FURTHER STEPS IN INTEGRATING THE PLATFORMS OF *WoS* AND *SCOPUS*: HISTORIOGRAPHY WITH *HISTCITE*™ AND MAIN-PATH ANALYSIS

Avances para integrar las plataformas *WoS* y *Scopus*: historiografía con *HistCite*™ y análisis de camino principal

Loet Leydesdorff, Andreas Thor and Lutz Bornmann

Loet Leydesdorff (Ph.D. Sociology, M.A. Philosophy, and M.Sc. Biochemistry) is professor emeritus at the *Amsterdam School of Communications Research* (*ASCoR*) of the *University of Amsterdam*. He is associate faculty at the *Science and Technology Policy Research Unit* (*SPRU*) of the *University of Sussex*, visiting professor of the *Institute of Scientific and Technical Information of China* (*ISTIC*) in Beijing, guest professor at *Zhejiang University in Hangzhou*, and visiting fellow at the *School of Management*, Birkbeck, *University of London*. He has published extensively in systems theory, social network analysis, scientometrics, and the sociology of innovation. With Henry Etzkowitz, he initiated a series of workshops, conferences, and special issues about the Triple Helix of University-Industry-Government Relations. *Thomson Reuters*/*Clarivate* lists him as a highly-cited author since 2014.
http://www.leydesdorff.net/list.htm
http://www.leydesdorff.net/th2/index.htm
http://highlycited.com
http://orcid.org/0000-0002-7835-3098

University of Amsterdam, *Amsterdam School of Communication Research* (*ASCoR*)
PO Box 15793, 1001 NG Amsterdam, The Netherlands
loet@leydesdorff.net

Andreas Thor is full professor at the *Leipzig University of Telecommunications* (*HfTL*), a private university owned by *Deutsche Telekom*. Before joining *HfTL* in 2013, he was an acting professor at the *University of Passau* and a visiting research scientist at the *University of Maryland*. In 2008, he received a PhD in Computer Science from the *University of Leipzig* where he worked as a researcher at the database group for several years. Andreas' research interests are in the area of scalable data management. More specifically, he has been working on approaches for entity resolution, ontology alignment, and optimizing cloud technologies. He is regularly applying advanced data management techniques to research problems in the life sciences and bibliometrics.
http://orcid.org/0000-0003-2575-2893

University of Applied Sciences for Telecommunications Leipzig
Gustav-Freytag-Str. 43-45, 04277 Leipzig, Germany
thor@hft-leipzig.de

Lutz Bornmann works as a sociologist of science at the *Division for Science and Innovation Studies* in the *Administrative Headquarters of the Max Planck Society* in Munich (Germany). His research interests include research evaluation, peer review, bibliometrics, and altmetrics. Since 2004, he has published more than 250 papers in journals covered by *Clarivate Analytics Web of Science* with a total of more than 4,000 citations. His publication list and citation metrics are available at: *http://www.researcherid.com/rid/A-3926-2008 Clarivate Analytics* lists him among the most-highly cited researchers worldwide over the last ten years (since the first release of this service in 2014):
http://highlycited.com
http://orcid.org/0000-0003-0810-7091

Administrative Headquarters of the Max Planck Society, *Division for Science and Innovation Studies*,
Hofgartenstr. 8, 80539 Munich, Germany
bornmann@gv.mpg.de

Manuscript received on 01-07-2017
Approved on 24-07-2017

Abstract

The program *HistCite*™ enables an analyst to identify significant works on a given topic using the citation links between them diachronically. However, using *Scopus* data for drawing historiograms with *HistCite*™ has hitherto been a problem. In the new version of the program *CRExplorer*, one can translate citation data from *Scopus* to *WoS* formats (or *vice versa*) and then import the data into *HistCite*™. In this brief communication, we demonstrate these options using the papers of Eugene Garfield (1925-2017) in *Scopus* for main-path analysis. The two historiograms are considerably different: unlike the *WoS* set, the networked connections between the time lines are sparse in the representation of the *Scopus* data; the secondary documents (e.g., editorials in *Current Contents*) not processed in *Scopus*, but included in *WoS* enrich the representation. Furthermore, *HistCite*™ has an option to export the citation network as a *Pajek* file that can be read by most network analysis and visualization programs. Garfield's texts are centered by him as a personality and entrepreneur in different domains more than intellectually, although there are a number of recurring themes. Our own main paths are shaped along a line along which components indicate longer-term projects.

Keywords

Scopus; *HistCite*™; *CRExplorer*; Historiography; Eugene Garfield; Main path.

Resumen

El programa *HistCite*™ permite identificar trabajos significativos sobre un tema dado usando los enlaces de citas entre ellos diacrónicamente. Sin embargo, el uso de los datos de *Scopus* para dibujar historiogramas con *HistCite*™ ha sido hasta ahora problemático. En la nueva versión del programa *CRExplorer* se pueden traducir datos de citas de *Scopus* a formatos *WoS* (o viceversa) e importar los datos a *HistCite*™. En esta breve comunicación demostramos estas opciones usando los documentos de Eugene Garfield (1925-2017) en *Scopus* para el análisis del camino principal. Los dos historiogramas son considerablemente diferentes: a diferencia del conjunto de *WoS*, las conexiones en red entre las líneas de tiempo son escasas en la representación de los datos de *Scopus*. Los documentos secundarios (por ejemplo, editoriales en *Current Contents*) no procesados en *Scopus*, pero incluidos en la *WoS* enriquecen la representación. Además, *HistCite*™ tiene la opción de exportar la red de citas como un archivo *Pajek* que puede ser leído por la mayoría de los programas de análisis y visualización de la red. Los textos de Garfield son centrados por él como personalidad y empresario en diferentes dominios más que intelectualmente, aunque hay una serie de temas recurrentes. Los caminos principales de los propios autores están formados a lo largo de una línea a lo largo de la cual los componentes indican proyectos a más largo plazo.

Palabras clave

Scopus; *HistCite*™; *CRExplorer*; Historiografía; Eugene Garfield; Camino principal.

