our observing in our attempt to describe and explain what we do" (Maturana, 1990, p. 12).

> and throw it; but the throw depends in turn on the hostility, speed, deftness or tactics of the others. At any point, the trajectory of the ball may be interrupted, deflected or diverted by the other team – playing here the role of the dissenters – and interrupted, deflected or diverted by the players of your own team. The total movement of the ball, of a statement, of an artefact, will depend to some extent on your action but to a much greater extent on that of a crowd over which you have little control. The construction of fact, like a game of rugby, is thus a collective process (Latour, 1987, p. 104).

And in which way is a scientific publication "taken up"? What is the underlying mechanism? What is assumed here is that the answer lies in the citation[77] phenomenon. Obviously, I am not implying that scientific communities are re-produced only through citations. The attempt here is to furnish just a contribution to the lacunous area of inquiry that is "citation theory", while being aware that the reaching of a complete theory, encompassing also the description of citers' behaviour, is far from achieved.

In Luhmann's theory human beings are not components of social systems. However, meaning is generated in human interactions, which means that the existence of social systems depends on human beings. This potential contradiction inside the theory is solved by Luhmann thanks to the concept of *interpenetration*. He adopted a revised version of Parsons' concept developed some years before (Parsons, 1951) in order to explain the structural relations between different types of systems. In this way, as previously mentioned, Luhmann adds an operational coupling to the structural one (Leydesdorff, 2010).

By using the concept of interpenetration Luhmann is able to conceive of human beings as part of multiple systems at the same time: the psychological (based on consciousness), the biological (human beings as biologic organisms) and the social (based on communication). Assuming the inextricability of the social and psychological system, it seems possible to reject Wouters' thesis[78] according to which we need two different theories for describing the citation phenomenon – "one set of theories should be focused on the citing behaviour of scientists and scholars […]; the other set of theories should be devoted to the function of the citation [..."] (Wouters, 1999b, p. 62). Of course, as Wouters stated, we can identify a technical differentiation between citations and bibliographic references but, as underlined by Moed: "Reference and citation theories, although analytically distinct, should not be separated from one another" (Moed, 2005, p. 16). Moreover, as underlined by Nicolaisen: "[...] if we are to understand the nature of the citation, we need to understand the nature of the reference. And if we are to understand the nature of the reference we need a theory of citing that explains why authors cite the way they do. Ignoring the reference [...] in order to understand the citation is logically impossible (Nicolaisen, 2007, p. 633).

Even if the practice of explicitly citing previous works goes back to the second half of the 19th century

[77] No distinction between citation and bibliographic references is assumed here (for the opposite opinion, see Wouters, 1999a).

[78] According to Wouters, the citation is the mirror image of the reference, thus they have different semiotic properties: "[…] the citation is the product of the citation indexer, not of the scientist" (Wouters 1999a:4); "If the citation is distinct from the reference, it seems natural to construct different citation theories" (Wouters, 1999b, p. 562).

(De Bellis, 2009; Leydesdorff & Wouters, 1999), the statement by Leydesdorff (1998) about the emergence of modern citation, thus fulfilling the function of code and medium of communication during the early 20th century, seems to be consistent with the dating estimated by Stichweh related to the achievement of the autopoietic status by Science. As already specified, such an achievement means that Science relies only on knowledge that is self-produced by the system itself. Considering citations as "codes and medium of communication" also means to confer on them a specific function. In order to reveal this function we need to clarify what we mean when talking about citations as "codes". What is the kind of communication conveyed by them? The answer is anything but trivial, as there is no agreement about the meaning(s) and the function(s) of citations. Nevertheless, Bibliometrics is based on the general assumption that a bibliographic reference stands for an intellectual link; quoting is an intellectual choice which establishes a relationship between who cites and who is cited. Such a statement relies on Robert Merton's conceptualisation of the academic reward system in Science, according to which a reference is the payment of an intellectual debt. It is used in Science to acknowledge colleagues' work.[79] But talking about "payment of debts" and "acknowledgement of colleagues' work" implies some specific features of scientific knowledge. Intellectual property in Science is anomalous as it is established only when given away, that is to say published, and recognition comes by the usage of published scientific results.[80] Furthermore, usage of scientific literature means to cite. Merton also detects institutional consequences of this process: the symbolic credit (conveyed in the bibliographic reference) flows naturally from the documents to the authors, as it is converted into institutional roles and positions of power, and from the authors to institutions, nations and magazines (Merton, 1979). But, of course, citations have also an important cognitive function: the transmission of knowledge (Merton & Gaston, 1977). The act of citing is the way in which, as Latour argues, publications are taken and spread in the scientific community. At this point another question needs to be asked: why did citations become a code and a medium of communication? The answer stands in the cumulative nature of modern scientific knowledge, which entails the need to cite back in time:

> Citations and references thus operate within a jointly cognitive and moral framework. In their cognitive aspect, they are designed to provide the historical lineage of knowledge and to guide readers of new work to sources they may want to check or draw upon for themselves. In their moral aspect, they are designed to repay intellectual debts in the only form in which this can be done: through open acknowledgment of them. Such repayment is no minor normative requirement. That is plain from the moral and sometimes legal sanctions visited upon those judged to have violated the norm through the kinds of grand and petty intellectual larceny which we know as plagiarism (Merton, 1979, p. VI).

---

[79] In this regard, Eugene Garfield asked Robert Merton to write the foreword to his book entitled "Citation indexing - Its theory and application" (Garfield, 1979).

[80] See Footnote 27.

Scientometricians found in the Mertonian norms a theoretical frame useful to describe citers' behaviour as rational and predictable because it is regulated by a professional ethics.[81] The existence of norms is proved in a counterintuitive way, that is to say, through the sanctions imposed by the community on deviant behaviours. As Giddens stated: "The communication of meaning in interaction, it should be stressed, is separable only analytically from the operation of normative sanctions" (Giddens, 1984, p. 28). Or to paraphrase Luhmann – who, differently from Giddens, does not believe that meanings are communicated only in interactions - sanctions are a consequence of disappointed expectations. In the context of citations sanctions occur as consequence of the violation of the so-called "humility norm" (Kaplan, 1965[82]; Merton, 1957), according to which authors are expected to acknowledge prior work, accurately respecting original author's intentions. The analysis relating to the institutional features of Science, therefore, seems to be fundamental for understanding its internal processes: "from an evolutionary perspective, [...], the institutions can be considered as the fingerprints of the communication patterns that have been functional for the reproduction of the system hitherto" (Leydesdorff, 2001, p.182).

Stating that citations allow the transmission of knowledge is not a trivial claim. How do they perform such a function? Citations work like symbols, or better, as Henry Small explains, if a citation does not derive from a fraud, it must be considered both as a sign (with objective character) and a symbol (symbol of the concept contained in the cited work) (Small, 1978). A symbol can be a "nonce" one or a "standard" one: the former is used to convey citer's ideas, while the latter produces meanings which are shared by a scientific community or a group of scientists and, therefore, are frequently cited (Small, 1978). Here we find a very strong similarity with Maturana's reasoning about the (re)production of scientific communities: "The use of scientific explanations by the members of a community of standard observers to directly or indirectly validate all their statements defines and constitutes Science as a cognitive domain that defines as a scientific community of those observes that use it" (Maturana, 1990, p. 24). Small also concentrated on the construction of the meaning of knowledge conveyed by the cited paper (the one the author wanted to communicate):

> [...] as a document is repeatedly cited, the citers engage in a dialogue on the document's significance. The verdict or consensus which emerges (if one does) from this dialogue is manifested as a uniform terminology in the contexts of citation. Meaning has been conferred through usage and what is regarded and accepted as currently valid theory or procedure has been socially selected and defined (Small, 1978, p. 338).[83]

In his conceptualisation Small combined agreement and construction: "It is here that normative and

[81] See Footnote 75.
[82] Even if Merton developed the "humility norm", it was Norman Kaplan (1965) who first applied it to Bibliometrics.
[83] Small stresses that such a process could result in a transformation of the meaning conveyed by the cited document. Thus, it comes to symbolise something different from what the author intended.

constructivist approaches can find common ground. […] Here the symbolism of reward and the symbolism of meaning are operating in tandem" (Small, 2004, p 76). Moreover, here we find those elements which according to Luhmann constitute communicative processes, that is to say, information (message content), utterance (form of production of the information plus sender's intentions) and understanding (generated meaning from receiver's point of view) (Luhmann, [1984] 1995). Luhmann model can be read as Shannon's communication model (Shannon, 1948) with the addition that the content of the (codified) message matters.[84]

On the basis of what has been argued so far, citations seem to have a double function; on the one hand, they constrain scientists' actions, as they have to cite properly, and on the other hand, they enable the accumulation of scientific knowledge. Actually, the constraining and enabling properties are received by citations from the norms regulating the system. As Leydesdorff explains:

> […] the *system* under study *is* accessible as expectations entertained in the reflexive systems of communication. Scientific discourses generate standards endogenously which function as codes of communication, but these codes themselves also remain discursive constructs. They operate as methodological programs on the substantive programs, and this co-evolution and mutual shaping can develop along a trajectory and thus stabilize otherwise volatile representations (Leydesdorff, 2007, pp. 15,16, italics mine).

Accordingly, citations work as devices by means of which the system is defined, produced and re-produced: the system is self-reproducing. To borrow Merton's terminology, the "latent function"[85] (Merton, 1949) of citations is the re-production of the intellectual and cognitive structure of scientific communities:

> [...] when scientists agree on what constitutes prior relevant literature, including what is significant in that literature, they are in fact defining the structures of their communities (Small, 2004, p. 72).

Thus, it seems possible to argue that bibliometric distributions inform about which kind of information (scientific knowledge) is reproduced in the system, and about the cognitive or intellectual structure[86] of

[84] In 1948 Shannon wrote: "these semantic aspects of communication are irrelevant to the engineering problem" (Shannon, 1948, p. 379).

[85] Merton distinguishes between "manifest" and "latent" functions. The first refers to the conscious and deliberate – the expected consequences produced by actors. The second refers to the structural consequences of the selective operations (Merton, 1949).

[86] Even if sometimes the terms "cognitive structure" and "intellectual structure" are used as synonyms, they refer to two different things. As explained by Chen (2006), a discipline can be divided into an intellectual base and a research front. The former can be defined as the complex of previous literature cited by current works; the latter is the complex of current works. Research fronts develop on the basis of previous literature. Thus, the reference to intellectual source establishes connections among new and old

scientific communities. This is also useful in highlighting new trends in knowledge development.

The term "structure" (and all its derivatives) is often used in Scientometrics, and more precisely in Relational Bibliometrics. The purpose of Relational Bibliometrics is that of intercepting and studying relationships within scientific literature, in order to discover the intellectual structure of research fields, the emergence of new research fronts or national and international co-authorship patterns (Thelwall, 2008). As Small and Griffith affirmed: "Our task is to depict […] Science relationships in ways that shed light on the structure of Science" (Small & Griffith, 1974, p. 40). Moreover: "[…] a map of bibliographic data is a useful heuristic device by providing a visible organizing structure to information" (Small, 1999, p. 812). Even if the detection of Science's structures seems to be a clear aim of Relational Bibliometrics, it is not possible to find the same clarity with reference to the meaning that the term structure has in this field. By conceiving scientific communities as cognitive domains it seems reasonable to think about the possibility to visualise the cognitive structure of scientific communities; as a consequence, it should also be possible to visualise disciplines, specialties or research areas boundaries. Thus, we can state that, depending on the technique employed (e.g. author-based, documents-based or semantic-based techniques),[87] Bibliometrics enables us to visualise boundaries as authors, documents or words. However, what do these structures detected by relational bibliometric mean? It seems that in the literature it is not possible to find clarifications about this subject.

In Sociology there are two approaches that conceptualise "structures": functionalism and structuralism/post-structuralism. As Giddens explains:

> […] structure is usually understood by functionalists […] as some kind of 'patterning' of social relations or social phenomena. This is often naively conceived of in terms of visual imagery, akin to the skeleton or morphology of an organism or the girders of a building. Such conceptions are closely connected to the dualism of subject and social object: 'structure' here appears as 'external' to human action, as a source of constraint on the free initiative of the independently constituted subject. As conceptualized in structuralist and post-structuralist thought, on the other hand, the notion of structure is more interesting. Here it is characteristically thought of not as a patterning of presences but as an intersection of presence and absence; underlying codes have to be inferred from surface manifestation (Giddens, 1984, p. 16).

In these (apparently) different conceptions of structures, Giddens sees a unitary whole, which takes shape in the formulation of the so-called "structuration theory". The possibility to apply Giddens' structuration theory to Scientometrics has already been suggested more than once by both Cronin and Leydesdorff (for example, see Cronin, 1984; Leydesdorff, 1998, 2001). In 1998 Cronin affirmed: "Structuration may offer

---

concepts. It is from the representations of these connections that the cognitive structure of research fields emerges (Small, 1981).

[87] See Footnote 68.

the possibility of carving out a middle ground between naïve normativism and extreme relativism" (Cronin, 1998, p. 53).

Giddens' theory has been widely used in organisational studies[88] focused on the production and reproduction of meaning through communication, with the aim of "examining simultaneously how meanings emerge from interaction and how they act to constrain subsequent interaction" (Monge & Contractor, 2003, p. 20). A complementarity between Gidden's and Luhmann's theoretical systems has already been identified by Mingers (2004). In 2010 Leydesdorff recombined the two theories in a new one consisting of the following factors: "structuration of expectations, interactions, organisation, and self-organisation of intentional communications" (Leydesdorff, 2010, p. 2138). In order to do this, he translated Giddens' action theory in a theory based on the structuration of expectations operating at the supraindividual level as horizons of meaning: "codes of expectations structurate the reproduced relations among expectations over time in addition to the structures that the communication networks contain at any given moment" (Leydesdorff, 2010, p. 2142). Furthermore, he stressed how the social order (made of expectations) and the individual expectations mutually interpenetrate underlining the dependence of both the autopoietic process on the reflexivity and of the communication on actors (reconnecting to Habermas, 1981). Finally, he modifies Luhmann's theory with reference to the conception of the codes as binary: "the codes are not given (as in DNA) but remain reflexively under constant reconstruction" (Leydesdorff, 2010, p. 2142).

Once performed the translation of the action-theoretical concepts into communication-theoretical ones, it seems possible to say that the structuration theory seems to have all those features required by an autopoietic system: self-organisation, self-reproduction, self-production and self-referentiality. Obviously, for the purpose of a unified approach, we have to substitute human agents with communicative events defined according to Luhmann, and accept his suggestion of a paradigm shift: "society should no longer be considered as composed of human beings, but as consisting of communication" (Luhmann, 1984/1995). As previously stated, this does not imply the exclusion of human beings from the theory, who are instead fundamental to the production and reproduction of the system. They are part of the psychological and the link between the two systems is enabled by the interpenetration mechanism, which allows us to consider an element as part of two (or more) systems at the same time (Luhmann, 1984/1995). As maintained by constructivist positions in Sociology of Science, citing is mainly a rhetorical activity employed by scientists to support their results and persuade readers of the worth of their assertions (Gilbert, 1977; Latour, 1987). Even if one does not totally agree with this view, to deny the fact that the citation phenomenon, which occurs in the social system, is affected by elements coming from other systems, such as the psychological one, can be euphemistically defined as an unwise thought. Further details about these statements are given in the following discussion. Components of the systems, besides communicative events, are also *structures*, which refer to the relations of transformation and mediation triggering and allowing the reproduction of the system (Giddens, 1984)

[88] For a review on the topic see Monge and Contractor (2003).

Giddens' concept of "structure" (with its property of duality)[89], bearing in mind Leydesdorff's translation mentioned above, captures all aspects discussed so far (normative, cognitive and re-productive dimension) and, as suggested above the way in which he conceptualises resources is fundamental to clarify what we see when watching bibliometric maps. Specifically, s*tructure* is recursively organised and consists of institutionalised sets of rules and resources implicated in social reproduction; it is out of time and space and is temporarily crystallised in its instantiation. Of particular importance here is the following statement: "structure can be conceptualized abstractly as two aspects of rules: normative elements and codes of signification" (Giddens, 1984, p. XXXI). In this way an immediate link to Mertonian conceptual apparatus and the reasoning referring to communication is established. In terms of "resources" Giddens means two types of factors: "allocative" and "authoritative" resources. The first refers to "forms of transformative capacity generating command over objects, goods or material phenomena" (Giddens, 1984, p. 33); the second relates to "types of transformative capacity generating command over persons or actors" (ibidem). What is suggested here is that bibliometric maps provide information with respect to *authoritative* resources, which are represented as nodes in bibliometric networks. Authors and documents which appear in these maps and have been selected because they received at least a certain amount of citations[90] – according to the way in which the re-production of scientific communities has been here described and the fact that the normative approach is assumed to be able to describe high citation counts – are supposed to exercise a big influence on the production of new knowledge.[91] Obviously, these considerations are also valid for the various semantic elements (e.g. terms or authors' names) carrying meaning in a scientific text.

Of course, the concept of influence can be very elusive. The way in which previous highly cited works - paradigmatic works (Small, 2004) influence the production of "new" knowledge in a system based on the accumulation of knowledge – but characterised by microrevolutions (Toulmin, 1970) - is here conceived of as follows: 1) what is not in line with more or less established traditions will be considered "deviant", and what is considered deviant will, presumably, occupy a marginal position with respect to the scientific mainstream. This is because of the following dynamics: in Science there are two opposite forces at work, namely conservatism and the norm of open-mindedness (Barber, 1961), and, as underlined by Bourdieu (1975), the success in introducing novelties in Science seems to be the privilege of those scholars who are rich of *scientific capital*[92]. In practical terms, we can suppose it is more convenient

---

[89] "The constitution of agents and structures are not two independently given sets of phenomena, a dualism, but represent a duality. According to the notion of the duality of structure, the structural properties of social systems are both medium and outcome of the practices they recursively organize. Structure is not external to individuals […]. Structure is not be equated with constraints but is always both constraints and enabling" (Giddens, 1984, p. 25).

[90] These, therefore, can be considered *relatively* highly cited.

[91] I must underline that it is one thing to detect the "relevant" items, and another to detect the direction of their importance. In particular, I am referring to those interpretation problems linked to very highly cited items, whose localisation in space is not always easy.

[92] For the notion of "scientific capital" see Bourdieu (1975).

(mainly for those who are not affirmed scholars) to follow the old path than engaging in a marginal one[93]; 2) criticisms, new paths of research, new techniques, new ideas, and new theories will raise from/against/for differentiation with respect to the knowledge developed until that moment. For example, in speaking about Kuhn's *The structure of scientific revolutions* (1962/1970), Bourdieu stated: "[...] Thomas Kuhn radically changed the space of theoretical possibles in sociology of science" (Bourdieu, 2004, p. 14).

6.4. IMPLICATIONS FOR SCIENTOMETRIC INDICATORS AND MAPS

Conceiving of Science as an autopoietic cognitive domain, and communicative events as the elements of this system, allows us to interpret citations as devices through which the *structuration* process in scientific communities is accomplished. According to the view sustained here, citations are codes and medium of communication incorporating the cognitive, intellectual and social dimension of scientific activity. Thus, citing means to reproduce (but also to change) the structural patterns of a scientific community (specialty, research area, etc):

> Thus, the modern citation is constitutionally complex, and therefore it can function in scientific practices by indicating both the cognitive and the social contexts of a knowledge claim. At a generalized level, citations, as potentially repeated operations, sustain communication in the sciences by drawing upon cognitive and social contexts (Leydesdorff, 1998, p. 9).

This implies that analysing bibliometric distributions, working with relational techniques as well as correlations between patterns of relations allows us to obtain information about social, intellectual and cognitive patterns in Science domains. It is fundamental to assign a meaning to citations in order to understand what it means for a paper or an author to be frequently cited or, conversely, to be rarely cited. Thus, we need to understand why authors cite, which in turn entails obtaining responses with reference to the *manifest functions*[94] (Merton, 1949) of citations. Traditionally, answers arrive from two opposite sociological approaches: the normative and the constructivist. In these pages it is suggested that normative theories on citations are useful in describing high citations counts, while constructivist ones help in understanding low frequencies. However, it must be underlined that, as already suggested (see, for example, Baldi, 1998; Moed, 2005; Zuckerman, 1987), rhetorical reasons and normative ones are not mutually exclusive; it is quite normal for a scientist to cite internationally recognised works/authors to support his/her findings and thus to persuade future readers of the worth and validity of his/her work. Persuasion plays a fundamental role in Science: partly depending on it manuscripts will be published or not and, in case of publication, authors will have the possibility to be rewarded. It is interesting to note

[93] This does not mean that marginal or new traditions are not pursued.
[94] See Note 85.

that such constructivist considerations are strictly linked to those by Robert Merton on the function of papers as means for establishing priority as well as for establishing and maintaining intellectual property (Merton, 1957; Merton, 1961). As clearly explained by Price:

> The prime object of the scientist is not, after all, the publication of scientific papers. Further, the paper is not for him purely and simply a means of communication knowledge. […] If, then, the prototype of the modern scientific paper is a social device rather than a technique for cumulating quanta of information, what strong force called it into being and kept it alive? Beyond a doubt, the motive was the establishment and maintenance of intellectual property. [...] the never-gentle art of establishing priority claims. […] For these reasons scientists have a strong urge to write papers but only a relatively mild one to read them (Price, 1963, pp. 62,65,70).

It seems possible to hypothesise that scientific norms work until a certain point; in other words, authors behave normatively with respect to the intellectual base (previous literature) and also with respect to certain current works constituting the research front and representing innovative knowledge (solution to problems or creative works) acknowledged by the community. Once a scientist has paid all his "intellectual debts" and has behaved humbly, it is possible to argue that other factors attributable to his personal inner (psychological) dimension affect the choice of whom to cite.

The validity of bibliometric tools in studying Science and producing indicators is supported by the approach presented in these pages to the extent that indicators and terms such as "impact" and "influence" are not used as synonyms for quality. In line with what Garfield stated in 1979 (Garfield, 1979), in this work is supported the idea that Scientometrics is able to detect those authors and works which *are considered* relevant in a scientific community as long as they are cited by scientists, influencing in this way the production of new knowledge. Obviously, relevance does not exclude quality,[95] but while we can affirm that quantity in usage reflects relevance[96], we cannot make a similar claim with reference to quality. As demonstrated by Shadish et al., for example, "citations are not interchangeable with quality for two reasons. First, citations clearly measure things in addition to quality, such as exemplar[97] status. […] Second, not all high-quality work is highly cited" (Shadish et al., 1995, p. 492).

On the basis of what has been argued so far, citations trends and words (as carriers of meaning) enable us to obtain information on intellectual, cognitive and social structures characterising disciplines or scientific communities in a certain temporal window. It means that observing citation trends as well as

[95] For a discussion of the topic see Nicolaisen (2002).
[96] The obliteration by incorporation phenomenon (Merton, 1968) is not neglected here. The position of the bibliometric community (MacRoberts & MacRoberts, 1989) is shared: obliteration by incorporation affects already highly cited authors or documents (Cole & Cole, 1983).
[97] The reference here is to the Kuhnian concept of "exemplars", scientific achievements of the past constituting a model for the solutions of scientific problems (called by Kuhn "puzzles") (Kuhn, 1962/1970).

word and co-word ones in time allows us to monitor those variations that Toulmin (revisiting Kuhn's micro-revolution concept) calls “microrevolutions” (Toulmin, 1970). Toulmin conceptualises the intellectual change made of the categories of *tradition*, *innovation* and *selection*; he conceives of the “micro-revolutions” as units of variations:

> [...] we will then be faced with a picture of science in which the theories currently accepted at each stage serve as starting-points for a large number of suggested variants; but in which only a small fraction of these variants in fact survive and become established within the body of ideas passed on to the next generation (Toulmin, 1970, p. 46).

Toulmin distinguishes among three aspects of this micro-revolutional evolution. The *quantity* of innovation (in a given field at any time), the *direction* of the innovation, and the *selection criteria* on which basis variants are chosen to be perpetuated (Toulmin, 1970). If for information about the selection criteria we need to refer directly to scientists (analysing for example the peer review process) as we are interested in knowing the reasons underlying a selection, regarding *quantity* and *direction* we can refer to bibliometric analysis of scientific products. In fact, assuming the correctness of what has been argued so far, it can be reasonable to think that bibliometric tools constitute a suitable technical apparatus to study the growth, or better the evolution, of scientific knowledge. Bibliometrics seems able to detect the variation in the “historical organisation of communications whose instantiation is continuously disturbed by new interactions” (Leydesdorff, 2010, p. 2148).

SECTION 2
THE CASE OF ITALIAN SOCIOLOGY

7. THE ORGANISATION OF SCIENTIFIC COMMUNITIES

7.1. THE INSTITUTIONALISATION OF SCIENTIFIC KNOWLEDGE

It is not possible to talk about Science and its structure without considering the context in which it is developed, that to say universities:

> [...] universities have most often provided the appropriate context in which a professional identity might be built. Their need of teaching cadres and their tolerance of research have permitted the growth of regular career structures directly or indirectly dependent on the discipline. Prizes, research grants, specially equipped facilities, prestigious positions, and honorific appointments […] the existence of career-related opportunities and rewards gives new meaning to an ability to perform well at the intellectual tasks of the discipline. The growth of such socially patterned arrangements inevitably transforms the enterprise on deeper levels. […] It also encompasses a transformation in the images of the discipline and its social and cultural functions as variously perceived by its practitioners, its patrons, and its larger public (Thackray & Merton, 1972, p. 474).

The consequence of the words contained in the passage above is the following: it is not possible to distinguish the emergence of cognitive identities from that of professional identities, as they are strictly interwoven (Thackray & Merton, 1972). Robert Merton was the first sociologist who took as object of study Science, becoming thus the founding father of what is called *Sociology of Science*. Actually, mertonian Sociology of Science is called *Institutional Sociology of Science* because of the attention that Merton gave to the institutional aspect of Science:

> We are here concerned in a preliminary fashion with the cultural structure of science, that is, with one limited aspect of science as an institution. Thus, we shall consider, not the methods of science, but the mores with which they are hedged about (Merton, 1973, p. 268).

Therefore, Science is considered as:

> [...] the organized social activity of men and women who are concerned with extending man's body of empirical knowledge [...]. The relationships among these people, guided by a set of shared norms, constitute the social characteristics of science (Storer, 1973, p. 3).

Thanks to Merton, the pure Science paradigm comes to an end. Science is not an autonomous activity, it is not conducted in a social vacuum and therefore its products can no longer be considered undisputed truths (Merton, 1949). It is considered as a social subsystem ruled by norms and values, which interacts with other subsystems preserving its own autonomy. The relationship between Science and the surrounding social structure is defined by Merton as *dynamically interdependent*:

> [...] the socially patterned interests, motivations, and behavior established in one institutional sphere […] are interdependent with the socially patterned interests, motivations, and behavior obtaining in other institutional spheres [...]. The same individuals have multiple social status and roles [...]. This fundamental linkage in social structure in itself makes for some interplay between otherwise distinct institutional sphere even when they are segregated into seemingly autonomous departments of life (Merton, 1973, p. 175).

According to Merton, Science is an autonomous self-regulating sub-system characterised by the so called major norms, or institutional imperatives, in which resides the ethos of Science, "a statement of their interdependence as well as their functional relationships to the formal goal of scientific work: the extension of certified knowledge" (Merton, 1973, p. 226). These norms are: *universalism* (scientific results are to be subject to preestablished impersonal criteria and are judged independently of author's personal or social attributes); *communism* (even if scientists have to be rewarded for their scientific contributes, these are products of social collaboration, they are not property of the researcher but constitute a common heritage both of the scientific community and of the whole society); *disinterestedness* (absence of economical or personal motivations in research activities, guaranteed by a pattern of institutional control); *organized skepticism* (it is both a methodological and institutional mandate which refers to the temporary suspension of judgment over all scientific result, and the detached scrutiny of beliefs in terms of empirical and logical criteria). These imperatives are internalised by scientists having as result the creation of a scientific conscience (Merton, 1949).

Mertonian Sociology of Science attracted many criticisms and has been considered a sort of prescriptive idealisation. But as Merton stressed, his norms refer to Science as an institution, and not to scientists' motivation system; the functions of such norms would be demonstrated by the sanctions to deviant behaviours imposed by the scientific community (Merton, 1973). Merton and his group, the so called *first circle*, focused on some mechanisms characterising scientific communities: acknowledgement as reward; the Matthew effect (cumulative effect rewarding exponentially those who occupy privileged institutional position); the establishment of intellectual priority by publication; acknowledgement through citations; the phenomenon of independent multiple discoveries (as evidence of the social determination of knowledge) (Merton, 1949, 1973). Merton was mainly focused on the macro and meso level:

> We are here concerned in a preliminary fashion with the cultural structure of science, that is, with one limited aspect of science as an institution. Thus, we shall consider, not the methods of science, but the mores with which they are hedged about (Merton, 1973, p. 268).

but it does not mean that he was unaware of the fact that there was also a micro level. In the late 1940s, for example, he stressed that methodological books present ideal patterns about the way in which scientists must think, feel and act, but these normative patterns do not reflect the way in which scientists really act. For these reasons, the official scientific documentation is not useful for a complete reconstruction of the scientific development (Merton, 1949).

In 1976 Richard Whitley wrote the paper "Umbrella and Polytheistic Scientific Disciplines and Their Elites" (Whitley, 1976), where the relationships among disciplines, specialties and research areas are studied taking in consideration the institutional context in which they are embedded: universities. Whitley distinguishes between *umbrella* and *polytheistic* disciplines: the former are characterised by a generalised epistemological consensus, a high level of institutionalisation, and well defined and autonomous (with respect to the discipline) specialties; the latter refers to disciplines characterised by different views about discipline's nature, not highly institutionalised specialties (and research areas), a low level of institutionalisation, and a dependence of specialties to the discipline. The degree of institutionalisation refers to the extent to which both the ordering principle is formulated and members of a discipline share a focus of commitment. As a consequence, in case of strongly institutionalised disciplines it is difficult that metaphysical debates start, unless they derive from serous inconsistencies and failures in key research approaches. Otherwise, weakly institutionalised disciplines are characterised by metaphysical debates not deriving from specific scientific results or scientific activities (Whitley, 1976).[98]

From the Institutional Sociology of Science perspective, Whitley's work is interesting as it shows how disciplinary structure is reproduced inside and by universities, and the consequences of the institutionalisation of disciplines in the latter (bearing in mind that not all scientific activities are institutionalised in universities). In polytheistic disciplines there is the tendency towards the identification of units of research with those of recruitment, education and careers. Such a mechanism tends to institutionalise organisational authority as intellectual authority; research mainly focuses on developing techniques and research practices which are not part of a general strategy. As a consequence, research is conducted with reference to a certain disciplinary view, general cultural values[99] and the availability of resources:

> As long as the discipline is the primary unit of social and scientific identity, and is also institutionalized in university departments, disciplinary authority is going to be strongly linked to organisational authority. [...] This implies that there is no coherent disciplinary elite, in the

[98] Whitley's conceptualisations about scientific disciplines are deepened in Section 7.3.

[99] As Fleck claimed: "In science, just as in art and in life, only that which is true to culture is true to nature" (Fleck, 1935/1979, p. 35).

> sense of a group of scientists acting as guardians and interpreters of the institutionalized ordering principle, legitimating specialty concerns and models and allocating resources. Instead, there is a plurality of views which are not institutionalized throughout the discipline; given the strong influence of organisational structures, we would anticipate these to be linked to particular organisational units. Particular conceptions of the discipline and appropriate research will tend to be localized in departmental structures, because these structures offer careers, social support and intellectual commonality. Particular types of work in the discipline tend to be conducted in certain organisations and not others […] (Whitley, 1976, p. 485).

This means that organisational and intellectual authority overlap in polytheistic disciplines. Elites are based on organisational positions through which they control the reproduction of labour force, the allocation of rewards and the most important disciplinary journals as well as other media. Resources and training are fundamental as they constitute the basis of hegemony. Due to the lack of an unshared disciplinary definition external sources of influence might become the criterion for establishing disciplinary status and authority. In this way, polytheistic disciplines will probably be characterised by different changing "coalitions and alliances of organisational elites, subject to considerable influence from the state and other goal-setting and resource allocation agencies" (Whitley, 1976, p. 486). It is important to underline that:

> It should be borne in mind that where divergent views are institutionalized in university departments they do not have equal resources, and they vary in their consonance with dominant conceptions of science and *commonsense* rationality. This works on two levels. First, there is the location of the department within the university, its claim on resources and facilities, and its legitimacy in the university. The effect of this location is obviously connected with the overall status and size of the university. Secondly, there are direct relations between the department and external agencies which provide resources and exercise influence. […] In a situation of competing disciplinary views, then, some will have more facilities than others and this may lead to particular conceptions becoming dominant (Whitley, 1976, pp. 486,487).

It is clear that in polytheistic disciplines teaching departments play a fundamental role with reference to both social control and career opportunities.

The basis of authority in umbrella disciplines, instead, is constituted by a shared ordering principle and thus the pursuit of common goals. Here organisational authority and general culture contribute in an indirect way to the formation of elites. Also disciplinary elite performs its influence in an indirect way:

> First, it legitimates, and is the main agent of allocating resources to, specialties; and, secondly, it is the main agent, not of reproducing the labour force directly, but of reproducing the

> recruitment stratum from which specialties recruit and socialize their labour force (Whitley, 1976, p. 488).

Unlike polytheistic discipline, here department universities and laboratories are not the privileged places where elites exert their influence. While in the former university departments are the major locus of influence and control of both work and scientific development, in umbrella disciplines such functions are performed by specialties. Here the system of influences plays in reverse: from specialties to discipline, and from them to organisational authority (Whitley, 1976).

7.2. THE SCIENTIFIC FIELD

Bourdieu's concept of scientific field focuses on the political, and thus strategic, nature of dynamics within scientific communities. According to Bourdieu, a field is that domain of specialised practice (e.g. art, literature, sociology) with its own logic and constituted by a unique combination of different types of capital, e.g. financial capital, symbolic capital (e.g. prestige and reputation) or social capital (Bourdieu, 1993, 1996, 2005). The result of this combination is called *scientific capital,* which can be defined as a set of properties which are the product of acts of knowledge and recognition performed by agents engaged in the scientific field (Bourdieu, 1975). The scientific field is that objective space defined by the opposing forces involved in the struggle for scientific stakes (Bourdieu, 1975):

> It is the scientific field which, as the locus of a political struggle for scientific domination, assigns each researcher, as a function of his position within it, his indissociably political and scientific problems and his methods - scientific strategies which, being expressly or objectively defined by reference to the system of political and scientific positions constituting the scientific field, are at the same time political strategies. Every scientific "choice" [...] is a political investment strategy, directed, objectively at least, towards maximisation of strictly scientific profit, i.e. of potential recognition by the agent's competitor-peers (Bourdieu, 1975, pp. 22,23).

Bourdieu's analysis of the internal dynamics of scientific communities emphasises the importance of the role occupied by scholars in academic hierarchies. The amount of scientific capital, which is fundamental to the strategies that scientists can adopt in the scientific activity, is strictly linked to the role occupied in the field. In this regards, Bourdieu distinguishes between *succession strategies* (strategies aimed at the maintaining of the scientific order – conservatives strategies) and *subversion strategies* (strategies aimed at overturning the established order) (Bourdieu, 1975). According to him, while the dominants (those rich in scientific capital) are principally orientated towards the former, the newcomers choose the former or the latter on the basis of their scientific capital. Specifically, "scientific revolution is the business not of the poorest but of the richest (in scientific capital) among the new entrants" (Bourdieu, 1975, p 33).

The distribution of power is what defines the structure of the scientific field, where *structure* means "the space of objective relations [...] of which the communicative exchanges we directly observe (interaction) are but the expression" (Bourdieu & Wacquant, 1992, p. 256). The structure of the field is the result of "previous struggles that is objectified in institutions and dispositions and commands the strategies and objective chances of the different agents or institutions in the present struggles" (Bourdieu, 1975, p. 27). Thus, the history of the scientific-political dynamics going on in the scientific field is crystallised in institutions.

According to Bourdieu, the official order of Science is composed of: 1) the complex of scientific resources inherited from the past existing in *the state of objectification* (i.e. objectified in instruments, text, institutions etc.); 2) *the state of incorporation* (encompassing the scientific habitus, and systems of generative schemes of perception, appreciation and action); 3) the educational system (having the purpose to perpetuate the official science by instilling it in the newcomers in the field of science); 4) the institutions responsible for consecration (for example academies, prizes, etc.); 5) the means of scientific communication, specifically scientific journals "which, by selecting their articles in terms of the dominant criteria, consecrate productions faithful to the principles of official science, thereby continuously holding out the example of what deserves the name of science, and exercise a de facto censorship of heretical productions, either by rejecting them outright or by simply discouraging the intention of even trying to publish them by means of the definition of the publishable which they set forward" (Bourdieu, 1975, p. 30).

Once again, the analysis of the institutional level can furnish interesting insights about the organisation of scientific activity inside universities as well as the consequences of such organisation with respect to the reproductive dynamics of the *scientific order*:

> The objective orchestration of the practical schemes inculcated by explicit instruction and familiarisation, which constitutes the basis of the practical consensus on what is at stake in the field, i.e. on the problems, methods and solutions immediately regarded as scientific, is itself based on the whole set of institutional mechanisms which ensure the social and academic selection of the researchers (through, for example, the established hierarchy of the disciplines), the training of the selected agents, control over access to the instruments of research and publication, etc. (Bourdieu, 1975, p. 34).

7.3. THE CONCEPTUALISATION OF SCIENTIFIC SPECIALTIES

> Specialization is the hallmark of modern science. With the institutionalization of science in universities, the fragmentation of knowledge into intellectual provinces called disciplines was legitimated (Chubin, 1976, p. 448).

Scientific specialties are the consequence of the division of scientific work: as a discipline grows and develops it fragments in fields of competence. Thus, specialisation seems to be the way in which disciplines naturally organise themselves (Stehr, 1974; Whitley, 1976). Specialisation is not an obstacle to universality, on the contrary: "[...] *the* divisions of labor in science […] are not detrimental to universality but are instead basic to its establishment" (Shinn, 2002, p. 99).

A specialty is a collection of individuals (scholars) smaller than a discipline or a field, thus often "specialty" and "subfield" (but also "area of inquiry") are used as synonyms. In 1966 Kadushin, after an analysis of the social organisation of research areas in Science, developed the concept of *social circle*: a social circle is characterised by the fact that relations among its members are denser than relations among members and non-members; social circles have *invisible colleges*[100], which help in unifying areas and in providing coherence and direction to their fields (Kadushin, 1966). Furthermore, he affirmed that a social circle is not a group. In particular: a circle may have a chain or network of indirect interactions such that most members of a circle are linked to other members, at least through a third party. It is, thus, not a pure face-to-face group; the network exists because members of the circle share common interests-political or cultural; the circle is not formal, thus there are no clear leaders, although there may be central figures, no clearly defined goals for the circle, though it almost always has some implicit functions, no definite rules which determine modes of interaction, though there are often customary relationships, and no distinct criteria of membership (Kadushin, 1966). Crane, who was one of the first to employ sociometric data to study scientific specialties, in 1969 published the paper "Social structure in a group of scientists: a test of the invisible college hypothesis" (Crane, 1969), where she tried to attest the existence of a social organisation within a research area on the bases of the following conditions: scientists who have published in the area have more social ties among them than with scientists who have not published; scientists who have published in the area can be differentiated according to the degree of social participation within the area (Crane, 1969). On the basis of her findings, Crane reflected on the fact that a social circle is something that can or cannot develop in a subfield, and if the latter, the size and importance of its members are mutable in time (ibid). Furthermore, referring to Fisher's claims related to the association between discontinuity in research areas personnel, lack of productivity, and failure in knowledge accumulation (Fisher, 1966, 1967), resulting in a continuous retreatment of research problems (Barton & Wilder, 1964), Crane developed the idea that the presence in the research area of scientists

[100] Kadushin refers to Price's concept of invisible college: unofficial organisation of scientists in contact with each other (Price, 1963).

sufficiently productive to make them visible to most of those who enter the field, even briefly, is the bases for the production of a social circle, which in turn plays a fundamental role in the normal growth of the research area (Crane, 1969).

Sociology started to focus on scientific specialties structure and organisation in the 1960s, but the 1970s were definitely more prolific. As Chubin claimed, "only recently have we recognized what the fragmentation of knowledge or the *ethnocentrism of disciplines* hath wrought" (Chubin, 1976, p. 448). Mertonian Sociology of Science mainly concentrated on the social organisation of scientific research, leaving aside what was instead the focus of historians, that is to say the conditions of discoveries, their consequences on scientific progress, and discoverers' careers (Chubin, 1976). One of the first sociological conceptualisations of the 1970s on scientific specialties is the following:

> The set of scientists in a discipline who engage in research along similar lines can be called the scientific specialty. It is reasonable to believe that scientists will communicate most often and intensively with others in their specialties, exchanging preprints with them, citing their work and exchanging reprints. […] scientific specialties are microenvironments for research, as frequently traversed regions on a blurred map of science (Hagstrom, 1970, pp. 91,93)

In 1966 Joseph Ben-David and Randall Collins developed a model in which social variables such as social status and competitive position were at the basis of an explanatory model of specialties formation (Ben-David & Collins, 1966). Another attempt in this direction is that by the Starnberg group[101], which developed a three-stage model inspired by Kuhn and Lakatos' works, involving the following phases: a pre-theoretical one without a definite theoretical or methodological program; paradigmatic one in which there is the birth of a field thanks to the development of a theoretical program; a post-paradigmatic or called "finalization" one of normal science, characterized by the maturity of the program (Schafer, 1983). In 1972 Nicholas Mullins presented his model consisting of a general theory made up of four stages of specialty development[102], according to which changes in specialty social structure are linked to theoretical and intellectual development. According to him, the first stage is identified with the *paradigm group*, in which a group of scientists is engaged in solving a specific problem, it is the minimal form of scientific group: "a paradigm group is thus a set of individuals, all of whom have moved into a similar cognitive situation with respect to the same, or similar, problems" (Mullins, 1972, p. 54). The second phase is the *communication network*, characterised by the expansion of the networks of those working on the problem, or better: "a set of pairs and triads of scientists engaged in regular communication, or colleagueship, over a period of time" (Mullins, 1972, p. 58). After this phase there is the cluster one: "a cluster forms when

[101] A research group created in Starnberg in Germany in the 1970s at the Max Planck Institute for the Study of the Conditions of Life in the Scientific-Technical World (Hess, 1997).