Leydesdorff, Loet; **Thor, Andreas**; **Bornmann, Lutz** (2017). "Further steps in integrating the platforms of *WoS* and *Scopus*: Historiography with *HistCite*™ and main-path analysis". *El profesional de la información*, v. 26, n. 4, pp. 662-670.

https://doi.org/10.3145/epi.2017.jul.10

1. Introduction

Following up on his initial work with Irving Sher in the 1960s (**Garfield**; **Sher**; **Torpie**, 1964), Eugene Garfield joined forces with Alexander Pudovkin in the early 2000s to further develop a program for the historiography of science, namely *HistCite*™ (**Garfield**; **Pudovkin**; **Istomin**, 2003).[1] *HistCite*™ reads records retrieved from the *Web of Science* (*WoS*) and generates a historiogram —that is, a visual representation of the historiography in terms of citation relations— based on the cited references among the papers in the sample. The program also allows for exporting the citation data in the network format of *Pajek* (**Leydesdorff**; **Bornmann**; **Comins**; **Marx**; **Thor**, 2016).[2] The *Pajek* format has become a kind of currency among programs for network analysis and visualizations. In social network analysis, **Hummon** & **Doreian** (1989); **Carley**, **Hummon**, & **Harty** (1993) have developed main-path analysis with a similar objective (**Liu**; **Lu**, 2016; **Batagelj**, 2003; **Batagelj**; **Doreian**; **Ferligoj**; **Kejzar**, 2014; **Lucio-Arias**; **Leydesdorff**, 2008).

Hitherto, data downloaded from *Scopus* could not easily be imported into *HistCite*™, *Pajek*, and other network analysis programs because of incompatibilities in the format. *Scopus* has provided an alternative to *WoS* since 2006. Like *WoS*, *Scopus* contains citation information, but based on a larger set of journals (**Leydesdorff**; **De-Moya-Anegón**; **De-Nooy**, 2016). The standardization of cited references (but also the address information) in *Scopus* is different from that of *WoS*. In the most recent version of *CRExplorer*,[3] however, the translation of *Scopus*' citation data into the *WoS* format and *vice versa* is flawless. In this brief communication, we show the possibilities for translating data across platforms using Garfield's œuvre as data; we compare the results with those based on our own publications.

The standardization of cited references (but also the address information) in *Scopus* is different from that of *WoS*

2. Eugene Garfield

Eugene Garfield (1925-2017) is best known for two major inventions: the *Science Citation Index* (**Garfield**, 1955; cf. **Cronin**; **Atkins**, 2000; **Wouters**, 2000) —later organized with similar citation indices into *WoS*— and the (2-year) journal impact factor (JIF) (**Garfield**, 1972; **Garfield**; **Sher**, 1963) —later extended to other journal indicators such as the 5-year JIF and the immediacy index (**Price**, 1970). In addition to these entrepreneurial initiatives, Garfield took a keen interest in developing the quantitative study of the sciences around the journal *Scientometrics* (**Garfield**, 1979; **Price**, 1978; cf. **Glänzel**; **Schubert**; **Schlemmer**, 2007) and by further developing tools for the mapping of the sciences (**Small**; **Garfield**, 1985) and their historical evolution (**Elkana** *et al*., 1978; **Leydesdorff**, 2010). He also served as President of the *American Society for Information Science and Technology* (1999-2000).

> Eugene Garfield (1925-2017) is best known for two major inventions: the *Science Citation Index* and the (2-year) journal impact factor (JIF)

Garfield's œuvre has been analyzed in a number of studies using *HistCite*™ and other routines[4] (**Bornmann**; **Haunschild**; **Leydesdorff**, in preparation; **Leydesdorff**, 2010; **Wouters**, 1999). Using data retrieved from *WoS*, **Bornmann** *et al.* (in preparation) note that 77.8% (n = 1,063) of Garfield's 1,558 papers included in *WoS* were editorials in *Current Contents*. Another 148 papers were published in *The Scientist*, a professional magazine for scientists founded by Garfield in 1986. Garfield's scholarly œuvre included in *WoS* contains the remaining 257 documents. Only 86 of these documents contain cited references. Documents which do not contain cited references can still be part of a citation network in their quality of being cited. Using *Scopus*, one retrieves approximately the same number of documents authored by Garfield as his scholarly contributions (n = 249).

Table 1. Summary statistics of the data used in the analysis.

		Scopus 8 author identities[5]	***WoS***
Garfield	N of documents (secondary documents)	249 (2,541)[6]	1,558 (257 in œuvre)
	Cited references after disambiguation	1,317 952	20,900 15,195
Leydesdorff	N of documents (secondary documents)	349 (1,062)	314
	Cited references after disambiguation	11,657 5,308	6,112 5,901
Bornmann	N of documents (secondary documents)	255 (438)	242
	Cited references after disambiguation	6,909 3,917	4,195 4,096

An additional 2,541 documents are indicated as "secondary documents;" these documents are cited in journals covered by *Scopus*, but not processed by *Scopus* as source documents of the database.

3. Import and export of data into and from *CRExplorer*

Data was downloaded from *Scopus* in the week of June 10, 2017. Table 1 provides summary statistics.

After exporting the retrieval (including the bibliographic references) from *Scopus* in csv-format, the *Scopus* output can be read into *CRExplorer*. By default, *CRExplorer* converts only those records which include cited references, but this can be changed in the Settings of the program (at File > Settings > Import/Export > Advanced Export Options > Include Publications without CRs in export). When *CRExplorer* writes the records into the *WoS* format, it modifies the header as follows: "FN Thomson Reuters *Web of Science* modified by *CRExplorer*". The resulting file can be read into *HistCite*™ after changing this header into "FN ISI Export Format".

CRExplorer was developed for disambiguation of the cited references (**Thor**; **Marx**; **Leydesdorff**; **Bornmann**, 2016a, 2016b). The program detects variants of the same cited reference, clusters them, and merges their occurrences. The clustering and merging of the data is especially important for *Scopus* data, because these are more heterogeneous than *WoS* data. The cited references contain more information in *Scopus* than *WoS*: all authors are listed, and the titles of the referenced publications are provided. However, these specifics increase the number of variants among the cited references. Furthermore, *Scopus* data sometimes contain fragmented cited references data which cannot be completely parsed into the bibliographic categories (for example when authors or volume numbers are missing in cited references).