[102] "The analytical distinction of the different stages is not intended to imply that the activities which constitute the preceding stages are no longer functioning; it is intended only to show that for a given intellectual problem, a more complex structure has been established" (Mullins, 1972, p. 54).

scientists become self-conscious about their patterns of communication and begin to set boundaries around those who are working on their common problem" (Mullins, 1972, p. 69). The last one, *specialty* or *discipline*, is that in which there is the emergence of a discipline entailing formal organisation, training institution, formal meetings and specialised journals: "a specialty is an institutionalised cluster which has developed regular processes for training and recruitment into roles which are institutionally defined as belonging to that specialty. Members are aware of each other's work, although not necessarily deeply involved in communications with one another. They may share a paradigm and a set of judgements about what general work should be done in the field, although the details of those ideas might differ" (Mullins, 1972, p. 74).

In 1973, one year later the publication by Mullins, John Law published the article "The Development of Specialties in Science: The Case of X-Ray Protein Crystallography" (Law, 1973), in which the intellectual structure of a scientific specialties is discussed in detail. As Law claimed, his work is not only compatible with Mullins' theory, but it also supplements it by focusing much more on cultural factors than Mullin's (Law, 1973). Law distinguished among three different types of specialties: *technique*- or *methods-based* specialties; *theory-based* specialties; s*ubject-matter* specialties, and analysed them also using Durkheim's category of "mechanical solidarity" and "organic solidarity"[103]:

> In science, mechanical solidarity may be defined as the development and maintenance of relationships which depend on shared standards and exemplars, and hence on a relatively high degree of consensus about theory and method. In the thirties, X-ray crystallography approximated to a specialty whose members were held in relationship to one another on a basis of mechanical solidarity. Thus they used broadly the same theories and methods, passed innovations round the community, and fiercely objected to what they felt were deviant developments in crystallographic methods. The interaction of protein crystallographers with interested non-crystallographers developed, in the first instance, on a basis of organic solidarity. In science, organic solidarity may be defined as an aspect of the division of labour in which scientists come into relationship with one another because one performs services which the other cannot easily carry out for himself (Law, 1973, pp. 278,279).

Technique- or *methods-based* specialties constitute interacting groups of scientists whose solidarity is a mechanical one, as they strictly refer to shared scientific techniques and methods. Misuses of the method result in negative sanctioning for the deviants. Subject matters are chosen in relation to the methodological standards. *Theory-based* specialties have the same characteristics of the first, but the glue and the point of reference is a certain theory. *Subject-matter* specialties are specialties defined with reference to a specific subject matter or problem; members employ various and different techniques and theories. Here we find an organic solidarity. Law, referring to Mullins' specialty development scheme,

[103] Durkheim, 1933/1984.

claims that *method* and *theory-based* specialties are at the “cluster” or “specialty” stages, as they are based on received and specified achievements constituting a clear guide to action. Subject-matter specialties, instead, are defined on the basis of a shared problem, thus they correspond to Mullins' “network” stage of specialty development (Law, 1973).

As previously mentioned, the 1970s constituted a prolific period for the field of Sociology of Science as many works focusing on specialties structure and dynamics were produced. According to Small and Griffith “the science is a mosaic of specialties” (Small & Griffith, 1974, p. 17), which in turn are seen as the “building blocks” of Science (Small & Griffith, 1974, p. 17). Crawford (1971) demonstrated that specialties are organised as communication systems centred on few key scientists and on few research centres. Goffman (1966) and Jahn (1972) demonstrated that specialties are defined by few but very important journals. Crane in 1972 suggested the idea of mapping the structure of specialties (Crane, 1972).

The studies conducted in the 1970s demonstrated that intellectual, cognitive or problem contents can generate different kinds of structures. Thus the s*ocial division of labour* (Etzkowitz & Leydesdorff, 2000) and the *intellectual division of labour* (Galison, 1997) are two analytically distinct but strictly interwoven phenomena.[104] As affirmed by Cole and Zuckerman in 1975:

> Development and elaboration of the cognitive structure of new specialties appear to depend in part on correlative development of their social structures - on the routinization of an evaluation and reward system, procedures of communication, acquisition of resources and the socialization of new recruits. In short, the tandem development of both cognitive and social structures of specialties seems central to their institutionalization and establishment as legitimate areas of inquiry (Cole & Zuckerman, 1975, p. 143)

In this period the topics of *intellectual solidarity* and *social solidarity* started to be the focus of interest of many sociologists of science. Communication turned out to be a fundamental element in the social dimension of specialties which “has been equated with purposeful communication among a collection of individuals” (Chubin, 1976, p. 451); and (formal) communication was taken as a device to detect specialties (and more in general discipline) structure. In 1972 there was one of the first attempts to map networks of scientific journals (Narin, Carpenter, & Berlt, 1972) by means of bibliometric tools, even if the first application was that by Price, who in 1965 employed this technique to map scientific papers (Price, 1965). From this moment onwards, many citational studies have been conducted with the aim of modelling the social organisation of Science, and many efforts have been made to improve citational techniques (see, for example, Small, 1973; Small, 1977; Small & Crane, 1979; Small and Garfield, 1985; Small & Griffith, 1974).

[104] In 1988 Shinn detected a third analytical sphere: the division of labor that regulates the interactions between the cognitive and social sphere (Shinn, 1988).

One important contribution in the conceptualisation of the structure of disciplines is that by Whitley. In his paper "Umbrella and polytheistic scientific disciplines and their elites" (Whitley, 1976) he investigated the various "ways in which scientific production is organized, and the concomitant variations in knowledge structures" (Whitley, 1976, p. 471). He proposed to differentiate among different types of organisation in Science by focusing on two aspects of variation, that is to say the degree and the type of specialisation. Thus, he detected different cognitive and social structures distinguishing between highly fragmented and homogenous ones. Furthermore, specialisation generates from different concerns, and there are different levels of specialisation that can be more or less specific. With the purpose of furnishing a classification of scientific social organisations with respect to the level of specialisation, Whitley (1976) developed a framework distinguishing among "research areas", "specialties" and "disciplines":

> *Research areas* are collectivities based on some degree of commitment to a set of research practices and techniques. Membership is defined in terms of agreed procedures for specifying research problems and for selecting appropriate techniques to operate on them. […] *Specialties* are focussed on explanatory models and definitions of the phenomena under consideration. Membership of specialties implies commitment to particular types of accounts and preferred ways of formulating the underlying object of concern. […] The intellectual foundation of a *discipline* is, I suggest, more removed from the day-to-day specification and transformation of cognitive objects than are specialties and research areas. It applies a particular set of scientific values to some domain or field of reality, and so orders that reality. [...] Disciplinary ideals refer to particular ways of doing science and integrating the results (Whitley, 1976, pp. 472,473).

Discipline organisation[105] depends on the level of articulation of the "ordering principles" which refer to concerns and explanatory models. Thus, the more an ordering principle is well defined, the more related specialties will be clearly articulated, "although it is possible to have a fairly clear idea of science in a domain without a concomitantly clear view of reality concerns and explanatory approaches which are uniquely connected to it" (Whitley, 1976, p. 474); this situation reflects that of many disciplines in the Social Sciences. The frequency of deviant behaviours depends on the level of institutionalisation[106] of a discipline or of a specialty besides their degree of organisation.

Whitley develops the concepts of *umbrella* and *polytheistic* scientific disciplines describing two ideal typical disciplinary organisations. Umbrella disciplines "act as a loose holding organisations for diverse specialties" (Whitley, 1976, p. 476), while polytheistic ones are characterised by the organisation of the scientific work around divergent views of the discipline:

[105] Obviously, not all scientific activities are institutionalised in disciplines or specialties (Whitley, 1976).
[106] The level of institutionalisation is defined with reference to the level of definition of the ordering principles of a discipline (Whitley, 1976).

> While umbrella disciplines largely provide a very general view of science, and a set of basic ideas for research which are refined and concretized in specialties, the intellectual foundations of polytheistic disciplines are much more closely associated with current debates and issues. Scientific activities in the former will be focussed mainly on theoretically defined phenomena which are amenable to a limited set of technical approaches, whereas research topics in the latter will be more influenced by commonsense understandings and by general conceptions of science. Research will not, here, be so specialized or self-contained (Whitley, 1976, p. 476).

Specialties in polytheistic disciplines can be seen as associations of researchers sharing a common focus on some vaguely defined object such as "science" or "family" and having little else in common. This does not mean that there is absolutely no overlap of goals, techniques or approaches, or that there are not more stable associations. Thus, for example, with reference to Sociology, we can find Criminology, which has its own clear identity thanks to external support and legitimation, or the sect-like ethnomethodological group. Nevertheless "even in these cases the development of distinct reality concerns and explanatory approaches is unlikely" (Whitley, 1976, p. 479). It is important to underline that often in polytheistic disciplines divergent views of Science are not linked to specialty topics or explanatory models. Furthermore, scientific development consists of technical improvements and addition of new problems, which usually are suggested by funding agencies. Finally, discipline is the main point of reference as source of cognitive and social identities (Whitley, 1976).

In umbrella disciplines specialties play a more important role as they are relatively autonomous from the discipline. They are more developed, and specialty concerns are central in research focuses. Umbrella disciplines are characterised by a diffuse epistemological consensus; conflicts occur among the different disciplinary articulations. Here we find specialties which play an important role in disciplinary development, which are also important at the institutional level: they play a fundamental role in recruitment and training, also performing control functions. Disciplines of this kind are mainly concerned with the function of institutionalisation rather than that of organising current work, which is instead fulfilled by the discipline. Whitley identified two patterns of intellectual organisation in umbrella disciplines. The first is ideal and/or old, as it contemplates a full mechanical solidarity and a pure functional division of labour, which is not likely to be found in a highly differentiate discipline. The second one is more likely to occur and is characterised by specialties competition on the interpretation of the disciplinary ordering principle, as well as on the relative importance of each interpretation. Finally, in umbrella disciplines there is a hierarchy of problems and topics implying a hierarchy of specialties (ibid).

There is another important difference between the two types of disciplines detected by Whitley:

> Whereas in umbrella disciplines postgraduate training is in specialties and deals with components peculiar, by and large, to individual specialties, it is a feature of polytheistic disciplines that techniques are applicable in all areas - and, indeed, sometimes in other disciplines (e.g., multiple regression analysis in economics and sociology; factor analysis in

> sociology and psychology) - so that, once acquired, they allow the researcher to work on most topics without much further development. Within some general perspective (usually inculcated, at least in outline, during undergraduate education), the scientist is free to study anything. (Whitley, 1976, p. 493).

As illustrated by Whitley, differences in training are big and fundamental with respect to the way in which the respective structure is perpetuated.

Studies like the one introduced above, which describe disciplinary internal organisation, furnished important insights helpful in understanding why scientific development does not have the same pattern in all disciplines. Of the same importance are those studies focused on the level of closure and openness of specialities borders, as they inform of the following aspects: communication among specialties; level of innovation (or conservatism) of a specialty/discipline (Chubin, 1976).

The end of this prolific period can be dated in the late 1970s. The work by David Edge and Michael Mulkay (1976) can be considered as a symbol of this ending: the authors demonstrated that none of the models developed up to that moment was suitable for their case study, namely the emergence of radio astronomy, suggesting the impossibility to develop a unique model describing the development of all specialties.

8. DESCRIBING ITALIAN SOCIOLOGY

8.1. SOCIOLOGY: A POLYTHEISTIC-DISINTEGRATED DISCIPLINE

> *Science, carved up into a host of detailed studies that have no link with one another, no longer forms a solid whole.*
>
> Èmile Durkheim[107]

Sociology has always had identity problems. Debates about both its objects of analysis and purposes are almost as old as the discipline (Bryman, 1984), and if someone conceives of it as a *multiparadigmatic* discipline, (see, for example, Friedrichs, 1970), others talks about *pre-paradigmaticity* (see, for example, Chiesi, 2005; Goldthorpe, 2000). As Kuhn witnessed during his study in 1958 at the Center for Advanced Studies in the Behavioral Sciences, there is a significant difference between Social Sciences scientific communities and Natural Sciences ones:

> Particularly, I was struck by the number and extent of the overt disagreements between social scientists about the nature of legitimate scientific problems and methods. Both history and acquaintance made me doubt that practitioners of the Natural Sciences possess firmer or more permanent answers to such questions than their colleagues in social science. Yet, somehow, the practice of astronomy, physics, chemistry, or biology normally fails to evoke the controversies over fundamentals that today often seen endemic among say, psychologist or sociologists (Kuhn, 1962/1970, p. X).

With the attempt to find the cause of this difference, Kuhn detected the role of the so called *paradigms*: "[...] universally recognized scientific achievements that for a time provided model problems and solutions to a community of practitioners" (Kuhn, 1962/1970, p. X). According to Kuhn, the Social Sciences have not yet reached a scientific maturity level, which is defined by the acceptance of one paradigm by the scientific community (Kuhn, 1962/1970). Thus, those who define Sociology preparadigmatic confer on it a not well defined status; those who define it multiparadigmatic confer on it a special status. Of course, these two views entail two different conceptions of Sociology, but of interest here is the fact that the consequence is the same: in both cases we are dealing with a discipline without centre. The lack of a hub turns specialisation, which is a natural process in every discipline (Stehr, 1974; Whitley, 1976), into tribalisation or, as stressed by Whitley (1976), disciplinary disintegration. Whitley (1976) distinguishes between fragmented (specialised with a centre) and disintegrated disciplines (specialised with no centre), or better between "umbrella" and "polytheistic" discipline:

---

[107] Durkheim, E., 1933/1984, p. 294.

> While there are a number of ways in which specialties are related to disciplines and to each other in the sciences, there remain important distinctions between disciplines which act as loose 'holding' organizations for diverse specialties, and those where scientific work is organized around divergent views of the discipline. In the first type of disciplinary organization, research production is predominantly organized at the specialty and research area levels without direct reference to, or influence from, the discipline. In the second case, research is primarily distinguished in terms of metaphysical commitments which tend to focus on disciplinary identities. I shall term the former type an 'umbrella' discipline and the latter a 'polytheistic' discipline. While umbrella disciplines largely provide a very general view of science, and a set of basic ideas for research which are refined and concretized in specialties, the intellectual foundations of polytheistic disciplines are much more closely associated with current debates and issues. Scientific activities in the former will be focussed mainly on theoretically defined phenomena which are amenable to a limited set of technical approaches, whereas research topics in the latter will be more influenced by commonsense understandings and by general conceptions of science (Whitley, 1976, p. 476).

According to Whitley, Sociology is a polytheistic discipline; as Boudon affirmed: Sociology can be considered a “house with many mansions” (Boudon, 2002, p. 372).

As affirmed by Stinchcombe (1994), disintegrated disciplines with many different and incompatible standards for what is good work tend to be precarious in academic settings:

> [...] Sociology has a dim future first because it is unlikely to develop much consensus on who best represents the sociologists' sociologist to be hired in elite departments. […] Second, it has a dim future because it is unlikely to be able to argue with one voice about what is *elementary*. This will be true even if some leading sociologists will be recognized as clearly elite by many (never all) people from other disciplines, as for example Pierre Bourdieu or James S. Coleman would be (Stinchcombe, 1994, p. 80).

The problem of the control of propositions is, obviously, a fundamental one (Chiesi, 2007). Which are the criteria used for the selection of those sociological contributes elected to participate at the process of the accumulation of sociological knowledge?

> The problem is that the absence of any common criterion to verify scientific claims leads to a sort of anarchy in the research, which is merely the prelude to the dissolution of the idea of scientific community. In this situation, the risk is that the accumulation of knowledge becomes impossible, since every author wants to self-evaluate his work (Magatti, 2007, p. 4,5, own translation).

But how did Sociology get to this point? After the decline of the high-consensus period (1960s-1970s), characterised by the sharing of the structural-functionalism model, the sociological scientific community did not regain its unity (Magatti, 2007). From the "chaos of ideas", typical of such a period of change, did not arise a new Sociology but various sociologies:

> Rather than actual schools of thought - anchored to their own competing theoretical approaches - the debate (if so we can define it) over the last decades has been characterised by personalisms. The best of what sociology seems to be able to produce is the thought of big authors, more or less isolated, each of whom offers a very personal interpretation of both the discipline and its main theoretical perspectives, usually incompatible – and never communicating - with that of others. In this way, sociology seems more a philosophy - as a systematic reflection detached from the empirical data on social life - than a science. It is the strength of the authors and of their arguing as well as fascinating abilities that matters (Magatti, 2007, p. 1, own translation).

The lack of an international public debate has been fostering centrifugal tendencies; the increasing distance between theory and empirical work hampered and hampers debates on theoretical topics (Magatti, 2007). The last effort to develop a global theory of society was that by Talcott Parsons (1951). Nowadays, at the centre of debates there are still questions such as "what is, or what should be the object of study of Sociology?"; "Which and how many Sociologies are there?"[108] This is not an easy question to answer as relativism (objectivised in the interpretative paradigm), the enemy of any Science, is inside the discipline.

8.2. SOCIOLOGY IN ITALY

As previously discussed, Sociology is a discipline in crisis, and this crisis is a feature characterising the discipline in the whole. However, as Sociology of Science teaches us, sociological knowledge is not the product of a standardised activity but it is the result of the institutional and cultural features related to the locus of production. The obvious consequence is that, focusing on the national level, Sociology in Italy has different characteristics than, for example, Sociology in France or in Germany. As previously mentioned, in this second part of the thesis an empirical study focused on Italian Sociology is introduced, therefore it is fundamental to understand both under which circumstances the discipline developed and its most important characteristics. Accordingly, the following two Sections deal with the institutionalisation process of Italian Sociology and the so-called tripartition phenomenon.

[108] In this respect see the proposals by Goldthorpe (2000), Boudon (2002) and Burawoy (2007).

8.2.1. INSTITUTIONALISATION

The institutionalisation of Sociology in Italy occurred in the second half of the 20th century, specifically between the 1960s and the 1970s. That was the time of the First Republic (1946-1994). The Italian political scene was dominated by two big parties: the Christian Democracy (Partito Democratico - DC), which was always at the majority; the Italian Communist Party (Partito Comunista Italiano - PCI), the opposition party (Massironi, 1975).

After the Second World War Italy, or better northern Italy, undertook a modernisation and industrialisation process which led to cultural and social changes. The USA were the protagonists of this process: Americanisation arrived before with the reconstruction after the war, and then with the *Marshall Plan* (officially approved by the American government in 1948) and the diffusion of the *Scientific Management* (skilful managers were necessary to guide Italian industries, thus American experts arrived in Italy bringing and diffusing knowledge in business management). While the North was americanised with the purpose of industrialisation and modernisation, the south, especially between 1949 and 1955, attracted American sociologists (such as E. C. Banfield, G. Peck and F.G. Friedmann), who saw in rural southern Italy an interesting opportunity for researches in the field of community studies (ibid).

Thus, Italian society was affected by many and important changes and the ruling class was governing a new Country: industrialisation changed the social structure as new social classes were emerging (working class and middle class); cities were attracting more and more people; consumerism and the bourgeois society were affirming; secularisation was an increasing phenomena. Furthermore, the Church was losing power on the working class:

> Investigations made in the industrial countries of Europe [...] document the painful separation between the world of workers and the Church. [...] The reasons of such a de-christianization lie in the fact that the Church of the twentieth century is no longer able to direct the direction of the promoting working class movement, on the contrary it seems to want to keep the people under both its rigid guardianship and that of the even more despotic ruling class (Kothen, 1954, pp. 207,214).

Catholics were afraid that their loss in consensus could be intercepted by the communist culture. Some DC and high ecclesiastical hierarchy exponents saw Sociology as a useful tool to understand and control the above-mentioned tendencies. A group of catholics-innovators made the Catholic University of Milan a promoting centre for the development of the Social Sciences ensuring their scientific nature. As discussed in Section 8.2.2., the Catholic University of Milan will become the point of reference of one of the three Italian sociological components, called “the Catholic component”. In the 1950s Adriano Olivetti furnished a big support to Sociology: he was the first to involve sociological research in company policy and his social engagement created many research opportunities pivoting on his cultural and political project called *Movement of Community* (Movimento di Comunità) (Massironi, 1975).

Notwithstanding these successful applications, Sociology at the time was still almost absent from the academic scene and in general from the Italian scenario. At this point it is necessary to take a step back and dwell on the history of Sociology in Italy before the 1950s. Italian Sociology can be divided into two periods: the first between 1860 and 1910 (dominated by positivism); the second after its Liberation until 1980s-1990s (the so called “new” Sociology period). In the middle there was a phase which in turn can be divided into two stages: a crisis phase (caused by the crisis of the so called *ésprit positif*), and a instrumentalised period under Fascism. During the second period, Sociology had to fight against two strong cultures for affirmation: the idealistic one developed on the basis of Benedetto Croce's thought and the communist one (Barbano, 1998).

Croce, who was a firm anti-positivist and anti-scientist, was harshly critical with respect to Sociology effectiveness in detecting social laws (he defined Sociology a pseudo-science):

> [...] "sociology": name that has been defined as an inelegant hybrid of Latin and Greek, which would be slight and pardonable sin compared to the thing itself, that is a sick science, arbitrary and incoherent. [...] whose abstract and empirical laws [...] must succumb to the living reality of historical thought, as the derivative compared to the original and the artificial compared to the genuine (Croce, 1950, pp. 21,22, own translation).

According to Antonio Gramsci, exponent of the second opposition current:

> The so-called sociological laws assumed as cause, actually have no causative capacity, they are almost always tautologies, paralogisms. An event or a series of events are described with a mechanical process of generalisation, a similarity relationship is derived, this is called law and it is taken as cause. But what novelty has been found? The novelty consists in collective names given to a series of facts, but names are not novelties (an example is the "charismatic leader"). It is not clear that this leads to a baroque form of Platonic idealism because these abstract laws strangely resemble Plato's pure ideas which are the essence of the true terrestrial facts (Gramsci, 1948, p. 128, own translation).

Besides the perplexities expressed by Gramsci in the passage above, the left wing was suspicious toward Sociology for another reason. During the two World Wars, Sociology suffered a strong setback in Europe but not in the USA, where instead the discipline expanded. When sociological studies were resumed in Europe, it was impossible not to look at the works of American sociologists such as Paul F. Lazarsfeld, Robert K. Merton, Talcott Parsons and Alfred Schütz (who moved from Vienna to the Unites States in 1939). Sociology was seen as a danger by the left because it was too much impregnated of American cultural imperialism, thus it was considered something to avoid: there was in stake the achievement of communism. As a matter of fact, the discipline was banned in the USSR (Burgalassi, 1996).

Notwithstanding such oppositions, Sociology was able to affirm itself in the academic world. Until the 1950s Sociology was taught only in the Faculties of Statistics, Economics and Demography (it was also included as complementary course in Political Science Faculties). The opening of universities to the Social Sciences was fostered by the will of reaching the same cultural level of the other western democracies. Among those working in the direction of affirming and legitimating Sociology we find Nicola Abbagnano[109] and Franco Ferrarotti. They founded the first sociological journal born after the fascism, *Quaderni di Sociologia* (1951), with the aim of promoting a critical debate on Sociology in Italy (Massironi, 1975).

An important event for the institutionalisation of Sociology was the conference "Philosophy and Sociology" held in 1954 in Bologna under the patronage of the University of Bologna and the support of the journal *Il Mulino*. It can be considered the event which officially opened the doors of the academic world to Sociology. The aims of the meeting[110] were mainly two: formulate a definition of the sociological discipline; define a professional model. The first competition for the first chair in Sociology[111] was held in 1961 by the Sapienza University of Rome; it was won by Franco Ferrarotti (in 1967 he will found the journal *La critica sociologica*, with the same purposes of *Quaderni di Sociologia*) (ibid). Later, in the 1980s, Rome will become the reference point of the third component of Italian Sociology, namely the so called Roman, Third or Staterian component (see Section 8.2.2.). The process of institutionalisation of Sociology in Italy occurred between 1957 and 1968. In this process two steps were fundamental, namely the foundation of the "Italian Association in Social Sciences" (Associazione Italiana di Scienze Sociali) in 1957, and the first Congress of Social Sciences held in Milan in 1958, where participants agreed only on one thing: Sociology is a Science. At the Congress participated, among the others, a group of *dissident marxists*, which saw in Sociology an effective tool for studying society. Sociologists supporting Communism divided into two factions: one promoting an "official" communist culture (represented by Guiducci's *organic sociology)*; a "heretical" one (represented by Pizzorno's *sociology-science*). According to the first view, research problems should have been set by the Party and the analyses should have been conducted with reference to the Marxists logic (i.e. class perspective). The second position sustained that sociological issues had to be set looking at the society and its changes (the phenomenon of organisations, new production relationships, automation, new political participation forms, etc). This debate concentrated in the North, and in particular in Bologna, Milan and Turin (Chiaretti, 1975). Here we find the roots of another of the three Italian sociological components, that is to say the Mi-To (Milan-Turin) component.

It was thanks to the left orientated "policy makers" (i.e. employees in public administration, political parties, industries, and trade unions which were politically oriented towards the left-democracy and the

---

[109] Together with the colleagues of his Centre Methodological Studies (Centro Studi Metodologico) of Turin.

[110]The sociologists who took part in the event were: Felice Battaglia, Nicola Abbagnano, Pietro Chiodi, Enzo Paci, Renato Treves, Nicola Matteucci, Luigi Pedrazzi, Antonio Santucci, Cesare Luporini and Renato Zangheri (Massironi, 1975).

[111] The first wave of sociologists hired in universities dates back to 1964 (Massironi, 1975).

Socialist party) if in the late 1950s and early 1960s Sociology developed and moved from a marginal position to a more central one. Sociologists and political forces had the same reformist aspirations. But Sociology did not developed inside universities: at that time sociological researches were mainly promoted by the "National Center for Prevention and Social Defence" (Centro Nazionale di Prevenzione e Difesa Sociale) (CNPDS) and the group "Il Mulino", which aimed to give scientific legitimation to the profession of the *social technician*. Notwithstanding these two organisations had the same final scope, they differed with regard to their nature and audience. The CNPDS, founded in 1947, played a role of political intermediary among the political forces, and his main scope was to inform the Parliament with reference to prevention and social defence matters;[112] the group "Il Mulino", founded in 1954,[113] was close to social-democratic and anti-communist political forces. It aimed to spread among people those values characterising western democracies. Furthermore, he worked in line with a program aiming to renovate the national culture. Il Mulino soon became a neuralgic cultural point of reference promoting conferences, congresses, studies and researches on many and different topics; it also became a big enterprise encompassing the homonymous journal and publisher as well as the "Carlo Cattaneo Institute"[114]. Its importance is also witnessed by the interest that the USA showed towards the group, which was seen as a means through which spread the American culture in Italy. For this reason the US decided to support the institute in its purpose of developing Italian Sociology (Chiaretti, 1975).

In the 1960s Sociology still had marginal positions in the academic world. It was subject to the other corporations and in general to institutional interests. This situation led to a decrease in the disciplinary development and to a generalised malcontent, but the university reform, planned by the first central-left government, gave to sociologists good expectations. Sociology was introduced in many faculties (political sciences and humanistic fields) and in the late 1950s the Catholic University of Milan decided to give space and resources to Sociology; in 1962 Catholics fostered the birth of the first Sociology Faculty in Trent (the Higher Institute of Social Sciences, officially recognised in 1966). However, this was an isolated case as the institution of Social Sciences faculties was opposed inside the universities by jurists, who did not want to lose power both inside the academic world and in the formation of future public officials. It was thanks to the Cospos (Board for political and Social Sciences)[115] if Sociology could proceed with the institutionalisation process. Between 1966 and 1968 many projects were realised, such as the creation of the Higher Institute of Sociology (Istituto Superiore di Sociologia) in Milan or the laboratory of Political Sciences directed by Norberto Bobbio at the University of Turin (ibid).

While the debate on the university reform was still going on (in 1967 was held a congress entitled "Social Sciences, university reform and Italian society"), students started to occupy universities. The first University to be occupied was in Trent (which had already been occupied twice in 1966) and the second

[112] The CNPDS was founded in 1947 by four (antifascist) magistrates: Banfi A; Beria di Argentine A; Dell'Oro A; Savini A (Chiaretti, 1975).
[113] It was founded by the same group which in 1951 created the journal "Il Mulino" (Chiaretti, 1975).
[114] The institute was founded in 1956 with the scope of promoting studies and researches on central social topics, such as education, depressed areas or mutualistic institutes (Chiaretti, 1975).
[115] It was founded by Ford Foundation and Olivetti Foundation (Chiaretti, 1975).

one was the Catholic University; the Faculty of Political Science at the University of Milan was the place where student protests lasted longer, until 1973. Not only Universities but the entire Science was in question: scientific knowledge was seen by students as a by-product of power, inside and outside universities. Sociology had always had to fight for its legitimacy, but this time was different: sociologists were under attack by the left, and thus they were considered in line with the system. With reference to Sociology, institutional responses to the protests differed, depending on the extent to which students participated to demonstrations. Thus, were Sociology turned out to be a problem, as in Trent and Milan (both at Catholic University and the Faculty of Political Science at the University of Milan) there were repressive responses: from Trent were removed two principals who demonstrated to be close to students (F. Alberoni and G. Baglioni) and, for the same reason, at the faulty of Political Science in Milan there was an opposition to A. Pagani; at Catholic University Sociology was practically deleted. Elsewhere, where the discipline did not cause big problems, thus were students demonstrated to be not so critical, Sociology was fostered. In these cases, the lack of criticisms to "the system" and to the bourgeois society, which instead characterised the demonstrations of those years, made possible the production and re-production of a "certain" sociological knowledge (ibid).

One of the major conquests of students' protests was the free access to universities. The growth of Sociology in universities was positive quantitatively speaking but not qualitatively, as one of the consequences of the "mass university" phenomenon was the deskilling (ibid).

Sociology in Italy suffered for several reasons. In addition to the above mentioned ones, we must add the "familistic" organisation of the discipline, which is a typical phenomenon of Italian academic system. As Laura Balbo demonstrated in a study published in 1975, sociologists were organised in families and the access to academic positions depended on close relationships with a "father"[116] of the Italian Sociology of the time:

> The recognition by a "father" [...] was the only existing legitimation to enter the profession and often the only link with the culture and practice of sociology (Balbo, 1975, p. 217, own translation).

At that time, the geographic distribution of the main sociological "families" was the following:

- Rome: this pole was formed around Ferrarotti and, among the others, there were Izzo, Lelli, F. Martinelli, Rizzo, Statera[117], Santolini, Viola;
- Milan (Political Science): from 1969 we find Balbo, Martinelli, Martinotti and Tomeo;

[116] According to her study, in the first 1970s the most powerful sociologists were Alberoni, Ardigò, Castellano, Ferrarotti, Acquaviva, Baglioni, Barbano, Gallino, Leonardi, Pizzorno, Spreafico and Treves (Balbo, 1975).

[117] The third or roman component is also called Staterian.

- Trent: here sociologists from Catholic University arrived such as Cella, Ferraresi, Romagnoli and Rusconi, joining Livolsi and Manoukian;
- Bologna (Political Science): from the Catholic University arrived Guidicini, Rescigno, Di Nallo, Bellasi, Morra, Piazzi e Sandri, joining Ardigò, Cesareo and Stroppa;
- Catania: the most central figures were Leonardi and Spreafico around whom gather Catanzaro, Cazzola, Gennaro, Talamo and Reyneri;
- Turin (Political Science): around the central figure of Barbano we find Farneti, Bonazzi, Marletti, Follis and Taglioli;
- Turin Magisterial: around the central figure of Gallino there were Bravo, Maggi, Pischierri and Baldissera;
- Florence: Cavalli is the central figure and around him we find Bettini, Giovannini, Marsiglia, Tinacci and Mannelli (Balbo, 1975).

As previously mentioned, the late 1960s, and somewhere also the first 1970s, were hard years for universities which had to face students' demonstrations and occupations. In that period the above mentioned mechanism regulating university access based on the relationship "father-son" was not working well and some "fathers" felt the need to find a way to regulate, according formal and public criteria, the professional affiliation on two occasions in 1971: the Turin Congress on the crisis of sociological method; the proposal to create the Italian Sociological Association[118] (Associazione Italiana di Sociologia - AIS). In both occasions Ardigò and Ferrarotti missed, thus the catholic and roman component missed (even if Ignazio Ughi, director of the Luigi Sturzo Institute took part in the event). According to Balbo, they missed because they covered important roles and were among the most important figures of Italian Sociology: they did not have reasons to change the system which made and was maintaining them so influential (Balbo, 1975).

The Congress turned out to be a way for the powerful to maintain their power: only the "sons" of organisers were admitted with the condition of accepting what was established by the latter (rules and purposes of the meeting). However, even if the Congress should have been on Sociology crisis and on the role of the sociologist (in relation with political and economic structures), it turned out in an ordered investiture of the second generation of sociologists, and debates were about general problems of the discipline (paradigms, how both Sociology and sociologists should be) rather than the crisis of Italian Sociology. Leaving aside the obvious fact that the way in which the meeting was organised (participants and participation rules) compared with its purposes constitutes a counter-sense, in talking about the autonomy of the discipline from the power they did not take into account both the fact that sociological knowledge is produced in universities and the working conditions of those sociologists not part of the higher levels (i.e. those fighting against the system and considered the cause of the crisis).[119] In this

[118] There was the need to replace the inactive Italian Association of Social Sciences founded in 1957.
[119] Even if Pizzorno mentioned this situation in his report, it occupied a marginal position in his analysis

regard it is meaningful that the book gathering the congress proceedings instead of being entitled *Sociology crisis* (as the title of the congress) was entitled *Sociological research and sociologist's role* (Ricerca sociologica e ruolo del sociologo) (Cavalli, della Porta, Donati, & Rositi, 2010). As Pietro Rossi, the organiser of the event, affirmed, the meeting was characterised by the worrying for the future of Sociology, which was mainly felt by the first generation.[120] As Pizzorno and Gallino claimed:

> We talk of Sociology crisis but we are sociologists, and we expect rewards and gratification if we continue in this job [...] the program of those talking of the crisis is almost always to continue to stay inside (Pizzorno, 1972, p. 352, own translation).

> The crisis of the speech is mistaken for a crisis of the subject under discussion [...] the crisis that has been talked about is, in hindsight, a crisis of the conception of science (Gallino, 1972, p. 314).

On the basis of what has been argued so far, it emerges that Italian Sociology since the beginnings of its second period (called by Barbano "new sociology") structured around poles which were not grounded on theoretical criteria. As Donati explains, there was an "original sin" at the re-birth of Sociology after Fascism, as it was dominated by the major political forces of that time (marxists, liberal-socialists and catholics). This distortion constituted the basis for further ones and gave to Italian Sociology a structure that is inevitably still characterising it (Cavalli et al., 2010). One example of the atavistic fragmentation of national Sociology is the failure of the attempt by the Italian Association of Sociology (AIS) of creating between 1995 and 1998 an autonomous national agency, the Cirs (Inter-university centre for social research), gathering the most important departments of Sociology. The agency lasted only from 1999 to 2002 as, according to Donati, some sectors of the discipline brought the project to failure. The same thing occurred some years later when another President of the AIS made a similar attempt with the same results (Cavalli et al., 2010).

In conclusion, it must be mentioned that this familistic model characterising the Italian Sociology was not passively accepted by everyone. The so called "group of young sociologists" or "Bologna group" constituted the first opposing attempt. It was founded in 1970 by left politically orientated young sociologists also joining extra-parliamentary groups. It was not a formal group but it made two significant

---

(Cavalli et al., 2010).

[120] It must be considered that Sociology of Science arrived in Italy only in the second half of the 1970s, with the exception of the early contributes by Statera "La sociologia della scienza di Robert K. Merton" in *La critica sociologica*, and the one by Barbano "La sociologia della scienza in Italia: situazione e problemi", in *Sociologia* (Barbano, 2003). In 1974 the publishing house of the Pci published the book *Attualità del materialismo dialettico* (Geymonat, 1974). It must be underlined that the Party decided to publish the book because on the one hand it wanted to hide its hostility towards the diffusion of a critical approach to Science and on the other hand it wanted to control (or better to limit) the diffusion of the book (as happened) (Barbano, 2003)

public actions: besides refusing to join the AIS, they subscribed a document about "urgent measures" that should have been taken in order to stop the decline of the Italian academic Sociology (Balbo, 1975).

8.2.2. THE TRIPARTITION OF ITALIAN SOCIOLOGY

One of the things wrong with sociology, in our country, is that we need a better country.
Harvey Molotch[121]

Even if Sociology crisis has been at the centre of debates for years, both at global and national level, recently in Italy it has become a neuralgic topic. It seems that the starting point was the article by Guido Martinotti "The disappearance of Sociology from Italian official science: suicide or homicide?"[122] (Martinotti, 2010), written after the proposal to dismiss Sociology as an autonomous discipline from the official list of those recognised by the State (Freschi & Santoro, 2010a). Martinotti's article (based on three points: sociologists attitude to forget their forebears; the institutional weakness of Sociology; the discipline theoretical and methodological uncertainty) gave birth to a debate fostered by various sources: the Treccani website, the ISA blog "Universities in crisis", the Italian on line journal "Socio*logica*", the journal "Il Mulino" and also by the newspaper "Repubblica". From the discussions emerged that one of the problems affecting national Sociology is its organisation in three components: "Mi-To", "Catholics", and the so called "Third component". Among the three camps the second and the third are formalised into specific professional associations called "Sociologia per la Persona" and AIS3, respectively; the first one is a rather loose group of scholars who share some views about selection and cooptation mechanisms within the discipline. This picture is completed by a number of social scientists who do not feel part of any component.

As Carola Freschi and Marco Santoro explained in "Italy: A troubled and Divided Academic Field under Neoliberal Pressure" (Freschi & Santoro, 2010a) on the ISA blog, the "Mi-To" (Milan-Turin) component is lay and left politically orientated, the "Catholic" component refers to the Roman Catholic tradition and the "Third" one is appealing to those sociologists of the Middle-Southern Italy which are not part of the Catholic wing and do not want to be part or are not accepted by the Mi-To component. Furthermore, each component seems to be specialised in one or more topics: Mi-To is focused on social inequalities, politics, economy, and social movements; the Catholics are interested in cultural processes and sociology of communication; the Third component shares with the Catholic one the interest on sociology of communication but it is very strong in methodology.

As previously mentioned, this threefold partition seems to have neither logical nor ontological reasons; its origins go back to the institutionalisation of Sociology in Italy:

---

[121] Molotch, 1994, p. 222.
[122] Original title: "La scomparsa della sociologia dalla scienza ufficiale italiana: suicidio od omicidio?"

> Since its inception [...] the field structured itself around two great poles: a lay pole, on the left, and a Christian (or better a Roman Catholic one), on the right. A great tension was therefore organizing and driving the emerging Italian sociological field in its early years (approximately 1960-1964), a tension which would be soon institutionalized in a true cleavage between two so-called "components," or better "camps": the Catholic camp (with the full leadership of Ardigò and at least two strongholds: the Catholic University in Milan and the Faculty of Political Science in Bologna), and a less organized and more polycentric camp of lay (usually left-oriented) sociologists. […] in the early eighties *there is the formation* of a new group, the so-called Third Wing, which was and still is appealing above all to sociologists in the Centre and South which were not part of the Catholic wing and not willing to stay in MiTo – indeed, were not really accepted by the latter (Freschi & Santoro, 2010a, italics mine).

This organisation has consequences on every kind of decision, such as public competitions or elective charges. The strength of this tripartition is attested by the implicit rule of the "components rotation" governing the presidential elections of the Italian Sociological Association (Freschi & Santoro, 2010a; Mora, 2007), the first professional association founded in 1982. As suggested by Santoro (2007), when referring to the case of Italian Sociology, we should add a category to Burawoy's classification[123] (Burawoy, 2005), that is to say the *baronial-bureaucratic* one:

> [...] what counts are not the scientific standards or the practical effectiveness, but the hierarchy of academic power, well-protected by a well-established institutional system (made not only of legally established norms but also of acquired practices and customs which are part of the academic habitus, transmitted from generation to generation) (Santoro, 2007, p. 11, own translation).

This situation generates worries about the production and re-production of the Italian sociological knowledge: if such mechanisms are active in the discipline, how is the control on sociological knowledge production performed? This worry seems to be shared by various scholars in Italy such as, Freschi and Santoro (2010b), Magatti (2007) and Santoro (2007). As Santoro claims:

> A hierarchy is the farthest thing from the peer institution that there can be. But without peers, there is no peer control, and this means no profession. The perverse consequences of such a weak, if no absent, professional statute are evident (Santoro, 2007, p. 11, own translation).

These considerations are very important as they question the effectiveness of the peer review institution, which is usually considered a guarantee of the value of scientific works published in journals. Bearing in

[123] He distinguishes among public, professional, policy and critical Sociology (Burawoy, 2005).

mind the familistic character and the threefold structuration of the Italian sociological field, besides the lack of shared criteria of scientific research quality and of a national as well as of an official group of peers over the parts (or better over the components), it is clear that the peer review control cannot be considered as a warranty of the quality of the sociological scientific production.[124] Under these conditions, the peer review turns to be another means to re-produce the underlying structure. The peers and a public arena of discussion are the main means through which control in scientific communities is exercised; but if these means fail in their purpose, who does control sociological propositions? How is the accumulation of sociological knowledge possible? As Santoro explains, be part of a component means much more than attending certain conferences and working in certain institutions; it also means reading certain journals and publishing in certain journals, having as a result the impairment of the scientific communication among the components (Santoro, 2011).

[124] "Most journals are not based on any clear cultural project or even the mutual choice of recognised and acknowledged peers, but on the belonging of the editors (and even the authors) to one of the (three) main national academic groups (or lobbies) which control the field" (Freschi & Santoro, 2010b, p. 5).

9. STRUCTURING ITALIAN SOCIOLOGY

This Section consists of the application of Bibliometrics to the case of Italian Sociology. As we have seen in the previous Chapter, Sociology in Italy seems to have a peculiar configuration, called *tripartition*, which generates worries with reference to the quality of the scientific work produced in that community. The previous Section ends with a sentence related to the status of the scientific communication in that fragmented context. According to the theoretical proposal illustrated in Chapter 6, in talking about scientific communication we refer to that dynamic and cognitive process *structuring* scientific communities. As discussed in Chapter 4 (Section 4.4.), the scope of Relational Bibliometrics is that of furnishing maps informing about the cognitive and intellectual structure of scientific communities. Generally speaking, Scientometrics consists of the quantitative study of scientific communication based on bibliometric data (Section 1.1.). Therefore, with the purpose of detecting the intellectual and cognitive structure of Italian Sociology, and thus obtaining information with reference to the scientific communication characterising it, an all author co-citation analysis is performed and two co-citation maps are developed. The theoretical presumptions at the basis of this work can be found in what has been conceptualised in Chapter 6. Results are supposed to confirm the effectiveness of both Relational Bibliometrics in furnishing informative maps about scientific communities and the normative approach in describing citing behaviour with respect to high citation counts, which in turn will give some information about the goodness of what has been hypothesised in the above-mentioned Chapter. The work contained in this Chapter is organised as follows. In the first Section the hypotheses and purposes related to the empirical work are introduced. The second Section refers to the data-collection process, and the third one consists of the analysis of the bibliometric data collected. In Sections 9.4. and 9.5. two co-citation maps are introduced and analysed, respectively. In Section 9.6. conclusions are given.

9.1. PURPOSES AND HYPOTHESES

The empirical work here introduced has multiple objectives. The first objective is to test the effectiveness of Bibliometrics in mapping scientific literature. The second one is to test the effectiveness of the normative approach in describing citer's behaviour as well as the effectiveness of what has been hypothesised in the theoretical proposal of Chapter 6. More specifically, only after the corroboration of the hypotheses introduced below we could conclude that in correspondence of the high frequencies citers behave in a normative way. Finally, we aim to furnish a bibliometric description of the scientific community of Italian Sociology with respect to its intellectual and cognitive structure. Obviously, the validity of the description relies on the effectiveness of both the techniques employed and the theoretical assumptions in which the latter are rooted. Therefore, before proceeding with the description of the intellectual structure of Italian sociology, two hypotheses must be tested. The first hypothesis relates to the effectiveness of the (all) author co-citation analysis (ACA) technique in detecting the intellectual and cognitive structure of scientific communities. More particularly, considering the following presumptions:

1) the organisation of scientific disciplines in specialties and research areas as a result of the differentiation of scientific work (see, for example, Chubin, 1976; Stehr, 1974; Whitley, 1976);

2) authors active in a research field will tend to cite those authors active in the same field more than those who are not (see, for example, Braam et al., 1991a,b; Hagstrom, 1970);

3) the main premise of the ACA technique is that the co-citation of couple of authors is based on the co-occurence of these latter in a certain number of bibliographic lists (McCain, 1990);

4) data are taken from authoritative journals selected by Italian sociologists;[125] specialties, research areas and fields of inquiry are defined by few but very important journals (Goffman, 1966; Jahn, 1972)

it seems reasonable to hypothesise that the result of the clustering and mapping process will be a certain number of clusters showing internal coherence with reference to the topics of interests of the authors constituting them. Furthermore, if we take in consideration the question related to the multiparadigmatic nature of Sociology (see Section 8.1.), we can expand the hypothesis above. As White underlined in 1990, usually in Science maps a qualitative-quantitative polarity appears (White, 1990, p. 103). Thus, the expectations about the findings of the empirical work here introduced are the following: we expect that the clusters show homogeneity with reference to the topic, but at the same time we expect that clusters division informs on the multiparadigmatic nature of Sociology, which is materialised in the polarity qualitative-quantitative.

The second hypothesis is strictly linked to the peculiar structure of Italian Sociology introduced in Chapter 8. For reasons that must be traced back to the period of the institutionalisation of the discipline (see Section 8.2.1.), Italian Sociology is characterised by the presence of 3 components (Mi-To, Catholics and Third component); even if the (political) premises which caused this fragmentation are now over, the tripartition seems to be still a reality. Of course, this has various implications, such as the tendency to both read and publish on certain journals having as a result the impairment of the scientific communication (Santoro, 2011). Furthermore, for the picture to be completed, we must mention the *intellectual parochialism* which seems to characterise the field: as underlined by Santoro (2011), each component is specialised in one or more topics: the component called Mi-To seems to be focused on social inequalities, politics, economy and social movements; the Catholics are interested in cultural processes and sociology of communication; the Third component shares with the Catholic one the interest in sociology of communication but it is very strong in methodology. According to this description, Italian Sociology shows a peculiar organisation making it an interesting case study for testing the effectiveness of the normative approach in describing citers' behaviour in correspondence of high citation frequencies because of the presence of a potentially strong constructivist reason for citing, namely the affiliation to a

[125] See Section 9.2.

component. Therefore, the main question to be asked is the following: will the co-citation maps reflect this tripartition, thus corroborating constructivist citation theories, or will they confirm the thesis here suggested about the strength of the normative approach in relation to high citation counts? Unfortunately, due to the lack of official lists informing on the affiliation of each author, it is not possible to work directly with the tripartition issue. However, the author co-citation technique enables us to observe the usage of scientific literature made by Italian sociologists. On the basis of what has been sustained so far, with reference to the ability of the co-citation method to generate topic-based clusters and according to the theory developed in this thesis, sustaining the usefulness of the normative approach in describing citers' behaviour with reference to the so called standard or paradigmatic works (Small, 2004) (i.e. works frequently cited induce scholars, according to the humility norm, to cite them when referring to the knowledge contained in them), we expect the intellectual parochialism will be detected. This will be interpreted in the following way: given the tripartition of the field, and given that each component is strong in certain topics, the detection of the intellectual parochialism demonstrates the acknowledgment of the scientific successes of each component by the entire community. I must underline, in fact, that being strong in a subject is not the same as being the only one working on it. Conversely, a crisis in the scientific communication is supposed to be reflected by a *double-level fragmentation* indicating that one or more topics are discussed by more than one cluster. This suggests that the citers did not cite in a normative way, and that the clusters differentiate not only with respect to the topic; in this case, it will be possible to hypothesise the presence of an additional factor originating clusters, namely the affiliation to one component.

9.2. COLLECTING DATA

Usually bibliometric studies are performed employing data taken from databases like Scopus or the Web of Science. Unfortunately, these databases do not contain data about Italian Sociology[126], thus I decided to collect them by hand creating my own database. In this way it was possible to avoid some of the most common biases and shortcomings usually occurring in big and automated databases (see Section 3.2.2.). One of these problems relates to the so called *clerical errors* made by authors when compiling bibliographies. Obviously, it is impossible to control for all errors made by authors in writing bibliographic lists, but it is possible to reduce them; in order to do this it is necessary a knowledge of the field under analysis. Furthermore, with the purpose of reducing problems in authors' name and journals' title, I created two .txt files, one about names and the other about journals. In the first case I was able to avoid problems related to multiple names/surnames, as I chose one variant and homogenised according to

[126] No Italian sociological journal shows the criteria established by Thomson Reuters (http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/). In Scopus it is possible to find some data; however, because of the lack of continuity in the storage of bibliographic information and because of the sparse coverage of Italian sociological journals, I preferred to collect all data by hand.

it[127]. The second .txt file allowed me to avoid mistakes in journals' title as I took them directly from my list.

Data collection covered a four-year period, from 2007 to 2010. The database has the same structure of those by the Web of Science; it consists of a .txt file including the following tags: PT (Publication Type); AU (Author); TI (Title); SO (source); LA (Language); DT (Document Type); PU (Publisher); JN (Journal); CD (Cited Documents); NR (Item Number). The database contains information from 1,564 papers; cited authors are 73,699[128].

Due to the lack of an official ranking of Italian academic sociological journals,[129] in order to select those periodicals to include in my analysis (from which to obtain my author co-citation data), I conducted a survey among the members of the Italian Association of Sociology (AIS)[130] in January 2011. More particularly, with the aim of obtaining a list of authoritative journals, I addressed by e-mail the coordinators and the secretaries (in total 25 scholars) of the AIS sections[131] asking them to select among the journals indicated by me those which according to them were both scientific and academic. Respondents had also the possibility to add journals which were not included in the list I set up. Due to the moderate response rate (13 out of 25 interviewed replied), I decided to choose a low cutting point, thus all journals that obtained at least 2 scores (i.e. selected by at least two respondents) were included in the analysis.[132]

The procedure introduced above resulted in the selection of the following 24 Italian journals listed in alphabetical order: Etnografia e ricerca qualitativa; Futuribili; Ikon; Il Mulino; La critica sociologica; Mondi migranti; Polis; Quaderni di sociologia; Rassegna italiana di sociologia; Rassegna italiana di valutazione; Rivista italiana di comunicazione pubblica; Rivista trimestrale di scienza dell'amministrazione; Salute e società; Sociologia della comunicazione; Sociologia del diritto; Sociologia del lavoro; Sociologia e politiche sociali; Sociologia e ricerca sociale; Sociologia urbana e rurale; Sociologica; Stato e mercato; Studi culturali; Studi organizzativi; Studi di sociologia (In Appendix A scores for each journal are available). However, I could extract data only from 23 journals, as articles in "Il Mulino" do not have any bibliography. Furthermore, I could not collect data from some of the above mentioned journals with reference to some numbers, as they have not yet been published during the data collection period; specifically:

---

127 For example: TOTA A (Tota Anna Lisa); SCHADEE H (Schadee HMA); for the author TORRE AT I kept both name initials for differentiate him from the author TORRE AR.
128 This number contains duplicates (for a definition of "duplicates" see Footnote 134).
129 A ranking about Italian sociological journals was developed in 2012 by the Italian Association of Sociology. The ranking is available here: http://cdn.ais-sociologia.it/uploads/2012/02/ranking-completo-indici.pdf
130 Associazione Italiana di Sociologia.
131 http://www.ais-sociologia.it/
132 Incomplete and evasive replies were considered as no responses.

- Etnografia e ricerca qualitativa: the year 2007 is lacking as the journal has been founded in 2008;
- Futuribili: the years 2009 and 2010 were not published;
- Ikon: the year 2010 has not yet been published;
- Rivista italiana di comunicazione pubblica: the year 2009 has not yet been published;
- Sociologia della comunicazione: the year 2007 and 2008 were not published; the year 2010 has not yet been published.

Only articles, essays, proceeding papers and research notes were included. This means that comments, debates, editorials and reviews were excluded.

9.3. ANALYSING DATA

After collecting data (the total amount of documents in the database is 1,564) I processed them through the software Bibexcel[133]. In order to conduct an *all author co-citation analysis* (Persson, 2001), the following steps were performed (I must anticipate that I performed them twice, as after a first analysis, in line with the purpose of analysing Italian sociology, I removed from the dataset the Italian scholars who are not sociologists): 1) extraction of cited works; 2) extraction of cited authors; 3) decompression of rows with multiple units (i.e. authors in case of co-authorship) in order to place each unit in one row (the total amount of authors first found is 73,699, after the second routine is 72,452); 4) removal of duplicates[134] (unique authors after the first routine = 32,092; unique authors after the second routine = 32,033); 5) authors citation-frequencies computation; 6) co-occurrences frequencies computation; 7) creation of a .net file to upload in the software VOSviewer for clustering and mapping. Only those authors cited at least 10 times (citation frequency ≥ 10) were included in the 6th step. "The goal of ACA is to produce empirical maps of prominent authors in an academic discipline. [...] An important purpose of ACA is an overall examination of the intellectual structure of an academic discipline. Therefore, it is critical to establish a diversified list of authors" (Eom, 2009, p. 145). However, a threshold must be set and there are two ways for doing this: by using a subjective approach (the researcher selects the authors to be included in the analysis); by using an objective approach (i.e. referring to authors' citations frequencies). As explained by Eom (2009), it is important to avoid personal judgements if the purpose is to examine the intellectual structure of a discipline but, of course, the subjectivity of the researcher comes out in the selection of the cutting point. In fact, "there are no quantitative tools that can be blindly applied in deciding the number of authors" (Eom, 2009, p. 148). The selection of the cutting point applied to the

133 http://www8.umu.se/inforsk/Bibexcel/; see also Persson et al., 2009.

134 Duplicates (i.e. items – here authors – cited more than once in the same article) were removed as they usually create technical problems in the analysis of co-occurrences (Persson et al., 2009).

work introduced in the following pages is based on the analysis of the distribution of citations frequencies:

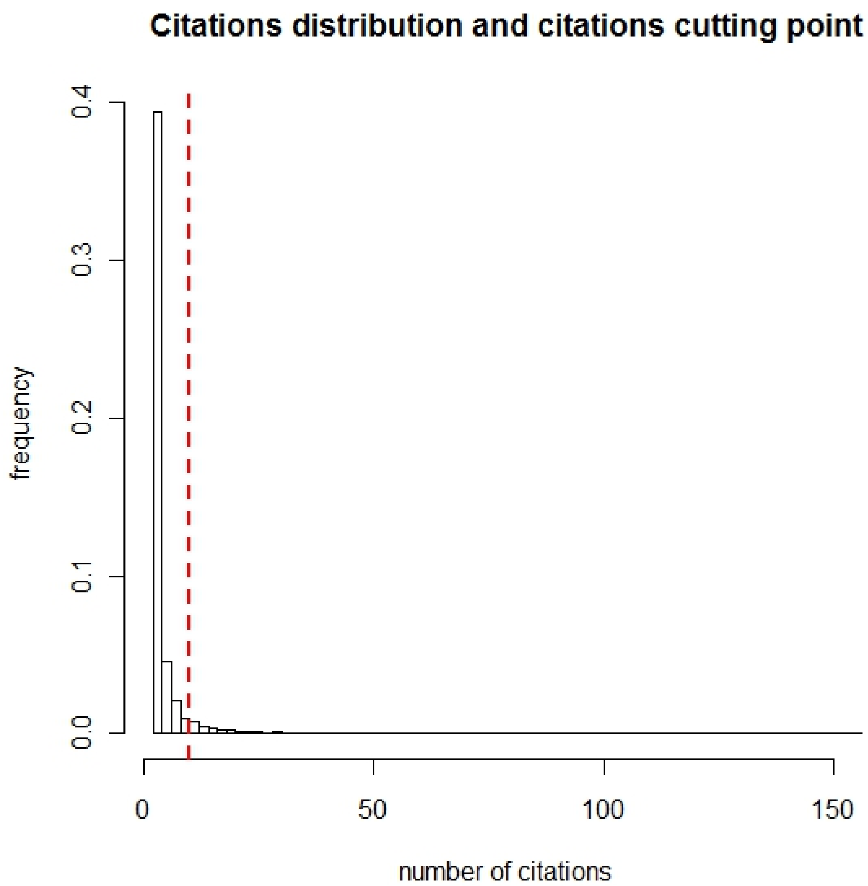


Figure 19: Authors' citations distribution. The red dotted line equals the 95th percentile.

In the selection of the citations threshold I proceeded as follows: bearing in mind the aim of detecting the most influential authors, I decided to divide the distribution in quintiles, more specifically in percentiles, and to refer to the 95th percentile for establishing the cutting point. Actually, the 95th percentile equals 12 but I chose the value 10 for being more generous in the inclusion of the authors. As mentioned above, the selection of the number of authors to include is very important as:

> In the aggregate, this author set defines the scholarly landscape being mapped. If the authors are not chosen to capture the full range of variability in subject specializations, methodologies, political orientations, etc., these aspects of structure cannot be determined (McCain, 1990, p. 433).

In the calculation above are not included those authors cited only once as their frequency is definitely high (24,356 out of a total of 32,024, namely 76%) constituting, in this case, a disturbing element in the calculus for the determination of a reasonable cutting point. With reference to the co-citations processed in step 7, I did not select any cutting point; the reason stands in the will of obtaining a map that was as informative as possible. Establishing a threshold is not a "structured process; rather, it is an unstructured process requiring the investigator's personal judgments. An exact quantitative basis for deciding the threshold cocitation has not been developed" (Eom, 2009, p. 149). Once again, it is the researcher taking a decision on the basis of his/her purposes.

As previously discussed (see Section 4.4.), the technique called *all author co-citation analysis* is a variant of the author co-citation (ACA). While the classic ACA technique takes in consideration only the first author of a scientific publication (McCain, 1990), the all author method takes in consideration all authors co-publishing a paper or a book included in a reference list. The reason why the all author form should be preferred to the classic one is clearly explained by its developer, namely Olle Persson:

> [...] all author citation counts should be preferred when visualizing the structure of research fields. *First author citation* studies distort the picture […] Supporters of *(first) author co-citation* studies may object to such a criticism by saying that the aim is not to rank authors but rather to identify research themes, and as representing such themes first cited authors might suffice. Well, that may very well be the case, but then one should also test if the structure of author co-citation maps is dependent on the counting method used (Persson, 2001, pp. 339, 343).

As in the classic version, also in this variant, a co-occurrence between two authors is established if, for a certain amount of times, they were cited together by thirds.

Once obtained the file with the co-occurrences information I processed my data in the software VOSviewer, thanks to which, by means of a unified approach for clustering and mapping, I obtained two maps. As discussed in Section 5.3., VOSviewer employs a variation of the Louvain Method for clustering, and a variation of the multidimensional scaling for mapping. The unified approach for mapping and clustering is based on the minimisation of Equation (12). Furthermore, co-citations are normalised employing the so called *association strength* measure given by Equation (11).[135] Thanks to this method I will produce groupings (clusters) of similar authors (with respect to co-citations) and visualise them in a two-dimensional map were similar authors are close to each other and dissimilar authors are placed farther apart.

In the first map, which we can call "international" or, more simply, Map1, the nodes of the network stand for both national and international[136] authors, while the links between each couple indicate the strength of their relation. The second map, which we can call "national" or Map2, refers only to national

[135] For a detailed description of the method see Section 5.3.

[136] The expression "international authors" refers to authors not working in Italian Universities. Thus, foreigner scholars working in Italian universities are considered part of the national Sociology, on the contrary Italian sociologists working abroad are considered international authors.

authors. The two maps differ with reference to the authors included. Thus, if in the international map are included international authors belonging to different disciplines as well as Italian sociologists, in the national map, we find only Italian sociologists. This means that from both maps Italian authors who are not sociologists were excluded.[137] In order to do this, I isolated Italian sociologists (i.e. sociologists whose affiliation is an Italian institution), then I searched for their characteristics with reference to affiliation and topics of interests/research areas. Information about the affiliation were collected as follows: first I searched for each author's information in the website of MIUR[138] (Ministry of Education, University and Research); in case of retired professors, scholars working with a collaboration contract, and adjunct professors, whom are not included in the MIUR database, I searched on the internet through Universities' websites. For all other scholars the rule followed for the identification of the affiliation is the following: the last affiliation is the one here considered (collaborators are not considered affiliated). I referred to internet also for searching for information about the area of interests of each author. During this process I was able to detect and remove those scholars who are not sociologists (for example, economists or demographers). Therefore, not only Italian scholars belonging to disciplinary fields different that Sociology were dropped, but, in map2, I did the same also with reference to Italian sociologists working abroad. The maps obtained at the end of this process are introduced in the following Sections.

9.4. GENERATING MAPS

In this Section the two maps developed with the techniques described in the previous one are introduced. The first one to be shown is the *international* map, or Map1; the *national* map, or Map2, is introduced at a later time. The reason behind the choice of generating two maps lies in the fact that the presence of foreign authors could hide a possible communication crisis among Italian scholars. More particularly, the majority of sociologists considered as founding fathers (of a method or of a school of thought) or research leaders are not Italian, consequently, a massive presence in Italian articles of foreign scholars acting like, we can say, *citations collectors* is quite obvious. Thus, for example, it can happen that one author belongs to a cluster only because it is linked to one of the above mentioned *citations collectors*, and not because it is linked to any Italian scholar. Therefore, it will be interesting to compare the two maps and observe if and which changes occur when removing international scholars. Is the number of clusters the same? What about their internal composition?

[137] International authors were not subjected to this type of analysis. On the contrary, due to the interdisciplinary nature of Sociology, the presence of authors belonging to different disciplines is considered important for the creation of clusters related to research topics.
[138] http://www.istruzione.it/web/hub/home

9.4.1. INTERNATIONAL MAP: THE STRUCTURE

The map here introduced is referred to as *international* because it includes, besides Italian sociologists, also international or foreign scholars. Both clustering and mapping were obtained by processing data in the software VOSviewer. As showed in Figure 20, the 511 authors in the map are split in 8 clusters. In order to furnish a concise description of the map (of which 3 different views are given; see Fig. 20, 21 and 22), for each cluster the following features are given: graph descriptive statistics (network centralisation; network density; density by groups); number of authors; topic(s) characterising the cluster (in brackets the TF-IDF values); position of the cluster with respect to the whole map (interpreted referring to the polarity qualitative-quantitative methods)[139]; international authors with high citation frequencies (i.e. those with citation counts ≥ 20;[140] the full list of the authors included in map1 is in Appendix B)[141]; Italian authors and citation frequencies. Both international and Italian scholars are listed according to the citation counts. With reference to the identification of the "topics characterising the clusters", I proceeded as follows: first I standardised and reduced to keywords the information relating to authors' research interests, then I calculated the TF-IDF (term frequency–inverse document frequency) index (Jones, 1972; Salton & McGill, 1983) related to each keyword in each cluster, in order to select those labels relating to topics describing the cluster. This index reflects how important a word is to a document in a collection of documents (or, in this case, clusters). It is used in information retrieval and text mining as a weighting factor. The reason why the TF-IDF is widely used is that even if its value increases proportionally to the number of times a word appears in the document, it is offset by the frequency of the word in the corpus. In this way it is possible to control for the fact that some words are generally more common than others (Salton & McGill, 1983). The TF-IDF is given by:

$$tf * idf(t,d,D) = tf(t,d) \times idf(t,D) \qquad (15)$$

where *tf(t,d)* is the frequency of the term *t* in document (here cluster) *d*, *idf(t,D)* is the logarithm of the quotient of the total number of documents *D* (clusters) containing the term *t* divided by the number of documents *d* (clusters) containing the term *t*:

[139] To establish if a cluster is characterised by quantitative or qualitative methods, I referred to the presence of authors who are representative of a certain approach (for example, Goffman, Garfinkel, Merton or Lazarsfeld) as well as the methods employed by Italian authors (in this case I referred to publications, information given by the authors on their web-pages and/or academic curricula, and my knowledge about authors' career).

[140] With the exception of cluster 8 where, due to the fact that only 2 of the authors belonging to it reached the threshold of 20 citations, international scholars with lower frequencies count were listed.

[141] Obviously, data about international authors are given only in the analysis of the international map.

$$idf_i = \log \frac{|D|}{|\{d : t_i \in d\}|} \quad (16)$$

[142]

The terms with the highest scores according to Formula 16 were extracted. The 5 highest scores for each cluster are listed in Appendix C.

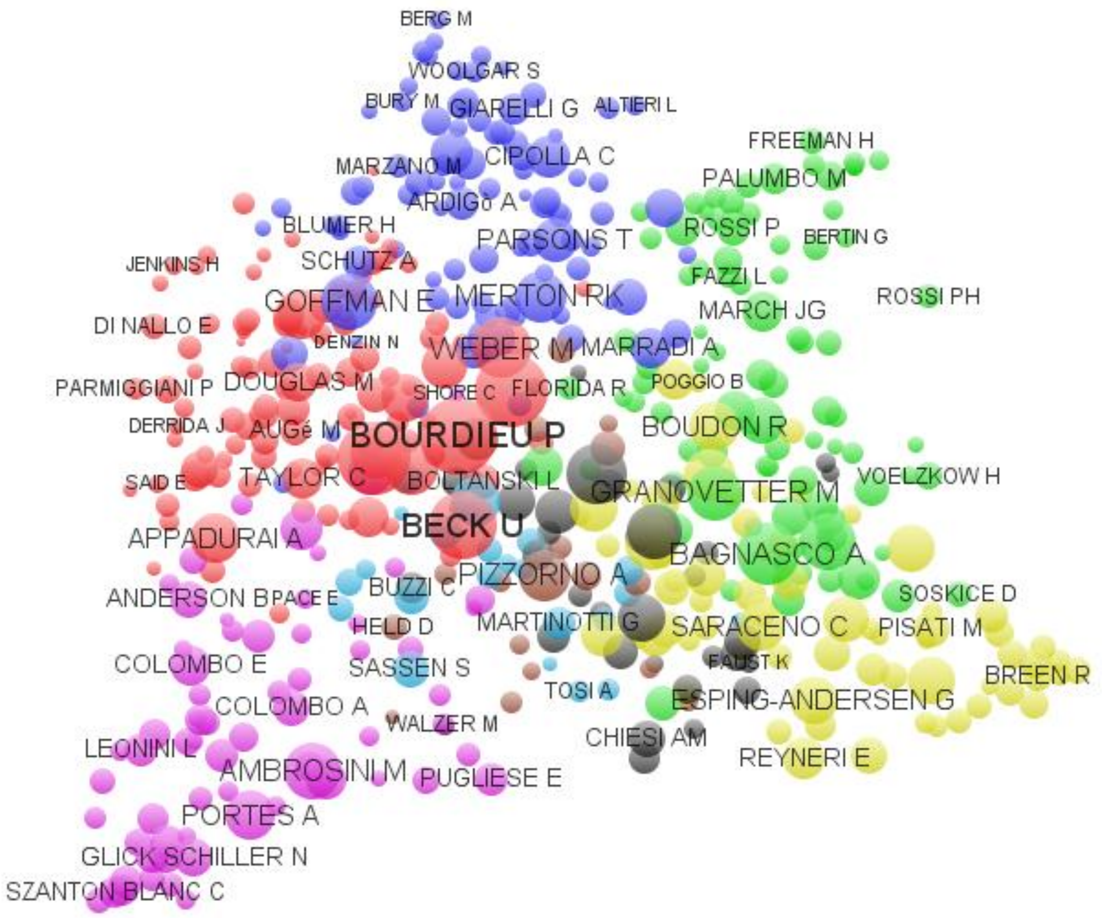


Figure 20: Label view of the international co-citation map of Italian Sociology. The nodes of the map are authors (citation threshold = 10). The map was developed processing data obtained by the so called "all author co-citation" technique (Persson, 2001). Data were collected from scientific papers published in 23 Italian journals during the period of time 2007-2010. Journals were selected through a survey among the members of the Italian Association of Sociology (section coordinators and secretaries). The map, generated by the unified approach implemented in the software VOSviewer, is made of 8 clusters identified by different colours. The size of the nodes is related to the sum of the strength of the links incident with them; the bigger the node, the greater the strength of its links.

[142] If the term is not in the corpus, this will lead to a division-by-zero. It is therefore common to adjust the formula adding (+1) to the denominator (Salton & McGill, 1983).

Figure 21: Cluster density view of the international co-citation map of Italian Sociology. This view, compared to the previous one, furnishes a clearer representation of the clusters.

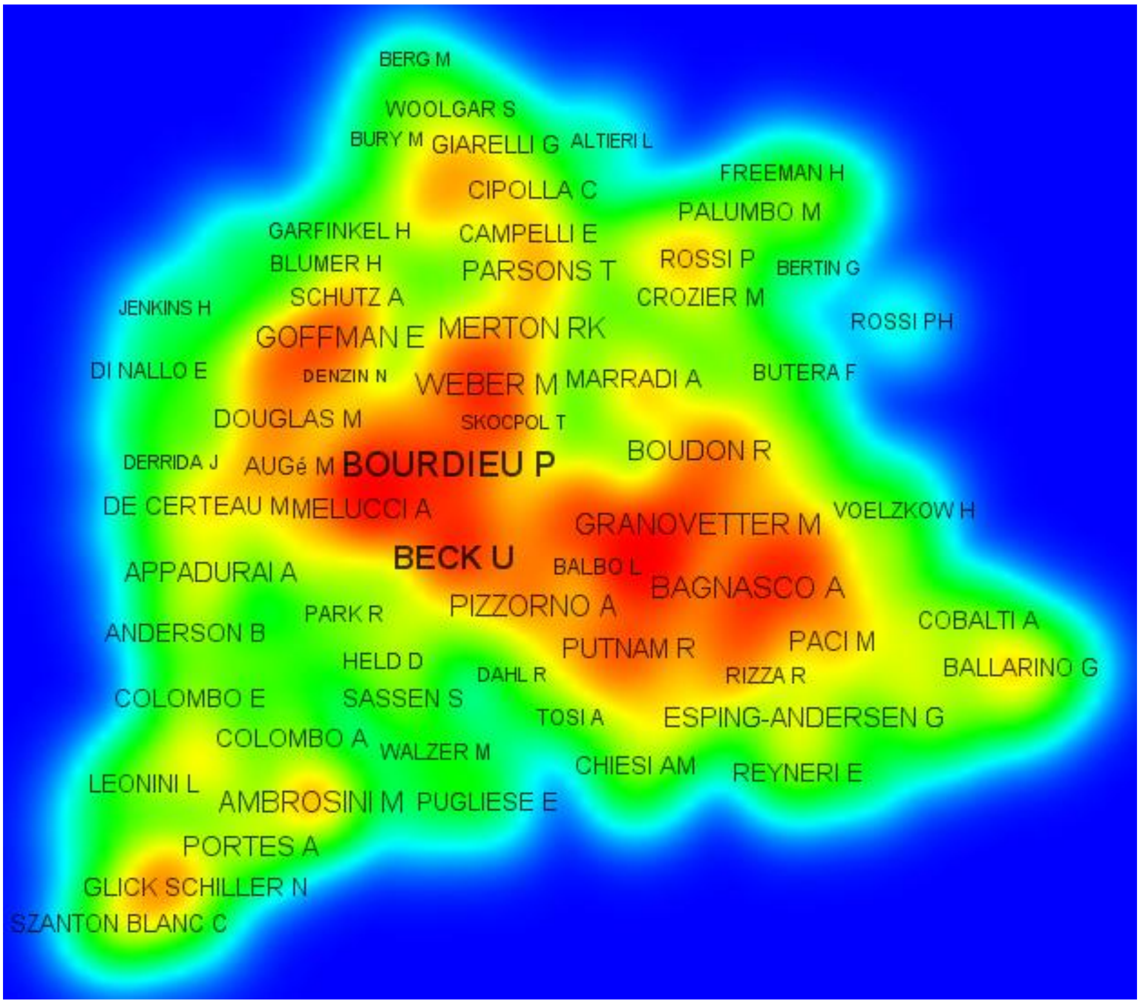


Figure 22: Density view of the international co-citation map of Italian Sociology. Moving from red to green the density decreases. Density depends on the number of nodes in the neighbourhood of a node as well as on the weights of such nodes. Thanks to this view it is possible to get an immediate idea of the most important zones, in terms of density, in the map.

Network centralisation: 6.483%

Network density: 0.8228

Density by clusters:

Cluster 1 (red)

Nr. of nodes: 113

Density: 1.087

Topics: sociology of cultural processes (14.29), sociology of communication (14.26), sociology of consumption (13.12), sociology of religion (4.14), gender studies (2.76);

Position: this cluster is located in the right part of the map; it is characterised by the presence of most of the classics of the sociological thought; it is characterised by the prevalence of qualitative methods applied to research;

International authors: Bourdieu P (155), Giddens A (136), Bauman Z (133), Beck U (122), Weber M (106), Foucault M (99), Castells M (78), Simmel G (66), Habermas J (65), Luhmann N (65), Appadurai A (56), Durkheim E (55), De Certeau M (43), Taylor C (43), Douglas M (39), Becker H (36), Hall S (31), Boltanski L (31), Lévi-Strauss C (31), Lash S (31), Augé M (30), Elias N (30), Mauss M (29), Morin E (29), Jedlowski P (28), Marx K (27), Ricoeur P (26), Arendt H (26), Berger PL (23), Baudrillard J (22), Said E (21), Cohen S (21), Latouche S (20), Butler J (20);
Italian authors: Cesareo V (38), Ferrarotti F (33), Codeluppi V (29), Sassatelli R (29), Paltrinieri R (20), Di Nallo E (19), Bovone L (17), Garelli F (16), Vaccarini I (15), Alberoni F (14), Mora E (14), Santoro M (14), Leccardi C (13), Boccia Artieri G (12), Morcellini M (12), Parmiggiani P (12), Ferrari V (12), Abruzzese A (10).

Cluster 2 (green)
Nr. of nodes: 91
Density: 1.300
Topics: economic sociology (12.22), local governance (9.51), public politics (9.51), sociology of organisation (7.33), sociology of work (6);
Position: it occupies the extreme right position in the upper and central part of the map; it expands towards the bottom-left side overlapping mainly with the cluster 4. Its position can be linked to the prevalence of quantitative methods employed in doing research;
International authors: Granovetter M (67), Powell WW (49), Di Maggio PJ (43), Crouch C (42), March JG (42), Hirschman AO (39), Simon H (38), Polanyi K (33), Le Galès P (27), Piore M (26), Sabel C (26), Streeck W (26), Olsen JP (23), Crozier M (21);
Italian authors: Bagnasco A (83), Trigilia C (77), Mutti A (33), Palumbo M (31), Ramella F (31), Pichierri A (29), La Rosa M (28), Barbera F (25), Cella GP (23), Rossi P (23), Bonazzi G (22), De Leonardis O (22), Butera F (20), Bifulco L (19), Donolo C (19), Fazzi L (19), Ceri P (17), Perulli P (15), Battistelli F (12), Bertin G (12), Cerase FP (12), Guala C[143] (12), Negrelli S (12), Burroni L (10), Minardi E (10).

Cluster 3 (blue)
Nr. of nodes: 91
Density: 1.058
Topics: methodology (25.36), sociology of health (22.49), sociology of cultural processes (9.9), sociology of knowledge (7.92), social policies (6.89);
Position: it is located in the top-centre in a central position, which is attributable to the topics characterising it, and specifically to that related to “methodology”. The cluster is oriented towards the

[143] Guala Alessandro, named “Chito”; in the map indicated as Guala C.

right in relation to those scholars mainly related to qualitative methods, while it is oriented towards the centre in correspondence to those authors linked to quantitative research;

International authors: Goffman E (92), Merton RK (62), Parsons T (48), Latour B (45), Geertz C (36), Lazarsfeld P (30), Strauss A (30), Schutz A (29), Kuhn T (28), Bateson G (24), Garfinkel H (23), Luckmann T (22), Illich I (22), Knorr-Cetina K (21);

Italian authors: Cipolla C (59), Marradi A (46), Gherardi S (37), Ardigò A (36), Giarelli G (31), Maturo A (28), Crespi F (25), Campelli E (24), Cardano M (23), Bucchi M (20), Neresini F (19), Cannavò L (17), Capecchi V (17), Bichi R (16), Marzano M (16), Altieri L (15), Ricolfi L (15), Gobo G (14), Bruni A (13), Manghi S (13), Fele G (12), Ingrosso M (12), Statera G (12), Bruschi A (11), Cipriani R (11), Moruzzi M (11), Rositi F (11), Agnoli MS (10), Fasanella A (10), Giglioli PP (10), Vicarelli G (10).

Cluster 4 (yellow)

Nr. of nodes: 70

Density: 2.339

Topics: social mobility (31.70), sociology of work (15), welfare (14.67), economic sociology (13.45), social policies (11.03), inequalities (8.27);

Position: it occupies the extreme bottom-right side of the map and it expands towards the top-centre overlapping with the cluster 3. Its position can be linked to the prevalence of quantitative methods employed in doing research;

International authors:

Sen A (71), Esping-Andersen (63), Boudon R (52), Goldthorpe J (46), Castel R (33), Collins R (31), Dahrendorf R (25), Becker G (22), Breen R (22), Blau P (21), Blossfeld HP (21), Lewis J (20), Erikson R (20);

Italian authors: Gallino L (66), Saraceno C (63), Schizzerotto A (55), Barbagli M (50), Magatti M (46), Paci M (46), Reyneri E (44), Corbetta P (40), De Lillo A (40), Pisati M (34), Ranci C (36), Regini M (28), Mingione E (27), Barbieri P (26), Borghi V (25), Ballarino G (20), Cobalti A (20), Accornero A (19), Gasperoni G (19), Schadee H (19), Negri N (19), Bianco ML (18), Naldini M (17), Scherer S (17), Rizza R (14), Ruspini E (15), Lodigiani R (13), Pavolini E (13), Fullin G (12), Gambardella D (12), Semenza R (12), Olagnero M (11), Piccone Stella S (11), Balbo L (11), Facchini C (10), Moscati R (10), Poggio B (10).

Cluster 5 (violet)

Nr. of nodes: 65

Density: 2.198

Topics: migration (28.95), sociology of cultural processes (11), deviance (7.92), multiculturalism (6.89), ethnic studies (6.34);

Position: it occupies the bottom-left side of the map. Its position is attributable to the prevalent use of qualitative techniques;

International authors: Portes A (54), Anderson B (38), Glick Schiller N (30), Massey D (28), Vertovec S (27), Hannerz U (26), Levitt P (23), Landolt P (21), Sayad A (21), Basch L (20), Guarnizo L (20), Marshall TH (20);

Italian authors: Ambrosini M (79), Melucci A (48), Dal Lago A (47), Zanfrini L (40), Colombo E (39), Colombo A (35), Sciortino G (30), Caponio T (25), Leonini L (25), Pugliese E (25), Queirolo Palmas L (24), Semi G (20), Pace E (17), Rebughini P (17), Lagomarsino F (15), Macioti MI (15), Palidda S (15), Zucchetti E (15), Besozzi E (14), Melossi D (14), Zincone G (14), Tognetti Bordogna M (13), Frisina A (12), Bosisio R (11), Decimo F (11), Allievi S (10), Cotesta V (10), Scidà G (10).

Cluster 6 (black)

Nr. of nodes: 35

Density: 2.800

Topics: social networks (9.37), social capital (8.27), civicness (6.97), third sector (6.34), social policies (6.34);

Position: it is localised in the bottom-central zone of the map expanding towards right; it overlaps almost entirely with other the clusters, namely clusters 1, 4, 6, 8. Its central position can be due to both the lack of prevalence of qualitative or quantitative methods and the linkages with the theorists placed in the left side of the map;

International authors: Putnam R (68), Coleman J (63), Archer M (33), Elster J (29), Burt R (24), Fukuyama F (23), Wellman B (21);

Italian authors: Donati P (95), Cavalli A (45), Piselli F (34), Sciolla L (31), Colozzi I (30), Chiesi AM (29), Prandini R (26), Pellizzoni L (25), Tronca L (20), Rossi G (16), Buzzi C (21), Di Nicola P (21), Osti G (14), Caselli M (12), Boccacin L (11).

Cluster 7 (light blue)

Nr. of nodes: 26

Density: 1.880

Topics: urban sociology (8.56), sociology of the environment (5.62), ethnic studies (3.17), social integration (3.17), sociology of international relations (3.17);

Position: it is located in the lower part of the map, in a central position making it overlapping with the clusters 1, 4, 5, 6, 8. Its central position can be due to the lack of prevalence of qualitative or quantitative methods, and/or to the linkage with theorists placed in the left side of the map;

International authors: Sennett R (54), Sassen S (38), Harvey D (33), Urry J (30), Wacquant L (23).

Italian authors: Martinotti G (34), Mela A (25), Amendola G (24), Guidicini P (19), Nuvolati G (18), Zajczyk F (14), Tosi A (12), Davico L (11), Gasparini A (11), Borlini B (10), Gasparini G (10).

Cluster 8 (brown)

Nr. of nodes: 20

Density: 1.579

Topics: deviance (6.34), political sociology (4.63), social capital (3.17), social movements (3.17), public policies (2.44);

Position: see position of cluster 7

International authors: Touraine A (37), Elster J (29), Norris P (18), Offe C (16), Held D (15), Tilly C (14), Cohen J (13), Dahl R (13), Eisenstadt SN (13), Inglehart R (13), Tarrow S (13), Rosanvallon P (11), Lipset SM (10), Mair P (10), Skocpol T (10);

Italian authors: Pizzorno A (53), Della Porta D (23), Catanzaro R (15), Diani M (13).

9.4.2. ITALIAN MAP: THE STRUCTURE

The following map is based on links among Italian authors, therefore it is referred to as *Italian*. It is based only on Italian sociologists (i.e. working in Italian universities, or those who worked there until the end of their career), thus Italian authors who are not sociologists were removed. Both clustering and mapping are obtained by processing data in the software VOSviewer. In Figure 23 it is possible to observe the 8 clusters making up the map (of which 3 different views are given; see Fig. 23, 24 and 25). For each cluster the following features are given: graph descriptive statistics (network centralisation; network density; density by groups); number of authors; topic(s) characterising the cluster;[144] position of the cluster with respect to the whole map (interpreted according to the duality qualitative-quantitative related to methods used in doing research)[145]; Italian authors and citation frequencies. The total amount of authors in the map is 170.

[144] See page 157.

[145] See Footnote 139.

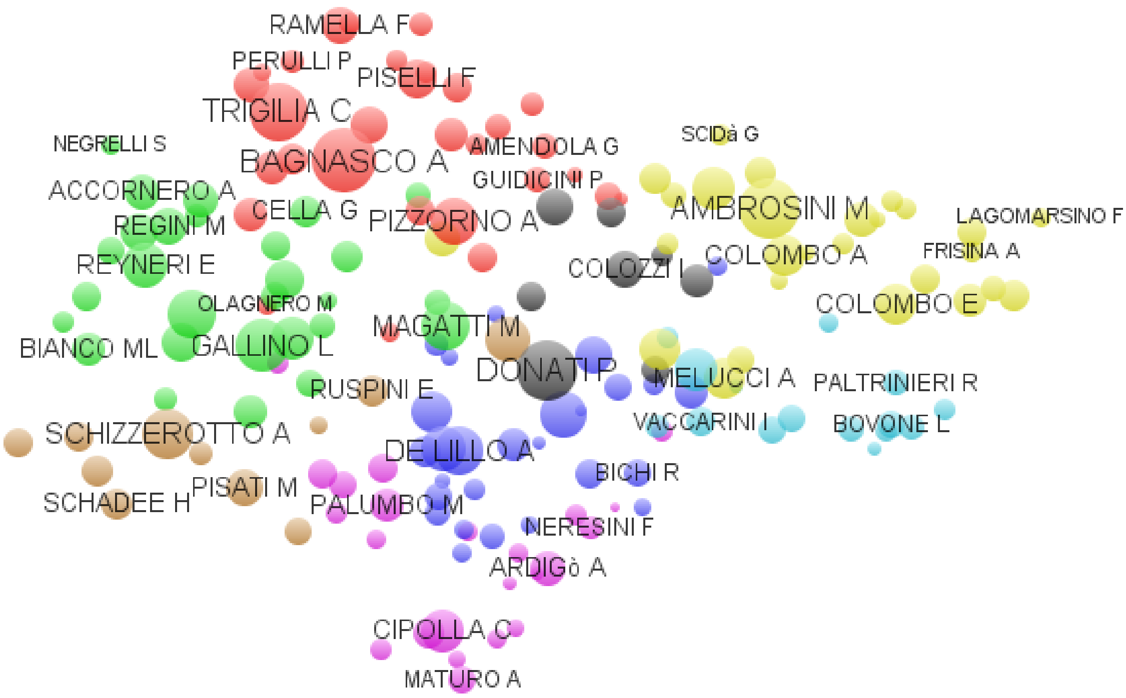


Figure 23: Label view of the national co-citation map of Italian Sociology. The nodes of the map are Italian authors (citation threshold = 10); the links represents co-citations. The map was developed processing data obtained by the so called "all author co-citation" technique (Persson, 2001). Data were collected from scientific papers published in 23 Italian journals during the period of time 2007-2010. Journals were selected through a survey among the members of the Italian Association of Sociology (section coordinators and secretaries). The map generated by the unified approach implemented in the software VOSviewer, is made of 8 clusters identified by different colours. The size of the nodes is related to the sum of the strength of the links incident with them. The bigger the node, the greater the strength of its links.