> 2,541 Garfield's documents, indicated as "secondary documents" are cited in journals covered by *Scopus*, but not processed by *Scopus* as source documents of the database

The disambiguation of cited references is important for error control in citation analysis. Citation scores can thus be made more reliable. However, the usage of *CRExplorer* in this project is different: citation scores are attributed to units of analysis (e.g., authors, departments, or journals). This leads to attribute statistics and rankings. *SPSS*, for example, can be used for this type of analysis. Our usage in this study leads to network statistics using relations as units of analysis. From a graph-analytical perspective, disambiguation is less important because the relations among texts in a graph are unique. Using the disambiguation in *CRExplorer*, one risks generating cycles in the citation flows, as we shall see below.

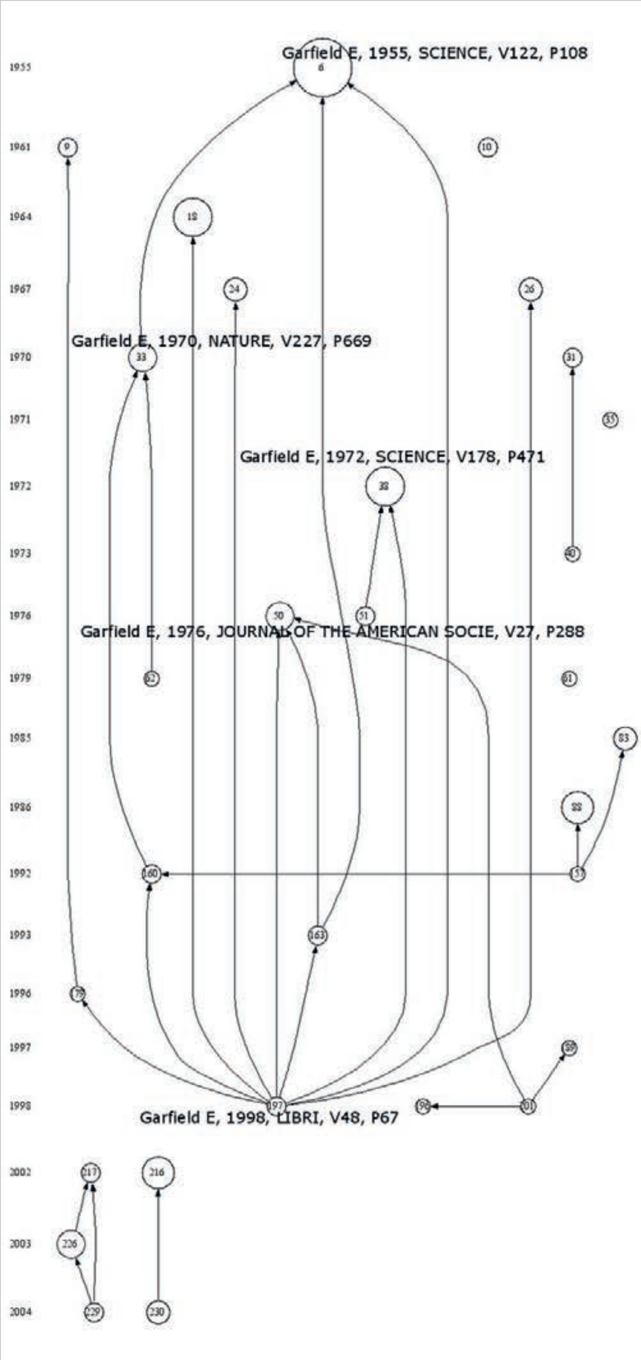


Figure 1. Historiogram of the 30 documents with highest Local Citation Scores in the œuvre of Eugene Garfield, using *Scopus* data.

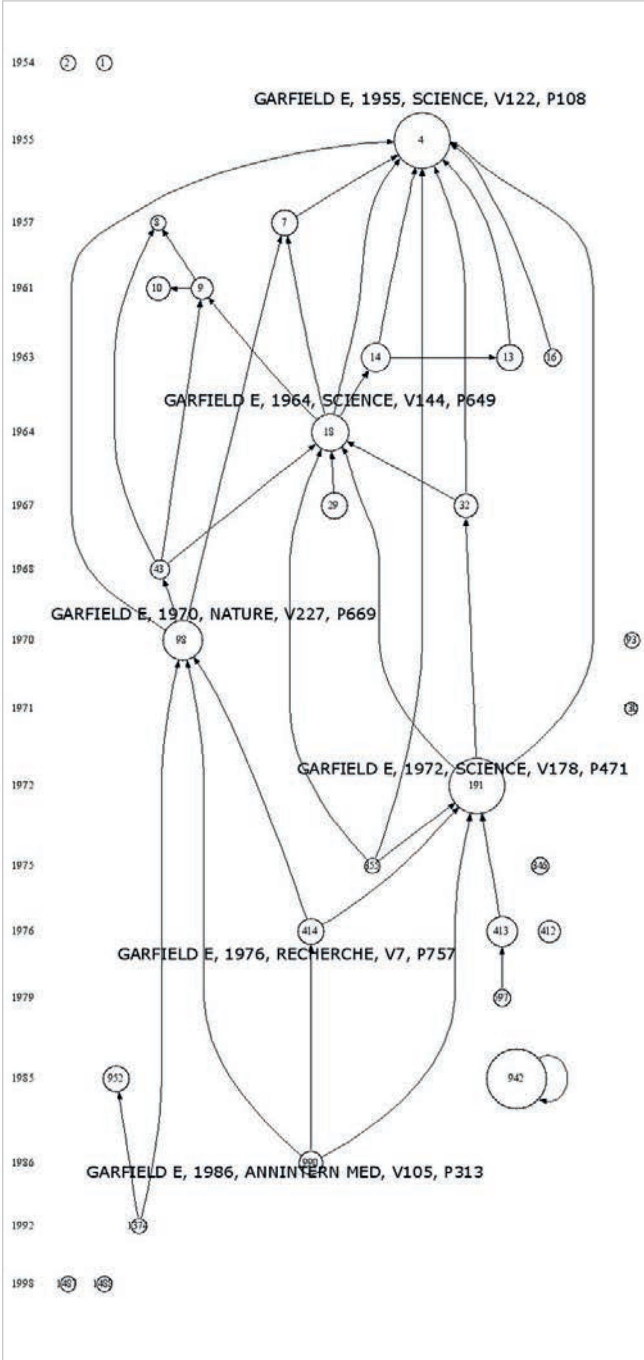


Figure 2. Historiogram of the 30 documents with highest Local Citation Scores in the œuvre of Eugene Garfield, using *WoS* data.