Figure 24: Cluster view of the national co-citation map of Italian Sociology. This view, compared to the previous one, furnishes a clearer representation of the clusters.

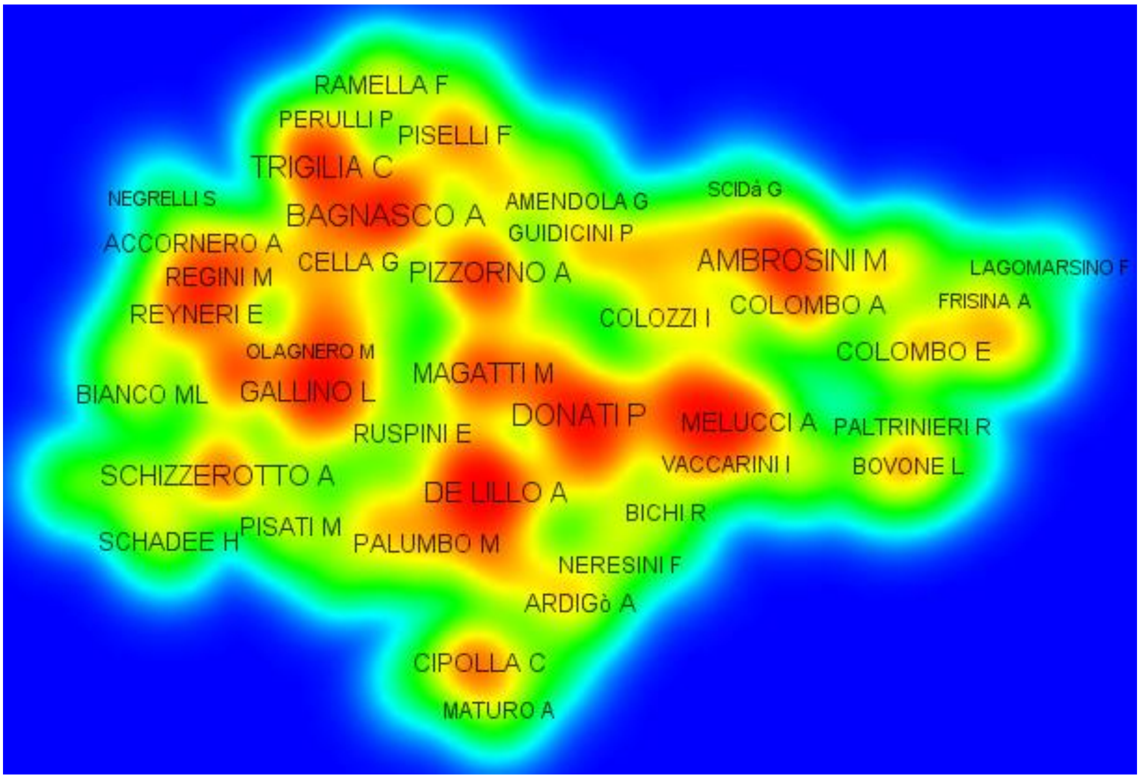


Figure 25: Density view of the national co-citation map of Italian Sociology. Moving from red to green the density decreases. Density depends on the number of nodes in the neighbourhood of a node as well as on the weights of such nodes. Thanks to this view is possible to get an immediate idea of the most important, in terms of density, zones in a map.

Network centralisation: 6,169%

Network density: 0.5559

Density by clusters:

Cluster 1 (red)

Nr of nodes: 31

Density: 2.501

Topic(s): urban sociology (19.02), local governance (13.93), sociology of the environment (12.68), local development (11.25), public policies (9.29);

Position: this cluster is located at the top of the map; it is oriented towards the left but it also occupies some space in the middle. This is due to the fact that even if the quantitative methods prevail, there is the presence of authors focused on qualitative research;

Italian authors: Bagnasco A (83), Trigilia C (77), Pizzorno A (53), Martinotti G (34), Piselli F (34), Mutti A (33), Ramella F (31), Pichierri A (29), Barbera F (25), Mela A (25), Pellizzoni L (25), Amendola G (24), Della Porta D (23), De Leonardis O (22), Donolo C (19), Guidicini P (19), Negri N (19), Nuvolati G (18), Catanzaro R (15), Perulli P (15), Osti G (14), Zajczyk F (14), Cerase FP (12), Guala C (12), Tosi

A (12), Davico L (11), Gasparini A (11), Olagnero M (11), Borlini B (10), Burroni L (10), Gasparini G (10).

Cluster 2 (green)
Nr of nodes: 29
Density: 2.643
Topic(s): economic sociology (18.34), sociology of work (18.00), welfare (9.9), social policies (8.56), labour policies (6.97);
Position: it is located in the left-middle side; there is a prevalence of quantitative methods;
Italian authors: Gallino L (66), Saraceno C (63), Magatti M (46), Paci M (46), Reyneri E (44), Gherardi S (37), Ranci C (36), La Rosa M (28), Regini M (28), Mingione E (27), Barbieri P (26), Borghi V (25), Cella GP (23), Bonazzi G (22), Butera F (20), Accornero A (19), Bifulco L (19), Bianco ML (18), Naldini M (17), Scherer S (17), Rizza R (14), Lodigiani R (13), Pavolini E (13), Fullin G (12), Negrelli S (12), Semenza R (12), Piccone Stella S (11), Balbo L (11), Minardi E (10), Poggio B (10).

Cluster 3 (blue)
Nr of nodes: 28
Density: 1.641
Topic(s): methodology (28.53), sociology of cultural processes (9.78), sociology of knowledge (7.92), migration (6.34), social theory (6.11);
Position: this cluster is located at the bottom of the map, in a middle-left position. We can say that it is located in a position that it is exactly in the middle between those clusters representing the quantitative pole (to the left) and those representing the qualitative one (to the right);
Italian authors: Marradi A (46), Cavalli A (45), Corbetta P (40), De Lillo A (40), Ferrarotti F (33), Sciolla L (31), Crespi F (25), Campelli E (24), Cardano M (23), Rossi P (23), Buzzi C (21), Cannavò L (17), Capecchi V (17), Ceri P (17), Bichi R (16), Garelli F (16), Ricolfi L (15), Gobo G (14), Battistelli F (12), Caselli M (12), Morcellini M (12), Statera G (12), Bruschi A (11), Cipriani R (11), Rositi F (11), Agnoli MS (10), Facchini C (10), Fasanella A (10).

Cluster 4 (yellow)
Nr of nodes: 27
Density: 3.014
Topic(s): migration (33.28), sociology of cultural processes (12.22), urban sociology (7.92), ethnic studied (7.5), multiculturalism (6.89);
Position: it is located at the extreme right of the map. It is characterised by the usage of qualitative methods;
Italian authors: Ambrosini M (79), Melucci A (48), Dal Lago Alessandro (47), Zanfrini L (40), Colombo E (39), Colombo A (35), Sciortino G (30), Caponio T (25), Leonini L (25), Pugliese E (25),

Queirolo Palmas L (24), Semi G (20), Pace E (17), Rebughini P (17), Lagomarsino F (15), Macioti MI (15), Palidda S (15), Zucchetti E (15), Besozzi E (14), Zincone G (14), Tognetti Bordogna M (13), Frisina A (12), Bosisio R (11), Decimo F (11), Allievi S (10), Cotesta V (10), Scidà G (10).

Cluster 5 (violet)

Nr of nodes: 21

Density: 1.708

Topic(s): sociology of health (22.49), sociology of communication (6.34), sociology of science (6.34), social policies (6.11), sociology of cultural processes (6.89);

Position: this cluster occupies the southern-most positioned; it is positioned at the centre. The central position can be related to the different techniques employed by the scholars constituting the cluster;

Italian authors: Cipolla C (59), Ardigò A (36), Giarelli G (31), Palumbo M (31), Stame N (31), Maturo A (28), Bucchi M (20), Fazzi L (19), Neresini F (19), Marzano M (16), Altieri L (15), Melossi D (14), Bruni A (13), Manghi S (13), Bertin G (12), Fele G (12), Ferrari V (12), Ingrosso M (12), Moruzzi M (11), Giglioli PP (10), Vicarelli G (10).

Cluster 6 (light blue)

Nr of nodes: 15

Density: 2.562

Topics: sociology of cultural processes (14.67), sociology of consumption (13.12), sociology of communication (12.68), gender studies (3.17), bibliometrics (3.17);

Position: it occupies one of the two (see cluster 4) rightmost positions in the map; as cluster 4, it is characterised by the usage of qualitative methods;

Italian authors: Cesareo V (38), Sassatelli R (29), Codeluppi V (29), Paltrinieri R (20), Di Nallo E (19), Bovone L (17), Vaccarini I (15), Alberoni F (14), Mora E (14), Santoro M (14), Diani M (13), Leccardi C (13), Parmiggiani P (12), Boccia Artieri G (12), Abruzzese A (10).

Cluster 7 (brown)

Nr of nodes: 10

Density: 5.490

Topics: social mobility (7.92), sociology of education (5), political sociology (4.14), inequalities (3.67), welfare (3.3);

Position: the cluster is located at the bottom, in the left-most position. It is characterised by the prevalent usage of quantitative methods; its position is specular to that of cluster 4, which is, instead, characterised by qualitative methods;

Italian authors: Schizzerotto A (55), Barbagli M (50), Pisati M (34), Ballarino G (20), Cobalti A (20), Gasperoni G (19), Schadee H (19), Ruspini E (15), Gambardella D (12), Moscati R (10).

Cluster 8 (black)
Nr of nodes: 8
Density: 8.571
Topics: social capital (6.89), third sector (6.34), social network (4.14), social policies (3.67), welfare (3.3);
Position: it is located in the centre of the map; it is a sparse cluster. It is possible to hypothesise that its position is due to both the lack of prevalence of qualitative or quantitative method and the linkages with the theorists placed in the left side of the map;
Italian authors: Donati P (95), Colozzi I (30), Chiesi AM (29), Prandini R (26), Di Nicola P (21), Tronca L (20), Rossi G (16), Boccacin L (11).

9.5. DISCUSSING RESULTS

The first considerations, about the empirical work previously introduced, relate to the following question: do results corroborate or not the hypotheses made? Before proceeding in that direction, I think it is useful to recapitulate the hypotheses made in Section 9.1. The first hypothesis is about the effectiveness of the ACA technique in intercepting clusters related to those research areas constituting the discipline. In particular, we expect that clusters have a certain homogeneity with reference to the topic, but at the same time we expect that cluster division informs about the qualitative-quantitative polarity. For this purpose, in the following pages the structural analysis of the networks will be also performed applying some basic network analysis measures, whose usefulness is fundamental to interpret the structure of the graphs (in Chapter 4 I have shown the affinity of relational bibliometric and social network analysis) and to make a comparison between them. The second hypothesis relates to the status of the scientific communication of Italian Sociology: according to what has been discussed in Section 8.2.2, the peculiar organisation of the Italian sociological community, which is characterised by the fragmentation into 3 components (Mi-To, Catholics and Third component) specialised in certain topics, would be reflected in the detection of what can be defined *intellectual parochialism*.

9.5.1. TESTING THE AUTHOR CO-CITATION TECHNIQUE

Analyses based on co-citation techniques (see Chapter 4) are widespread in science studies. They are among the major quantitative methods used with the aim of mapping the structure and dynamics of scientific research. According to the supporters of these techniques, thanks to them it is possible to detect "research foci" and their relations at the level of research specialities (Braam et al., 1991a):

> Co-citation analysis is in fact an attempt to identify [...] "high density areas" in a citation network by clustering highly co-cited documents, thus indicating the existence of these research

fronts. [...] The cluster of co-cited documents is considered to represent the knowledge base of the specialty [...] (Braam et al., 1991a, p. 233).

As discussed in Chapter 4, the co-citation approach comprises a variety of techniques, which, even if differ for the unit of analysis (authors, documents or words), have in common the scope, namely mapping scientific communities by identifying clusters relating to scientific domains (Griffith et al., 1974; Small & Crane, 1979; Small & Sweeney, 1985; Small & Griffith, 1974; Narin, 1976; Leydesdorff, 1987b; Todorov & Glänzel, 1988). The technique employed in this empirical study is called author co-citation analysis (ACA); intuitively, the unit of analysis are authors. Specifically, the version called all author co-citation analysis (Persson, 2001) including, in case of co-authorship, not only the first but all authors co-publishing, is used. White (1990) listed the main features of the ACA technique, namely: 1) author maps reveal the "cognitive" or "intellectual structure" of a field by showing the consensus of citers to important contributors and works; 2) maps based on co-cited authors show those scholars who are peripheral, and conversely, those who are central in a field; 3) the maps enable us to detect who is central and who is peripheral within clusters; 4) a quantitative-qualitative dimension usually is supposed to appear. As one of the purposes of the empirical work here developed is that of testing the effectiveness of the ACA technique, we should also be able to detect in our maps all the above mentioned elements.

With relation to the first point, in Section 7.3. we have seen that disciplines consist of specialties or research areas - actually the two terms, even if indicate two different domain configurations (see, for example Whitley, 1976), are often used for indicating scientific domains characterised by the commitment to a certain topic (or related topics); hereafter, when referring to these latter we will use the term research areas. Furthermore, we have just seen in point 1 that the main scope of Relational Bibliometrics is the detection of the intellectual and cognitive structure of scientific communities, therefore we expect that the networks are characterised by a topic-based cluster division criterion. In order to verify the occurence of this structure, we must refer to the labels extracted for identifying the nature of the clusters (see Appendix C and D). Thanks to them it is possible to notice that each cluster is characterised by topics not characterising the others. Specifically, even if for each cluster are listed 5 labels, in most cases only few of them show high TF-IDF values. These highly scored labels are those clearly defining the "identity" of the clusters. Accordingly, it seems possible to confirm the expectation about the topic-based cluster division criterion.

With reference to point 2, we can use some basic descriptive statistics that are effective in detecting who is central and who is peripheral in a network. One of them relates to the amount of citations each author received in the period considered (2007-2010). The 10 most cited authors of map1 (see Appendix E) are Bourdieu (155), Giddens (136), Bauman (133), Beck (122), Weber (106), Foucault (99), Donati (95), Goffman (92), Bagnasco (83), Ambrosini (79), for the less cited see directly Appendix E as the number of those obtaining score 10 is too high for being listed here. The 10 most cited of map2 are Donati (95), Bagnasco (83), Ambrosini (79), Trigilia (77), Gallino (66), Saraceno (63), Cipolla (59), Schizzerotto (55), Pizzorno (53) and Barbagli (50), while the less 10 cited authros are, Abruzzese (10), Agnoli (10), Allievi (10), Borlini (10), Burroni (10), Cotesta (10), Facchini (10), Fasanella (10), Gasparini (10), Giglioli (10), Minardi (10), Moscati (10), Poggio (10) and Viarelli (10). Secondly, we can use the statistics developed in the field of Network Analysis called degree centrality (local centrality): both graphs are valued (links are weighed according to the times a pair of authors was co-cited by thirds), thus the degrees consist of the sums of the weights of the links incident with each node (Wasserman & Faust, 1994). According to the nodal degree measure, the 10 most central authors of map1 are (see Appendix J): Bourdieu (1,979); Giddens (1,807); Bauman (1,799); Beck (1,593); Bagnasco (1,301); Weber (1,226); Donati (1,212); Ambrosini (1,121); Trigilia (1,099); Goffman (1,086). The most central authors of map2 are (see Appendix K): Bagnasco (593); Donati P (531); Ambrosini (527); Trigilia (499); Gallino (413); Schizzerotto (372); Paci (351); De Lillo (350); Magatti (347); Cavalli (318). As regards the most peripheral ones, in map1 we find Freeman (71), Denzin (70), Facchini (70), Moruzzi (69), De Sousa Santos (61), Pitch (55), Zagrebelsky (51), Ferrari (51), De Kerckhove (48), Kelsen (37) (see Appendix J); in map2 we find Allievi (38), Facchini (38), Boccia Artieri (33), Fele (32), Rositi (30), Gasparini G (30), Giglioli (30), Abruzzese (27), Morcellini (23), Ferrari (13) (see Appendix K).

Point 3 relates to considerations about the most and less central authors within each cluster. Statistics based on the nodal degree can also be used for obtaining this information. Appendix L and M show the percentage of authors' co-citations (degree centrality) divided by the total amount of co-citations within the cluster with reference to the international map and the national one, respectively. Formally:

$$(d_{ni} / \Sigma_{dc})*100 \tag{17}$$

where d_{ni} stands for the degree of a node, and $\sum_{dc}$ stands for the sum of the nodal degrees in a cluster. The measure is not standardised, therefore it is not possible to make a comparison across clusters. Looking at Appendix L, for example, we can observe that in cluster 1 of map1 the most central author is Bourdieu, whose nodal degree corresponds to the 6.10% of the cluster degree; the most peripheral is Kelsen with 0.11%. In cluster 1 of map2 the most central is Bagnasco with the percentage of 13.03%, while the less central is Gasparini G. with 0.66% (see Appendix M). Furthermore, we can, once again, look at the composition of the clusters (see Section 9.4.1 and 9.4.2) where each author is listed together with his/her citation count. So, for example, in cluster 8 of map2 Donati is the most central author with value 95, while the less central one is Boccacin with value 11.

For testing point 4 it is sufficient to observe the maps introduced in the previous Section having in mind the information about the position of each cluster. Thanks to this latter, it is possible to notice that in both maps the clusters are arranged in the two-dimensional space according to the qualitative-quantitative polarity. As it is indicated in the descriptive characteristics of each cluster (see Section 9.4.1. and 9.4.2.): in map1 moving from left to right we find authors mainly involved in qualitative and quantitative research, respectively; in map2 the qualitative pole is placed to the right while the quantitative is placed to the left.

In conclusion, according to the analysis performed it seems possible to say that the author co-citation analysis is effective in detecting the intellectual and cognitive structure of scientific fields.

9.5.2. TESTING THE INTELLECTUAL PAROCHIALISM HYPOTHESIS

What can be inferred from the maps with reference to the intellectual and cognitive structure of Italian Sociology? Which kind of information can we elicit with reference to the scientific communication characterising the Italian sociological community? Is it an intellectual tribalised community or a double-fragmented one? As mentioned in the Introduction to this Chapter, the second hypothesis tested in this work is the following: according to the normative approach and the conceptualisations of Chapter 6, we do not expect to detect in our co-citation maps serious communication problems, which would be caused by the *tripartition* characterising the Italian sociological field. This lack of communication should be reflected in a fragmentation attested by the presence of a *double-level fragmentation*, thus the maps should be characterised by two or more clusters dealing with the same topic(s), demonstrating that the clusters do not differentiate only with reference to the research topic but also to one or more other factors, such as the affiliation to a component. Conversely, we expect to detect traces of an intellectual parochialism supporting the assumption about a normative-oriented behaviour related to high citation frequencies.

In line with our expectations, the maps do not show signs of a communication crisis affecting Italian Sociology. This statement is supported by the same factor that allowed us to corroborate the first hypothesis. Specifically, each cluster is uniquely characterised by one or more topics (see Appendix C and Appendix D); it seems possible to exclude the *double-level fragmentation* hypothesis. However, as Santoro (2011) explained, each component seems to be specialised in one or more topics: the component called Mi-To seems to be focused on social inequalities, politics, economy and social movements; the Catholics are interested in cultural processes and sociology of communication; the Third component is very strong in methodology. This means that a further step in our analysis is necessary: we must search for this pattern in our maps. Unfortunately, it is not possible to establish the affiliation of every author in the maps to one component; this is because of different reasons: there is only one component furnishing a list of its members, namely the Catholic one (see www.sociologiaperlapersona.it); not all Italian sociologists adhere to a component; not all those adhering to a component make it explicit. As a consequence, to test the hypothesis of intellectual parochialism we must proceed in an indirect way. We can, for example, analyse the composition of the clusters in terms of the institutional affiliation[146] (University) of each author, observing if there are over-represented Universities in the clusters. Information about clusters composition is shown in Appendix F and G from which it is possible to derive the following tables showing the occurrence of each University in each cluster:

[146] For the method used to establish the institutional affiliation of the authors see page 156.

Table 4: Occurrence of each University in the clusters of Map1.

UNIVERSITY	**CL 1**	**CL 2**	**CL 3**	**CL 4**	**CL 5**	**CL 6**	**CL 7**	**CL 8**
BERGAMO	0	0	1	0	0	0	0	0
BICOCCA	1	5	0	9	1	0	3	0
BOLOGNA[147]	1	0	2	3	2	0	0	1
BOLOGNA (SPO)[148]	3	1	3	1	0	3	1	0
BOLOGNA (SPO Roberto Ruffili)[149]	0	0	2	1	1	0	0	0
CATANZARO	0	0	1	0	0	0	0	0
CATTOLICA	4	0	1	2	3	3	1	0
FERRARA	0	0	1	0	0	0	0	0
FIRENZE	0	2	2	0	0	0	1	1
GENOVA	0	1	0	0	4	0	0	0
IUE	0	0	0	0	0	0	0	1
IULM	1	0	0	0	0	0	0	0
LUSPIO	1	0	0	0	0	0	0	0
MACERATA	0	0	0	1	0	0	0	0
MILANO	2	1	1	3	5	1	0	0
NAPOLI-FEDERICO II	0	1	0	1	0	1	0	0
PADOVA	0	0	1	1	3	0	0	0
PARMA	0	0	1	0	0	0	0	0
PAVIA	0	1	1	0	0	1	0	0
PIEMONTE ORIENTALE	0	1	0	1	0	0	0	0
POLITECNICO MARCHE	0	0	1	0	0	0	0	0
POLITECNICO MILANO	0	0	0	1	0	0	1	0
POLITECNICO TORINO	0	0	0	0	0	0	2	0
PERUGIA	0	0	1	0	0	0	0	0
ROMA3	0	0	1	0	1	0	0	0
ROMA-LA SAPIENZA	2	2	5	3	2	1	0	0
TERAMO	0	1	0	0	0	0	0	0
TORINO	1	5	2	5	3	1	0	0
TRENTO	0	1	4	5	2	1	0	1
TRIESTE	0	0	0	0	0	2	1	0
UNIMORE (Modena-Reggio Emilia)	1	0	0	0	0	0	0	0
URBINO	1	1	0	0	0	0	0	0
VENEZIA (CA' FOSCARI)	0	1	0	0	0	0	0	0
VERONA	0	0	0	0	0	1	0	0

Legend of the clusters: sociology of cultural processes/sociology of communication/sociology of consumption/sociology of religion/gender studies; 2) methodology/sociology of health/sociology of cultural processes/sociology of knowledge/social policies; 3) economic sociology/local governance/public politics/sociology of organisation/sociology of work; 4) social mobility/sociology of work/welfare/economic sociology/social policies/inequalities; 5) migration/sociology of cultural processes/deviance/ multiculturalism/ethnic studies; 6) social networks/social capital/civicness/third sector/social policies; 7) urban sociology/sociology of the environment/ethnic studies/social integration/sociology of international relations; 8) deviance/political sociology/social capital/social movements/public policies.

[147] With reference to the University of Bologna I specified the Faculty as, according to what has been said in Section 8.2.2, the Faculty of Political Science used to be strongly related to the Catholic component.
[148] "SPO" stands for Faculty of Political Science.
[149] Faculty of Political Science "Roberto Ruffili".

Table 5: Occurrence of each University in the clusters of Map2.

UNIVERSITY	**CL 1**	**CL 2**	**CL 3**	**CL 4**	**CL 5**	**CL 6**	**CL 7**	**CL 8**
BERGAMO	0	0	0	0	1	0	0	0
BICOCCA	4	6	3	1	0	1	4	0
BOLOGNA[150]	1	0	2	1	2	1	2	0
BOLOGNA (SPO)[151]	1	2	0	0	3	3	0	3
BOLOGNA (SPO Roberto Ruffili)[152]	0	1	0	1	2	0	0	0
CATANZARO	0	0	0	0	1	0	0	0
CATTOLICA	1	1	2	3	0	4	0	2
FERRARA	0	0	0	0	1	0	0	0
FIRENZE	2	0	3	0	0	0	0	0
GENOVA	0	0	0	4	1	0	0	0
IUE	1	0	0	0	0	0	0	0
IULM	0	0	0	0	0	1	0	0
LUSPIO	0	0	0	0	0	1	0	0
MACERATA	0	1	0	0	0	0	0	0
MILANO	0	3	1	5	1	1	1	1
NAPOLI-FEDERICO II	2	0	0	0	0	0	1	0
PADOVA	0	1	0	3	1	0	0	0
PARMA	0	0	0	0	1	0	0	0
PAVIA	1	0	2	0	0	0	0	0
PIEMONTE ORIENTALE	1	1	0	0	0	0	0	0
POLITECNICO MARCHE	0	0	0	0	1	0	0	0
POLITECNICO MILANO	1	1	0	0	0	0	0	0
POLITECNICO TORINO	2	0	0	0	0	0	0	0
PERUGIA	0	0	1	0	0	0	0	0
ROMA3	0	0	1	1	0	0	0	0
ROMA-LA SAPIENZA	1	3	8	2	1	0	0	1
TERAMO	1	1	0	0	0	0	0	0
TORINO	5	3	4	3	0	0	0	0
TRENTO	0	4	1	2	4	1	2	0
TRIESTE	3	0	0	0	0	0	0	0
UNIMORE (Modena-Reggio Emilia)	0	0	0	0	0	1	0	0
URBINO	1	0	0	0	0	1	0	0
VENEZIA (CA' FOSCARI)	0	0	0	0	1	0	0	0
VERONA	0	0	0	0	0	0	0	1

Legend of the clusters: 1) urban sociology/local governance/sociology of the environment/local development/public policies; 2) economic sociology/sociology of work/welfare/social policies/labour policies; 3) methodology/sociology of cultural processes/sociology of knowledge/migration/social theory; 4) migration/sociology of cultural processes/urban sociology/ethnic studied/multiculturalism; 5) sociology of health/sociology of communication/sociology of science/social policies/sociology of cultural processes; 6) sociology of cultural processes/Sociology of consumption/sociology of communication/gender studies/bibliometrics; 7) social mobility/sociology of education/political sociology/inequalities/welfare; 8) social capital/third sector/social network/social policies/welfare.

[150] See Footnote 147.
[151] See Footnote 148.
[152] See Footnote 149.

As previously mentioned, the field of communication and cultural processes should be ruled by the Catholics, the Third component should dominate the methodological field, and Mi-To should be focused on inequalities, social mobility, economic sociology and politics (Santoro, 2011). Therefore, with the objective of searching for some form of intellectual parochialism, it seems reasonable to concentrate on these topics and observe if the occurrence pattern follows the above mentioned considerations. In order to do this, we must refer to the Section dedicated to the description of the Italian sociological field (Section 8.2.2.) and recall the Universities historically related to the components. In summary: Mi-To refers to the Universities in Milan, Turin, and Bologna; the Catholic component is linked to the Catholic University of Milan and the Faculty of Political Science in Bologna; the Third, or Roman, camp is quite generically related to the central and southern Italy (Freschi & Santoro, 2010). Bearing in mind this information, we can check in the tables above if it is possible to detect the subject-component pattern above described:

1) **Methodology**: the cluster number 3 of Map2 (the one with only Italian sociologists) is about methodology and it is characterised by one of the highest occurrence frequency in both maps: the Sapienza University in Rome with value 8. In Map1 methodology shares the cluster with the topics of sociology of science and sociology of health, therefore it is not possible to make a comparison between the two networks;
2) **Economic sociology/social mobility**: in both maps these fields are definitely the prerogative of Northern Italy with the University of Milano-Bicocca, Torino and Trento;
3) **Cultural processes/sociology of communication**: in both maps the Catholic University and the Faculty of Political Science at the University of Bologna seem to play an important role in the fields.

On the basis of the data shown in tables 4 and 5 it seems possible to talk about intellectual parochialism when referring to Italian Sociology: there are some topics which are the prerogative of some Universities. However, for example, with reference to point 1, we cannot exclude that summing the institutions referring to authors adhering to the other two components (Mi-To and Catholics) we could obtain a value equal, or even higher, to that obtained by the Sapienza. It must be underlined that the institutional affiliation cannot be used for obtaining information relating to the component adherence of the authors, as working in a university, a department or a faculty does not imply (at least officially) to be a member of one of the three components. Actually, as previously mentioned, on the website of the official association representing the Catholic component, namely “Sociologia per la persona”, it is possible to find a public list of the members. Using this source we can make some considerations on the composition of cluster 6 in Map2, which is characterised by those topics in which the Catholic group seems to be specialised (sociology of consumption, sociology of cultural processes and sociology of communication). The composition of the cluster is the following (in brackets degree centrality and institutional affiliation):

Cesareo V (38, Cattolica), Sassatelli R (29; Milano), Codeluppi V (29; Unimore), Paltrinieri R (20; Bologna-Political Science), Di Nallo E (19; Bologna-Political Science), Bovone L (17; Cattolica), Vaccarini I (15; Cattolica), Alberoni F (14; LUSPIO), Mora E (14; Cattolica), Santoro M (14; Bologna-Scienza Statistiche), Diani M (13; Trento), Leccardi C (13; Bicocca), Parmiggiani P (12; Bologna SPO), Boccia Artieri G (12; Urbino), Abruzzese A (10; IULM). Of these 15 authors only the following 4, according to Persona's website, belong to the Catholic component: Cesareo, Mora, Paltrinieri and Parmiggiani. But we do not know if also the other three scholars affiliated to the Universities historically linked to this component, that is to say Di Nallo, Bovone and Vaccarini, are affiliated to it or not.

In conclusion, as hypothesised in Chapter 6, the analysis here performed seems to support the idea relating to the effectiveness of the normative approach in describing citers' behaviour in relation to high citation frequencies. According to our expectations, we found traces of intellectual parochialism and no traces of tripartition. However, we cannot conclude that the tripartition does not have any effect on the scientific communication of Italian Sociology: we worked with high citations frequencies; we do not know what happens under the cutting point here chosen. Finally, it must be underlined that the intellectual parochialism cannot be used as a proof of the existence of the components; it can be considered as a proof of the existence of intellectual traditions linked to certain Universities.

9.5.3. SOME ADDITIONAL CONSIDERATIONS

As we previously discussed, the two networks show a similarity with reference to the number and nature of the clusters. Thanks to some structural measures[153] we can obtain more information about the structure of the graphs, besides making a comparison between them. This similarity is formally expressed by the structural measures related to each network. The network centralisation[154] of map1 is 6.483%, that of map2 is 6.169%; the density[155] of map1 and map2 are 1.111 and 0.5559 respectively.[156] The first measure relates to the variation in the co-citation pattern of the nodes, the second one, instead, relates to the average of the links weights in a graph. The maps show a small difference in the two measures, confirming the similarity of their structural characteristics. Going back to the Section dedicated to the description of the clusters, it can be interesting to dwell for a moment on the density by group values. Comparing the values between the two groups it is immediately clear that the networks show a difference

[153] Social network analysis is also referred to as structural analysis (Wellman & Berkowitz, 1988).

[154] Network centralisation measures relate to the variation among the centrality of the nodes in a graph; the value 0 indicates that all nodes occupy the same structural position. In valued graphs the centralisation measure loses its original meaning as it expresses the variation among nodal centrality values which, in turn, relate to the sum of the weighs of the links incident with each node. Therefore, in valued graphs, centralisation measures based on nodal degrees relate to the variation among the links weights in a graph (Chiesi, 1999).

[155] The density of a graph relates to the probability that a couple of nodes is connected. In case of valued graphs this measure can be understood as indicating the average of the links weights in the graph, therefore it can results in values bigger than 1 (Chiesi, 1999).

[156] The structural analysis was performed using the software Ucinet (Borgatti, Everett, & Freeman, 2002).

in clusters density. However, the density is inversely related to the network size (the formula for the density in a valued and not oriented graph is 2L/n(n-1)), therefore this measure is not useful for comparison purposes. As suggested by de Nooy et al. (2005), it can be useful to refer to the *average degree* of the nodes to measure the structural cohesion of a network as it is not dependent on the graph size. The values related to each cluster in both networks are the following:

Map1 → cluster (1) 2.429; cluster (2) 2.526; cluster (3) 2.213; cluster (4) 3.425; cluster (5) 3.148; cluster (6) 3.786; cluster (7) 2.483; cluster (8) 2.290
Map2 → cluster (1) 3.663; cluster (2) 3.233; cluster (3) 2.504; cluster (4) 3.792; cluster (5) 3.375; cluster (6) 3.091; cluster (7) 5.698; cluster (8) 8.889

We can now compare the total average degree in the networks summing the average measure of each cluster, which in map1 is 2.788 while in map2 is 4.280. The values suggest that in map2 scholars belonging to the same research area are co-cited more often than in map1. Once again it seems possible to notice what can be defined as a “disturbing” factor related to the presence of the foreign-classic scholars: as already suggested, they seem to act like *citations collectors* gathering citations from authors belonging to different and various sub-fields generating co-citation patterns such that often authors in a cluster has nothing in common but one or more of these *citations collectors*.

For having information about the “qualitative” characteristics of the maps we must perform an analysis focused on the composition of the clusters. Even if both the maps show 8 clusters, there are differences with reference to their composition: the presence of foreign authors produces co-citation patterns that disappear once removed. For example, Vincenzo Ferrari passes from cluster 1 in Map1 to cluster 5 in Map2. The reason for this change stands in his co-citation pattern: in Map1 he is in the first cluster as he was frequently cited together with those classics of the sociological thought such as Foucault, Giddens, Mead GH and Schutz as well as with philosophers such as Kant and Deridda who are in cluster 1 too. It seems possible to confirm what has been hypothesised in Section 9.4. about the role of these foreign authors defined as *citations collectors*: their presence is very important for obtaining information about the Italian sociological field (as they give information about the theoretical approach and/or the methodological one) but, as we are interested in extrapolating information related to the scientific communication of Italian Sociology, it constitutes a disturbing element in the detection of relations among Italian sociologists.

At an overall level, we can observe the following changes occurring in the passage from Map1 to Map2: we observe the decomposition of cluster 2, which split into two clusters in Map2 (clusters 1 and 2) as well as the decomposition of cluster 3 (it split in map2 in clusters 3 and 5); there is the disappearance of cluster 8. The following lists enable us to make an easy comparison between the maps:

A) MAP1 (list numbers relate to the clusters):

1. Sociology of cultural processes (14.29), sociology of communication (14.26), sociology of consumption (13.12), sociology of religion (4.14), gender studies (2.76);
2. Economic sociology (12.22), local governance (9.51), public politics (9.51), sociology of organisation (7.33), sociology of work (6);
3. Methodology (25.36), sociology of health (22.49), sociology of cultural processes (9.9), sociology of knowledge (7.92), social policies (6.89);
4. Social mobility (31.70), sociology of work (15), welfare (14.67), economic sociology (13.45), social policies (11.03), inequalities (8.27);
5. Migration (28.95), sociology of cultural processes (11), deviance (7.92), multiculturalism (6.89), ethnic studies (6.34);
6. Social networks (9.37), social capital (8.27), civicness (6.97), third sector (6.34), social policies (6.34);
7. Urban sociology (8.56), sociology of the environment (5.62), ethnic studies (3.17), social integration (3.17), Sociology of international relations (3.17);
8. Deviance (6.34), political sociology (4.63), social capital (3.17), social movements (3.17), public policies (2.44).

B) MAP2

1. Urban sociology (19.02), local governance (13.93), sociology of the environment (12.68), local development (11.25), public policies (9.29);
2. Economic sociology (18.34), sociology of work (18.00), welfare, (9.9), social policies (8.56), labour policies (6.97);
3. Methodology (28.53), sociology of cultural processes (9.78), sociology of knowledge (7.92), migration (6.34), social theory (6.11);
4. Migration (33.28), sociology of cultural processes (12.22), urban Sociology (7.92), ethnic studied (7.5), multiculturalism (6.89);
5. Sociology of health (22.49), sociology of communication (6.34), sociology of science (6.34), social policies (6.11), sociology of cultural processes (6.89);
6. Sociology of cultural processes (14.67), sociology of consumption (13.12), sociology of communication (12.68), gender studies (3.17), bibliometrics (3.17);
7. Social mobility (7.92), sociology of education (5), political Sociology (4.14), inequalities (3.67), welfare (3.30);
8. Social capital (6.89), third sector (6.34), social network (4.14), social policies (3.67), welfare (3.30).

I think it is interesting to dwell on the change occurred to cluster 3 of Map1, which in Map2 split into 2 clusters. Considering the composition of these groupings, it is clear that the removal of foreign-classical authors highly cited by Italian sociologists engaged in sociology of medicine, sociology of health, and sociology of Science (such as, for example, Robert Merton), had as a consequence the split of the cluster labelled methodology/sociology of health/sociology of knowledge in Map1 into two clusters of Map2 labelled methodology (cluster 3) and sociology of health (cluster 5), respectively.

Notwithstanding the fact that the clusters seem to be characterised by certain topics which, therefore, define their identity, it is possible to find a few authors apparently misplaced. The reason why I talk about "apparent misplacing" stands in the nature of the data processed by VOSviewer: co-occurrences among pairs of authors. We must keep in mind that what we see in the maps is the *usage of literature*. This implies different things. First of all, it implies that the presence of an author in a cluster does not mean that she/he is active only in that field. It means that her/his publications about that topic were highly cited by those scholars active in that specific field. To give a practical example of this phenomenon we can refer to the case of Antonio Cobalti: in both maps he belongs to the cluster related to the topics of social mobility but he also wrote about social capital and civicness. Thus, the fact that he is not a member of the cluster related to social capital and civicness does not mean that he is unrelated to those topics. In the second place, we must consider that the belonging of two (or more) authors to the same cluster can be caused by a co-publication of these authors. Intuitively, the co-publishing pattern is an important factor in the creation of clusters. Thus, potentially, for an author it is sufficient one co-publication which is highly cited for being "misplaced".

Finally, looking at the maps it is easily understandable the reason why I chose the unified method for mapping and clustering developed by Van Eck and Waltman (2010) among the available ones. Even if it does not allow the authors to belong to more than one cluster, the variety of their interests (here objectivised in publications cited at least 10 times in the period 2007-2010) is captured by overlapping areas as well as by the position in the map with respect to both the other members of her/his cluster and the other authors in the network. Thus, for example, Marco Caselli in Map2 belongs to cluster 3 but, due to its publications about topics such as migration and collective phenomena, he is placed in the area between the clusters 4 and 8.

9.6. SUMMARISING RESULTS

We have started this Section with two cognitive demands: is relational bibliometrics effective in mapping the intellectual and cognitive structure of science? Is the normative approach effective in describing citers' behaviour with reference to high citations counts? In order to answer these questions I developed two co-citation maps informing on the scientific communication of Italian Sociology as well as on its intellectual and cognitive structure. With reference to the first question, I hypothesised a positive response in case of clusters internal homogeneity with reference to the topic(s); we hypothesised the topic to be the fundamentum divisionis. As regards the second question, we hypothesised a positive response if traces of intellectual parochialism are detected: since there are institutions traditionally linked to certain topics, we expect the Italian sociologists behave normatively paying the intellectual debts even in case of membership to different components. The opposite situation would be that in which one or more topics concur to determine the "identity" of more than one cluster. After performing the analyses it seems possible to say that empirical results corroborate the hypotheses: according to the label-extraction, each cluster is clearly characterised by one or more topics which do not characterise the others; those Universities traditionally linked to certain topics seem to concentrate in those clusters characterised by them. The empirical study performed do not shows traces of a communication crisis in the field of Italian Sociology. Of course, this consideration relies on a study based on high citation frequencies; we do not know what happens in correspondence to low frequencies. Such findings are in accordance with what has been hypothesised in Chapter 6, where the effectiveness of the normative approach for describing citers' behaviour in relation to high citation counts is suggested. The constructivist approach is considered effective in describing citers' behaviour with reference to low citation counts. In accordance with the conceptualisations developed in Chapter 6, the maps show the authoritative resources of Italian Sociology (with reference to the period 2007-2010).

CONCLUSION

The work introduced in the previous pages had the purpose of proposing a sociologically integrated approach to Scientometrics as well as to demonstrate the effectiveness of Relational Bibliometrics in furnishing maps showing the intellectual and cognitive structure of disciplines, specialties or research areas. In order to do this, Bibliometrics was applied to the case of Italian Sociology and two co-citation maps were developed and analysed through both bibliometric and network analysis tools.

The thesis is structured into two Sections, one dealing with theoretical aspects and the other one dealing with empirical ones. Section 1 introduced Bibliometrics and Scientometrics. In the first Chapter, where the characteristics of Bibliometrics were introduced, we saw the reason why the terms Bibliometrics and Scientometrics are often used as synonyms: when the latter, for its purposes of quantification of scientific activity, uses bibliometric data and bibliometric techniques there is an overlap of the two fields. Bibliometrics roots date back to the 19^{th} century but its development occurred in the second half of the 1960s, after the introduction of the Science Citation Index by Eugene Garfield. The SCI, besides constituting an important source for the application of the 3 bibliometric laws, which are still considered the pillars of the discipline, was crucial for the development of Scientometrics. Furthermore, the SCI was instrumental for all those bibliometric measures, such as the impact factor or the h-index, which give information on the scientific performance of scholars, institutions or Countries. In the second Chapter the social and theoretical factors at the basis of the development of Scientometrics were introduced. The birth of the discipline is linked to the works by Derek John de Solla Price and Robert Merton - thanks to whom Science finally became subject to study: the first demonstrated the descriptive potential of Bibliometrics applied to Science; the second developed a theoretical frame which attracted Scientometrics (and, in particular, citationalists), as the scientific ethos allowed them to conceive of scientists’ behaviour as rational and predictable. Also Kuhn’s conceptualisations about the development of scientific knowledge were useful to scientomterists. In particular, Kuhn suggested referring to scientific literature (for example, manuals or manuscripts) as well as citation patterns to detect the so called “scientific revolutions”. Besides the theoretical background, the social background played a fundamental role in the development of Scientometrics, as the events of World War II, the uncontrolled growth of Science and the will of planning scientific activity constituted the perfect substratum for the development of a Science of Science. Citations, after the development of the SCI, immediately became the object of study of scientomterists who conferred to them a meaning and certain functions referring to Merton’s conceptualisations. In Chapter 3 the theoretical proposals on citation theory developed so far were introduced. They can be classified with respect to the level of adherence to the Mertonian normative approach in Sociology of Science or to the constructivist one. So, on the one hand we find those theories conceiving of citations as reward tools (used by scientists to pay intellectual debts) and intellectual links, on the other there are those theories stressing the rhetorical and strategic aspect in the act of citing (the stress in on the social and personal characteristics of the cited). Moreover, two proposals trying to build a bridge between the two approaches were introduced. Even if citations are more and more employed for descriptive and evaluative purposes in science policy they are not infallible indicators. The second part of Chapter 3 illustrated the limits and shortcomings of citation analysis, stressing those mistakes made by scholars in compiling bibliographic lists as well as those mistakes and problems relating to the way in which databases, such as the SCI or Scopus, are built. Furthermore, special attention is given to the application of Bibliometrics to the Social Sciences and Humanities, which are less and less bibliometrically inaccessible. Citations and other bibliometric elements (such as documents or semantic elements) are analysed through relational techniques with the purpose of furnishing maps of Science. Chapter 4 introduced the subfield called Relational Bibliometrics, stressing the importance of being aware of its collocation in the sociological relational paradigm as well as the potentialities of the application of Social Network Analysis to bibliometrics networks. The way in which the bibliometric maps of Science obtained from the analysis of scientific literature are developed was shown in Chapter 5, where the most used clustering and mapping techniques were introduced. Particular attention was given to the unified method for clustering and mapping called VOSviewer, as it is the method used for the creation of the maps developed in the second Section of this thesis. The last Chapter of this first Section was dedicated to the introduction of a theoretical proposal in citation theory. The increasing employment of bibliometric measures for assessing, describing and mapping Science inevitably led to the increasing need for a citation theory constituting a theoretical frame for both citation analysis and the description of citers' behaviour. Therefore, a theoretical model, encompassing both normative and constructivist approaches, is suggested. The conceptualisation of scientific communities as autopoietic systems, whose components are communicative events as well as structures, allow us to observe the re-productive function of citations conceived of as codes and medium of scientific communication. Citations, thanks to their constraining and enabling properties, constitute the engine of the structurating process ensuring the re-production of scientific communities. By referring to Gidden's structuration theory, Luhmann's theory about social systems as communicative networks, Merton's Sociology of Science and his conceptualisations about the functions of citations, as well as Small's proposal about citations as concept-symbols, a sociologically integrated approach to Scientometrics was proposed. Finally, it was suggested that while the normative approach is useful for describing the citing behaviour with respect to high citation counts, the constructivist one is useful for describing citers' behaviour with reference to low citation ones. Bibliometric measures and maps furnish information about those resources that Giddens call authoritative.