4. Use of data from *CRExplorer* in *HistCite™*

By default, *HistCite*™ generates a historiogram of the 30 papers with the highest citation scores within the set (the so-called "Local Citation Scores" or LCS). Within the drawing screen, one can also select "full" under Size for exporting the resulting graph as .png-file. A .png-file can be edited, for example, using *InkScape*, a freeware program that allows for embellishing the output. *InkScape* is available at
https://inkscape.org/en

In Figure 1, for example, the central nodes were labeled (by us) on the basis of the legends to the *HistCite*™ output. Figure 1 shows the historiogram of Garfield's work on the basis of *Scopus* data. Although the number of cited references is reduced by disambiguation from 1,317 to 952, the resulting historiogram is precisely the same as the one based on all cited references without disambiguation. Figure 2 shows the corresponding historiogram on the basis of the much larger set of *WoS* data.

The two historiograms are considerably different. Both originate from **Garfield**'s (1955) initial publication in *Science* proposing the *Science Citation Index*. **Garfield**'s (1972) article in *Science* introducing the JIF is indicated in both historiograms as well. Unlike the *WoS* set, however, the networked connections between the time lines are sparse in the representation of the *Scopus* data. The secondary documents (e.g.,

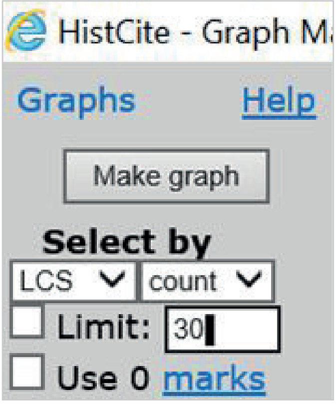


Figure 3. Unlimited export of *HistCite*™ results into *Pajek* format

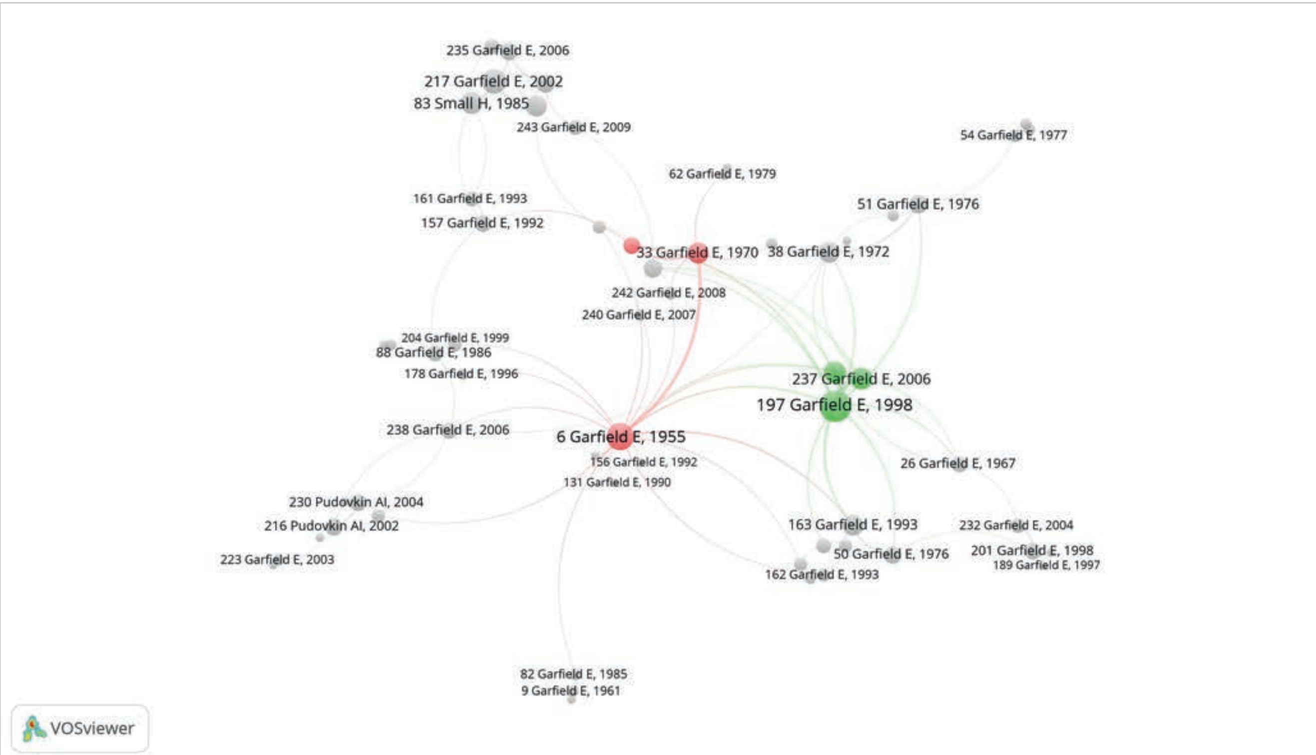
editorials in *Current Contents*) not processed in *Scopus*, but included in *WoS* enrich the representation.

5. Export from *HistCite*™ to *Pajek*: main-path analysis

HistCite™ graphs can be exported in the *Pajek* format. By default the historiogram shown on the screen, is exported. However, one can untick the limit in the left-top corner of the pane and ask for the full set (Figure 3). Even in the case of moderately large sets such as the ones under study above (in section 3), *HistCite*™ fails to generate a graph,[7] but one nevertheless can export the full set in the network format.