Section 2 consists of an empirical work having as object of analysis Italian Sociology, which was treated as a case study. Before developing two co-citation maps informing on the intellectual and cognitive structure of the field, the literature relating to the institutionalisation of scientific knowledge as well as the conceptualisation of scientific specialties and, more generally, the organisation of the scientific field, were introduced. The technique used for the empirical purposes of this thesis is called author co-citation analysis and relies on a certain conception of the structure of scientific communities. The institutionalisation of disciplines in universities is an important factor that must be considered if, as in this case, there is the will of analysing a scientific community: certain dynamics and mechanisms can be understood only if the institutional context is considered. Thus, for example, the Bourdieusian conceptualisations of the scientific field relate to the strategies implemented by scholars for being successful according to their role in the hierarchy of scientific communities; this hierarchy is determined by the institutional organisation of disciplines. As previously mentioned, the object of study of the bibliometric research performed was Italian Sociology, therefore Chapter 8 dealt with the structure of the Italian sociological field. The history of its institutionalisation was traced as in that process we find the reason of its peculiar organisation. Italian Sociology, in fact, seems to be divided into three camps or components. The reasons of this division are neither logical nor ontological, as they are rooted in the political forces that led the institutionalisation of the discipline in the 1960s and 1970s. Worries about the impairment of the scientific communication arose. Italian Sociology constituted, thus, a perfect case for testing what was hypothesised in Chapter 6, as the belonging to one component can be seen as a potentially strong constructivist reason for citing. The bibliometric maps developed in Chapter 9 (one related to both Italian and foreign scholars and the other relating only to Italian sociologists), besides providing a picture of the intellectual and cognitive structure of Italian Sociology (in the period 2007-2010), were useful for testing the two following hypotheses: is the technique called author co-citation analysis effective in furnishing structural snapshots of scientific domains? Is the normative approach effective in describing citers' behaviour in relation to high citation frequencies? Results point in the direction of a confirmation of the two hypotheses. The maps show clusters whose fundamentum divisionis seems to be the topic(s) and not the belonging to different components, therefore a crisis in the scientific communication of Italian Sociology was not detected. According to the normative approach in citation theory and as hypothesised in Chapter 6, most Italian scholars seem to cite normatively, paying their intellectual debts and acknowledging the work of the colleagues even when these latter are not part of their component. Considering the tripartition of the field and the fact that each component is supposed to be strong in some topics, I hypothesised to detect traces of intellectual parochialism. This hypothesis was corroborated by the frequent occurence of those Universities historically related to certain topics in those clusters dealing with these latter. The intellectual parochialism cannot be used as a proof of the existence of the components; it is a proof of the existence of intellectual traditions linked to certain Universities. The impossibility to work directly on the components issue is due to the difficulty of establishing the affiliation of each author: besides the lack of official affiliation lists (except for the Catholic component formalised in the association "Sociologia per la persona"), it must be considered that some scholars do not declare or deny the affiliation to one component. Of course, the two maps are the result of the data employed in the research: on the basis of the selections made by 25 members of the Italian Association of Sociology, I selected 23 journals of Sociology that I used for building a dataset from which I extracted data relating to bibliographic references in order to perform the author co-citation analysis. The elements in the maps must be read according to the theory developed in Chapter 6, therefore authors (and the documents authored by them that were cited at least 10 times in the period considered) must be understood as the authoritative resources, for the period analysed, of the Italian (academic) sociological community.

APPENDIX A

Table 6: Scores received by Italian journals.

Italian sociological journals	**Scores**
Etnografia e ricerca qualitativa	8
Futuribili	2
Ikon	3
Il Mulino	6
La critica sociologica	8
Mondi migrant	5
Polis	9
Quaderni di sociologia	11
Rassegna italiana di sociologia	12
Rassegna italiana di valutazione (RIV)	5
Rivista italiana di comunicazione pubblica	6
Rivista trimestrale di scienza dell'amministrazione	4
Salute e società	8
Sociologia della comunicazione	9
Sociologia del diritto	10
Sociologia del lavoro	9
Sociologia e politiche sociali	8
Sociologia e ricerca sociale	10
Sociologia urbana e rurale	9
Sociologica	11
Stato e mercato	11
Studi culturali	8
Studi organizzativi	10
Studi di Sociologia	9

APPENDIX B

Table 7: Composition (authors and clusters) of map1.

NAME	CLUSTER	NAME	CLUSTER	NAME	CLUSTER
BOURDIEU P	1	HEDSTROM P	2	LEWIS J	4
GIDDENS A	1	PATTON MQ	2	BALBO L	4
BAUMAN Z	1	PIERSON P	2	SAVAGE M	4
BECK U	1	WEICK KE	2	DUBET F	4
WEBER M	1	STORPER M	2	JONSSON JO	4
CASTELLS M	1	STIGLITZ J	2	LODIGIANI R	4
SIMMEL G	1	WEISS CH	2	GANZEBOOM HBG	4
FOUCAULT M	1	PFEFFER J	2	PICCONE STELLA S	4
APPADURAI A	1	SCHON D	2	BISON I	4
HABERMAS J	1	HALL P	2	GINTIS H	4
LUHMANN N	1	BOYER R	2	HOUT M	4
TAYLOR C	1	PORTER M	2	POGGIO B	4
CESAREO V	1	BERTIN G	2	CROMPTON R	4
DOUGLAS M	1	POLANYI M	2	GAMBARDELLA D	4
DE CERTEAU M	1	LIPSEY MW	2	O'CONNOR J	4
LASH S	1	LAVILLE JL	2	PAVOLINI E	4
BOLTANSKI L	1	GUBA EG	2	BOWLES S	4
SASSATELLI R	1	LINCOLN YS	2	SEMENZA R	4
BECKER H	1	BURRONI L	2	OLAGNERO M	4
HALL S	1	NEGRELLI S	2	GERSHUNY J	4
JEDLOWSKI P	1	ROWAN B	2	LAYTE R	4
MORIN E	1	GUALA C	2	MOSCATI R	4
ELIAS N	1	FLIGSTEIN N	2	FACCHINI C	4
CODELUPPI V	1	BATTISTELLI F	2	AMBROSINI M	5
AUGé M	1	KRUGMAN P	2	PORTES A	5
BOVONE L	1	NONAKA I	2	MELUCCI A	5
MAUSS M	1	FRIEDMAN M	2	DAL LAGO A	5
FERRAROTTI F	1	GAMBETTA D	2	ZANFRINI L	5
VACCARINI I	1	WOLLEB G	2	COLOMBO E	5
PALTRINIERI R	1	OLSON M	2	GLICK SCHILLER N	5
ISHERWOOD B	1	CERASE FP	2	ANDERSON B	5
ARENDT H	1	MINARDI E	2	COLOMBO A	5
MARX K	1	FREEMAN RE	2	LANDOLT P	5
LéVI-STRAUSS C	1	GOFFMAN E	3	VERTOVEC S	5
MORA E	1	MERTON RK	3	LEONINI L	5
BALIBAR E	1	DURKHEIM E	3	GUARNIZO L	5
COHEN S	1	PARSONS T	3	SCIORTINO G	5
BAUDRILLARD J	1	CIPOLLA C	3	PUGLIESE E	5

DI NALLO E	1	LATOUR B	3	LEVITT P	5
MILLER D	1	MARRADI A	3	HANNERZ U	5
RICOEUR P	1	GHERARDI S	3	BASCH L	5
VEBLEN T	1	LAZARSFELD P	3	SEMI G	5
MAFFESOLI M	1	GEERTZ C	3	SZANTON BLANC C	5
LATOUCHE S	1	CRESPI F	3	REBUGHINI P	5
LECCARDI C	1	ARDIGò A	3	QUEIROLO PALMAS L	5
ADORNO T	1	SCHUTZ A	3	CAPONIO T	5
LIPOVETSKY G	1	KUHN T	3	MASSEY D	5
ALEXANDER JC	1	GIARELLI G	3	MARSHALL TH	5
TODOROV T	1	LUCKMANN T	3	ZUCCHETTI E	5
GARELLI F	1	STRAUSS A	3	SAYAD A	5
RITZER G	1	CAPECCHI V	3	RUMBAUT R	5
LASCH C	1	CAMPELLI E	3	CLIFFORD J	5
FEATHERSTONE M	1	KNORR-CETINA K	3	FAIST T	5
BELL D	1	GARFINKEL H	3	BESOZZI E	5
ALBERONI F	1	ABBOTT A	3	CASTLES S	5
WARDE A	1	WENGER E	3	HALLER W	5
BENJAMIN W	1	MATURO A	3	BOSISIO R	5
WALLERSTEIN I	1	NERESINI F	3	WALDINGER R	5
SANTORO M	1	CARDANO M	3	BRUBAKER R	5
BUTLER J	1	ILLICH I	3	MANNHEIM K	5
RIFKIN J	1	BERGER PL	3	NEE V	5
DELEUZE G	1	BATESON G	3	ZHOU M	5
HONNETH A	1	LAW J	3	HOCHSCHILD AR	5
PARMIGGIANI P	1	CALLON M	3	MELOSSI D	5
SILVERSTONE R	1	BERGER P	3	BAUMANN G	5
WIEVIORKA M	1	BRUNI A	3	LAGOMARSINO F	5
AGAMBEN G	1	WITTGENSTEIN L	3	WALZER M	5
SWIDLER A	1	RICOLFI L	3	ZINCONE G	5
CAMPBELL C	1	GOBO G	3	FRISINA A	5
ROSE N	1	FREIDSON E	3	PACE E	5
KANT I	1	DEWEY J	3	MARCUS G	5
ROBERTSON R	1	LAVE J	3	TOGNETTI BORDOGNA M	5
BOUTANG YM	1	MILLS CW	3	ANDALL J	5
BELLAH RN	1	MARZANO M	3	PALIDDA S	5
FINE GA	1	STATERA G	3	MILLER MJ	5
JENKINS H	1	BICHI R	3	SCIDà G	5
HEBDIGE D	1	BUCCHI M	3	PARRENAS R	5
SAID E	1	SMELSER NJ	3	BURAWOY M	5
HORKHEIMER M	1	SMITH A	3	HUNTINGTON SP	5

KATZ E	1	WOOLGAR S	3	DECIMO F	5
LEFEBVRE H	1	POPPER KR	3	ONG A	5
CRANE D	1	GIGLIOLI PP	3	JOPPKE C	5
THOMPSON JB	1	BLUMER H	3	MACIOTI MI	5
HOBSBAWM EJ	1	NICOLINI D	3	NUSSBAUM M	5
HALBWACHS M	1	HUSSERL E	3	COTESTA V	5
SCHMITT C	1	GOOD B	3	LISTER R	5
TEUBNER G	1	CICOUREL A	3	SHORE C	5
HARDT M	1	STAR SL	3	ALLIEVI S	5
MILLER P	1	BRUSCHI A	3	DONATI P	6
RAWLS J	1	ALTIERI L	3	COLEMAN J	6
BOCCIA ARTIERI G	1	CANNAVò L	3	PUTNAM R	6
DERRIDA J	1	FELE G	3	CAVALLI A	6
McLUHAN M	1	SUCHMAN L	3	SCIOLLA L	6
GUATTARI F	1	THOMAS WI	3	CHIESI AM	6
BARTHES R	1	BIJKER WE	3	PISELLI F	6
ANDERSON C	1	MEAD GH	3	COLOZZI I	6
WILLIAMS R	1	BURY M	3	BURT R	6
GILROY P	1	MALINOWSKI B	3	ARCHER M	6
FREUD S	1	CZARNIAWSKA B	3	DI NICOLA P	6
BRAUDEL F	1	MANGHI S	3	LIN N	6
MEYROWITZ J	1	GADAMER HG	3	FUKUYAMA F	6
HERVIEU-LéGER D	1	FASANELLA A	3	BUZZI C	6
ABRUZZESE A	1	TIMMERMANS S	3	TRONCA L	6
JONAS H	1	HUGHES E	3	PRANDINI R	6
MORCELLINI M	1	TURNER B	3	PELLIZZONI L	6
SPIVAK GC	1	BROWN P	3	WELLMAN B	6
TURNER V	1	ROSITI F	3	FORSé M	6
DE SOUSA SANTOS B	1	VICARELLI G	3	JACOBS J	6
PITCH T	1	CIPRIANI R	3	ROSSI G	6
ZAGREBELSKY G	1	GALTUNG J	3	SCOTT J	6
FERRARI V	1	BERG M	3	BORGATTI SP	6
DE KERCKHOVE D	1	BRUNER J	3	WASSERMAN S	6
KELSEN H	1	COSER L	3	PARETO V	6
BAGNASCO A	2	CONRAD P	3	OSTI G	6
TRIGILIA C	2	SONTAG S	3	FAUST K	6
GRANOVETTER M	2	BERTAUX D	3	MAYNTZ R	6
POWELL WW	2	SILVERMAN D	3	ERIKSON EH	6
CROUCH C	2	INGROSSO M	3	LOCKWOOD D	6
DI MAGGIO PJ	2	MOSCOVICI S	3	RHODES R	6
MARCH JG	2	AGNOLI MS	3	BOCCACIN L	6

HIRSCHMAN AO	2	SACKETT DL	3	CASELLI M	6
MUTTI A	2	MINTZBERG H	3	KOOIMAN J	6
POLANYI K	2	DENZIN N	3	FOLGHERAITER F	6
PIORE M	2	MORUZZI M	3	SENNETT R	7
PICHIERRI A	2	SEN A	4	SASSEN S	7
SIMON H	2	GALLINO L	4	HARVEY D	7
LE GALèS P	2	BOUDON R	4	MARTINOTTI G	7
ROSSI P	2	ESPING-ANDERSEN G	4	URRY J	7
RAMELLA F	2	SARACENO C	4	MELA A	7
CELLA G	2	GOLDTHORPE J	4	WACQUANT L	7
STREECK W	2	SCHIZZEROTTO A	4	VITALE T	7
BARBERA F	2	MAGATTI M	4	PARK R	7
PALUMBO M	2	PACI M	4	AMIN A	7
SABEL C	2	DE LILLO A	4	ZUKIN S	7
BONAZZI G	2	REYNERI E	4	NUVOLATI G	7
DE LEONARDIS O	2	BARBAGLI M	4	AMENDOLA G	7
STAME N	2	RANCI C	4	ZAJCZYK F	7
OLSEN JP	2	COLLINS R	4	GUIDICINI P	7
CROZIER M	2	CORBETTA P	4	TOSI A	7
BIFULCO L	2	MINGIONE E	4	BURGESS E	7
DONOLO C	2	REGINI M	4	BORLINI B	7
LA ROSA M	2	FERRERA M	4	THRIFT N	7
VOELZKOW H	2	BARBIERI P	4	RAUTY R	7
FREEMAN H	2	PISATI M	4	WILSON WJ	7
SOSKICE D	2	CASTEL R	4	DAVICO L	7
PERULLI P	2	DAHRENDORF R	4	DONZELOT J	7
DORE R	2	BORGHI V	4	LYNCH K	7
WEICK K	2	NEGRI N	4	GASPARINI G	7
BUTERA F	2	COBALTI A	4	GASPARINI A	7
PAWSON R	2	BIANCO ML	4	PIZZORNO A	8
LIPPI A	2	ACCORNERO A	4	TOURAINE A	8
CERI P	2	BALLARINO G	4	ELSTER J	8
SWEDBERG R	2	BREEN R	4	DELLA PORTA D	8
ROSSI PH	2	BLOSSFELD HP	4	OFFE C	8
FAZZI L	2	PASSERON JC	4	ETZIONI A	8
MEYER JW	2	SCHADEE H	4	EISENSTADT SN	8
FLORIDA R	2	BLAU P	4	HELD D	8
OSTROM E	2	SHAVIT Y	4	TILLY C	8
WILLIAMSON OE	2	BERNARDI F	4	NORRIS P	8
FRIEDBERG E	2	NALDINI M	4	DIANI M	8
STINCHCOMBE A	2	BECKER G	4	ROSANVALLON P	8
THéVENOT L	2	ERIKSON R	4	DAHL R	8

ARGYRIS C	2	GASPERONI G	4	CATANZARO R	8
POWER M	2	TREIMAN D	4	TARROW S	8
LEWIN K	2	SUPIOT A	4	INGLEHART R	8
SALAIS R	2	DUNCAN O	4	LIPSET SM	8
SCHUMPETER J	2	MULLER W	4	MAIR P	8
SCHON DA	2	RUSPINI E	4	SKOCPOL T	8
SCRIVEN M	2	FULLIN G	4	COHEN J	8
SELZNICK P	2	SCHERER S	4		
AXELROD R	2	RIZZA R	4		

APPENDIX C (international map labels extraction)[157]

CLUSTER1

TERM	TF[158]	TF_IDF[159]
sociology of cultural processes	13	14.29
sociology of communication	9	14.26
sociology consumption	7	13.12
sociology of religion	3	4.14
gender studies	2	2.76

CLUSTER2

TERM	TF	TF_IDF
economic sociology	10	12.22
local governance	6	9.51
public policies	3	9.51
sociology of organisation	6	7.33
sociology of work	6	6.00

CLUSTER3

TERM	TF	TF_IDF
methodology	16	25.36
sociology of health	12	22.49
sociology of cultural processes	9	9.90
sociology of knowledge	5	7.92
social policies	5	6.89

CLUSTER4

TERM	TF	TF_IDF
social mobility	10	31.7
sociology of work	15	15.00
welfare	12	14.67
economic sociology	11	13.45
social policies	8	11.03
inequalities	6	8.27

[157] The 5 labels with the highest TF-IDF values are shown.
[158] TF counts refer to the times a term appears in a cluster.
[159] TF-IDF counts are the result of the number of times a word appears in a cluster divided by the frequency of the word in all clusters.

CLUSTER5

TERM	TF	TF_IDF
migration	21	28.95
sociology of cultural processes	10	11.00
deviance	5	7.92
multiculturalism	5	6.89
ethnic studies	4	6.34

CLUSTER6

TERM	TF	TF_IDF
social networks	5	9.37
social capital	6	8.27
civicness	3	6.97
third sector	4	6.34
social policies	3	4.14

CLUSTER7

TERM	TF	TF_IDF
urban sociology	7	8.56
sociology of the environment	3	5.62
ethnic studies	2	3.17
social integration	1	3.17
sociology of international relations	1	3,17

CLUSTER8

TERM	TF	TF_IDF
deviance	3	6.34
political sociology	4	4.63
social capital	3	3.17
social movements	3	3.17
public policies	2	2.44

APPENDIX D (national map labels extraction)[160]

CLUSTER1	TERM	TF[161]	TF_IDF[162]
	urban sociology	13	19.02
	local governance	6	13.93
	sociology of the environment	4	12.68
	local development	6	11.25
	public policies	4	9.29

CLUSTER2	TERM	TF	TF_IDF
	economic sociology	15	18.34
	sociology of work	19	18.00
	welfare	9	9.90
	social policies	7	8.56
	labour policies	3	6.97

CLUSTER3	TERM	TF	TFIDF
	methodology	18	28.53
	sociology of cultural processes	8	9.78
	sociology of knowledge	5	7.92
	migration	4	6.34
	social theory	5	6.11

CLUSTER4	TERM	TF	TF_IDF
	migration	21	33.28
	sociology of cultural processes	10	12.22
	urban sociology	5	7.92
	ethnic studies	4	7.50
	multiculturalism	5	6.89

[160] See Footnote 157.
[161] See Footnote 158.
[162] See Footnote 159.

CLUSTER5	TERM	TF	TF_IDF
	sociology of health	12	22.49
	sociology of communication	4	6.34
	sociology of science	2	6.34
	social policies	5	6.11
	sociology of cultural processes	4	4.89

CLUSTER6	TERM	TF	TF_IDF
	sociology of cultural processes	12	14.67
	sociology of consumption	7	13.12
	sociology of communication	8	12.68
	gender studies	2	3.17
	bibliometrics	1	3.17

CLUSTER7	TERM	TF	TF_IDF
	social mobility	5	7.92
	sociology of education	5	5.00
	political sociology	3	4.14
	inequalities	3	3.67
	welfare	3	3.30

CLUSTER8	TERM	TF	TF_IDF
	social capital	5	6.89
	third sector	4	6.34
	social networks	3	4.14
	social policies	3	3.67
	welfare	3	3.30

APPENDIX E

Table 8: Times each author was cited in the period 2007-2010.

CITATION COUNTS	NAME	CITATION COUNTS	NAME	CITATION COUNTS	NAME
155	BOURDIEU P	13	FRIEDMAN M	12	TOSI A
136	GIDDENS A	13	WOLLEB G	12	SWEDBERG R
133	BAUMAN Z	20	ERIKSON R	12	STINCHCOMBE A
122	BECK U	20	GUARNIZO L	12	DERRIDA J
106	WEBER M	20	BUTERA F	12	HALBWACHS M
99	FOUCAULT M	20	MARSHALL TH	12	CERASE FP
95	DONATI P	19	DI NALLO E	12	BRUBAKER R
92	GOFFMAN E	19	SILVERSTONE R	12	SWIDLER A
83	BAGNASCO A	19	BIFULCO L	12	CASELLI M
79	AMBROSINI M	19	BALIBAR E	12	WEISS CH
78	CASTELLS M	19	GUIDICINI P	12	MORCELLINI M
77	TRIGILIA C	19	DELEUZE G	12	BELLAH RN
71	SEN A	19	WENGER E	12	HUNTINGTON SP
68	PUTNAM R	19	MAFFESOLI M	12	NONAKA I
67	GRANOVETTER M	19	ACCORNERO A	12	SMELSER NJ
66	GALLINO L	19	SCHADEE H	12	FRISINA A
66	SIMMEL G	19	FLORIDA R	12	TURNER B
65	LUHMANN N	19	NERESINI F	12	THéVENOT L
65	HABERMAS J	19	GASPERONI G	12	MEYROWITZ J
63	SARACENO C	19	NEGRI N	12	SILVERMAN D
63	COLEMAN J	19	FAZZI L	12	STATERA G
63	ESPING-ANDERSEN G	19	SOSKICE D	12	INGROSSO M
62	MERTON RK	19	DONOLO C	12	FULLIN G
59	CIPOLLA C	19	PASSERON JC	12	DUNCAN O
56	APPADURAI A	18	NUVOLATI G	12	BIJKER WE
55	DURKHEIM E	18	ETZIONI A	12	NEGRELLI S
55	SCHIZZEROTTO A	18	BENJAMIN W	12	FERRARI V
54	SENNETT R	18	CALLON M	12	HARDT M
54	PORTES A	18	VOELZKOW H	12	JONAS H
53	PIZZORNO A	18	BIANCO ML	12	TIMMERMANS S
52	BOUDON R	18	ABBOTT A	12	GAMBARDELLA D
50	BARBAGLI M	18	SHAVIT Y	12	BERTIN G
49	POWELL WW	18	SZANTON BLANC C	12	BOCCIA ARTIERI G
48	MELUCCI A	18	RIFKIN J	12	PARMIGGIANI P

48	PARSONS T	18	JENKINS H	12	GUALA C
47	DAL LAGO A	18	SMITH A	12	WALLERSTEIN I
46	GOLDTHORPE J	18	PARK R	12	BOYER R
46	PACI M	18	BLUMER H	12	GERSHUNY J
46	MARRADI A	18	NORRIS P	12	WOOLGAR S
46	MAGATTI M	17	CANNAVò L	12	MAYNTZ R
45	LATOUR B	17	PAWSON R	12	FREUD S
45	CAVALLI A	17	CAPECCHI V	12	SEMENZA R
44	REYNERI E	17	BOVONE L	12	OLSON M
43	TAYLOR C	17	MILLER D	12	FELE G
43	DI MAGGIO PJ	17	REBUGHINI P	12	BATTISTELLI F
43	DE CERTEAU M	17	DORE R	12	PARETO V
42	MARCH JG	17	BELL D	12	PFEFFER J
42	CROUCH C	17	VITALE T	11	HOBSBAWM EJ
40	CORBETTA P	17	ROSE N	11	ANDERSON C
40	FERRERA M	17	FREIDSON E	11	HEDSTROM P
40	ZANFRINI L	17	NALDINI M	11	SUCHMAN L
40	DE LILLO A	17	SCHERER S	11	SAVAGE M
39	DOUGLAS M	17	WILLIAMSON OE	11	GINTIS H
39	COLOMBO E	17	PACE E	11	SACKETT DL
39	HIRSCHMAN AO	17	CLIFFORD J	11	BRUSCHI A
38	CESAREO V	17	LIPOVETSKY G	11	KATZ E
38	SIMON H	17	RITZER G	11	TURNER V
38	ANDERSON B	17	LIN N	11	HUSSERL E
38	SASSEN S	17	GOOD B	11	CIPRIANI R
37	GHERARDI S	17	LIPPI A	11	OLAGNERO M
37	TOURAINE A	17	RUMBAUT R	11	RAUTY R
36	BECKER H	17	ADORNO T	11	MEAD GH
36	RANCI C	17	AGAMBEN G	11	GASPARINI A
36	ARDIGò A	17	CERI P	11	BOUTANG YM
36	GEERTZ C	16	BERNARDI F	11	HUGHES E
35	COLOMBO A	16	BERGER P	11	BOSISIO R
34	PISELLI F	16	ROSSI G	11	MILLER MJ
34	MARTINOTTI G	16	POPPER KR	11	CICOUREL A
34	PISATI M	16	OSTROM E	11	FAIST T
33	CASTEL R	16	OFFE C	11	BISON I
33	ARCHER M	16	STIGLITZ J	11	GAMBETTA D
33	POLANYI K	16	DUBET F	11	GADAMER HG
33	FERRAROTTI F	16	ALEXANDER JC	11	BOCCACIN L
33	HARVEY D	16	ISHERWOOD B	11	NUSSBAUM M

33	MUTTI A	16	GARELLI F	11	STAR SL
31	HALL S	16	MILLS CW	11	ROSITI F
31	LéVI-STRAUSS C	16	LAW J	11	DAVICO L
31	COLLINS R	16	LAVILLE JL	11	BERTAUX D
31	RAMELLA F	16	WEICK K	11	GANZEBOOM HBG
31	SCIOLLA L	16	BICHI R	11	ROWAN B
31	PALUMBO M	16	ROSSI PH	11	MORUZZI M
31	BOLTANSKI L	16	FREEMAN H	11	GUATTARI F
31	LASH S	16	CASTLES S	11	HEBDIGE D
31	STAME N	16	MARZANO M	11	DEWEY J
31	GIARELLI G	15	HALL P	11	WIEVIORKA M
30	ELIAS N	15	POLANYI M	11	SCHMITT C
30	SCIORTINO G	15	MULLER W	11	BROWN P
30	LAZARSFELD P	15	TREIMAN D	11	WILSON WJ
30	AUGé M	15	HELD D	11	SPIVAK GC
30	COLOZZI I	15	POWER M	11	ROSANVALLON P
30	GLICK SCHILLER N	15	MEYER JW	11	SONTAG S
30	URRY J	15	PERULLI P	11	KRUGMAN P
30	STRAUSS A	15	CONRAD P	11	SCRIVEN M
29	MORIN E	15	RICOLFI L	11	WASSERMAN S
29	SCHUTZ A	15	WALZER M	11	DECIMO F
29	CODELUPPI V	15	CATANZARO R	11	THOMPSON JB
29	PICHIERRI A	15	PALIDDA S	11	HOUT M
29	CHIESI AM	15	SCOTT J	11	LIPSEY MW
29	ELSTER J	15	ERIKSON EH	11	BALBO L
29	MAUSS M	15	AMIN A	11	DENZIN N
29	SASSATELLI R	15	MACIOTI MI	11	RHODES R
28	MATURO A	15	LAGOMARSINO F	11	HORKHEIMER M
28	JEDLOWSKI P	15	RUSPINI E	11	PICCONE STELLA S
28	REGINI M	15	BURY M	11	NICOLINI D
28	LA ROSA M	15	ARGYRIS C	11	PARRENAS R
28	MASSEY D	15	ZUCCHETTI E	11	SALAIS R
28	KUHN T	15	VACCARINI I	11	ZAGREBELSKY G
27	MINGIONE E	15	ALTIERI L	11	MANNHEIM K
27	MARX K	15	JACOBS J	10	JOPPKE C
27	VERTOVEC S	15	WITTGENSTEIN L	10	THRIFT N
27	LE GALèS P	14	BARTHES R	10	SCIDà G
26	PRANDINI R	14	PATTON MQ	10	KOOIMAN J

26	HANNERZ U	14	FEATHERSTONE M	10	BURRONI L
26	STREECK W	14	MORA E	10	LOCKWOOD D
26	PIORE M	14	ZUKIN S	10	LINCOLN YS
26	SABEL C	14	ALBERONI F	10	DE SOUSA SANTOS B
26	RICOEUR P	14	KANT I	10	ONG A
26	BARBIERI P	14	SCHON DA	10	GUBA EG
26	ARENDT H	14	CROMPTON R	10	POGGIO B
25	CAPONIO T	14	FREEMAN RE	10	MAIR P
25	CRESPI F	14	OSTI G	10	LIPSET SM
25	LEONINI L	14	SCHUMPETER J	10	LISTER R
25	BARBERA F	14	VEBLEN T	10	GILROY P
25	BORGHI V	14	MELOSSI D	10	ANDALL J
25	DAHRENDORF R	14	TILLY C	10	LYNCH K
25	MELA A	14	SANTORO M	10	FINE GA
25	PELLIZZONI L	14	RIZZA R	10	DONZELOT J
25	PUGLIESE E	14	BESOZZI E	10	ALLIEVI S
24	CAMPELLI E	14	ZAJCZYK F	10	STORPER M
24	QUEIROLO PALMAS L	14	HOCHSCHILD AR	10	ROBERTSON R
24	BATESON G	14	GOBO G	10	GIGLIOLI PP
24	AMENDOLA G	14	ZINCONE G	10	SCHON D
24	BURT R	14	CAMPBELL C	10	FOLGHERAITER F
23	CELLA G	14	MARCUS G	10	FLIGSTEIN N
23	LEVITT P	14	AXELROD R	10	MOSCATI R
23	OLSEN JP	14	PIERSON P	10	MOSCOVICI S
23	ROSSI P	13	BURAWOY M	10	AGNOLI MS
23	BERGER PL	13	LECCARDI C	10	BORGATTI SP
23	DELLA PORTA D	13	DAHL R	10	SHORE C
23	GARFINKEL H	13	JONSSON JO	10	MALINOWSKI B
23	FUKUYAMA F	13	LEWIN K	10	VICARELLI G
23	WACQUANT L	13	LAVE J	10	PITCH T
23	CARDANO M	13	HONNETH A	10	COTESTA V
22	BONAZZI G	13	McLUHAN M	10	BORLINI B
22	BECKER G	13	LODIGIANI R	10	FACCHINI C
22	DE LEONARDIS O	13	WEICK KE	10	LEFEBVRE H
22	BAUDRILLARD J	13	DIANI M	10	LAYTE R
22	LUCKMANN T	13	TARROW S	10	CZARNIAWSKA B
22	BREEN R	13	LASCH C	10	SKOCPOL T
22	ILLICH I	13	TEUBNER G	10	GALTUNG J
21	SAYAD A	13	HALLER W	10	CRANE D

21	LANDOLT P	13	INGLEHART R	10	O'CONNOR J
21	BLAU P	13	FRIEDBERG E	10	GASPARINI G
21	BUZZI C	13	TODOROV T	10	DE KERCKHOVE D
21	WELLMAN B	13	MANGHI S	10	MINTZBERG H
21	CROZIER M	13	PAVOLINI E	10	BRAUDEL F
21	BLOSSFELD HP	13	PORTER M	10	SELZNICK P
21	SAID E	13	TOGNETTI BORDOGNA M	10	WALDINGER R
21	COHEN S	13	SUPIOT A	10	BAUMANN G
21	KNORR-CETINA K	13	THOMAS WI	10	ABRUZZESE A
21	DI NICOLA P	13	FORSé M	10	FASANELLA A
20	BUCCHI M	13	EISENSTADT SN	10	MILLER P
20	BUTLER J	13	COHEN J	10	COSER L
20	LATOUCHE S	13	ZHOU M	10	BERG M
20	PALTRINIERI R	13	NEE V	10	FAUST K
20	LEWIS J	13	WARDE A	10	BOWLES S
20	TRONCA L	13	RAWLS J	10	HERVIEU-LéGER D
20	BASCH L	13	BRUNER J	10	MINARDI E
20	SEMI G	13	BRUNI A	10	KELSEN H
20	BALLARINO G	12	BURGESS E		
20	COBALTI A	12	WILLIAMS R		

It must be considered that the frequencies were calculated after the removal of duplicates (see Footnote 134)

APPENDIX F

Table 9: Clusters information (number of cluster, topics of reference, scholars, and scholars' affiliation) related to the international map.

Cluster	Surname	Name	Affiliation[163]
Sociology of cultural processes (14.29); sociology of communication (14.26); sociology of consumption (13.12); sociology of religion (4.14); gender studies (2.76)			
1	ABRUZZESE	Alberto	IULM
1	ALBERONI	Francesco	LUSPIO
1	BOCCIA ARTIERI	Giovanni	URBINO
1	BOVONE	Laura	CATTOLICA
1	CESAREO	Vincenzo	CATTOLICA
1	CODELUPPI	Vanni	UNIMORE
1	DI NALLO	Egeria	BOLOGNA (Scienze Politiche)
1	FERRARI	Vincenzo	MILANO
1	FERRAROTTI	Franco	ROMA-LA SAPIENZA
1	GARELLI	Franco	TORINO
1	LECCARDI	Carmen	BICOCCA
1	MORA	Emanuela	CATTOLICA
1	MORCELLINI	Mario	ROMA- LA SAPIENZA
1	PALTRINIERI	Roberta	BOLOGNA (Scienze Politiche)
1	PARMIGGIANI	Paola	BOLOGNA (Scienze Politiche)
1	SANTORO	Marco	BOLOGNA (Scienze Statistiche)
1	SASSATELLI	Roberta	MILANO
1	VACCARINI	Italo	CATTOLICA
Economic sociology (12.22); local governance (9.51); public politics (9.51); sociology of organisation (7.33); sociology of work (6)			
2	BAGNASCO	Arnaldo	TORINO
2	BARBERA	Filippo	TORINO
2	BATTISTELLI	Fabrizio	ROMA- LA SAPIENZA
2	BERTIN	Giovanni	VENEZIA (CA' FOSCARI)
2	BIFULCO	Lavinia	BICOCCA
2	BONAZZI	Giuseppe	TORINO
2	BURRONI	Luigi	TERAMO
2	BUTERA	Federico	BICOCCA
2	CELLA	Gian Primo	MILANO
2	CERASE	Francesco Paolo	NAPOLI - FEDERICO II
2	CERI	Paolo	FIRENZE

[163] For each Author it is indicated the last affiliation. For the method used for establishing the institutional affiliation of the authors see page 156.

2	DE LEONARDIS	Ota	BICOCCA
2	DONOLO	Carlo	ROMA- LA SAPIENZA
2	FAZZI	Luca	TRENTO
2	GUALA	Alessandro	TORINO
2	LA ROSA	Michele	BOLOGNA (Scienze Politiche)
2	MINARDI	Everardo	TERAMO
2	MUTTI	Antonio	PAVIA
2	NEGRELLI	Serafino	BICOCCA
2	PALUMBO	Mauro	GENOVA
2	PERULLI	Paolo	PIEMONTE ORIENTALE
2	PICHIERRI	Angelo	TORINO
2	RAMELLA	Francesco	URBINO
2	ROSSI	Paolo	BICOCCA
2	TRIGILIA	Carlo	FIRENZE

Methodology (25.36); sociology of health (22.49); sociology of cultural processes (9.9); sociology of knowledge (7.92); social policies (6.89)

3	AGNOLI	Maria Stella	ROMA- LA SAPIENZA
3	ALTIERI	Leonardo	BOLOGNA (Scienze Politiche)
3	ARDIGò	Achille	BOLOGNA (Scienze Politiche)
3	BICHI	Rita	CATTOLICA
3	BRUNI	Attila	TRENTO
3	BRUSCHI	Alessandro	FIRENZE
3	BUCCHI	Massimiano	TRENTO
3	CAMPELLI	Enzo	ROMA- LA SAPIENZA
3	CANNAVò	Leonardo	ROMA- LA SAPIENZA
3	CAPECCHI	Vittorio	BOLOGNA (Scienze della Formazione)
3	CARDANO	Mario	TORINO
3	CIPOLLA	Costantino	BOLOGNA (Scienze Politiche II)
3	CIPRIANI	Roberto	ROMA3
3	CRESPI	Franco	PERUGIA
3	FASANELLA	Antonio	ROMA- LA SAPIENZA
3	FELE	Giolo	TRENTO
3	GHERARDI	Silvia	TRENTO
3	GIARELLI	Guido	CATANZARO-MAGNA GRECIA
3	GIGLIOLI	Pier Paolo	BOLOGNA (Scienze della comunicazione)
3	GOBO	Giampiero	MILANO
3	INGROSSO	Marco	FERRARA
3	MANGHI	Sergio	PARMA
3	MARRADI	Alberto	FIRENZE
3	MARZANO	Marco	BERGAMO
3	MATURO	Antonio Francesco	BOLOGNA (Scienze Politiche II)

3	MORUZZI	Mauro	BOLOGNA (Scienze Politiche)
3	NERESINI	Federico	PADOVA
3	RICOLFI	Luca	TORINO
3	ROSITI	Franco	PAVIA
3	STATERA	Gianni	ROMA- LA SAPIENZA
3	VICARELLI	Giovanna	MARCHE POLITECNICO

Social mobility (31.70); sociology of work (15); welfare (14.67); economic sociology (13.45); social policies (11.03); inequalities (8.27)

4	ACCORNERO	Aris	ROMA- LA SAPIENZA
4	BALBO	Laura	PADOVA
4	BALLARINO	Gabriele	MILANO
4	BARBAGLI	Marzio	BOLOGNA (Scienze Statistiche)
4	BARBIERI	Paolo	TRENTO
4	BIANCO	Maria Luisa	PIEMONTE ORIENTALE
4	BORGHI	Vando	BOLOGNA (Scienze Politiche)
4	COBALTI	Antonio	TRENTO
4	CORBETTA	Piergiorgio	BOLOGNA (Scienze della Formazione)
4	DE LILLO	Antonio	BICOCCA
4	FACCHINI	Carla	BICOCCA
4	FULLIN	Giovanna	BICOCCA
4	GALLINO	Luciano	TORINO
4	GAMBARDELLA	Dora	NAPOLI-FEDERICO II
4	GASPERONI	Giancarlo	BOLOGNA (Lettere e Filosofia)
4	LODIGIANI	Rosangela	CATTOLICA
4	MAGATTI	Mauro	CATTOLICA
4	MINGIONE	Enzo	BICOCCA
4	MOSCATI	Roberto	BICOCCA
4	NALDINI	Manuela	TORINO
4	NEGRI	Nicola	TORINO
4	OLAGNERO	Manuela	TORINO
4	PACI	Massimo	ROMA- LA SAPIENZA
4	PAVOLINI	Emmanuele	MACERATA
4	PICCONE STELLA	Simonetta	ROMA- LA SAPIENZA
4	PISATI	Maurizio	BICOCCA
4	POGGIO	Barbara	TRENTO
4	RANCI	Costanzo	MILANO POLITECNICO
4	REGINI	Marino	MILANO
4	REYNERI	Emilio	BICOCCA
4	RIZZA	Roberto	BOLOGNA (Scienze Politiche II)
4	RUSPINI	Elisabetta	BICOCCA
4	SARACENO	Chiara	TORINO

4	SCHADEE	Hans	BICOCCA
4	SCHERER	Stefani	TRENTO
4	SCHIZZEROTTO	Antonio	TRENTO
4	SEMENZA	Renata	MILANO

Migration (28.95); sociology of cultural processes (11); deviance (7.92); multiculturalism (6.89); ethnic studies (6.34)

5	ALLIEVI	Stefano	PADOVA
5	AMBROSINI	Maurizio	MILANO
5	BESOZZI	Elena	CATTOLICA
5	BOSISIO	Roberta	NA[164]
5	CAPONIO	Tiziana	TORINO
5	COLOMBO	Asher	BOLOGNA (Scienze della Formazione)
5	COLOMBO	Enzo	MILANO
5	COTESTA	Vittorio	ROMA3
5	DAL LAGO	Alessandro	GENOVA
5	DECIMO	Francesca	TRENTO
5	FRISINA	Annalisa	PADOVA
5	LAGOMARSINO	Francesca	GENOVA
5	LEONINI	Luisa	MILANO
5	MACIOTI	Maria Immacolata	ROMA- LA SAPIENZA
5	MELOSSI	Dario	BOLOGNA (Giurisprudenza)
5	MELUCCI	Alberto	MILANO
5	PACE	Enzo	PADOVA
5	PALIDDA	Salvatore	GENOVA
5	PUGLIESE	Enrico	ROMA- LA SAPIENZA
5	QUEIROLO PALMAS	Luca	GENOVA
5	REBUGHINI	Paola	MILANO
5	SCIDà	Giuseppe	BOLOGNA (Scienze Politiche II)
5	SCIORTINO	Giuseppe	TRENTO
5	SEMI	Giovanni	TORINO
5	TOGNETTI BORDOGNA	Mara	BICOCCA
5	ZANFRINI	Laura	CATTOLICA
5	ZINCONE	Giovanna	TORINO
5	ZUCCHETTI	Eugenio	CATTOLICA

[164] In this case the affiliation cannot be established as the author is a collaborator.

Social networks (9.37); social capital (8.27); civicness (6.97); third sector (6.34); social policies (6.34)

6	BOCCACIN	Lucia	CATTOLICA
6	BUZZI	Carlo	TRENTO
6	CASELLI	Marco	CATTOLICA
6	CAVALLI	Alessandro	PAVIA
6	CHIESI	Antonio Maria	MILANO
6	COLOZZI	Ivo	BOLOGNA (Scienze Politiche)
6	DI NICOLA	Patrizio	ROMA- LA SAPIENZA
6	DONATI	Pierpaolo	BOLOGNA (Scienze Politiche)
6	OSTI	Giorgio	TRIESTE
6	PELLIZZONI	Luigi	TRIESTE
6	PISELLI	Fortunata	NAPOLI FEDERICO II
6	PRANDINI	Riccardo	BOLOGNA (Scienze Politiche)
6	ROSSI	Giovanna	CATTOLICA
6	SCIOLLA	Loredana	TORINO
6	TRONCA	Luigi	VERONA

Urban sociology (8.56); sociology of the environment (5.62); ethnic studies (3.17); social integration (3.17); sociology of international relations (3.17)

7	AMENDOLA	Giandomenico	FIRENZE
7	BORLINI	Barbara	NA[165]
7	DAVICO	Luca	TORINO POLITECNICO
7	GASPARINI	Alberto	TRIESTE
7	GUIDICINI	Paolo	BOLOGNA (Scienze Politiche)
7	MARTINOTTI	Guido	BICOCCA
7	MELA	Alfredo	TORINO POLITECNICO
7	NUVOLATI	Giampaolo	BICOCCA
7	TOSI	Antonio	MILANO POLITECNICO
7	GASPARINI	Giovanni	CATTOLICA
7	ZAJCZYK	Francesca	BICOCCA

Deviance (6.34); political sociology (4.63); social capital (3.17); social movements (3.17); public policies (2.44)

8	CATANZARO	Raimondo	BOLOGNA (Lettere e Filosofia)
8	DELLA PORTA	Donatella	FIRENZE
8	DIANI	Mario	TRENTO
8	PIZZORNO	Alessandro	IUE

[165] See Footnote 164.

APPENDIX G

Table 10: Clusters information (number of cluster, topics of reference, scholars, and scholars' affiliation) related to the national map.

Cluster	Surname	Name	Affiliation[166]
Urban sociology (19.02); local governance (13.93); sociology of the environment (12.68); local development (11.25); public policies (9.29)			
1	AMENDOLA	Giandomenico	FIRENZE
1	BAGNASCO	Arnaldo	TORINO
1	BARBERA	Filippo	TORINO
1	BORLINI	Barbara	NA[167]
1	BURRONI	Luigi	TERAMO
1	CATANZARO	Raimondo	BOLOGNA (Lettere e Filosofia)
1	CERASE	Francesco Paolo	NAPOLI - FEDERICO II
1	DAVICO	Luca	TORINO POLITECNICO
1	DE LEONARDIS	Ota	BICOCCA
1	DELLA PORTA	Donatella	FIRENZE
1	DONOLO	Carlo	ROMA- LA SAPIENZA
1	GASPARINI	Alberto	TRIESTE
1	GASPARINI	Giovanni	CATTOLICA
1	GUALA	Alessandro	TORINO
1	GUIDICINI	Paolo	BOLOGNA (Scienze Politiche)
1	MARTINOTTI	Guido	BICOCCA
1	MELA	Alfredo	TORINO POLITECNICO
1	MUTTI	Antonio	PAVIA
1	NEGRI	Nicola	TORINO
1	NUVOLATI	Giampaolo	BICOCCA
1	OLAGNERO	Manuela	TORINO
1	OSTI	Giorgio	TRIESTE
1	PELLIZZONI	Luigi	TRIESTE
1	PERULLI	Paolo	PIEMONTE ORIENTALE
1	PICHIERRI	Angelo	TORINO
1	PISELLI	Fortunata	NAPOLI FEDERICO II
1	PIZZORNO	Alessandro	IUE
1	RAMELLA	Francesco	URBINO
1	TOSI	Antonio	MILANO POLITECNICO
1	TRIGILIA	Carlo	FIRENZE
1	ZAJCZYK	Francesca	BICOCCA

[166] For each author is indicated the last affiliation.
[167] See Footnote 164.

Economic sociology (18.34); sociology of work (18.00); social policies (8.56); labour policies (6.97)

2	ACCORNERO	Aris	ROMA- LA SAPIENZA
2	BALBO	Laura	PADOVA
2	BARBIERI	Paolo	TRENTO
2	BIANCO	Maria Luisa	PIEMONTE ORIENTALE
2	BIFULCO	Lavinia	BICOCCA
2	BONAZZI	Giuseppe	TORINO
2	BORGHI	Vando	BOLOGNA (Scienze Politiche)
2	BUTERA	Federico	BICOCCA
2	CELLA	Gian Primo	MILANO
2	FULLIN	Giovanna	BICOCCA
2	GALLINO	Luciano	TORINO
2	GHERARDI	Silvia	TRENTO
2	LA ROSA	Michele	BOLOGNA (Scienze Politiche)
2	LODIGIANI	Rosangela	CATTOLICA
2	MAGATTI	Mauro	CATTOLICA
2	MINARDI	Everardo	TERAMO
2	MINGIONE	Enzo	BICOCCA
2	NALDINI	Manuela	TORINO
2	NEGRELLI	Serafino	BICOCCA
2	PACI	Massimo	ROMA- LA SAPIENZA
2	PAVOLINI	Emmanuele	MACERATA
2	PICCONE STELLA	Simonetta	ROMA- LA SAPIENZA
2	POGGIO	Barbara	TRENTO
2	RANCI	Costanzo	MILANO POLITECNICO
2	REGINI	Marino	MILANO
2	REYNERI	Emilio	BICOCCA
2	RIZZA	Roberto	BOLOGNA (Scienze Politiche II)
2	SARACENO	Chiara	TORINO
2	SCHERER	Stefani	TRENTO
2	SEMENZA	Renata	MILANO

Methodology (28.53); sociology of cultural processes (9.78); sociology of knowledge (7.92); migration (6.34); social theory (6.11)

3	AGNOLI	Maria Stella	ROMA- LA SAPIENZA
3	BATTISTELLI	Fabrizio	ROMA- LA SAPIENZA
3	BICHI	Rita	CATTOLICA
3	BRUSCHI	Alessandro	FIRENZE
3	BUZZI	Carlo	TRENTO
3	CAMPELLI	Enzo	ROMA- LA SAPIENZA
3	CANNAVò	Leonardo	ROMA- LA SAPIENZA

3	CAPECCHI	Vittorio	BOLOGNA (Scienze della Formazione)
3	CARDANO	Mario	TORINO
3	CASELLI	Marco	CATTOLICA
3	CAVALLI	Alessandro	PAVIA
3	CERI	Paolo	FIRENZE
3	CIPRIANI	Roberto	ROMA3
3	CORBETTA	Piergiorgio	BOLOGNA (Scienze della Formazione)
3	CRESPI	Franco	PERUGIA
3	DE LILLO	Antonio	BICOCCA
3	FACCHINI	Carla	BICOCCA
3	FASANELLA	Antonio	ROMA- LA SAPIENZA
3	FERRAROTTI	Franco	ROMA- LA SAPIENZA
3	GARELLI	Franco	TORINO
3	GOBO	Giampiero	MILANO
3	MARRADI	Alberto	FIRENZE
3	MORCELLINI	Mario	ROMA- LA SAPIENZA
3	RICOLFI	Luca	TORINO
3	ROSITI	Franco	PAVIA
3	ROSSI	Paolo	BICOCCA
3	SCIOLLA	Loredana	TORINO
3	STATERA	Gianni	ROMA- LA SAPIENZA
Migration (33.28); sociology of cultural processes (12.22); urban sociology (7.92); ethnic studies (7.5); multiculturalism (6.89)			
4	ALLIEVI	Stefano	PADOVA
4	AMBROSINI	Maurizio	MILANO
4	BESOZZI	Elena	CATTOLICA
4	BOSISIO	Roberta	NA[168]
4	CAPONIO	Tiziana	TORINO
4	COLOMBO	Asher	BOLOGNA (Scienze della Formazione)
4	COLOMBO	Enzo	MILANO
4	COTESTA	Vittorio	ROMA3
4	DAL LAGO	Alessandro	GENOVA
4	DECIMO	Francesca	TRENTO
4	FRISINA	Annalisa	PADOVA
4	LAGOMARSINO	Francesca	GENOVA
4	LEONINI	Luisa	MILANO
4	MACIOTI	Maria Immacolata	ROMA- LA SAPIENZA
4	MELUCCI	Alberto	MILANO

[168] See note 164.

4	PACE	Enzo	PADOVA
4	PALIDDA	Salvatore	GENOVA
4	PUGLIESE	Enrico	ROMA- LA SAPIENZA
4	QUEIROLO PALMAS	Luca	GENOVA
4	REBUGHINI	Paola	MILANO
4	SCIDà	Giuseppe	BOLOGNA (Scienze Politiche II)
4	SCIORTINO	Giuseppe	TRENTO
4	SEMI	Giovanni	TORINO
4	TOGNETTI BORDOGNA	Mara	BICOCCA
4	ZANFRINI	Laura	CATTOLICA
4	ZINCONE	Giovanna	TORINO
4	ZUCCHETTI	Eugenio	CATTOLICA

Sociology of health (22.49); sociology of communication (6.34); sociology of science (6.34); social policies (6.11); sociology of cultural processes (6.89)

5	ALTIERI	Leonardo	BOLOGNA (Scienze Politiche)
5	ARDIGò	Achille	BOLOGNA (Scienze Politiche)
5	BERTIN	Giovanni	VENEZIA (CA' FOSCARI)
5	BRUNI	Attila	TRENTO
5	BUCCHI	Massimiano	TRENTO
5	CIPOLLA	Costantino	BOLOGNA (Scienze Politiche II)
5	FAZZI	Luca	TRENTO
5	FELE	Giolo	TRENTO
5	FERRARI	Vincenzo	MILANO
5	GIARELLI	Guido	CATANZARO-MAGNA GRECIA
5	GIGLIOLI	Pier Paolo	BOLOGNA (Scienze della comunicazione)
5	INGROSSO	Marco	FERRARA
5	MANGHI	Sergio	PARMA
5	MARZANO	Marco	BERGAMO
5	MATURO	Antonio Francesco	BOLOGNA (Scienze Politiche II)
5	MELOSSI	Dario	BOLOGNA (Giurisprudenza)
5	MORUZZI	Mauro	BOLOGNA (Scienze Politiche)
5	NERESINI	Federico	PADOVA
5	PALUMBO	Mauro	GENOVA
5	STAME	Nicoletta	ROMA- LA SAPIENZA
5	VICARELLI	Giovanna	POLITECNICO DELLE MARCHE

Sociology of cultural processes (14.67); sociology of consumption (13.12); sociology of communication (12.68); gender studies (3.17); bibliometrics (3.17)

6	ABRUZZESE	Alberto	IULM
6	ALBERONI	Francesco	LUPSIO
6	BOCCIA ARTIERI	Giovanni	URBINO
6	BOVONE	Laura	CATTOLICA
6	CESAREO	Vincenzo	CATTOLICA
6	CODELUPPI	Vanni	UNIMORE
6	DI NALLO	Egeria	BOLOGNA (Scienze Politiche)
6	DIANI	Mario	TRENTO
6	LECCARDI	Carmen	BICOCCA
6	MORA	Emanuela	CATTOLICA
6	PALTRINIERI	Roberta	BOLOGNA (Scienze Politiche)
6	PARMIGGIANI	Paola	BOLOGNA (Scienze Politiche)
6	SANTORO	Marco	BOLOGNA (Scienze Statistiche)
6	SASSATELLI	Roberta	MILANO
6	VACCARINI	Italo	CATTOLICA

Social mobility (7.92); sociology of education (5); political sociology (4.14); inequalities (3.67); welfare (3.30)

7	BALLARINO	Gabriele	MILANO
7	BARBAGLI	Marzio	BOLOGNA (Scienze Statistiche)
7	COBALTI	Antonio	TRENTO
7	GAMBARDELLA	Dora	NAPOLI-FEDERICO II
7	GASPERONI	Giancarlo	BOLOGNA (Lettere e Filosofia)
7	MOSCATI	Roberto	BICOCCA
7	PISATI	Maurizio	BICOCCA
7	RUSPINI	Elisabetta	BICOCCA
7	SCHADEE	Hans	BICOCCA
7	SCHIZZEROTTO	Antonio	TRENTO

Social capital (6.89); third sector (6.34); social network (4.14); social policies (3.67); welfare (3.30)

8	BOCCACIN	Lucia	CATTOLICA
8	CHIESI	Antonio Maria	MILANO
8	COLOZZI	Ivo	BOLOGNA (Scienze Politiche)
8	DI NICOLA	Patrizio	ROMA- LA SAPIENZA
8	DONATI	Pierpaolo	BOLOGNA (Scienze Politiche)
8	PRANDINI	Riccardo	BOLOGNA (Scienze Politiche)
8	ROSSI	Giovanna	CATTOLICA
8	TRONCA	Luigi	VERONA

APPENDIX H

Table 11: Occurrence frequencies of each University in each cluster with reference to the international map (map1).

UNIVERSITY	**CL 1**	**CL 2**	**CL 3**	**CL 4**	**CL 5**	**CL 6**	**CL 7**	**CL 8**
BERGAMO	0	1	0	0	0	0	0	0
BICOCCA	1	0	5	8	1	0	3	0
BOLOGNA	1	2	0	3	2	0	0	1
BOLOGNA SPO	3	3	1	1	0	3	1	0
BOLOGNA SPO II	0	2	0	1	1	0	0	0
CATANZARO	0	1	0	0	0	0	0	0
CATTOLICA	4	1	0	3	3	3	1	0
FERRARA	0	1	0	0	0	0	0	0
FIRENZE	0	2	2	0	0	0	1	1
GENOVA	0	0	1	0	4	0	0	0
IUE	0	0	0	0	0	0	0	1
IULM	1	0	0	0	0	0	0	0
LUSPIO	1	0	0	0	0	0	0	0
MACERATA	0	0	0	1	0	0	0	0
MILANO	2	1	1	3	5	1	0	0
NAPOLI-FEDERICO II	0	0	1	1	0	1	0	0
PADOVA	0	1	0	1	3	0	0	0
PARMA	0	1	0	0	0	0	0	0
PAVIA	0	1	1	0	0	1	0	0
PIEMONTE ORIENTALE	0	0	1	1	0	0	0	0
POLITECNICO MARCHE	0	1	0	0	0	0	0	0
POLITECNICO MILANO	0	0	0	1	0	0	1	0
POLITECNICO TORINO	0	0	0	0	0	0	2	0
PERUGIA	0	1	0	0	0	0	0	0
ROMA3	0	1	0	0	1	0	0	0
ROMA-LA SAPIENZA	2	5	2	3	2	1	0	0
TERAMO	0	0	2	0	0	0	0	0
TORINO	1	2	5	5	3	1	0	0
TRENTO	0	4	1	5	2	1	0	1
TRIESTE	0	0	0	0	0	2	1	0
UNIMORE (Modena e Reggio Emilia)	1	0	0	0	0	0	0	0
URBINO	1	0	1	0	0	0	0	1
VENEZIA (CA' FOSCARI)	0	0	1	0	0	0	0	0
VERONA	0	0	0	0	0	1	0	0

APPENDIX I

Table 12: Occurrence frequencies of each University in each cluster with reference to the national map (map2).

UNIVERSITY	CL 1	CL 2	CL 3	CL 4	CL 5	CL 6	CL 7	CL 8
BERGAMO	0	0	0	0	1	0	0	0
BICOCCA	4	3	5	1	0	1	4	0
BOLOGNA	1	2	0	1	2	1	2	0
BOLOGNA SPO	1	0	2	0	3	3	0	3
BOLOGNA SPO II	0	0	1	1	2	0	0	0
CATANZARO	0	0	0	0	1	0	0	0
CATTOLICA	1	3	3	3	0	4	0	2
FERRARA	0	0	0	0	1	0	0	0
FIRENZE	2	3	0	0	0	0	0	0
GENOVA	0	0	0	4	1	0	0	0
IUE	1	0	0	0	0	0	0	0
IULM	0	0	0	0	0	1	0	0
LUSPIO	0	0	0	0	0	1	0	0
MACERATA	0	0	1	0	0	0	0	0
MILANO	0	1	3	5	1	1	1	1
NAPOLI-FEDERICO II	2	0	0	0	0	0	1	0
PADOVA	0	0	1	3	1	0	0	0
PARMA	0	0	0	0	1	0	0	0
PAVIA	1	2	0	0	0	0	0	0
PIEMONTE ORIENTALE	1	0	1	0	0	0	0	0
POLITECNICO MARCHE	0	0	0	0	1	0	0	0
POLITECNICO MILANO	1	0	1	0	0	0	0	0
POLITECNICO TORINO	2	0	0	0	0	0	0	0
PERUGIA	0	1	0	0	0	0	0	0
ROMA3	0	1	0	1	0	0	0	0
ROMA-LA SAPIENZA	1	8	3	2	1	0	0	1
TERAMO	1	0	1	0	0	0	0	0
TORINO	5	4	3	3	0	0	0	0
TRENTO	0	1	4	2	4	1	2	0
TRIESTE	3	0	0	0	0	0	0	0
UNIMORE (Modena e Reggio Emilia)	0	0	0	0	0	1	0	0
URBINO	1	0	0	0	0	2	0	0
VENEZIA (CA' FOSCARI)	0	0	0	0	1	0	0	0
VERONA	0	0	0	0	0	0	0	1

APPENDIX J

Table 13: Degree centrality of the authors included in map1 (international map).

DEGREE CENTRALITY		DEGREE CENTRALITY		DEGREE CENTRALITY	
BOURDIEU P	1,979	GARFINKEL H	268	WIEVIORKA M	159
GIDDENS A	1,807	FORSé M	267	CICOUREL A	159
BAUMAN Z	1,797	ABBOTT A	267	FRISINA A	159
BECK U	1,593	NALDINI M	265	AGAMBEN G	158
BAGNASCO A	1.301	WENGER E	265	WEISS CH	158
WEBER M	1,226	BECKER G	264	SWIDLER A	157
DONATI P	1,212	PALTRINIERI R	263	STAR SL	157
AMBROSINI M	1,121	MATURO A	263	BRUSCHI A	157
TRIGILIA C	1,099	ERIKSON R	263	CAMPBELL C	156
GOFFMAN E	1,086	PARK R	260	POGGIO B	156
COLEMAN J	1,051	NERESINI F	260	PFEFFER J	155
GRANOVETTER M	1,028	RUMBAUT R	258	PACE E	155
CASTELLS M	964	ETZIONI A	256	CROMPTON R	155
SEN A	960	ISHERWOOD B	254	GAMBARDELLA D	155
GALLINO L	908	VOELZKOW H	254	MARCUS G	155
PIZZORNO A	899	ARENDT H	252	ALTIERI L	154
MERTON RK	887	CLIFFORD J	251	SCHON D	154
SIMMEL G	876	AMIN A	250	OGNETTI BORDOGNA M	154
FOUCAULT M	860	CARDANO M	250	CANNAVò L	154
BOUDON R	840	ILLICH I	249	O'CONNOR J	153
SENNETT R	836	EISENSTADT SN	249	HALL P	153
PORTES A	835	GASPERONI G	249	PAVOLINI E	153
APPADURAI A	823	ZUKIN S	246	FELE G	152
ESPING-ANDERSEN G	802	TREIMAN D	245	ANDALL J	151
SARACENO C	799	BERGER PL	245	SUCHMAN L	150
PUTNAM R	798	MARX K	245	BOYER R	150
GOLDTHORPE J	788	FAIST T	244	PORTER M	149
SCHIZZEROTTO A	785	NUVOLATI G	243	BERTIN G	149
MAGATTI M	780	LéVI-STRAUSS C	242	THOMAS WI	148
HABERMAS J	775	FREEMAN H	242	BIJKER WE	148
DURKHEIM E	752	SUPIOT A	241	FAUST K	148
PACI M	710	BATESON G	239	RAUTY R	148
LUHMANN N	702	DUNCAN O	238	MAYNTZ R	147
POWELL WW	693	SOSKICE D	236	DAHL R	147
MELUCCI A	674	PERULLI P	236	ROSE N	147
PARSONS T	674	MORA E	235	BOWLES S	146
DE LILLO A	655	MULLER W	235	MEAD GH	145
CAVALLI A	645	BESOZZI E	235	CATANZARO R	145
CROUCH C	626	RUSPINI E	235	POLANYI M	145

CIPOLLA C	617	CASTLES S	233	PALIDDA S	144
LATOUR B	617	DORE R	232	LIPSEY MW	144
REYNERI E	614	BALIBAR E	231	KANT I	144
BARBAGLI M	608	HALLER W	230	ROBERTSON R	143
TOURAINE A	598	FULLIN G	228	TARROW S	143
DAL LAGO A	595	WEICK K	228	WILSON WJ	143
DI MAGGIO PJ	585	LAW J	228	BURY M	143
ZANFRINI L	583	BUTERA F	227	BOUTANG YM	143
TAYLOR C	579	CALLON M	226	SEMENZA R	143
COLOMBO E	577	AMENDOLA G	224	ERIKSON EH	142
RANCI C	560	COHEN S	224	MALINOWSKI B	142
MARRADI A	558	BAUDRILLARD J	223	MILLER MJ	141
CESAREO V	557	DI NALLO E	222	SCIDà G	140
COLLINS R	549	SCHERER S	222	CZARNIAWSKA B	140
MARCH JG	538	PAWSON R	221	PARRENAS R	140
HIRSCHMAN AO	538	BERGER P	221	BELLAH RN	140
MUTTI A	535	RIZZA R	221	OLAGNERO M	140
GHERARDI S	527	LIPPI A	220	LOCKWOOD D	139
DOUGLAS M	525	HELD D	220	FINE GA	138
POLANYI K	522	MILLER D	218	BURAWOY M	137
CORBETTA P	520	BOSISIO R	218	LAVILLE JL	136
SCIOLLA L	492	RICOEUR P	218	INGLEHART R	136
MINGIONE E	491	JACOBS J	217	GUBA EG	134
GLICK SCHILLER N	490	ROSSI G	217	LINCOLN YS	134
ANDERSON B	485	BRUNI A	215	MANGHI S	134
LAZARSFELD P	482	VEBLEN T	214	JENKINS H	133
REGINI M	480	SCOTT J	213	GADAMER HG	132
SASSEN S	478	CERI P	213	FASANELLA A	132
GEERTZ C	470	SWEDBERG R	212	HUNTINGTON SP	132
DE CERTEAU M	469	ROSSI PH	212	DECIMO F	132
COLOMBO A	469	FAZZI L	211	ONG A	132
FERRERA M	459	LEWIS J	210	HEBDIGE D	131
CHIESI AM	455	MAFFESOLI M	210	BURRONI L	131
BARBIERI P	454	WITTGENSTEIN L	209	SAID E	130
PISATI M	452	RICOLFI L	209	DAVICO L	129
PISELLI F	451	WALDINGER R	208	JOPPKE C	127
CASTEL R	445	BALBO L	207	NEGRELLI S	126
HARVEY D	433	LATOUCHE S	207	LIPSET SM	125
PIORE M	433	SAVAGE M	206	HORKHEIMER M	124
LANDOLT P	432	MEYER JW	205	MACIOTI MI	124
CRESPI F	431	GOBO G	205	TIMMERMANS S	123
PICHIERRI A	429	DUBET F	205	HUGHES E	122
DAHRENDORF R	422	FLORIDA R	204	TURNER B	121

ELSTER J	420	LECCARDI C	204	MAIR P	121
SIMON H	419	OSTROM E	203	KATZ E	121
LASH S	418	WILLIAMSON OE	202	RHODES R	121
COLOZZI I	416	BRUBAKER R	201	BROWN P	120
LE GALèS P	416	FREIDSON E	201	ROSITI F	120
VERTOVEC S	414	FRIEDBERG E	200	VICARELLI G	119
BOLTANSKI L	412	ZAJCZYK F	199	LEFEBVRE H	118
BURT R	396	DEWEY J	199	CIPRIANI R	117
ARDIGò A	396	TILLY C	199	GALTUNG J	117
SASSATELLI R	391	MANNHEIM K	198	BOCCACIN L	116
ROSSI P	390	LAVE J	198	CRANE D	116
ARCHER M	390	JONSSON JO	198	CASELLI M	115
LEONINI L	388	LODIGIANI R	197	BERG M	115
RAMELLA F	388	ADORNO T	197	ROWAN B	115
BORGHI V	387	LIPOVETSKY G	197	THOMPSON JB	114
GUARNIZO L	385	ALEXANDER JC	196	GUALA C	114
SCHUTZ A	383	STINCHCOMBE A	196	FLIGSTEIN N	114
CELLA G	382	MILLS CW	195	BATTISTELLI F	113
MARTINOTTI G	380	TODOROV T	194	GERSHUNY J	113
STREECK W	377	GARELLI F	194	KRUGMAN P	112
SCIORTINO G	376	NEE V	193	BRUNER J	112
NEGRI N	375	RITZER G	192	COSER L	112
COBALTI A	375	MARZANO M	192	NONAKA I	111
URRY J	373	LASCH C	191	HOBSBAWM EJ	111
PUGLIESE E	372	FEATHERSTONE M	191	LAYTE R	111
KUHN T	369	BELL D	191	HALBWACHS M	110
BECKER H	364	THéVENOT L	190	CONRAD P	110
LEVITT P	363	ARGYRIS C	190	SONTAG S	110
HANNERZ U	362	ALBERONI F	189	NUSSBAUM M	110
DI NICOLA P	361	POWER M	188	SCHMITT C	109
BIANCO ML	359	ZHOU M	188	BERTAUX D	108
ACCORNERO A	357	GANZEBOOM HBG	187	FRIEDMAN M	108
BARBERA F	353	STATERA G	187	GAMBETTA D	108
BASCH L	352	WARDE A	186	TEUBNER G	106
GIARELLI G	351	BICHI R	186	SILVERMAN D	106
LIN N	350	BUCCHI M	185	WOLLEB G	106
HALL S	348	GUIDICINI P	185	INGROSSO M	103
JEDLOWSKI P	347	BENJAMIN W	183	KOOIMAN J	103
BALLARINO G	342	LEWIN K	182	HARDT M	100
SEMI G	337	SALAIS R	182	OLSON M	100
FUKUYAMA F	334	WALLERSTEIN I	182	MILLER P	100
PALUMBO M	333	SCHUMPETER J	182	RAWLS J	99
SABEL C	329	SCHON DA	182	CERASE FP	99

BONAZZI G	325	PICCONE STELLA S	181	BOCCIA ARTIERI G	99
SZANTON BLANC C	324	SMELSER NJ	181	COTESTA V	99
REBUGHINI P	323	SMITH A	181	MOSCOVICI S	98
DE LEONARDIS O	322	HOCHSCHILD AR	180	MOSCATI R	97
BUZZI C	322	MELOSSI D	179	MINARDI E	97
QUEIROLO PALMAS L	321	BISON I	179	SKOCPOL T	97
BREEN R	318	TOSI A	179	DERRIDA J	96
CAPONIO T	318	SCRIVEN M	178	FOLGHERAITER F	96
BLOSSFELD HP	317	BAUMANN G	178	McLUHAN M	93
STAME N	316	NORRIS P	177	GUATTARI F	93
MORIN E	315	SELZNICK P	176	LISTER R	93
LUCKMANN T	315	WOOLGAR S	176	DONZELOT J	92
MASSEY D	314	AXELROD R	175	BARTHES R	90
ELIAS N	314	HEDSTROM P	175	ANDERSON C	90
PASSERON JC	314	GINTIS H	175	WILLIAMS R	89
SCHADEE H	314	LAGOMARSINO F	174	COHEN J	89
BLAU P	309	SANTORO M	174	LYNCH K	88
OLSEN JP	306	BUTLER J	173	GILROY P	87
CODELUPPI V	305	POPPER KR	173	SHORE C	86
TRONCA L	304	HOUT M	172	AGNOLI MS	85
CROZIER M	304	PATTON MQ	172	FREUD S	85
PRANDINI R	303	GIGLIOLI PP	172	BRAUDEL F	85
AUGé M	301	BLUMER H	171	GASPARINI G	84
STRAUSS A	300	RIFKIN J	171	MEYROWITZ J	84
SHAVIT Y	296	BURGESS E	171	HERVIEU-LéGER D	83
MARSHALL TH	292	WALZER M	171	SACKETT DL	80
PELLIZZONI L	291	NICOLINI D	170	MINTZBERG H	79
CAPECCHI V	291	DELEUZE G	169	GASPARINI A	78
BOVONE L	290	BORGATTI SP	169	ABRUZZESE A	78
WELLMAN B	290	HONNETH A	169	JONAS H	77
CAMPELLI E	289	PIERSON P	169	MORCELLINI M	77
BIFULCO L	289	DIANI M	168	ALLIEVI S	77
KNORR-CETINA K	289	BORLINI B	167	SPIVAK GC	73
MAUSS M	288	HUSSERL E	166	TURNER V	71
MELA A	288	THRIFT N	164	FREEMAN RE	71
BERNARDI F	286	PARMIGGIANI P	163	DENZIN N	70
DELLA PORTA D	286	SILVERSTONE R	163	FACCHINI C	70
DONOLO C	286	WEICK KE	163	MORUZZI M	69
WACQUANT L	282	ZINCONE G	163	DE SOUSA SANTOS B	61
FERRAROTTI F	282	ROSANVALLON P	162	PITCH T	55
ZUCCHETTI E	280	WASSERMAN S	162	ZAGREBELSKY G	51
OFFE C	278	PARETO V	162	FERRARI V	51
VITALE T	277	STORPER M	161	DE KERCKHOVE D	48

LA ROSA M	272	GOOD B	160	KELSEN H	37
VACCARINI I	271	OSTI G	160		
SAYAD A	269	STIGLITZ J	160		

APPENDIX K

Table 14: Degree centrality of the authors included in map2 (national map).

DEGREE CENTRALITY		DEGREE CENTRALITY		DEGREE CENTRALITY	
BAGNASCO A	593	LA ROSA M	145	POGGIO B	79
DONATI P	531	REBUGHINI P	143	DI NALLO E	77
AMBROSINI M	527	DONOLO C	141	GARELLI F	77
TRIGILIA C	499	SCHADEE H	140	SEMENZA R	77
GALLINO L	413	ROSSI P	137	BERTIN G	76
SCHIZZEROTTO A	372	PRANDINI R	136	FRISINA A	74
PACI M	351	DE LEONARDIS O	135	ALTIERI L	73
DE LILLO A	350	VACCARINI I	135	BURRONI L	72
MAGATTI M	347	NALDINI M	135	PARMIGGIANI P	71
CAVALLI A	318	SEMI G	134	STATERA G	71
PIZZORNO A	317	TRONCA L	134	TOGNETTI BORDOGNA M	71
BARBAGLI M	310	FULLIN G	130	BORLINI B	70
REYNERI E	294	QUEIROLO PALMAS L	129	SANTORO M	70
MARRADI A	286	PELLIZZONI L	129	OSTI G	69
RANCI C	284	BALLARINO G	129	PALIDDA S	69
CESAREO V	279	STAME N	126	MELOSSI D	69
CIPOLLA C	271	SCHERER S	125	BUCCHI M	69
ZANFRINI L	263	MELA A	125	DECIMO F	67
COLOMBO E	247	BIFULCO L	125	PACE E	67
CORBETTA P	246	RICOLFI L	122	BOCCACIN L	66
DAL LAGO A	244	ZAJCZYK F	122	SCIDà G	66
COLOMBO A	242	ROSSI G	121	DAVICO L	66
MELUCCI A	242	CODELUPPI V	120	LAGOMARSINO F	65
FERRERA M	226	SASSATELLI R	119	CANNAVò L	65
PISELLI F	224	GASPERONI G	118	DIANI M	65
MINGIONE E	221	MATURO A	116	OLAGNERO M	64
REGINI M	211	RIZZA R	114	GUALA C	64
COLOZZI I	210	CARDANO M	112	BRUNI A	63
MUTTI A	210	BOVONE L	111	FASANELLA A	61
SCIOLLA L	207	FERRAROTTI F	111	MARZANO M	61
PISATI M	206	DELLA PORTA D	111	CASELLI M	61
CHIESI AM	203	BONAZZI G	110	VICARELLI G	59
RAMELLA F	200	BESOZZI E	110	INGROSSO M	57
BARBIERI P	197	FAZZI L	108	NEGRELLI S	56
PUGLIESE E	192	PICCONE STELLA S	108	CERASE FP	53
BORGHI V	190	BICHI R	108	BRUSCHI A	53
ARDIGò A	183	PALTRINIERI R	105	MANGHI S	52
PICHIERRI A	182	GOBO G	105	MINARDI E	52
ACCORNERO A	182	MORA E	102	BATTISTELLI F	49
SCIORTINO G	179	BALBO L	102	AGNOLI MS	48
BIANCO ML	171	AMENDOLA G	97	MOSCATI R	48
NEGRI N	168	ZINCONE G	97	MORUZZI M	48

BUZZI C	168	LODIGIANI R	97	COTESTA V	47
GHERARDI S	167	PAVOLINI E	96	CIPRIANI R	47
LEONINI L	164	GUIDICINI P	95	MACIOTI MI	43
MARTINOTTI G	163	ALBERONI F	93	GASPARINI A	39
BARBERA F	163	CAMPELLI E	93	ALLIEVI S	38
CRESPI F	161	TOSI A	93	FACCHINI C	38
PALUMBO M	160	BOSISIO R	92	BOCCIA ARTIERI G	33
DI NICOLA P	158	NUVOLATI G	90	FELE G	32
COBALTI A	153	CERI P	90	ROSITI F	30
ZUCCHETTI E	152	LECCARDI C	87	GASPARINI G	30
GIARELLI G	151	BUTERA F	86	GIGLIOLI PP	30
CAPONIO T	148	CATANZARO R	85	ABRUZZESE A	27
RUSPINI E	148	GAMBARDELLA D	85	MORCELLINI M	23
CAPECCHI V	147	PERULLI P	81	FERRARI V	13
CELLA G	146	NERESINI F	79		

APPENDIX L

Table 15: Nodal degree of each author; nodal degree expressed in percentage and relativised within each cluster. Data relate to the international map. Clusters are delimited by lines.

NAME	DEGREE CENTRALITY[169]	% DEGREE CENTRALITY WITHIN CLUSTER	CLUSTER
BOURDIEU P	1,979	6.10	1
GIDDENS A	1,807	5.57	1
BAUMAN Z	1,797	5.54	1
BECK U	1,593	4.91	1
WEBER M	1,226	3.78	1
CASTELLS M	964	2.97	1
SIMMEL G	876	2.70	1
FOUCAULT M	860	2.65	1
APPADURAI A	823	2.54	1
HABERMAS J	775	2.39	1
LUHMANN N	702	2.16	1
TAYLOR C	579	1.78	1
CESAREO V	557	1.72	1
DOUGLAS M	525	1.62	1
DE CERTEAU M	469	1.44	1
LASH S	418	1.29	1
BOLTANSKI L	412	1.27	1
SASSATELLI R	391	1.20	1
BECKER H	364	1.12	1
HALL S	348	1.07	1
JEDLOWSKI P	347	1.07	1
MORIN E	315	0.97	1
ELIAS N	314	0.97	1
CODELUPPI V	305	0.94	1
AUGé M	301	0.93	1
BOVONE L	290	0.89	1
MAUSS M	288	0.89	1
FERRAROTTI F	282	0.87	1
VACCARINI I	271	0.83	1
PALTRINIERI R	263	0.81	1
ISHERWOOD B	254	0.78	1
ARENDT H	252	0.78	1
MARX K	245	0.75	1
LéVI-STRAUSS C	242	0.75	1
MORA E	235	0.72	1

[169] In a weighed graph the degree centrality of a node consists of the sum of the weighs of the links incident with it.