Within *Pajek* (or a similar program) one can pursue network analysis and visualization. Using the largest component of networked papers (n = 61) within the *Scopus* set (n = 249), one can, for example, reconstruct the so-called "main path." Main-path analysis calculates the extent to which a particular citation is needed for linking articles as a transversal weight divided by the total number of paths between a citing and cited document in the citation network (**De-Nooy** *et al*., pp. 281 ff.). Since the analysis of the main path was originally proposed by **Hummon** & **Doreian** (1989); **Carley**, **Hummon**, & **Harty** (1993), a large number of variants of main-path analysis have been elaborated (**Batagelj** *et al*., 2014; **Liu**; **Lu**, 2012). For example, one can also count transversal weights of nodes (instead of links).

Figure 4, for example shows the standard (global) main paths for this set after weighing the citations using the SPC (search path counts) algorithm.[8] We used *VOSviewer* for the visualization of the *Pajek* output.[9] The differences with other methods of main-path analysis are often marginal,



6 Garfield, E. (1955). "Citation indexes for science". *Science*, v. 122, n. 3159, pp. 108-111.

33 Garfield, E., (1970). "Citation indexing for studying science". *Nature*, v. 227, n. 5259, pp. 669-671.

160 Garfield, E., Welljams-Dorof A. (1992). "Of Nobel class: A citation perspective on high impact research authors". *Theoretical medicine*, v. 13, n. 2, pp. 117-135.

197 Garfield, E. (1998). "From citation indexes to informetrics: Is the tail now wagging the dog?". *Libri*, v. 48, n. 2, pp. 67-80.

237 Garfield, E. (2006). "Commentary: Fifty years of citation indexing". *International journal of epidemiology*, v. 35, n. 5, pp. 1127-1128.

248 Garfield, E. (2016). "The evolution of the Web of Science from the Science Citation Index". *BiD*, n. 37.

Figure 4. The main path in the largest component (N = 61) of Garfield's œuvre based on *Scopus* data; two components distinguished (*Q* = 0.279); *VOSviewer* used for the layout and visualization.[10]

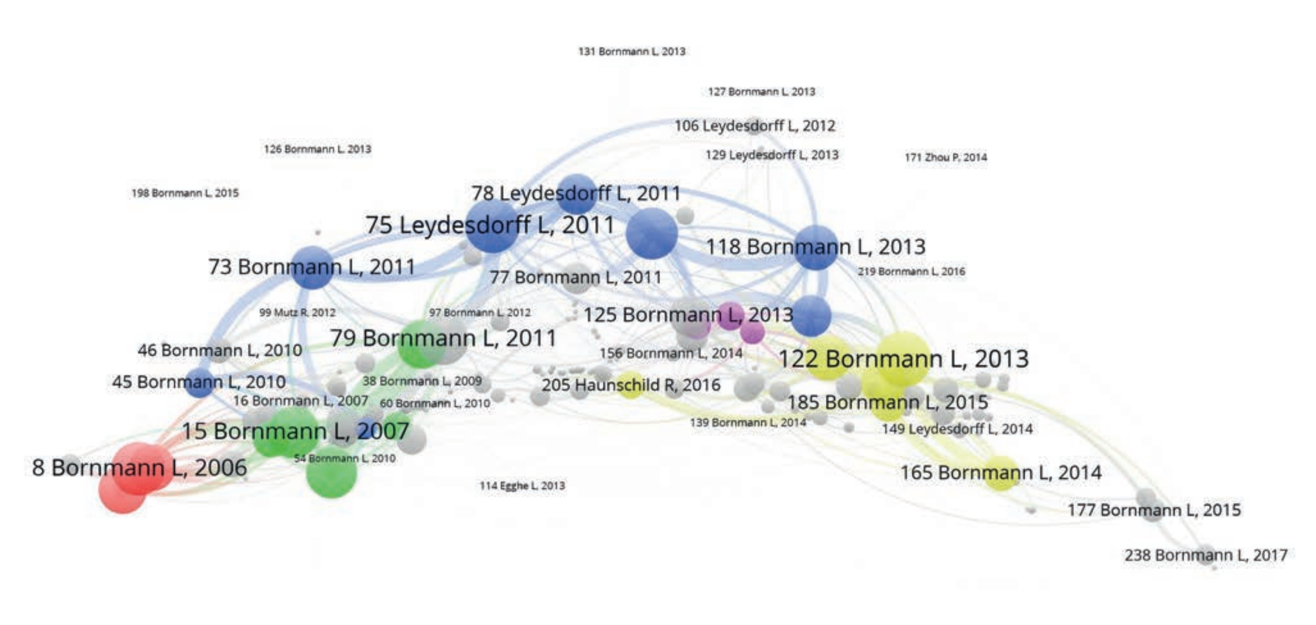


Figure 5. 25 documents on the main path among 233 documents (co-)authored by Bornmann and forming the largest component among 255 documents retrieved from *Scopus* (Q = 0.69).

and discussing them would lead us beyond the scope of this communication.

In accordance with the strongly decentralized patterns in this set, the main path is indicated by only six papers (among the 249 in total). Two components are distinguished when using a community-finding algorithm (Q = 0.28;[11] **Blondel** *et al*., 2008): the early part is oriented to the shaping and legitimation of citation analysis, and the later part (since 1998) is reflexive on the use of citation analysis in research and research evaluations. In summary, Garfield's texts are centered by him as a personality and entrepreneur more than intellectually, although there are a number of recurring themes. The largest component in the citation network contains only 61 of the 249 (24.5%) documents (without disambiguation of the cited references, the largest component contains 60 documents.)