BALIBAR E	231	0.71	1
COHEN S	224	0.69	1
BAUDRILLARD J	223	0.69	1
DI NALLO E	222	0.68	1
MILLER D	218	0.67	1
RICOEUR P	218	0.67	1
VEBLEN T	214	0.66	1
MAFFESOLI M	210	0.65	1
LATOUCHE S	207	0.64	1
LECCARDI C	204	0.63	1
ADORNO T	197	0.61	1
LIPOVETSKY G	197	0.61	1
ALEXANDER JC	196	0.60	1
TODOROV T	194	0.60	1
GARELLI F	194	0.60	1
RITZER G	192	0.59	1
LASCH C	191	0.59	1
FEATHERSTONE M	191	0.59	1
BELL D	191	0.59	1
ALBERONI F	189	0.58	1
WARDE A	186	0.57	1
BENJAMIN W	183	0.56	1
WALLERSTEIN I	182	0.56	1
SANTORO M	174	0.54	1
BUTLER J	173	0.53	1
RIFKIN J	171	0.53	1
DELEUZE G	169	0.52	1
HONNETH A	169	0.52	1
PARMIGGIANI P	163	0.50	1
SILVERSTONE R	163	0.50	1
WIEVIORKA M	159	0.49	1
AGAMBEN G	158	0.49	1
SWIDLER A	157	0.48	1
CAMPBELL C	156	0.48	1
ROSE N	147	0.45	1
KANT I	144	0.44	1
ROBERTSON R	143	0.44	1
BOUTANG YM	143	0.44	1
BELLAH RN	140	0.43	1
FINE GA	138	0.43	1
JENKINS H	133	0.41	1
HEBDIGE D	131	0.40	1
SAID E	130	0.40	1

HORKHEIMER M	124	0.38	1
KATZ E	121	0.37	1
LEFEBVRE H	118	0.36	1
CRANE D	116	0.36	1
THOMPSON JB	114	0.35	1
HOBSBAWM EJ	111	0.34	1
HALBWACHS M	110	0.34	1
SCHMITT C	109	0.34	1
TEUBNER G	106	0.33	1
HARDT M	100	0.31	1
MILLER P	100	0.31	1
RAWLS J	99	0.30	1
BOCCIA ARTIERI G	99	0.30	1
DERRIDA J	96	0.30	1
McLUHAN M	93	0.29	1
GUATTARI F	93	0.29	1
BARTHES R	90	0.28	1
ANDERSON C	90	0.28	1
WILLIAMS R	89	0.27	1
GILROY P	87	0.27	1
FREUD S	85	0.26	1
BRAUDEL F	85	0.26	1
MEYROWITZ J	84	0.26	1
HERVIEU-LéGER D	83	0.26	1
ABRUZZESE A	78	0.24	1
JONAS H	77	0.24	1
MORCELLINI M	77	0.24	1
SPIVAK GC	73	0.22	1
TURNER V	71	0.22	1
DE SOUSA SANTOS B	61	0.19	1
PITCH T	55	0.17	1
ZAGREBELSKY G	51	0.16	1
FERRARI V	51	0.16	1
DE KERCKHOVE D	48	0.15	1
KELSEN H	37	0.11	1
BAGNASCO A	1.301	5.31	2
TRIGILIA C	1.099	4.49	2
GRANOVETTER M	1.028	4.20	2
POWELL WW	693	2.83	2
CROUCH C	626	2.56	2
DI MAGGIO PJ	585	2.39	2
MARCH JG	538	2.20	2
HIRSCHMAN AO	538	2.20	2

MUTTI A	535	2.18	2
POLANYI K	522	2.13	2
PIORE M	433	1.77	2
PICHIERRI A	429	1.75	2
SIMON H	419	1.71	2
LE GALèS P	416	1.70	2
ROSSI P	390	1.59	2
RAMELLA F	388	1.58	2
CELLA G	382	1.56	2
STREECK W	377	1.54	2
BARBERA F	353	1.44	2
PALUMBO M	333	1.36	2
SABEL C	329	1.34	2
BONAZZI G	325	1.33	2
DE LEONARDIS O	322	1.31	2
STAME N	316	1.29	2
OLSEN JP	306	1.25	2
CROZIER M	304	1.24	2
BIFULCO L	289	1.18	2
DONOLO C	286	1.17	2
LA ROSA M	272	1.11	2
VOELZKOW H	254	1.04	2
FREEMAN H	242	0.99	2
SOSKICE D	236	0.96	2
PERULLI P	236	0.96	2
DORE R	232	0.95	2
WEICK K	228	0.93	2
BUTERA F	227	0.93	2
PAWSON R	221	0.90	2
LIPPI A	220	0.90	2
CERI P	213	0.87	2
SWEDBERG R	212	0.87	2
ROSSI PH	212	0.87	2
FAZZI L	211	0.86	2
MEYER JW	205	0.84	2
FLORIDA R	204	0.83	2
OSTROM E	203	0.83	2
WILLIAMSON OE	202	0.82	2
FRIEDBERG E	200	0.82	2
STINCHCOMBE A	196	0.80	2
THéVENOT L	190	0.78	2
ARGYRIS C	190	0.78	2
POWER M	188	0.77	2

LEWIN K	182	0.74	2
SALAIS R	182	0.74	2
SCHUMPETER J	182	0.74	2
SCHON DA	182	0.74	2
SCRIVEN M	178	0.73	2
SELZNICK P	176	0.72	2
AXELROD R	175	0.71	2
HEDSTROM P	175	0.71	2
PATTON MQ	172	0.70	2
PIERSON P	169	0.69	2
WEICK KE	163	0.67	2
STORPER M	161	0.66	2
STIGLITZ J	160	0.65	2
WEISS CH	158	0.64	2
PFEFFER J	155	0.63	2
SCHON D	154	0.63	2
HALL P	153	0.62	2
BOYER R	150	0.61	2
PORTER M	149	0.61	2
BERTIN G	149	0.61	2
POLANYI M	145	0.59	2
LIPSEY MW	144	0.59	2
LAVILLE JL	136	0.56	2
GUBA EG	134	0.55	2
LINCOLN YS	134	0.55	2
BURRONI L	131	0.53	2
NEGRELLI S	126	0.51	2
ROWAN B	115	0.47	2
GUALA C	114	0.47	2
FLIGSTEIN N	114	0.47	2
BATTISTELLI F	113	0.46	2
KRUGMAN P	112	0.46	2
NONAKA I	111	0.45	2
FRIEDMAN M	108	0.44	2
GAMBETTA D	108	0.44	2
WOLLEB G	106	0.43	2
OLSON M	100	0.41	2
CERASE FP	99	0.40	2
MINARDI E	97	0.40	2
FREEMAN RE	71	0.29	2
GOFFMAN E	1.086	5.01	3
MERTON RK	887	4.10	3
DURKHEIM E	752	3.47	3

PARSONS T	674	3.11	3
CIPOLLA C	617	2.85	3
LATOUR B	617	2.85	3
MARRADI A	558	2.58	3
GHERARDI S	527	2.43	3
LAZARSFELD P	482	2.23	3
GEERTZ C	470	2.17	3
CRESPI F	431	1.99	3
ARDIGò A	396	1.83	3
SCHUTZ A	383	1.77	3
KUHN T	369	1.70	3
GIARELLI G	351	1.62	3
LUCKMANN T	315	1.45	3
STRAUSS A	300	1.39	3
CAPECCHI V	291	1.34	3
CAMPELLI E	289	1.33	3
KNORR-CETINA K	289	1.33	3
GARFINKEL H	268	1.24	3
ABBOTT A	267	1.23	3
WENGER E	265	1.22	3
MATURO A	263	1.21	3
NERESINI F	260	1.20	3
CARDANO M	250	1.15	3
ILLICH I	249	1.15	3
BERGER PL	245	1.13	3
BATESON G	239	1.10	3
LAW J	228	1.05	3
CALLON M	226	1.04	3
BERGER P	221	1.02	3
BRUNI A	215	0.99	3
WITTGENSTEIN L	209	0.96	3
RICOLFI L	209	0.96	3
GOBO G	205	0.95	3
FREIDSON E	201	0.93	3
DEWEY J	199	0.92	3
LAVE J	198	0.91	3
MILLS CW	195	0.90	3
MARZANO M	192	0.89	3
STATERA G	187	0.86	3
BICHI R	186	0.86	3
BUCCHI M	185	0.85	3
SMELSER NJ	181	0.84	3
SMITH A	181	0.84	3

WOOLGAR S	176	0.81	3
POPPER KR	173	0.80	3
GIGLIOLI PP	172	0.79	3
BLUMER H	171	0.79	3
NICOLINI D	170	0.78	3
HUSSERL E	166	0.77	3
GOOD B	160	0.74	3
CICOUREL A	159	0.73	3
STAR SL	157	0.72	3
BRUSCHI A	157	0.72	3
ALTIERI L	154	0.71	3
CANNAVò L	154	0.71	3
FELE G	152	0.70	3
SUCHMAN L	150	0.69	3
THOMAS WI	148	0.68	3
BIJKER WE	148	0.68	3
MEAD GH	145	0.67	3
BURY M	143	0.66	3
MALINOWSKI B	142	0.66	3
CZARNIAWSKA B	140	0.65	3
MANGHI S	134	0.62	3
GADAMER HG	132	0.61	3
FASANELLA A	132	0.61	3
TIMMERMANS S	123	0.57	3
HUGHES E	122	0.56	3
TURNER B	121	0.56	3
BROWN P	120	0.55	3
ROSITI F	120	0.55	3
VICARELLI G	119	0.55	3
CIPRIANI R	117	0.54	3
GALTUNG J	117	0.54	3
BERG M	115	0.53	3
BRUNER J	112	0.52	3
COSER L	112	0.52	3
CONRAD P	110	0.51	3
SONTAG S	110	0.51	3
BERTAUX D	108	0.50	3
SILVERMAN D	106	0.49	3
INGROSSO M	103	0.48	3
MOSCOVICI S	98	0.45	3
AGNOLI MS	85	0.39	3
SACKETT DL	80	0.37	3
MINTZBERG H	79	0.36	3

DENZIN N	70	0.32	3
MORUZZI M	69	0.32	3
SEN A	960	3.83	4
GALLINO L	908	3.63	4
BOUDON R	840	3.35	4
ESPING-ANDERSEN G	802	3.20	4
SARACENO C	799	3.19	4
GOLDTHORPE J	788	3.15	4
SCHIZZEROTTO A	785	3.13	4
MAGATTI M	780	3.11	4
PACI M	710	2.83	4
DE LILLO A	655	2.62	4
REYNERI E	614	2.45	4
BARBAGLI M	608	2.43	4
RANCI C	560	2.24	4
COLLINS R	549	2.19	4
CORBETTA P	520	2.08	4
MINGIONE E	491	1.96	4
REGINI M	480	1.92	4
FERRERA M	459	1.83	4
BARBIERI P	454	1.81	4
PISATI M	452	1.80	4
CASTEL R	445	1.78	4
DAHRENDORF R	422	1.68	4
BORGHI V	387	1.55	4
NEGRI N	375	1.50	4
COBALTI A	375	1.50	4
BIANCO ML	359	1.43	4
ACCORNERO A	357	1.43	4
BALLARINO G	342	1.37	4
BREEN R	318	1.27	4
BLOSSFELD HP	317	1.27	4
PASSERON JC	314	1.25	4
SCHADEE H	314	1.25	4
BLAU P	309	1.23	4
SHAVIT Y	296	1.18	4
BERNARDI F	286	1.14	4
NALDINI M	265	1.06	4
BECKER G	264	1.05	4
ERIKSON R	263	1.05	4
GASPERONI G	249	0.99	4
TREIMAN D	245	0.98	4
SUPIOT A	241	0.96	4

DUNCAN O	238	0.95	4
MULLER W	235	0.94	4
RUSPINI E	235	0.94	4
FULLIN G	228	0.91	4
SCHERER S	222	0.89	4
RIZZA R	221	0.88	4
LEWIS J	210	0.84	4
BALBO L	207	0.83	4
SAVAGE M	206	0.82	4
DUBET F	205	0.82	4
JONSSON JO	198	0.79	4
LODIGIANI R	197	0.79	4
GANZEBOOM HBG	187	0.75	4
PICCONE STELLA S	181	0.72	4
BISON I	179	0.71	4
GINTIS H	175	0.70	4
HOUT M	172	0.69	4
POGGIO B	156	0.62	4
CROMPTON R	155	0.62	4
GAMBARDELLA D	155	0.62	4
O'CONNOR J	153	0.61	4
PAVOLINI E	153	0.61	4
BOWLES S	146	0.58	4
SEMENZA R	143	0.57	4
OLAGNERO M	140	0.56	4
GERSHUNY J	113	0.45	4
LAYTE R	111	0.44	4
MOSCATI R	97	0.39	4
FACCHINI C	70	0.28	4
AMBROSINI M	1.121	6.11	5
PORTES A	835	4.55	5
MELUCCI A	674	3.67	5
DAL LAGO A	595	3.24	5
ZANFRINI L	583	3.18	5
COLOMBO E	577	3.15	5
GLICK SCHILLER N	490	2.67	5
ANDERSON B	485	2.64	5
COLOMBO A	469	2.56	5
LANDOLT P	432	2.36	5
VERTOVEC S	414	2.26	5
LEONINI L	388	2.12	5
GUARNIZO L	385	2.10	5
SCIORTINO G	376	2.05	5

PUGLIESE E	372	2.03	5
LEVITT P	363	1.98	5
HANNERZ U	362	1.97	5
BASCH L	352	1.92	5
SEMI G	337	1.84	5
SZANTON BLANC C	324	1.77	5
REBUGHINI P	323	1.76	5
QUEIROLO PALMAS L	321	1.75	5
CAPONIO T	318	1.73	5
MASSEY D	314	1.71	5
MARSHALL TH	292	1.59	5
ZUCCHETTI E	280	1.53	5
SAYAD A	269	1.47	5
RUMBAUT R	258	1.41	5
CLIFFORD J	251	1.37	5
FAIST T	244	1.33	5
BESOZZI E	235	1.28	5
CASTLES S	233	1.27	5
HALLER W	230	1.25	5
BOSISIO R	218	1.19	5
WALDINGER R	208	1.13	5
BRUBAKER R	201	1.10	5
MANNHEIM K	198	1.08	5
NEE V	193	1.05	5
ZHOU M	188	1.03	5
HOCHSCHILD AR	180	0.98	5
MELOSSI D	179	0.98	5
BAUMANN G	178	0.97	5
LAGOMARSINO F	174	0.95	5
WALZER M	171	0.93	5
ZINCONE G	163	0.89	5
FRISINA A	159	0.87	5
PACE E	155	0.85	5
MARCUS G	155	0.85	5
TOGNETTI BORDOGNA M	154	0.84	5
ANDALL J	151	0.82	5
PALIDDA S	144	0.79	5
MILLER MJ	141	0.77	5
SCIDà G	140	0.76	5
PARRENAS R	140	0.76	5
BURAWOY M	137	0.75	5
HUNTINGTON SP	132	0.72	5

DECIMO F	132	0.72	5
ONG A	132	0.72	5
JOPPKE C	127	0.69	5
MACIOTI MI	124	0.68	5
NUSSBAUM M	110	0.60	5
COTESTA V	99	0.54	5
LISTER R	93	0.51	5
SHORE C	86	0.47	5
ALLIEVI S	77	0.42	5
DONATI P	1.212	10.49	6
COLEMAN J	1.051	9.10	6
PUTNAM R	798	6.91	6
CAVALLI A	645	5.58	6
SCIOLLA L	492	4.26	6
CHIESI AM	455	3.94	6
PISELLI F	451	3.90	6
COLOZZI I	416	3.60	6
BURT R	396	3.43	6
ARCHER M	390	3.38	6
DI NICOLA P	361	3.12	6
LIN N	350	3.03	6
FUKUYAMA F	334	2.89	6
BUZZI C	322	2.79	6
TRONCA L	304	2.63	6
PRANDINI R	303	2.62	6
PELLIZZONI L	291	2.52	6
WELLMAN B	290	2.51	6
FORSé M	267	2.31	6
JACOBS J	217	1.88	6
ROSSI G	217	1.88	6
SCOTT J	213	1.84	6
BORGATTI SP	169	1.46	6
WASSERMAN S	162	1.40	6
PARETO V	162	1.40	6
OSTI G	160	1.38	6
FAUST K	148	1.28	6
MAYNTZ R	147	1.27	6
ERIKSON EH	142	1.23	6
LOCKWOOD D	139	1.20	6
RHODES R	121	1.05	6
BOCCACIN L	116	1.00	6
CASELLI M	115	1.00	6
KOOIMAN J	103	0.89	6

FOLGHERAITER F	96	0.83	6
SENNETT R	836	13.07	7
SASSEN S	478	7.47	7
HARVEY D	433	6.77	7
MARTINOTTI G	380	5.94	7
URRY J	373	5.83	7
MELA A	288	4.50	7
WACQUANT L	282	4.41	7
VITALE T	277	4.33	7
PARK R	260	4.06	7
AMIN A	250	3.91	7
ZUKIN S	246	3.85	7
NUVOLATI G	243	3.80	7
AMENDOLA G	224	3.50	7
ZAJCZYK F	199	3.11	7
GUIDICINI P	185	2.89	7
TOSI A	179	2.80	7
BURGESS E	171	2.67	7
BORLINI B	167	2.61	7
THRIFT N	164	2.56	7
RAUTY R	148	2.31	7
WILSON WJ	143	2.24	7
DAVICO L	129	2.02	7
DONZELOT J	92	1.44	7
LYNCH K	88	1.38	7
GASPARINI G	84	1.31	7
GASPARINI A	78	1.22	7
PIZZORNO A	899	18.29	8
TOURAINE A	598	12.17	8
ELSTER J	420	8.55	8
DELLA PORTA D	286	5.82	8
OFFE C	278	5.66	8
ETZIONI A	256	5.21	8
EISENSTADT SN	249	5.07	8
HELD D	220	4.48	8
TILLY C	199	4.05	8
NORRIS P	177	3.60	8
DIANI M	168	3.42	8
ROSANVALLON P	162	3.30	8
DAHL R	147	2.99	8
CATANZARO R	145	2.95	8
TARROW S	143	2.91	8
INGLEHART R	136	2.77	8

LIPSET SM	125	2.54	8
MAIR P	121	2.46	8
SKOCPOL T	97	1.97	8
COHEN J	89	1.81	8

APPENDIX M

Table 16: Nodal degree of each author; nodal degree expressed in percentage and relativised within each cluster. Data relate to the national map. Clusters are delimited by lines.

NAME	DEGREE CENTRALITY	% DEGREE CENTRALITY WITHIN CLUSTER	CLUSTER
BAGNASCO A	593	13.03	1
TRIGILIA C	499	10.97	1
PIZZORNO A	317	6.97	1
PISELLI F	224	4.92	1
MUTTI A	210	4.62	1
RAMELLA F	200	4.40	1
PICHIERRI A	182	4.00	1
NEGRI N	168	3.69	1
MARTINOTTI G	163	3.58	1
BARBERA F	163	3.58	1
DONOLO C	141	3.10	1
DE LEONARDIS O	135	2.97	1
PELLIZZONI L	129	2.84	1
MELA A	125	2.75	1
ZAJCZYK F	122	2.68	1
DELLA PORTA D	111	2.44	1
AMENDOLA G	97	2.13	1
GUIDICINI P	95	2.09	1
TOSI A	93	2.04	1
NUVOLATI G	90	1.98	1
CATANZARO R	85	1.87	1
PERULLI P	81	1.78	1
BURRONI L	72	1.58	1
BORLINI B	70	1.54	1
OSTI G	69	1.52	1
DAVICO L	66	1.45	1
OLAGNERO M	64	1.41	1
GUALA C	64	1.41	1
CERASE FP	53	1.16	1
GASPARINI A	39	0.86	1
GASPARINI G	30	0.66	1
GALLINO L	413	8.20	2
PACI M	351	6.97	2
MAGATTI M	347	6.89	2
REYNERI E	294	5.84	2
RANCI C	284	5.64	2
FERRERA M	226	4.49	2
MINGIONE E	221	4.39	2

REGINI M	211	4.19	2
BARBIERI P	197	3.91	2
BORGHI V	190	3.77	2
ACCORNERO A	182	3.61	2
BIANCO ML	171	3.39	2
GHERARDI S	167	3.32	2
CELLA G	146	2.90	2
LA ROSA M	145	2.88	2
NALDINI M	135	2.68	2
FULLIN G	130	2.58	2
SCHERER S	125	2.48	2
BIFULCO L	125	2.48	2
RIZZA R	114	2.26	2
BONAZZI G	110	2.18	2
PICCONE STELLA S	108	2.14	2
BALBO L	102	2.03	2
LODIGIANI R	97	1.93	2
PAVOLINI E	96	1.91	2
BUTERA F	86	1.71	2
POGGIO B	79	1.57	2
SEMENZA R	77	1.53	2
NEGRELLI S	56	1.11	2
MINARDI E	52	1.03	2
DE LILLO A	350	10.78	3
CAVALLI A	318	9.79	3
MARRADI A	286	8.81	3
CORBETTA P	246	7.58	3
SCIOLLA L	207	6.38	3
BUZZI C	168	5.17	3
CRESPI F	161	4.96	3
CAPECCHI V	147	4.53	3
RICOLFI L	122	3.76	3
CARDANO M	112	3.45	3
FERRAROTTI F	111	3.42	3
BICHI R	108	3.33	3
GOBO G	105	3.23	3
CAMPELLI E	93	2.86	3
CERI P	90	2.77	3
GARELLI F	77	2.37	3
STATERA G	71	2.19	3
CANNAVò L	65	2.00	3
FASANELLA A	61	1.88	3

CASELLI M	61	1.88	3
BRUSCHI A	53	1.63	3
BATTISTELLI F	49	1.51	3
AGNOLI MS	48	1.48	3
CIPRIANI R	47	1.45	3
FACCHINI C	38	1.17	3
ROSITI F	30	0.92	3
MORCELLINI M	23	0.71	3
AMBROSINI M	527	13.47	4
ZANFRINI L	263	6.72	4
COLOMBO E	247	6.31	4
DAL LAGO A	244	6.24	4
COLOMBO A	242	6.19	4
MELUCCI A	242	6.19	4
PUGLIESE E	192	4.91	4
SCIORTINO G	179	4.58	4
LEONINI L	164	4.19	4
ZUCCHETTI E	152	3.89	4
CAPONIO T	148	3.78	4
REBUGHINI P	143	3.66	4
SEMI G	134	3.43	4
QUEIROLO PALMAS L	129	3.30	4
BESOZZI E	110	2.81	4
ZINCONE G	97	2.48	4
BOSISIO R	92	2.35	4
FRISINA A	74	1.89	4
TOGNETTI BORDOGNA M	71	1.81	4
PALIDDA S	69	1.76	4
DECIMO F	67	1.71	4
PACE E	67	1.71	4
SCIDà G	66	1.69	4
LAGOMARSINO F	65	1.66	4
COTESTA V	47	1.20	4
MACIOTI MI	43	1.10	4
ALLIEVI S	38	0.97	4
CIPOLLA C	271	13.33	5
ARDIGò A	183	9.00	5
PALUMBO M	160	7.87	5
GIARELLI G	151	7.43	5
ROSSI P	137	6.74	5
STAME N	126	6.20	5
MATURO A	116	5.71	5

FAZZI L	108	5.31	5
NERESINI F	79	3.89	5
BERTIN G	76	3.74	5
ALTIERI L	73	3.59	5
MELOSSI D	69	3.39	5
BUCCHI M	69	3.39	5
BRUNI A	63	3.10	5
MARZANO M	61	3.00	5
VICARELLI G	59	2.90	5
INGROSSO M	57	2.80	5
MANGHI S	52	2.56	5
MORUZZI M	48	2.36	5
FELE G	32	1.57	5
GIGLIOLI PP	30	1.48	5
FERRARI V	13	0.64	5
CESAREO V	279	18.67	6
VACCARINI I	135	9.04	6
CODELUPPI V	120	8.03	6
SASSATELLI R	119	7.97	6
BOVONE L	111	7.43	6
PALTRINIERI R	105	7.03	6
MORA E	102	6.83	6
ALBERONI F	93	6.22	6
LECCARDI C	87	5.82	6
DI NALLO E	77	5.15	6
PARMIGGIANI P	71	4.75	6
SANTORO M	70	4.69	6
DIANI M	65	4.35	6
BOCCIA ARTIERI G	33	2.21	6
ABRUZZESE A	27	1.81	6
SCHIZZEROTTO A	372	21.77	7
BARBAGLI M	310	18.14	7
PISATI M	206	12.05	7
COBALTI A	153	8.95	7
RUSPINI E	148	8.66	7
SCHADEE H	140	8.19	7
BALLARINO G	129	7.55	7
GASPERONI G	118	6.90	7
GAMBARDELLA D	85	4.97	7
MOSCATI R	48	2.81	7
DONATI P	531	34.06	8
COLOZZI I	210	13.47	8

CHIESI AM	203	13.02	8
DI NICOLA P	158	10.13	8
PRANDINI R	136	8.72	8
TRONCA L	134	8.60	8
ROSSI G	121	7.76	8
BOCCACIN L	66	4.23	8

REFERENCES

Abbott, A. (2001). *Chaos of disciplines*. Chicago: University of Chicago Press.

Acedo, F. J., & Casillas, J. C. (2005). Current paradigms in the international management field: An author co-citation analysis. *International Business Review, 14*(5), 619-639.

Adair, W. C. (1955). Citation indexes for scientific literature? *American documentation, 6*(1), 31-32.

Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology, 54*(6), 550-560.

Ahlgren, P., Jarneving, B., & Rousseau, R. (2004). Author cocitation analysis and Pearson's r. *Journal of the American Society for Information Science and Technology, 55*(9), 843-843.

Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the World Wide Web: methodological approaches to 'Webometrics'. *Journal of Documentation*, *53*(4), 404-426.

Anderson, T. R., Hankin, R. K. S., & Killworth, P. D. (2008). Beyond the Durfee square: Enhancing the h-index to score total publication output. *Scientometrics, 76*(3), 577-588.

Angermüller, J. (2010). Beyond excellence. *Sociologica, 3*, 1-16. Retrieved from doi:10.2383/33636.

Anthonisse, J. M. (1971). The rush in a directed graph *Technical Report BN9/71*. Amsterdam: Stichting Mahtematisch Centrum.

Archambault, É., Vignola-Gagne, É., Côté, G., Larivière, V., & Gingrasb, Y. (2006). Benchmarking scientific output in the Social Sciences and humanities: The limits of existing databases. *Scientometrics, 68*(3), 329-342.

Asai, I. (1981). Adjusted age distribution and its application to impact factor and immediacy index. *Journal of the American Society for Information Science, 32*(3), 172-174.

Balbo, L., (1975). *Gli usi della sociologia nell'università di massa*. In L. Balbo, G. Chiaretti & G. Massironi (Eds.), *L'inferma scienza*. Bologna: Il Mulino.

Baldi, S. (1998). Normative versus social constructivist processes in the allocation of citations: A network-analytic model. *American Sociological Review, 63*(6), 829-846.

Barbano, F. (1998). *La sociologia in Italia. Storia, temi e problemi 1945-1960*. Roma: Carocci.

Barbano, F. (2003). *La sociologia in Italia. Le trasformazioni degli anni '70*. Milano: Franco Angeli.

Barber, B. (1961). Resistance by Scientists to Scientific Discovery. *Science, 134*, 596-602.

Barr, L. J., McMullen, H., & Leach, S. G. (1983). *Libraries in American periodicals before 1876: a bibliography with abstracts and an index*. Jefferson, NC: McFarland & Company.

Bartolini, M. (2007). In ordine sparso. *Sociologica, 1*, 1-13. Retrieved from doi:10.2383/24196

Barton, A. H., & Wilder, D. E. (1964). Research and practice in the teaching of reading: A progress report. In M. B. Miles (Ed.), *Innovation in education* (pp. 361-398). New York: Bureau of Publications, Teachers College, Columbia University.

Bateson, G. (1972). *Steps to an ecology of mind*. New York: Ballantine.

Ben-David, J., & Collins, R. (1966). Social factors in the origins of a new science: The case of psychology. *American Sociological Review, 31*(4), 451-465.

Bellardo, T. (1980). The Use of Co-Citations to Study Science. *Library Research, 2*, 231-237.

Bensman, S. J. (2004). Pearson's r and author cocitation analysis: A commentary on the controversy. *Journal of the American Society for Information Science and Technology, 55*(10), 935-935.

Bernal, J. D. (1939). The Social Function of Science. *The Social Function of Science.*

Björneborn, L. (2004). *Small-world Link Structures Across an Academic Web Space: A Library and Information Science Approach.* Ph.D. Thesis, Royal School of Library and Information Science, Copenhagen.

Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology, 55*(14), 1216-1227.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P10008.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology, 2*(1), 113-120.

Borg, I., & Groenen, P. J. F. (2005). *Modern multidimensional scaling: Theory and applications*. Berlin: Springer Verlag.

Borgatti, S.P., Everett, M.G., & Freeman, L.C. (2002). *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.

Borgman, C. L. (1990). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology, 36*, 3-72.

Bornmann, L., & Daniel, H-D. (2008). What do citation counts measure? A review of studies on citing behaviour. *Journal of Documentation*, *64*(1), 45-80.

Boudon, R. (2002). Sociology that Really Matters. *European Sociological Review, 18*(3), 371-378.

Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the progress of reason. *Social Science Information, 14*(6), 19-47.

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.

Bourdieu, P. (1993). *The field of cultural production*. Cambridge: Polity Press.

Bourdieu, P. (1996). *The rules of art: Genesis and structure of the literary field*. Stanford: Stanford University Press.

Bourdieu, P. (2004). *Science of science and reflexivity*. Cambridge: Polity Press.

Bourdieu, P. (2005). The political field, the social science field, and the journalistic field. In R. Benson & E. Neveu (Eds.), *Bourdieu and the Journalistic Field* (pp. 29-47). Cambridge: Polity Press.

Bourdieu, P., & Wacquant, L. J. D. (1992). *An invitation to reflexive sociology*. Chicago: University of Chicago Press.

Bouyssou, D., & Marchant, T. (2011). Ranking scientists and departments in a consistent manner. *Journal of the American Society for Information Science and Technology, 62*(9), 1761-1769.

Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics, 64*(3), 351-374.

Boyack, K. W., Börner, K., & Klavans, R. (2009). Mapping the structure and evolution of chemistry research. *Scientometrics, 79*(1), 45-60.

Börner, K. (2009). Visual conceptualisations and models of science. *Journal of Informetrics, 3*(3), 161-172.

Börner, K., & Scharnhorst, A. (2009). Visual conceptualizations and models of science. *Journal of Informetrics, 3*(3), 161-172.

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information science and technology, 37*(1), 179-255.

Braam, R. R., Moed, H. F., & Van Raan, A. F. J. (1988). Mapping of Science: Critical Elaboration and New Approaches, a Case Study in Agricultural Related Biochemistry. In L. Egghe & R. Rousseau (Eds.), *Informetrics 87/88* (pp. 15-28). Amsterdam: Elsevier.

Braam, R.R.; Moed, H.F.; & Van Raan, A.F.J. (1989). Comparison and combination of co-citation and co-word clustering. In A. F. J Van Raan, A. J. Nederhof, & H. F. Moed (Eds.), *Science and technology indicators: Their use in science policy and their role in science studies* (pp. 307-337), Leiden, University of Leiden: DSWO Press.

Braam, R. R., Moed, H. F., & Van Raan, A. F. J. (1991a). Mapping of science by combined co-citation and word analysis, I. Structural aspects. *Journal of the American Society for Information Science, 42*(4), 233-251.

Braam, R. R., Moed, H. F., & Van Raan, A. F. J. (1991b). Mapping of science by combined co-citation and word analysis, I. Structural aspects. *Journal of the American Society for Information Science, 42*(4), 233-251.

Bradford, S.C. (1934) Sources of Information on Specific Subject. *Engineering, 137,* 85-86

Bradford, S.C. (1948). *Documentation.* London: Lockwood.

Breiger, R.L. (1974), The duality of persons and groups. *Social Forces, 53,* 181-190.

Breiger, R.L. (1990). *Social mobility and social structure.* Cambridge: Cambridge University Press.

Breiger, R. L., Boorman, S. A., & Arabie, P. (1975). An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling. *Journal of Mathematical Psychology, 12*(3), 328-383.

Broadus, R. N. (1971). The literature of the Social Sciences: a survey of citation studies. *International Social Sciences Journal, 23*(2), 236-243.

Broadus, R. N. (1987). Early approaches to bibliometrics. *Journal of the American Society for Information Science, 38*(2), 127-129.

Brookes, B.C. (1977). Theory of the Bradford law. *Journal of Documentation, 33*(3), 173-250.

Bryman, A. (1984). The debate about quantitative and qualitative research: a question of method or epistemology? *British Journal of Sociology, 35*(1), 75-92.

Burawoy, M. (2005). For public sociology. *American Sociological Review, 70*(1), 4-28.

Burgalassi, M. M. (1996) *Itinerari di una scienza*, Milano: Franco Angeli.

Bush, V. (1980). *Science, the endless frontier: A report to the President: Arno Press*. (Original work published 1945).

Callon, M. (1986). Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay. In J. Law (Ed.), *Power, Action and Belief: A New Sociology of Knowledge* (pp. 196-233). London: Routledge & Kegan Paul.

Callon, M. (1998). *The laws of the markets*. London: Macmillan

Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information, 22*(2), 191-235.

Campanario, J. M. (1995). Using neural networks to study networks of scientific journals. *Scientometrics, 33*(1), 23-40.

Campbell, D. T. (1969). Ethnocentrism of disciplines and the fish-scale model of omniscience. In M. Sherif & C. W. Sherif (Eds.), *Interdisciplinary Relationships in the Social Sciences*. Chicago: Aldine.

Cannavò, L. (2008) Qualità nella ricerca, qualità della ricerca. *Quaderni di sociologia, 52*(47), 147-156.

Cardona, M., Chamberlin, R. V., & Marx, W. (2007). The history of the stretched exponential function. *Annalen der Physik, 16*(12), 842-845.

Cardona, M., & Marx, W. (2008) Max Born and his legacy to condensed matter physics. *Annalen der Physik, 17*(7), 497-518.

Cavalli, A., della Porta, D., Donati, P., & Rositi, F. (2010). Sulla sociologia italiana. *Il Mulino, 59*(4), 655-669.

Chen, C. (1997). Structuring and visualizing the WWW with Generalized Similarity Analysis *Hypertext '97. Proceedings of the 8th ACM Conference on Hypertext* (pp. 177-186). Southampton: ACM.

Chen, C. (1998a). Bridging the gap: The use of pathfinder networks in visual navigation. *Journal of Visual Languages and Computing, 9*(3), 267-286.

Chen, C. (1998b). Generalised similarity analysis and pathfinder network scaling. *Interacting with Computers, 10*(2), 107-128.

Chen. C. (1999a) *Information visualisation and virtual environments.* Springer-Verlag: London.

Chen, C. (1999b). Visualising semantic spaces and author co-citation networks in digital libraries. *Information Processing and Management, 35*(3), 401-420.

Chen, C., & Carr, L. (1999). Trailblazing the literature of hypertext: author co-citation analysis (1989–1998) *Hyoertext '99. Proceedings of the tenth ACM Conference on Hypertext and hypermedia: returning to our diverse roots: returning to our diverse roots* (pp. 51-60). New York: ACM.

Chen, C., & Kuljis, J. (2003). The rising landscape: A visual exploration of superstring revolutions in physics. *Journal of the American Society for Information Science and Technology, 54*(5), 435-446.

Chen, C., & Paul, R. J. (2001). Visualizing a knowledge domain's intellectual structure. *Computer, 34*(3), 65-71.

Chiaretti, G. (1975). Un caso di organizzazione della scienza: la sociologia in Italia nel decennio 1958-1968. In L. Balbo, G. Chiaretti & G. Massironi (Eds.), *L'inferma scienza*. Bologna: Il Mulino

Chiesi, A. M. (1999) L'analisi dei reticoli Milano: Franco Angeli.

Chiesi, A. M. (2005). Tre dicotomie del metodo sociologico. In M. Borlandi & L. Sciolla (Eds.), *La spiegazione sociologica*. Bologna: Il Mulino.

Chiesi, A. M. (2007). Le sociologie, il controllo delle loro affermazioni e le loro degenerazioni, *Sociologica, 2,*1-14, doi: 10.2383/24760.

Chiesi, A. M. (2008) La valutazione della produzione sociologica. *Quaderni di Sociologia, 52*(47), 123-134.

Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences of the United States of America, 101*(Suppl 1), 5303-5310.

Chubin, D. E. (1976). State of the Field The Conceptualization of Scientific Specialties. *The Sociological Quarterly, 17*(4), 448-476.

Chubin, D. E., & Moitra, S. D. (1975). Content analysis of references: adjunct or alternative to citation counting? *Social Studies of Science, 5*(4), 423-441.

Clemens, E. S., Powell, W. W., McIlwaine, K., & Okamoto, D. (1995). Careers in print: Books, journals, and scholarly reputations. *American Journal of Sociology, 101*, 433-494.

Cole, S. (1983). The hierarchy of the sciences? *American Journal of Sociology, 89*, 111-139.

Cole, J. R., & Cole, S. (1972). The Ortega hypothesis. *Science, 178*(4059), 368-375.

Cole, J.R. & Cole, S. (1973). *Social stratification in science.* Chicago: University of Chicago Press

Collins, R. (1988). *Theoretical sociology.* San Diego: Harcourt Brace Jovanovich.

Collins, H.M. (2004). *Gravity's shadow: the search for gravitational waves.* Chicago University of Chicago Press.

Corcuff, P, (2007). *Les nouvelles sociologies.* Paris: Armand Colin.

Coutrot, L. (2009). La valutazione nelle scienze sociali: avete detto quantificare? *Quaderni di sociologia*, *49*(3), 183-204.

Cox, T.F. & Cox, M.A.A. (2001). *Multidimensional scaling* Boca Raton, Fla.:Chapman & Hall/CRC.

Cozzens, S. E. (1981). Taking the measure of science: A review of citation theories. *International Society for the Sociology of Knowledge Newsletter, 7*(1–2), 16-20.

Cozzens, S. E. (1982). Split citation identity: a case study from economics. *Journal of the American Society for Information Science, 33*(4), 233-236.

Cozzens, S. E. (1989). What do citations count? The rhetoric-first model. *Scientometrics, 15*(5-6), 437-447.

Crane, D. (1967). The gatekeepers of science: Some factors affecting the selection of articles for scientific journals. *The American Sociologist, 2*(4), 195-201.

Crane, D. (1969). Social structure in a group of scientists: A test of the" invisible college" hypothesis. *American Sociological Review, 34*(3), 335-352.

Crane, D. (1976). Reward systems in art, science, and religion. *American Behavioral Scientist, 19*(6), 719-734.

Crawford, S. (1971). Informal communication among scientists in sleep research. *Journal of the American Society for Information Science, 22*(5), 301-310.

Croce, B. (1950) Problemi di metodologia storica. *Quaderni della critica*, *16*, 6-38.

Cronin, B. (1981). The need for a theory of citing. *Journal of Documentation, 37*(1), 16-24.

Cronin, B. (1984). *The citation process: the role and significance of citations in scientific communication.* London: Taylor Graham.

Cronin, B. (2000). Semiotics and evaluative bibliometrics. *Journal of Documentation, 56*(4), 440-453.

Cronin, B., Snyder, H., & Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: A study of sociology. *Journal of Documentation, 53*(3), 263-273.

Daipha, P. (2001). The intellectual and social organisation of ASA 1990–1997: exploring the interface between the discipline of sociology and its practitioners. *The American Sociologist, 32*(3), 73-90.

Danon, L., Díaz-Guilera, A., Duch, J., & Arenas, A. (2005). Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment,* (9), 219-228.

Davis, J. A. (1994). What's wrong with sociology? *Sociological Forum, 9*(2), 179-197.

De Bellis, N. (2009). *Bibliometrics and citation analysis.* Lanham: Scarecrow Press.

De Bruin, R.E., Kint, A., Luwel, M., & Moed, H.F. (1993). A study of research evaluation and planning: The University of Ghent. *Research Evaluation*, *3*(1), 25–41.

de Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. Cambridge: Cambridge University Press.

Deineko, V. G., & Woeginger, G. J. (2009). A new family of scientific impact measures: The generalized Kosmulski-indices. *Scientometrics, 80*(3), 819-826.

Diani, M. (2008) Indicatori bibliometrici e sociologia italiana. *Quaderni di Sociologia, 53*(47), 135-146.

Doreian, P., Batagelj, V., & Ferligoj, A. (2005). *Generalized blockbusting*. Cambridge: Cambridge University Press.

Durkheim, É. (1984).*The division of labor in society*. New York: Free Press. (Original work published 1933).

Edge, D. (1979). Quantitative measures of communication in science: a critical review. *History of science; an annual review of literature, research and teaching*, *17*(36 Pt 2), 102-134.

Edge, D.O. & Mulkay, M. (1976). *Astronomy transformed: the emergence of radio astronomy in Britain.* New York: Wiley.

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics, 69*(1), 131-152.

Egghe, L. & Rousseau, R. (1990). *Introduction to Informetrics: Quantitative methods in library, documentation and information science*. Amsterdam: Elsevier Science Publishers.

Egghe, L., & Rousseau, R. (1995). Generalized success-breeds-success principle leading to time-dependent informetric distributions. *Journal of the American Society for Information Science, 46*(6), 426-445.

Egghe, L., Rousseau, R., & Van Hooydonk, G. (2000). Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for Information Science and Technology, 51*(2), 145-157.

Eom, S. B. (2009). *Author cocitation analysis: quantitative methods for mapping the intellectual structure of an academic discipline*: Information Science Publishing.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From National Systems and "mode 2" to a Triple Helix of university-industry-government relations. *Research Policy, 29*(2), 109-123.

Fairthorne, R. A. (1969). Empirical hyperbolic distributions (Bradford-Zipf-Mandelbrot) for bibliometric description and prediction. *Journal of Documentation, 25*(4), 319-343.

Farkas, I. J., Ábel, D., Palla, G., & Vicsek, T. (2007). Weighted network modules. *New Journal of Physics, 9*, 1-18.

Feyerabend, P. (2011). *The tyranny of science*. Cambridge: Polity Press.

Fisher, C. S. (1966). The death of a mathematical theory: a study in the sociology of knowledge. *Archive for History of Exact Sciences, 3*(2), 137-159.

Fisher, C.S. (1967) The last invariant theorists: a sociological study of the collective biographies of mathematical specialists. *European Journal of Sociology, 8*:216-244.

Fleck, L. (1979). Genesis and development of a scientific fact (F. Bradley & T. J. Trenn, Trans.). Chicago: The University of Chicago Press. (Original work published 1935).

Fonseca, E. (1973). Bibliografia Estatistica e Bibliometria: Uma Reivindicacao de Prioridades. *Ciencia da Informacao, 2*(1):5–7.

Fortunato, S., & Barthélemy, M. (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(1), 36-41.

Franceschet, M. (2009). A cluster analysis of scholar and journal bibliometric indicators. *Journal of the American Society for Information Science and Technology, 60*(10), 1950-1964.

Franklin, J. J. (1988). Testing and using quantitative methods in science policy contexts: a response to Hicks. *Social Studies of Science, 18*(2), 365-375.

Fredriksson, E. H. (2001) *A Century of Science Publishing:* A *Collection of Essays*. Amsterdam: IOS Press.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*, 35-41.

Freschi, C., & Santoro, M. (2010a). Italy: A Troubled and Divided Academic Field under Neoliberal Pressure. Retrieved from http://www.isa-sociology.org/universities-in-crisis/?p=373

Freschi, C., & Santoro, M. (2010b). Introduction. *Sociologica*, *3*:1-7, doi: 10.2383/33635.

Friedrichs, R. W. (1970). *A sociology of sociology.* New York: The Free Press.

Frost, C. O. (1979). The use of citations in literary research: A preliminary classification of citation functions. *The Library Quarterly, 49*, 399-414.

Galison, P. (1997). *Image and logic: A material culture of microphysics*. Chicago: University of Chicago Press.

Gallino, L. (1972). Crisi della sociologia, ricerca sociologica, ruolo del sociologo. In P. Rossi (Ed.), *Ricerca sociologica e ruolo del sociologo* (pp. 309-325). Bologna: Il Mulino.

Garfield, E. (1954). *Shepardizing the scientific literature*. Unpublished paper. Columbia University Library School.

Garfield, E. (1955) Citation indexes for science. *Science, 122*(3159), 108-111.

Garfield, E. (1959). A unified index to science *Proceedings of the International Conference on Scientific Information* (Vol. I, pp. 461-474). Washington: National Academy of Sciences-National Research Council Washington.

Garfield, E. (1960a) Citation Index to Genetics and General Science Literature. Research Proposal by the Institute for Scientific Information. Submitted July 15, 1960 to the National Science Foundation.

Garfield, E. (1960b). Citation index to genetics literature. Application for Research Grant. Philadelphia: Personal Archive Eugene Garfield.

Garfield, E. (1963). Citation indexes in sociological and historical research. *American documentation, 14*(4), 289-291.

Garfield, E. (1964). Science Citation Index-A new dimension in indexing. *Science, 144*(3619), 649-654.

Garfield, E. (1970). Citation indexing for studying science. *Nature, 227*(5259), 669-671.

Garfield, E. (1979). *Citation indexing: its theory and application in science, technology, and humanities*. New York: Wiley.

Garfield, E. (1980). *Essays on information scientists* (Vol. III). Philadelphia: ISI Press.

Garfield, E. (1986) The evolution to physical chemistry to chemical physics *Current Contents, 3*:3-12.

Garfield, E., (1990). How ISI Selects Journals for Coverage: Quantitative and Qualitative Considerations. *Essays of an Information Scientist, 13*:185-193.

Garfield, E. (1991). In truth, the 'flood'of scientific literature is only a myth. *The Scientist, 5*(17), 11-25.

Garfield, E. (1996a). How can impact factors be improved? *British Medical Journal, 313*(7054), 411-413.

Garfield, E. (1996b). Significant Scientific Literature Appears In a Small Core of Journals. *The Scientist, 10*(17), 13-16.

Garfield, E. (2001). *From bibliographic coupling to co-citation analysis via algorithmic historio-bibliography*. Speech delivered at Drexel University, Philadelphia.

Garfield, E. (2005). *The Agony and the Ecstasy: the History and the Meaning of the Journal Impact Factor*. Paper presented at the International Congress on Peer Review and Biomedical Publication, Chicago.

Garfield, E., & Sher. I. (1963). New Factors in the Evaluation of Scientific Literature Through Citation Indexing. *American Documentation*, *14*(3), p.195-201.

Garfield, E., Sher, I. H., & Torpie, R. J. (1964). *The use of citation data in writing the history of science*. Philadelphia: ISI Press.

Geymonat, L. (1974). *Attualità del materialismo dialettico*. Roma: Editori Riuniti.

Geymonat, L., & Giorello, G. (1986). *Le ragioni della scienza*. Bari: Laterza.

Gibbons, M., Limoges, H., Nowotny, S., Schwartzman, P. S., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: Sage.

Giddens, A. (1984). *The constitution of society: outline of the theory of structuration*. Cambridge: Polity Press.

Gilbert, G. N. (1977). Referencing as persuasion. *Social Studies of Science*, *7*(1), 113-122.

Gingras, Y. (2008). La Fièvre de l'évaluation de la recherche du mauvais usage de faux indicateurs. *Bulletin de Méthodologie Sociologique, 100*(1), 41-44.

Gingras, Y. (2010). Mapping the structure of the intellectual field using citation and co-citation analysis of correspondences. *History of European Ideas, 36*(3), 330-339.

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences, 99*(12), 7821-7826.

Glänzel, W. (1996). A bibliometric approach to Social Sciences. National research performances in 6 selected social science areas, 1990–1992. *Scientometrics, 35*(3), 291-307.

Glänzel, W. (2003). Bibliometrics as a research field. A course on theory and application of bibliometric indicators. Course handouts.

Glänzel, W., & Schoepflin, U. (1994). Little scientometrics, big scientometrics... and beyond? *Scientometrics, 30*(2), 375-384.

Glänzel, W., & Schoepflin, U. (1995). A bibliometric study on ageing and reception processes of scientific literature. *Journal of information science, 21*(1), 37-53.