As a further exploration, Figures 5 and 6 show similarly constructed main paths among the publications of two of the authors of this paper: Bornmann and Leydesdorff, respectively. Both main paths show a flow of related and overlapping research lines evolving over time. The largest components contain 92.4% and 95.07% of the sets, respectively, as against 24.5% for Garfield. The linear sequencing within both sets makes it possible to distinguish components indi-

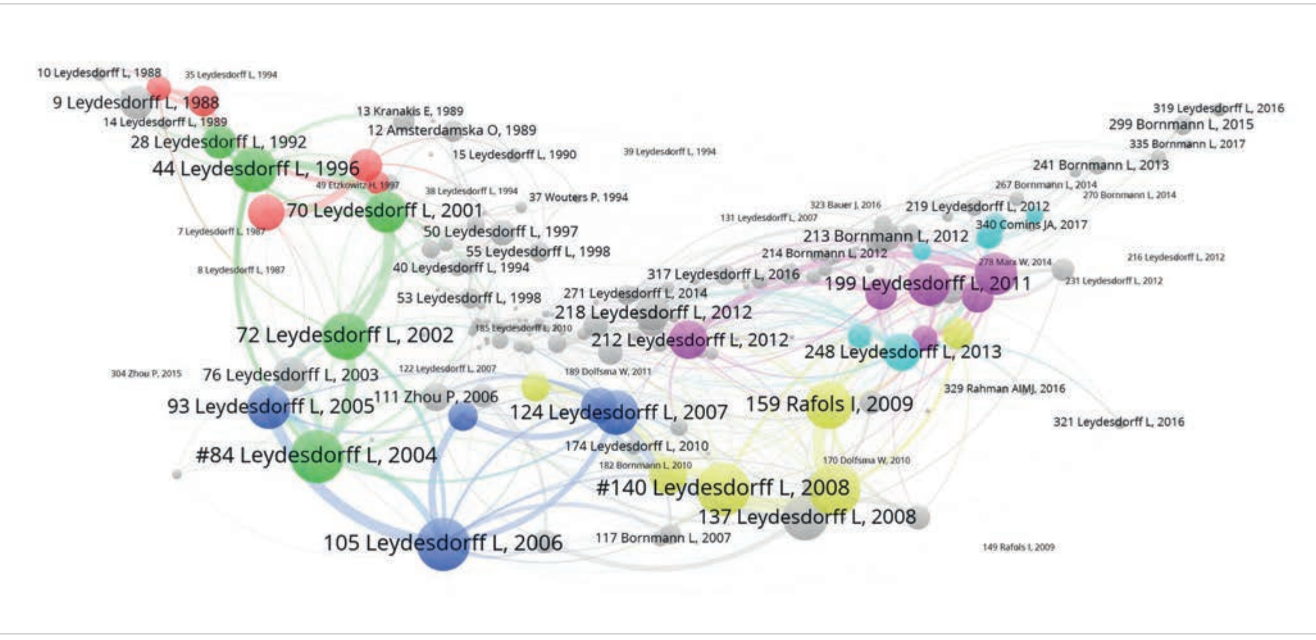


Figure 6. 34 documents on the main path among 326 documents (co-)authored by Leydesdorff and forming the largest component among 349 documents retrieved from *Scopus* (Q = 0.71).

cating periods with a modularity of 0.69 and 0.71, respectively, as against 0.28 for the star-shaped pattern in the case of Garfield's œuvre.

In Figure 6, documents #84 and #140 are marked with a hash, indicating that these references represent cycles which were shrunk (in *Pajek*; **Batagelj** *et al*., 2014). These cycles are generated by the disembiguation in *CRExplorer*.[12] Without disambiguation, however, the citation network is acyclic, the largest component in the set is 326, and the main path contains 33 (instead of 34) documents. In other words, the two functions of *CRExplorer* (disambiguation and translation between *Scopus* and *WoS*) serve different purposes and should not be confused. The overall purpose of *CRExplorer* is data analysis which is supported by several import/export formats as well as data cleaning and transformation routines (e.g., disambiguation, filtering / deleting, etc.).

> *CRExplorer* was developed for disambiguation of the cited references, detecting variants of the same cited reference, clustering them, and merging their occurrences

6. Limitations

One major limitation of using *HistCite*™ for generating network data is that the exported graphs are non-valued; each link is counted as one. In citation analysis, one may wish to work with valued networks. The same data in *WoS* format — either taken directly from *WoS* or processed via *CRExplorer* retrieved from *Scopus*— can alternatively be used as input to *CitNetw.exe* for generating a valued graph. Available at:
http://www.leydesdorff.net/software/citnetw

In the current version of *CRExplorer*, the translation of address information from *Scopus* into the *WoS* format is not yet fully reliable. We expect to improve this translation further in a next version of *CRExplorer*.

7. Conclusion

We have demonstrated how the new option in *CRExplorer* enables users to cross the divide between *Scopus* and *WoS*. The possibility to use both *WoS* and *Scopus* data is relevant for users who have access to only one of the databases, but may also be important for studying the reliability of results, the sensitivity to data choices (**Leydesdorff**; **De-Moya-Anegón**; **De-Nooy**, 2016), and the exchange and reproducibility of results among researchers (e.g., **Leydesdorff**; **Wagner**; **Bornmann**, 2016; **Vilchez-Román**, 2017).

Using the data for three researchers, we have shown that one of them (Garfield) entertained a star-shaped citation network. The main path was short: from one cluster including important papers in the early stage of the author's carreer to another containing more reflexive and programmatic publications. Bornmann and Leydesdorff's main paths tend to be shaped along a line leading to a much higher modularity ($Q \sim 0.7$ instead of 0.3). The main path organizes the majority of the papers along a line connecting longer-term research projects that can be visualized as components. However, Garfield's attention is distributed given his roles as an entrepreneur, a scholar, a journalist and editor, and a science-policy maker.

In our opinion, the relative ease with which the network characteristics can be retrieved makes it interesting to add this second-order perspective of network analysis in terms of links to the analysis and ranking of nodes in terms of attributes such as citation scores or the *h*-index values. The unit of analysis is then the communication as different from the communicators (**Luhmann**, 1996).

> By default, *HistCite*™ generates a historiogram of the 30 papers with the highest citation scores within a set (the so-called "Local Citation Scores" or LCS)

Notes

1. *HistCite*™ is freely available upon registration at
http://ip-science.thomsonreuters.com/thanks/HistCite

2. *Pajek* is a network analysis and visualization program, freely available for non-commercial usage at
http://mrvar.fdv.uni-lj.si/pajek

3. *CRExplorer* is freely available at
http://www.crexplorer.net

4. *CitNetExplorer* is software allowing for similar reconstructions on the basis of citation data. The program is freely available at
http://www.citnetexplorer.nl

5. The eight author identities of Eugene Garfield in *Scopus* are combined in the following search string: AU-ID ("Garfield, Eugene" 7005088140) OR AU-ID ("Garfield, Eugene" 56216352000) OR AU-ID ("Garfield, Eugene" 16065644600) OR AU-ID ("Garfield, Eugene" 56216352300) OR AU-ID ("Garfield, Eugene" 56907144200) OR AU-ID ("Garfield, Eugene" 57189553610) OR AU-ID ("Garfield, Eugene" 56904699100) OR AU-ID ("Garfield, Eugene" 57025772400)

6. Secondary documents in *Scopus* are cited, but not processed as source documents ("citing").

7. An alternative to *HistCite*™ is provided by *CitNetExplorer* at
http://www.citnetexplorer.nl

8. SPC is one among the algorithms which can be used for main path analysis (**Batagelj** *et al.*, 2014, pp. 76 ff.). Three algorithms are included in *Pajek* (**Batagelj**, 2003): (i) Search Path Counts (SPC) accounts for the highest traversal weight of the arcs connecting a source with a sink; (ii) the Search Path Link Count (SPLC) accounts for the number of all possible search paths through the network emanating from an origin; (iii) the Search Path Node Pair (SPNC) accounts for all connected vertex pairs along the paths (**Hummon**; **Doreian**, 1989, pp. 50-51).

9. *VOSviewer* is a network visualization program available at
http://www.vosviewer.com

VOSviewer foregrounds and backgrounds labels so that cluttering of labels is prevented and readability enhanced.

10. *VOSviewer* prevents cluttering of the labels by back- and foregrounding them.

11. *Q* is a measure of modularity provided, among others, by the *Louvain algorithm* (**Blondel** *et al.*, 2008).

12. Node 84, for example, contains a reference to Node 88 as "in print," and vice versa Node 88 refers to Node 84 as "forthcoming." *CRExplorer* makes these references complete by merging them correctly (!) into the corresponding clusters as (**Leydesdorff**, 2004a) and (**Leydesdorff**, 2004b), respectively. However, this mutual citation generates a cycle whereas main-path analysis is only possible on acyclic networks. Similarly, Node 140 (**Leydesdorff**; **Schank**, 2008) and Node 145 (**Leydesdorff**; **Schank**; **Scharnhorst**; **De-Nooy**, 2008) generate a cycle after completion of the references by *CRExplorer*.

> Since the analysis of the main path was originally proposed by Hummon & Doreian (1989); Carley, Hummon, & Harty (1993), a large number of variants of main-path analysis have been elaborated

8. References

Batagelj, Vladimir (2003). *Efficient algorithms for citation network analysis*.
http://arxiv.org/abs/cs/0309023

Batagelj, Vladimir; **Doreian, Patrick**; **Ferligoj, Anuška**; **Kejzar, Nataša** (2014). *Understanding large temporal networks and spatial networks: Exploration, pattern searching, visualization and network evolution*. Chichester, UK: John Wiley & Sons. ISBN: 978 1 118922279

Blondel, Vincent D.; **Guillaume, Jean-Loup**; **Lambiotte, Renaud**; **Lefebvre, Etienne** (2008). "Fast unfolding of communities in large networks". *Journal of statistical mechanics: Theory and experiment*, v. 8, n. 10, p. 10008.
http://iopscience.iop.org/article/10.1088/1742-5468/2008/10/P10008/pdf

Bornmann, Lutz; **Haunschild, Robin**; **Leydesdorff, Loet** (in preparation). Reference Publication Year Spectroscopy (RPYS) of Eugene Garfield's publications.
https://doi.org/10.7717/peerj-cs.32

Carley, Kathleen M.; **Hummon, Norman P.**; **Harty, Martha** (1993). Scientific influence: An analysis of the main path structure in the Journal of conflict resolution. *Science communication*, v. 14, n. 4, pp. 417-447.
https://doi.org/10.1177/107554709301400406

Cronin, Blaise; **Atkins, Helen B.** (eds.) (2000). *The Web of Knowledge: A festschrift in honor of Eugene Garfield*. Medford, NJ: Information Today. ISBN: 1 57387 099 4

De Nooy, Wouter; **Mrvar, Andrej**; **Batgelj, Vladimir** (2011). Exploratory social network analysis with *Pajek* (2nd ed.). New York, NY: Cambridge University Press. ISBN: 0 521174805

Elkana, Yehuda; **Lederberg, Joshua**; **Merton, Robert K.**; **Thackray, Arnold**; **Zuckerman, Harriet** (1978). *Toward a metric of science: The advent of science indicators*. New York, etc.: Wiley & Sons, Inc. ISBN: 0 471 98435 3

Garfield, Eugene (1955). "Citation indexes for science: A new dimension in documentation through association of ideas". *Science*, v. 122, n. 3159, pp. 108-111.
https://doi.org/10.1126/science.122.3159.108

Garfield, Eugene (1972). "Citation analysis as a tool in journal evaluation". *Science*, v. 178, n. 4060, pp. 471-479.
http://www.garfield.library.upenn.edu/essays/V1p527y1962-73.pdf

Garfield, Eugene (1979). *Citation indexing: Its theory and application in science, technology, and humanities*. New York: John Wiley. ISBN: 978 0 471025597

Garfield, Eugene; **Pudovkin, Alexander I.**; **Istomin, V. I.** (2003). "Mapping the output of topical searches in the Web of Knowledge and the case of Watson-Crick". *Information technology and libraries*, v. 22, n. 4, pp. 183-187.
http://garfield.library.upenn.edu/papers/casewatsoncrick2003.pdf

Garfield, Eugene; **Sher, Irving H.** (1963). "New factors in the evaluation of scientific literature through citation indexing". *American documentation*, v. 14, n. 3, pp. 195-201.
http://garfield.library.upenn.edu/essays/v6p492y1983.pdf

Garfield, Eugene; **Sher, Irving H.**; **Torpie, Richard J.** (1964). *The use of citation data in writing the history of science*. Philadelphia, PA: Institute for Scientific Information.
http://www.scimaps.org/exhibit/docs/Garfield1964use.pdf

Glänzel, Wolfgang; **Schubert, András**; **Schlemmer, Balázs** (eds.) (2007). *The multidimensional world of Tibor Braun*. Leuven: International Society for Scientometrics and Informetrics (ISSI).