Glänzel, W., & Schoepflin, U. (1999). A bibliometric study of reference literature in the sciences and Social Sciences. *Information processing & management, 35*(1), 31-44.

Glänzel, W., Thijs, B., & Schlemmer, B. (2004). A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics, 59*(1), 63-77.

Gmür, M. (2003). Co-citation analysis and the search for invisible colleges: A methodological evaluation. *Scientometrics, 57*(1), 27-57.

Goffman, W. (1966). Mathematical approach to the spread of scientific ideas-the history of mast cell research. *Nature, 212*(5061), 449-452.

Goldthorpe, J. H. (2000). *On Sociology*. Oxford: Oxford University Press.

Gorkova, V. I. (1988). *Informetriya [Informetrics]*. Moscow: VINITI.

Gross, P. L. K., & Gross, E. (1927). College libraries and chemical education. *Science, 66*(1713), 385-389.

Gramsci, A. (1948). *Materialismo storico e la filosofia di B. Croce*. Torino: Einaudi.

Griffith, B. C., Small, H. G., Stonehill, J. A., & Dey, S. (1974). The structure of scientific literatures II: Toward a macro- and micro-structure for science. *Science Studies, 4*(4), 339-365.

Guimerà, R., & Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature, 433*(7028), 895-900.

Habermas, J. (1981). Theories des kommunikativen handelns [The theory of communicative action]. Frankfurt, Germany: Suhrkamp.

Hagstrom, W. O. (1970). Factors related to the use of different modes of publishing research in four scientific fields. In C. E. Nelson & D. K. Pollock (Eds.), *Communication among Scientists and engineers*. Lexington, Mass.: Lexington Books.

Hammarfelt, B. (2012). Following the Footnotes: A Bibliometric Analysis of Citation Patterns in Literary Studies. (Doctoral dissertation). Uppsala: Acta Universitatis Upsaliensis.

Haythomthwaite, C., & Wellman, B. (1998). Work, friendship, and media use for information exchange in a networked organisation. *Journal of the American Society for Information Science, 49*(12), 1101-1114.

Hertzel, D. (2003). Bibliometrics History. In M. Drake (Ed.), *Encyclopedia of library and information science* (pp. 288-328). New York: Maecel Dekker.

Hess, D. J. (1997). *Science studies: An advanced introduction*. New York: New York University Press.

Hicks, D. (1987). Limitations of co-citation analysis as a tool for science policy. *Social Studies of Science, 17*(2), 295-316.

Hicks, D. (1999). The difficulty of achieving the full coverage of international social science literature and the bibliometric consequences. *Scientometrics, 44*(2), 193-215.

Hicks, D. M. (2004). The four literatures of social science. *Journal of management and social sciences, 1*(1), 1.

Hicks, D., & Potter, J. (1991). Sociology of scientific knowledge: a reflexive citation analysis or science disciplines and disciplining science. *Social Studies of Science, 21*(3), 459-501.

Hicks, D., & Wang, J. (2009). Toward a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project *A Report to the Project Board of the Scoping Study 'Towards a Bibliometric Database for the Social Sciences and the Humanities*: Under the auspices of the European Science Foundation and funded by ESRC, AHRC, ANR, NWO and DFG

Hinton, G. E. (1989). Connectionist learning procedures. *Artificial intelligence, 40*(1), 185-234.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569-16572.

Hodges, T. L. (1972). *Citation indexing: Its potential for bibliographical control.* Ph.D. thesis, University of California, Berkeley.

Holton, G. (1978). Can science be measured. In Y. Elkana , J. Ledeberg, R. K. Merton, A. Thackray & H. Zuckerman (Eds.), *Toward a metric of science: the advent of science indicators* (pp. 39-68). New York: Wiley.

Hood, W. W., & Wilson, C. S. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics, 52*(2), 291-314.

Hsieh, M. H., & Magee, C. L. (2008). An algorithm and metric for network decomposition from similarity matrices: Application to positional analysis. *Social Networks, 30*(2), 146-158.

Huffman, D. A. (1952). A method for the construction of minimum-redundancy codes. *Proceedings of the IRE, 40*(9), 1098-1101.

Hulme, E. W. (1923). *Statistical bibliography in relation to the growth of modern civilization*. London: Grafton.

Husserl, E. (1973). *Cartesian meditations and the Paris lectures*. (D. Cairns, Trans.). The Hague, Netherlands: Martinus Nijhoff. (Original work published 1929).

Iribarren-Maestro, I.; Lascurain-Sánchez, M. L., & Sanz-Casado, E. (2009). The use of bibliometric techniques in evaluating Social Sciences and Humanities. In: Åström, F.; Danell, R.; Larsen, B.; Schneider, J. W. *Celebrating Scholarly Communication Studies: A Festschrift for Olle Persson at his 60th Birthday*, (pp. 25-37). International Society for Scientometrics and Informetrics.

Irvine, J., & Martin, B. R. (1985). Basic research in the East and West: A comparison of the scientific performance of high-energy physics accelerators. *Social Studies of Science, 15*(2), 293-341.

Jahn, M. (1972). *Changes with growth of the scientific literatures of two biomedical specialties.* MS thesis, Drexel University.

Jones, K. S. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation, 28*(1), 11-21.

Jewett, C. C. (1849). Report of the Assistant Secretary Relative to the Library, Presented December 13, 1848. In *Third Annual Report of the Board of Regents of the Smithsonian Institution to the Senate and House of Representatives* (pp. 39-47). Washington: Tippin & Streeper Printers.

Kadushin, C. (1966). The friends and supporters of psychotherapy: on social circles in urban life. *American Sociological Review, 31*(6), 786-802.

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information processing letters, 31*(1), 7-15.

Kannan, R., Vempala, S., & Vetta, A. (2004). On clusterings: Good, bad and spectral. *Journal of the ACM, 51*(3), 497-515.

Kant, I. (1980). *The critique of judgem*ent. (J. C. Meredith, Trans.). Oxford: Clarendon. (Original work published 1790).

Kaplan, N. (1965). The norms of citation behavior: Prolegomena to the footnote. *American documentation, 16*(3), 179-184.

Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics, 31*(1), 31-43.

Katz, J. S. (1999). *Bibliometric indicators and the Social Sciences*. Brighton: University of Sussex.

Kessler, M. M. (1958). *Concerning Some Problems of Intrascience Communication*. Massachusetts: Massachusetts Institute of Technology, Lincoln Laboratory.

Kessler, M. M. (1963a). Bibliographic coupling between scientific papers. *American documentation, 14*(1), 10-25.

Kessler, M. M. (1963b). An experimental study of bibliographic coupling between technical papers. *IEEE Transactions on Information Theory, 9*(1), 49-51.

Kessler, M. M. (1965). Comparison of the results of bibliographic coupling and analytic subject indexing. *American documentation, 16*(3), 223-233.

King, J. (1987). A review of bibliometric and other science indicators and their role in research evaluation. *Journal of information science, 13*(5), 261-276.

Klavans, R., & Boyack, K. W. (2006). Quantitative evaluation of large maps of science. *Scientometrics, 68*(3), 475-499.

Knorr-Cetina, K. (1981). *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press.

Koefoed, P. A. (1976). SISWO groep wetenschapsonderzoekers nieuwsbrief nr. 10. SISWO, OZ Achterburgwal 128, Amsterdam. Copies from Arie Rip's Personal Archive.

Kohonen, T. (1990). The self-organising map. *Proceedings of the IEEE, 73*(9), 1551-1558.

Kohonen, T. (1995). Self-organising maps. In K. Lagus, T. Honkela, S. Kaski & T. Kohonen (Eds.), *Self-organizing maps of document collections: A new approach to interactive exploration* (pp. 238-243). New York: Springer.

Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original h-index. *ISSI newsletter, 2*(3), 4-6.

Kothen, R. (1954). Come sviluppare il senso della chiesa nella classe lavoratrice. *Quaderni di azione sociale*.

Krohn, W., Küppers, G., & Nowotny, H. (1990). *Selforganization: portrait of a scientific revolution*. Dordrecht: Kluwe

Kruskal, J., & Wish, M. (1978). *Multidimensional Scaling*. Beverly Hills, CA: Sage.

Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press. (Original work published 1962).

Kyvik, S. (1988). Internationality of the Social Sciences: the Norwegian case. *International Social Science Journal*, *115*, 163-172.

Lakatos, I., & Musgrave, A. (1965). *Criticism and the growth of knowledge: Proceedings of the International Colloquium in the Philosophy of Science* (Vol. IV). Cambridge University Press.

Lange, L. (1985). Effects of disciplines and countries on citation habits. An analysis of empirical papers in behavioural sciences. *Scientometrics*, *8*(3-4), 205-215.

Latour, B. (1987). *Science in action: how to follow scientists and engineers through society*. Cambridge, Mass.: Harvard University Press.

Latour, B. (1988). *The pasteurization of France.* Cambridge, Mass.: Harvard University Press.

Latour, B., & Woolgar, S. (1979). *Laboratory life: The construction of scientific facts*. London: Sage.

Law, J. (1973). The development of specialties in science: The case of x-ray protein crystallography. *Social Studies of Science, 3*(3), 275-303.

Leach, E. R. (1976). *Culture & communication: the logic by which symbols are connected: an introduction to the use of structuralist analysis in social anthropology*: Cambridge University Press.

Lewison, G., & Cunningham, P. (1991). Bibliometric studies for the evaluation of trans-national research. *Scientometrics, 21*(2), 223-244.

Leydesdorff, L. (1987a). Towards a theory of citation? *Scientometrics, 12*(5-6), 305-309.

Leydesdorff, L. (1987b). Various methods for the mapping of science. *Scientometrics, 11*(5-6), 295-324.

Leydesdorff, L. (1995). *The Challenge of Scientometrics: The development, measurement and self-organisation of scientific communications*. Leiden: DSWO Press.

Leydesdorff, L. (1998). Theories of citation? *Scientometrics*, 43(1), 5-25.

Leydesdorff, L. (2001). *A sociological theory of communication: the self-organisation of the knowledge-based society*. Parkland, FL: Universal Publishers.

Leydesdorff, L. (2005a). Similarity measures, author cocitation analysis, and information theory. *Journal of the American Society for Information Science and Technology, 56*(7), 769-772.

Leydesdorff, L. (2005b). Evaluation of research and evolution of science indicators. *Current science, 89*(9), 1510-1517.

Leydesdorff, L. (2007). Scientific communication and cognitive codification: Social systems theory and the sociology of scientific knowledge. *European Journal of Social Theory*, 10(3), 375-388.

Leydesdorff, L. (2008). On the normalization and visualization of author co-citation data: Salton's Cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology, 59*(1), 77-85.

Leydesdorff, L. (2010). The communication of meaning and the structuration of the expectations: Gidden's "structuration theory" and "Luhmann's "self-organisation". *Journal of the American Society for Information Science and Technology, 61*(10), 2138-2150.

Leydesdorff, L., & Amsterdamska, O. (1990). Dimensions of citation analysis. *Science, Technology & Human Values*, *15*(3), 305-335.

Leydesdorff, L., & Etzkowitz, H. (1996). Emergence of a triple helix of university-industry-government relations. *Science and Public Policy, 23*(5), 279-286.

Leydesdorff, L., & Etzkowitz, H. (1998). The triple helix as a model for innovation studies. *Science and Public Policy, 25*(3), 195-203.

Leydesdorff, L., & Vaughan, L. (2006). Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment. *Journal of the American Society for Information Science and Technology, 57*(12), 1616-1628.

Leydesdorff, L., & Wouters, P. (1999). Between texts and contexts: Advances in theories of citation? (A rejoinder). *Scientometrics, 44*(2), 169-182.

Logan, E., & Pao, M. (1990). Analytic and empirical measures of key authors in schistosomiasis. *Proceedings of the American Society for Information Science, 27*, 213-219.

Logan, E. L., & Pao, M. L. (1991). Identification of Key Authors in a Collaborative Network. *Proceedings of the American Society for Information Science, 28*, 261-266.

Lorrain, F., & White, H. C. (1971). Structural equivalence of individuals in social networks. *Journal of Mathematical Sociology, 1*(1), 49-80.

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences, 16,* 317-323.

Luhmann, N. (1982). *The differentiation of society*. (S. Holmes & C. Larmore, Trans.). New York: Columbia University Press. (Original work published 1981).

Luhmann, N. (1986a). *Love as passion: the codification of intimacy*. (J. Gaines & D. L. Jones, Trans.). Stanford: Stanford University Press. (Original work published 1982).

Luhmann, N. (1986b). The autopoiesis of social systems. In F. Geyer & J. van der Zouwen (Eds.), *Sociocybernetic Paradoxes*. London: SAGE Publications.

Luhmann, N. (1989). *Ecological communication*. (J. Bednarz, Trans.). Cambridge: Polity Press. (Original work published 1986).

Luhmann, N. (1990). The cognitive program of constructivism and a reality that remains unknown. In W. Krohn, G. Küppers & H. Nowotny (Eds.), *Portrait of a scientific revolution*. Dordrecht: Kluwer Academic Publishers.

Luhmann, N. (1995). *Social systems*. (J. Bednarz & D. Baecker, Trans.). Stanford: Stanford University Press. (Original work published 1984).

Luukkonen, T. (1997). Why has Latour's theory of citations been ignored by the bibliometric community? Discussion of sociological interpretations of citation analysis. *Scientometrics, 38*(1), 27-37.

Luwel, M., Moed, H. F., Nederhof, A. J., De Samblanx, V., Verbrugghen, K., & Van Der Wurff, L. J. (1999). Towards Indicators of Research Performance in the Social Sciences and Humanities. An Exploratory Study in the Fields of Law and Linguistics at Flemish Universities *Report of the Flemish Inter-University Council (V.L.I.R.)*. Brussels: Centre for Science and Technology Studies (CWTS), Leiden University, the Netherlands / Ministry of the Flemish Community.

MacRoberts, M. H., & MacRoberts, B. R. (1982). A Re-evaluation of Lotka's Law of scientific productivity. *Social Studies of Science, 12*(3), 443-450.

MacRoberts, M. H., & MacRoberts, B. R. (1986). Quantitative measures of communication in science: A study of the formal level. *Social Studies of Science, 16*(1), 151-172.

MacRoberts, M. H., & MacRoberts, B. R. (1987). Testing the Ortega hypothesis: Facts and artifacts. *Scientometrics, 12*(5), 293-295.

MacRoberts, M. H., & MacRoberts, B. R. (1989). Problems of citation analysis: A critical review. *Journal of the American Society for Information Science, 40*(5), 342-349.

Magatti, M. (2007). Sulla crisi della sociologia. *Sociologica, 2*. Retrieved from doi: 10.2383/24763

Mandelbrot, B. (1965). Information Theory and Psycholinguistics. In Wolman B.B; & Nagel E. (Eds.), *Scientific psychology*. New York: Basic Books.

Mannheim, K. (1952). *Essays on the Sociology of Knowledge*. London: Routledg.

Marchant, T. (2009). An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors. *Scientometrics, 80*(2), 325-342.

Marshakova, I. (1973). System of connections between documents based on references. *Nauchno-Tekhnicheskaya Informatsiya, 2*(6), 3-8.

Marshakova, I. (1996). The standard impact factor as an evaluation tool of science fields and scientific journals. *Scientometrics, 35*(2), 283-290.

Martyn, J. (1965). An examination of citation indexes. *Aslib Proceedings, 17*(6), 184-196.

Marx, W. (2011). Special features of historical papers from the viewpoint of bibliometrics. *Journal of the American Society for Information Science and Technology, 62*(3), 433-439.

Massironi, G. (1975). "Americanate". In L. Balbo, G. Chiaretti & G. Massironi (Eds.), *L'inferma scienza*. Bologna: Il Mulino.

Maturana, H. (1975a). Communication and representation functions. In J. Piaget (Ed.), *Encyclopédie de la Pléiade, Séries Méthodique*. Paris: Gallimard.

Maturana, H. (1975b). The organisation of the living: a theory of the living organisation. *International Journal of Man-Machine Studies, 7*(3), 313-332.

Maturana, H. (1990). Science and daily life: the ontology of scientific explanations. In W. Krohn, G. Küppers & H. Nowotny (Eds.), *Selforganisation. Portrait of a scientific revolution* (Vol. 12-35). Dordrecht: Kluwer Academic Oublishers.

Maturana, H. (2002). Autopoiesis, structural coupling and cognition: a history of these and other notions in the biology of cognition. *Cybernetics & Human Knowing, 9*(3-4), 5-34.

Maturana, H., & Varela, F. (1980). *Autopoiesis and cognition: the realization of the living*. Dordrecht: Reidel.

Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge. The biological roots of human understanding.* Boston: Shambhala Publications

McCain, K. W. (1986). Cocited Author Mapping as a Valid Representation of Intellectual Structure. *Journal of the American Society for Information Science, 37*(3), 111-122.

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

McVeigh, M. (2002). Journal self-citation in the journal citation reports. Retrieved from http://thomsonreuters.com/products_services/science/free/essays/journal_self_citation_jcr/

Meadows, A. J. (1990). Theory in information science. *Journal of Information Science, 16*(1), 59-63.

Merton, R. K. (1949). *Social theory and social structure*. New York: Free Press.

Merton, R. K. (1957). Priorities in scientific discovery: a chapter in the sociology of science. *American Sociological Review, 22*(6), 635-659.

Merton, R. K. (1961). Singletons and multiples in scientific discovery: A chapter in the sociology of science. *Proceedings of the American Philosophical Society, 105*(5), 470-486.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56-63.

Merton, R. K. (1973). *The sociology of science: theoretical and empirical investigations*. Chicago: University of Chicago Press.

Merton, R. K. (1979). Foreword. In E. Garfield (Ed.), *Citation indexing—its theory and application in science, technology, and the humanities*. Philadelphia: ISI Press.

Merton, R. K., & Gaston, J. (1977). *The sociology of science in Europe*. Carbondale: Southern Illinois University Press.

Meadows, A. J. (1990). Theory in information science. *Journal of information science, 16*(1), 59-63.

Moed, H. F. (1989). Bibliometric measurement of research performance and Price's theory of differences among the sciences. *Scientometrics, 15*(5), 473-483.

Moed, H. F. (2005). *Citation analysis in research evaluation*. New York: Springer.

Moed, H. F. (2011). The source normalized impact per paper is a valid and sophisticated indicator of journal citation impact. *Journal of the American Society for Information Science and Technology, 62*(1), 211-213.

Moed, H. F., Burger, W. J. M., Frankfort, J., & Van Raan, A. F. J. (1985a). The application of bibliometric indicators: important field-and time-dependent factors to be considered. *Scientometrics, 8*(3), 177-203.

Moed, H. F., Burger, W., Frankfort, J., & Van Raan, A. F. J. (1985b). The use of bibliometric data for the measurement of university research performance. *Research Policy, 14*(3), 131-149.

Moed, H. F., Linmas, J., Nederhof, A. J., Zuccala, A., Illescas, C. L., & de Moya, F. (2009). Options for a comprehensive database of research outputs in Social Sciences and humanities *A Report to the Project Board of the Scoping Study 'Towards a Bibliometric Database for the Social Sciences and the Humanities*: Under the auspices of the Science Foundation and funded by ESRC, AHRC, ANR, NWO and DFG.

Moed, H. F., Luwel, M., & Nederhof, A. (2002). Towards research performance in the humanities. *Library trends, 50*(3), 498-520.

Moed, H. F., & Van Leeuwen, T. N. (1995). Improving the accuracy of Institute for Scientific Information's journal impact factors. *Journal of the American Society for Information Science, 46*(6), 461-467.

Moed, H. F., & Van Leeuwen, T. N. (1996). Impact factors can mislead. *Nature a-z index, 381*(6579), 186.

Moed, H. F., Van Leeuwen, T. N., & Reedijk, J. (1998). A new classification system to describe the ageing of scientific journals and their impact factors. *Journal of Documentation, 54*(4), 387-419.

Moed, H. F., Van Leeuwen, T. N., & Reedijk, J. (1999). Towards appropriate indicators of journal impact. *Scientometrics, 46*(3), 575-589.

Molotch, H. (1994). Going out. *Sociological Forum, 9*(2), 221-239.

Molyneux, R. E. (1986). Patterns, Processes of Growth, and the Projection of Library Size: A Critical Review of the Literature on Academic Library Growth. *Library and Information Science Research, 8*(1), 5-28.

Monge, P. R., & Contractor, N. S. (2003). *Theories of communication networks*. New York: Oxford University Press.

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American Sociological Review, 69*(2), 213-238.

Mora, E. (2007). Ma che cos' è la sociologia espressiva? *Sociologica, 1,* 1-9. doi: 10.2383/24198.

Moravcsik, M. J., & Murugesan, P. (1975). Some results on the function and quality of citations. *Social Studies of Science, 5*(1), 86-92.

Moya-Anegón, F., Herrero-Solana, V., & Jiménez-Contreras, E. (2006). A connectionist and multivariate approach to science maps: the SOM, clustering and MDS applied to library and information science research. *Journal of information science*, *32*(1), 63-77.

Mullins, N. C. (1972). The development of a scientific specialty: The phage group and the origins of molecular biology. *Minerva, 10*(1), 51-82.

Murray, D. (1917). *Bibliography: it's Scope and Method*. Glasgow: James Maclehose and Sons.

Murugesan, P., & Moravcsik, M. J. (1978). Variation of the nature of citation measures with journals and scientific specialties. *Journal of the American Society for Information Science, 29*(3), 141-147.

Nalimov, V. V., & Mulchenko, Z. M. (1969). *Naukometriya. Izucheniye razvitiya nauki kak informatsionnogo protsessa (Scientometrics. The Study of Science as an Information Process)*. Moskva: Nauka. (English translation: 1971. Washington, D.C.: Foreign Technology Division. U.S. Air Force Systems Command, Wright-Patterson AFB, Ohio).

Narin, F., Carpenter, M., & Berlt, N. C. (1972). Interrelationships of scientific journals. *Journal of the American Society for Information Science*, *23*(5), 323-331.

Narin, F. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity*. Cherry Hill, N.J: Computer Horizons.

Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics, 66*(1), 81-100.

Nederhof, A. J. (1989). Books and chapters are not to be neglected in measuring research productivity. *American Psychologist, 44*(4), 734–735.

Nederhof, A., & Erlings, C. (1993). A bibliometric study of productivity and impact of modern language and literature research in the Netherlands, 1982-1991 *Report CWTS 93-09*. Leiden.

Nederhof, A., & Van Wijk, E. (1997). Mapping the social and behavioral sciences world-wide: Use of maps in portfolio analysis of national research efforts. *Scientometrics, 40*(2), 237-276.

Nederhof, A. J., & Zwaan, R. A. (1991). Quality judgments of journals as indicators of research performance in the humanities and the social and behavioral sciences. *Journal of the American Society for Information Science, 42*(5), 332-340.

Nederhof, A. J., Zwaan, R. A., De Bruin, R., & Dekker, P. (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social and beha vioural sciences: A comparative study. *Scientometrics, 15*(5), 423-435.

Newman, M. E. J. (2004a). Detecting community structure in networks. *The European Physical Journal B-Condensed Matter and Complex Systems, 38*(2), 321-330.

Newman, M. E. J. (2004b). Fast algorithm for detecting community structure in networks. *Physical Review E, 69*(6), 066133.

Newman, M. E. J. (2004c). Analysis of weighted networks. *Physical Review E, 70*(5), 056131.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences, 103*(23), 8577-8582.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69*(2), 026113.

Nicholas, D., & Ritchie, M. (1978). *Literature and bibliometrics*. London: Clive Bingley.

Nicolaisen, J. (2002). The J-shaped distribution of citedness. *Journal of Documentation*, 58, 383-395.

Nicolaisen, J. (2007). Citation analysis. *Annual review of information science and technology, 41*(1), 609-641.

O'Connor, D., & Voos, H. (1981). Empirical laws, theory construction, and bibliometrics. *Library trends, 30*(1), 9-20.

Oppenheim, C., & Renn, S. P. (1978). Highly cited old papers and the reasons why they continue to be cited. *Journal of the American Society for Information Science, 29*(5), 225-231.

Otlet, P. (1934). *Traité de Documentation: le livre sur le livre; théorie et pratique*. Bruxelles: Mundaneum.

Otte, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *Journal of information science, 28*(6), 441-453.

Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature, 435*(7043), 814-818.

Parsons, T. (1951). *The social system*. Glencoe: Free Press.

Peritz, B. C. (1983). Are methodological papers more cited than theoretical or empirical ones? The case of sociology. *Scientometrics, 5*(4), 211-218.

Persson, O. (2001). All author citations versus first author citations. *Scientometrics, 50*(2), 339-344.

Persson, O., Danell, R., & Schneider, J. W. (2009). How to use Bibexcel for various types of bibliometric analysis. In F. Åström, R. Danell, B. Larsen & J. W. Schneider (Eds.), *Celebrating scholarly communication studies: A Festschrift for Olle Persson at his 60th Birthday* (pp. 9-24). Leuven, Belgium: International Society for Scientometrics and Informetrics.

Peters, H., & Van Raan, A. F. J. (1993). Co-word-based science maps of chemical engineering. Part I: Representations by direct multidimensional scaling. *Research Policy, 22*(1), 23-45.

Peters, H., & Van Raan, A. F. J. (1995). Cognitive resemblance of citation scores: A case study in chemical engineering. *Journal of the American Society for Information Science, 46*, 9-21.

Phillips, D. L. (1974). Epistemology and the sociology of knowledge: The contributions of Mannheim, Mills, and Merton. *Theory and Society, 1*(1), 59-88.

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.

Pizzorno, A. (1972). Una crisi che non importa superare. In P. Rossi (Ed.), *Ricerca sociologica e ruolo del sociologo* (pp. 327-353). Bologna: Il Mulino.

Polanco, X. (1995). Aux sources de la scientométrie. *Solaris, 2*, 13-79.

Polanco, X., François, C., & Lamirel, J. C. (2001). Using artificial neural networks for mapping of scienceand technology: A multi-self-organising-maps approach. *Scientometrics, 51*(1), 267-292.

Porter, A. L. (1977). Citation analysis: Queries and caveats. *Social Studies of Science, 7*(2), 257-267.

Potter, W. G. (1981a). Introduction. *Library trends, 30*(1), 5-7.

Potter, W. G. (1981b). Lotka's law revisited. *Library trends, 30*(1), 21-39.

Press, W., Flannery, B., & Teukolsky, V. WT (1989) Numerical recipes in Pascal. Cambridge: Cambridge University Press.

Price, D. J. D. (1963). *Little Science, Big Science*. New York: Columbia University Press.

Price, D. J. D. (1965). Networks of scientific papers. *Science, 149*(3683), 510-515.

Price, D. J. D. (1970). Citation measures of hard science, soft science, technology and non-science. In C. E. Nelson & D. K. Pollock (Eds.), *Communication among Scientists and Engineers* (pp. 3-22). Lexington, MA: Heath Lexington Books.

Price, D. J. D. (1975). *Science since Babylon*. New Haven: Yale University Press. (Original work published 1961).

Price, D. J. D. (1976). A general theory of bibliometric and other cumulative advantage process. *Journal of the American Society of Information Science, 27*(5-6), 292-306.

Price, D. J. D. (1981). Multiple authorship. *Science, 212*(4498), 986.

Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation, 25*(4), 348-349.

Pritchard, A. (1972). Bibliometrics and information transfer. *Research in Librarianship, 4*, 37-46.

Pritchard, A. (1984). *On the structure of information transfer networks.* M.Phil. thesis, Polytechnic of North London (UK), London.

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., & Parisi, D. (2004). Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America, 101*(9), 2658-2663.

Ranganathan, S. (1947). Classification and international documentation. *Review of Documentation, 14*(4), 154-177.

Rider, F. (1944). *The scholar and the future of the research library: a problem and its solution*. New York: Hadham.

Rip, A. J., & Courtial, J. P. (1984). Co-word maps of biotechnology: An example of cognitive scientometrics. *Scientometrics, 6*(6), 381-400.

Rip, A. J. (1988). Mapping of science: Possibilities and limitations. In A. F. J. Van Raan (Ed.), *Handbook of Quantitative Studies of Science and Technology* (pp. 253-273). Amsterdam: Elsevier.

Riviera, E. (in press). Scientific communities as autopoietic systems: the re-productive function of citations. *Journal of the American Society for Information Science and Technology.*

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, *105*(4), 1118-1123.

Roth, G. (1986). Selbstorganisation - Selbsterhaltung - Selbestreferentialitat. In A.Dress, H. Hendrichs & G. Küppers (Eds.), *Selbstorganisation – Die entstehung von ordnung in nature und gesellschaft*. Munich: Piper.

Rousseau, R. (1988). Citation distribution of pure mathematics journals. In L. Egghe & R. Rousseau (Eds.), *Informetrics 87/88.* Amsterdam: Elsevier.

Rousseau, R., & Zuccala, A. (2004). A classification of author co-citations: Definitions and search strategies. *Journal of the American Society for Information Science and Technology, 55*(6), 513-529.

Rubio, A. V. (1992). Scientific production of Spanish universities in the fields of Social Sciences and language. *Scientometrics, 24*(1), 3-19.

Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*. New York: McGraw Hill.

Santoro, M. (2007). Per una sociologia professionale e riflessiva (solo così anche pubblica). *Sociologica, 1*, 1-19. Retrieved from doi:10.2383/24199

Santoro, M. (2011). Esiste una «comunità scientifica» per la sociologia italiana. *Rassegna Italiana di Sociologia, 52*(2), 253-282.

Shadish, W. R., Tolliver, D., Gray, M., & Sen Gupta, S. K. (1995). Author judgements about works they cite: three studies from psychology journals. *Social Studies of Science, 25*(3), 477-498.

Schafer, W. (1983). *Finalization in Science: the Social Organisation of Scientific Progress*. Dordrecht: Reidel.

Schaeffer, S. E. (2007). Graph clustering. *Computer Science Review, 1*(1), 27-64.

Schneider, J. W., & Borlund, P. (2007a). Matrix comparison, Part 1: Motivation and important issues for measuring the resemblance between proximity measures or ordination results. *Journal of the American Society for Information Science and Technology, 58*(11), 1586-1595.

Schneider, J. W., & Borlund, P. (2007b). Matrix comparison, Part 2: Measuring the resemblance between proximity measures or ordination results by use of the Mantel and Procrustes statistics. *Journal of the American Society for Information Science and Technology, 58*(11), 1596-1609.

Schoepflin, U. (1990). *Problems of representativity in the Social Sciences citation index*. Paper presented at the Representations of Science and Technology, Proceedings of the International Conference on Science and Technology Indicators, Bielefeld, Germany.

Schubert, A., & Braun, T. (1996). Scientopography—World maps and charts based on scientometric indicators. A bird's eye view on the metropolis and beyond. In R. Waast (Ed.), *20th Century sciences: Beyond the metropolis* (Vol. 6, pp. 65-72). Paris: ORSTOM.

Schvaneveldt, R. W. (1990). *Pathfinder associative networks: Studies in knowledge organisation*. Nonvood, NJ: Ablex Publishing.

SciVerse Scopus. (2011). Content coverage guide. Retrieved from http://www.info.sciverse.com/scopus/scopus-in-detail/facts

Sengupta, I. (1992). Bibliometrics, informetrics, scientometrics and librametrics: an overview. *Libri, 42*(2), 75-98.

Shannon, C. E. (1948). A mathematical theory of communications, I and II. *The Bell System Technical Journal, 27,* 379-423.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

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

Shils, E. (1972). *The intellectuals and the powers and other essays*. Chicago: University of Chicago Press.

Shinn, T. (1988). Hiérarchies des chercheurs et formes des recherches. *Actes de la recherche en sciences sociales, 74*(1), 2-22.

Shinn, T. (2002). Intellectual cohesion and organisational divisions in science. *Revue francaise de sociologie, 43*(Supplement), 99-122.

Simmel, G. (1950). The field of sociology. In K. H. Wolff (Ed.), *The sociology of Georg Simmel* (pp. 3-23).

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.

Small, H. G. (1977). A co-citation model of a scientific specialty: A longitudinal study of collagen research. *Social Studies of Science, 7*(2), 139-166.

Small, H. G. (1978). Cited documents as concept symbols. *Social Studies of Science, 8*(3), 327-340.

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

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.

Small, H. (2004). On the shoulders of Robert Merton: Towards a normative theory of citation. *Scientometrics, 60*(1), 71-79.

Small, H. (2006). Tracking and predicting growth areas in science. *Scientometrics, 68*(3), 595-610.

Small, H. G., & Crane, D. (1979). Specialties and disciplines in science and social science: An examination of their structure using citation indexes. *Scientometrics, 1*(5), 445-461.

Small, H., & Garfield, E. (1985). The geography of science: disciplinary and national mappings. *Journal of information science, 11*(4), 147-159.

Small, H., & Griffith, B. C. (1974). The structure of scientific literatures I: Identifying and graphing specialties. *Science studies, 4*(1), 17-40.

Small, H., & Sweeney, E. (1985). Clustering the science citation index® using co-citations I. A comparison of methods. *Scientometrics, 7*(3), 391-409.

Smith, L. C. (1981). Citation analysis. Library Trends, *30,* 83-106.

Sosteric, M. (1999). Endowing mediocrity: Neoliberalism, information technology, and the decline of radical pedagogy. *Radical Pedagogy, 1*(1).

Stehr, N. (1974). Paradigmatic crystallization: Patterns of interrelations among areas of competence in sociology. *Social Science Information, 13*(1), 19-137.

Stevens, R. E. (1953). Characteristics of subject literatures. In Association of College and Reference Libraries (Ed.), *Publications Committee* (Vol. 6, pp. 10-21): Publications Committee of the Association of College and Reference Libraries.

Stichweh, R. (1990). Self-organising and autopoiesis in the development of modern Science. In W. Krohn, G. Küppers & H. Nowotny (Eds.), *Selforganisation. Portrait of a scientific revolutions* (pp. 195-207). Dordreht: Kluwer Academic Publishers.

Stinchcombe, A. L. (1994). Disintegrated disciplines and the future of sociology. *9*, 279-291.

Storer, N. W. (1966). *The social system of science*: Holt, Rinehart and Winston.

Storer, N. W. (1967). The hard sciences and the soft: some sociological observations. *Bulletin of the Medical Library Association, 55*(1), 75-84.

Storer, N. W. (1973). Introduction. In R. K. Merton (Ed.), *The sociology of science: Theoretical and empirical investigations*. Chicago: University of Chicago Press.

Sullivan, D., White, D. H., & Barboni, E. J. (1977). Co-citation analyses of science: An evaluation. *Social Studies of Science, 7*(2), 223-240.

Tague-Sutcliffe, J. (1992). An introduction to Informetrics. *Information processing & management, 28*(1), 1-3.

Thackray, A., & Merton, R. K. (1972). On discipline building: The paradoxes of George Sarton. *Isis, 63*(4), 473-495.

Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of information science, 34*(4), 605-621.

Thompson, J. W. (2002). The death of the scholarly monograph in the humanities? Citation patterns in literary scholarship. *Libri, 52*(3), 121-136.

Thomson Reuters, www.thomsonreuters.com

Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review, 38*(5), 406-427.

Todorov, R. (1990). Representing Canadian geophysics: A bibliometric approach. In L. Egghe & R. Rousseau (Eds.), *Informetrics 89/90: Selection of Papers Submitted for Second International Conference on Bibliometrics, Scientometrics, and Informetrics, London, Ontario, Canada, 5-7 July 1989* (pp. 291-307). Amsterdam: Elsevier Science Ltd.

Todorov, R., & Glänzel, W. (1988). Journal citation measures: a concise review. *Journal of information science, 14*(1), 47-56.

Tomer, C. (1986). A statistical assessment of two measures of citation: the impact factor and the immediacy index. *Information processing & management, 22*(3), 251-258.

Traag, V., Van Dooren, P., & Nesterov, Y. (2011). Narrow scope for resolution-limit-free community detection. *Physical Review E, 84*(1), 016114.

Tryon, R. C. (1939). *Cluster analysis*. New York: McGraw-Hill.

Turner, S. P., & Turner, J. H. (1990). *The impossible science: An institutional analysis of American sociology*. Newbury Park, CA: Sage.

Van den Berghe, H., Houben, J. A., de Bruin, R. E., Moed, H. F., Kint, A., Luwel, M., & Spruyt, E. H. J. (1998). Bibliometric indicators of university research performance in Flanders. *Journal of the American Society for Information Science, 49*(1), 59-67.

Van Eck, N. J. P., & Waltman, L. (2007a). Bibliometric mapping of the computational intelligence field. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 15*(5), 625-645.

Van Eck, N., & Waltman, L. (2007b). *VOS: A new method for visualizing similarities between objects.* Paper presented at the Advances in data analysis: Proceedings of the 30th annual conference of the German Classification Society Heidelberg.

Van Eck, N. J., & Waltman, L. (2008a). Generalizing the h-and g-indices. *Journal of Informetrics, 2*(4), 263-271.

Van Eck, N. J., & Waltman, L. (2008b). Appropriate similarity measures for author co-citation analysis. *Journal of the American Society for Information Science and Technology, 59*(10), 1653-1661.

Van Eck, N. J., & Waltman, L. (2009). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology, 60*(8), 1635-1651.

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84*(2), 523-538.

Van Eck, N.J., Waltman, L., Van den Berg, J., & Kaymak, U. (2006). Visualizing the computational intelligence field. *IEEE Computational Intelligence Magazine*, *1*(4), 6-10.

Van der Meulen, B., & Leydesdorff, L. (1991). Has the study of philosophy at Dutch universities changed under economic and political pressures? *Science, Technology & Human Values, 16*(3), 288-321.

Van Hooydonk, G. (1997). Fractional counting of multi-authored publications: consequences for the impact of authors. *Journal of the American Society for Information Science, 48*(10), 944–945.

Van Raan, A. F. J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics, 36*(3), 397-420.

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.

Van Raan, A. F. J. (2003). The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technikfolgenabschätzung–Theorie und Praxis, 1*(12), 20-29.

Van Raan, A. F. J. (2005). Reference-based publication networks with episodic memories. *Scientometrics, 63*(3), 549-566.

VINITI. http://www2.viniti.ru/

Waltman, L., Berg, J., Kaymak, U., & Eck, N. J. P. (2006). Visualizing the Computational Intelligence Field. *IEEE Computational Intelligence Magazine, 1*(4), 6-10.

Waltman, L., Van Eck, N. J. (2011). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology*, *63*(2), 406-415.

Waltman, L., Van Eck, N. J., & Noyons, E. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics, 4*(4), 629-635.

Wassermann, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* Cambridge: Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *Nature, 393*(6684), 440-442.

Weinstock, N. (1971). Citation indexes. In A. Kent (Ed.), *Encyclopedia of Library and Information Science* (Vol. 5, pp. 16-41). New York: Marcel Dekker.

Wellman, B., & Berkowitz, S. D. (1988). *Social structures: A network approach* (Vol. 2): Cambridge University Press.

Wells, H. G. (1938). *World brain*. Garden City: Doubleday Doran.

White, H. D. (1986). Co-cited author retrieval. *Information Technology and Libraries, 5*, 93-99.

White, H. D. (1990). Author co-citation analysis: Overview and defense. In C. L. Borgman (Ed.), *Scholarly Communication and Bibliometrics* (pp. 84-106). Newbury Park: Sage Publication.

White, H. D. (2003). Author cocitation analysis and Pearson's r. *Journal of the American Society for Information Science and Technology, 54*(13), 1250-1259.

White, H. D. (2003). Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists. *Journal of the American Society for Information Science and Technology*, *54*(5), 423-434.

White, D. R., & Harary, F. (2001). The cohesiveness of blocks in social networks: Node connectivity and conditional density. *Sociological Methodology, 31*(1), 305-359.

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

White, H. D., & Griffith, B. C. (1982). Authors and markers of intellectual space: Co-citation in studies of science, technology and society. *Journal of Documentation*, *38*, 255–272.

White, H. D., & McCain, K. W. (1997). Visualization of Literatures. *Annual Review of Information Science and Technology (ARIST), 32*, 99-168.

White, H. D., & McCain, K. W. (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.

Whitley, R. (1976). Umbrella and polytheistic scientific disciplines and their elites. *Social Studies of Science, 6*(3/4), 471-497.

Wilkinson, D. M., & Huberman, B. A. (2004). A method for finding communities of related genes. *Proceedings of the National Academy of Sciences of the United States of America, 101*(Suppl 1), 5241-5248.

Willems, J. (2003). Bringing down the barriers. *Nature, 422*(6931), 470-470.

Williams, N. (1997). UK Universities—Jostling for Rank. *Science, 275*(5296), 18-19.

Wilson, C. S. (1999). Informetrics. *Annual Review of Information Science and Technology (ARIST), 34*, 107-247.

Woolgar, S. (1991). Beyond the citation debate: towards a sociology of measurement technologies and their use in science policy. *Science and Public Policy, 18*(5), 319-326.

Wouters, P. (1999a). The citation culture. Ph.D. Thesis, Universiteit van Amsterdam, Amsterdam.

Wouters, P. (1999b). Beyond the holy grail: from citation theory to indicator theories. *Scientometrics, 44*(3), 561-580.

Wu, Q. (2010). The w-index: A measure to assess scientific impact by focusing on widely cited papers. *Journal of the American Society for Information Science and Technology, 61*(3), 609-614.

Yermish, I. (1975). *A citation based interactive associative information retrieval system.* Ph.D. Thesis, University of Pennsylvania, Philadelphia.

Yitzhaki, M. (1998). The language preference in sociology. *Scientometrics, 41*(1-2), 243-254.

Zhao, D. (2006). Towards all-author co-citation analysis. *Information Processing and Management, 42*, 1578-1591.

Zipf, G. K. (1935). *The psycho-biology of language: an introduction to dynamic philology*. Boston: Houghton Mifflin.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley: Cambridge

Zitt, M., Bassecoulard, E., & Okubo, Y. (2000). Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science. *Scientometrics, 47*(3), 627-657.

Zuccala, A. (2004). *Revisiting the invisible college: A case study of the intellectual structure and social process of singularity theory research in mathematics.* University of Toronto.

Zuccala, A. (2006). Author cocitation analysis is to intellectual structure as web colink analysis is to…? *Journal of the American Society for Information Science and Technology, 57*(11), 1487-1502.

Zuckerman, H. A. (1968). Patterns of name ordering among authors of scientific papers: A study of social symbolism and its ambiguity. *American Journal of Sociology*, *74*(3), 276-291.

Zuckerman, H. (1987). Citation analysis and the complex problem of intellectual influence. *Scientometrics*, *12*(5), 329-338.

Zuckerman, H., & Merton, R. K. (1972). Age, aging, and age structure in science. *Aging and society*, *3*, 292-356.