Hummon, Norman P.; **Doreian, Patrick** (1989). "Connectivity in a citation network: The development of DNA theory". *Social networks*, v. 11, n. 1, pp. 39-63.
http://garfield.library.upenn.edu/papers/hummondoreian1989.pdf

Leydesdorff, Loet (2004a). "Clusters and maps of science journals based on bi-connected graphs in the Journal Citation Reports". *Journal of documentation*, v. 60, n. 4, pp. 371-427.
https://arxiv.org/abs/0912.1221
https://doi.org/10.1108/00220410410548144

Leydesdorff, Loet (2004b). "Top-down decomposition of the Journal Citation Reports of the Social Sciences Citation Index: Graph- and factor-analytical approaches". *Scientometrics*, v. 60, n. 2, pp. 159-180.
https://goo.gl/RKAM2f

Leydesdorff, Loet (2010). "Eugene Garfield and algorithmic historiography: Co-words, co-authors, and journal names". *Annals of library and information studies*, v. 57, n. 3, pp. 248-260.

https://goo.gl/7V3xdq

Leydesdorff, Loet; **Bornmann, Lutz**; **Comins, Jordan A.**; **Marx, Werner**; **Thor, Andreas** (2016). Referenced publication year spectrography (RPYS) and algorithmic historiography: A bibliometric reconstruction of András Schubert's *oeuvre*. In: W. Glänzel & B. Schlemmer (eds.), *András Schubert -A world of models and metrics*. Louvain: ISSI, pp. 79-96.
https://arxiv.org/abs/1604.04705

Leydesdorff, Loet; **De-Moya-Anegón, Félix**; **De-Nooy, Wouter** (2016). "Aggregated journal-journal citation relations in *Scopus* and Web-of-Science matched and compared in terms of networks, maps, and interactive overlays". *Journal of the Association for Information Science and Technology*, v. 67, n. 9, pp. 2194-2211.
https://doi.org/10.1002/asi.23372

Leydesdorff, Loet; **Schank, Thomas** (2008). "Dynamic animations of journal maps: Indicators of structural changes and interdisciplinary developments". *Journal of the American Society for Information Science and Technology*, v. 59, n. 11, pp. 1810-1818.
https://doi.org/10.1002/asi.20891

Leydesdorff, Loet; **Schank, Thomas**; **Scharnhorst, Andrea**; **De-Nooy, Wouter D.** (2008). "Animating the development of social networks over time using a dynamic extension of multidimensional scaling". *El profesional de la información*, v. 17, n. 6, pp. 611-626.
https://doi.org/10.3145/epi.2008.nov.04

Leydesdorff, Loet; **Wagner, Caroline**; **Bornmann, Lutz** (2016). "Replicability and the public/private divide" [Letter]. *Journal of the Association for Information Science and Technology*, v. 67, n. 7, pp. 1777-1778.
https://doi.org/10.1002/asi.23672

Liu, John S.; **Lu, Louis Y. Y.** (2012). "An integrated approach for main path analysis: Development of the Hirsch index as an example". *Journal of the American Society for Information Science and Technology*, v. 63, n. 3, pp. 528-542.
https://doi.org/10.1002/asi.21692

Lucio-Arias, Diana; **Leydesdorff, Loet** (2008). "Main-path analysis and path-dependent transitions in *HistCite*™-based historiograms". *Journal of the American Society for Information Science and Technology*, v. 59, n. 12, pp. 1948-1962.
https://doi.org/10.1002/asi.20903

Luhmann, Niklas (1996). "On the scientific context of the concept of communication". *Social science information*, v. 35, n. 2, pp. 257-267.
https://doi.org/10.1177/053901896035002005

Price, Derek J. de Solla (1970). "Citation measures of hard science, soft science, technology, and nonscience". In: C. E. Nelson & D. K. Pollock (eds.), *Communication among scientists and engineers* (pp. 3-22). Lexington, MA: Heath.

Price, Derek J. de Solla (1978). "Editorial statement". *Scientometrics*, v. 1, n. 1, pp. 7-8.
https://doi.org/10.1007/BF02016836

Small, Henry; **Garfield, Eugene** (1985). "The geography of science: disciplinary and national mappings". *Journal of information science*, v. 11, n. 4, pp. 147-159.
https://doi.org/10.1177/016555158501100402

Thor, Andreas; **Marx, Werner**; **Leydesdorff, Loet**; **Bornmann, Lutz** (2016a). "Introducing CitedReferencesExplorer: A program for reference publication year spectroscopy with cited references disambiguation". *Journal of informetrics*, v. 10, n. 2, pp. 503-515.
https://arxiv.org/abs/1601.01199
http://doi.org/10.1016/j.joi.2016.02.005

Thor, Andreas; **Marx, Werner**; **Leydesdorff, Loet**; **Bornmann, Lutz** (2016b). "New features of CitedReferencesExplorer (*CRExplorer*)". *Scientometrics*, v. 109, n. 3, pp. 2049-2051.
https://arxiv.org/abs/1607.01266
https://doi.org/10.1007/s11192-016-2082-3

Vilchez-Román, Carlos (2017). "BRICS scientific excellence and the search for relevance and replicability". *Journal of the Association for Information Science and Technology*, v. 68, n. 2, p. 539.
https://doi.org/10.1002/asi.23654

Wouters, Paul (1999). *The citation culture*. Amsterdam: Unpublished Ph.D. Thesis, University of Amsterdam.
http://garfield.library.upenn.edu/wouters/wouters.pdf

Wouters, Paul (2000). "Garfield as alchemist". In: B. Cronin & H. B. Atkins (eds.), *The Web of Knowledge: A festschrift in honor of Eugene Garfield* (vol. ASIS Monograph series, pp. 65-72). Medford, NJ: Information Today. ISBN: 1 57387 099 4