Fakultät VII
Wirtschaft und Management
Einstein Center Digital Future
Trust in Digital Services
Prof. Dr. Timm Teubner

Bachelorarbeit

Applied bibliometric analysis for the assessment of literature corpora

Vorgelegt von:

Edison Alain von Matt
Matrikelnummer: 383430
Tel.: +49 1525 3772172
E-Mail: Edisonvonmatt@gmail.com
Studiengang: Wirtschaftsingenieurwesen (B.Sc.)

Berlin, den 15. März 2022

Erstprüfer:
Prof. Dr. Timm Teubner (TU Berlin)

Zweitprüfer:
M.Sc. Florence Jaouat (TU Berlin)

Betreuer:
Prof. Dr. Timm Teubner (TU Berlin)

Eidesstattliche Erklärung

Hiermit gebe ich eine eidesstattliche Erklärung ab, dass ich die vorliegende Arbeit selbstständig verfasst und dabei keine anderen als die angegebenen Hilfsmittel benutzt habe. Sämtliche Stellen der Arbeit, die im Wortlaut oder dem Sinn nach Publikationen oder Vorträgen anderer Autoren entnommen sind, habe ich als solche kenntlich gemacht.

Die Arbeit wurde bisher weder gesamt noch in Teilen einer anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Berlin, den 7. März 2022

....

Unterschrift

Abstract

A growing number of systematic literature reviews employ bibliometric analyses to summarize the intellectual structure of a research field by analyzing the social and structural relationships between research constituents using network graph representations. This study aims to provide reference values for future research defining the density, connectivity, and clustering of network graphs computed from bibliometric data of literature reviews.

Therefore, a quantitative analysis was conducted summarizing relevant macro-level metrics of different graph samples created from the corpora of 14 selected literature reviews within the information science field. Confidence intervals for the median were employed to draw inferences on the population's central tendency of utilized metrics. The several metric distributions inherited varying dispersions and thus led to wider or narrower confidence intervals. Regarding the main results, it can be assumed that the publications included in a literature review within the field of information science have a common reference share of 1.1% to 2.6%, and their reciprocal citing density ranges from 3.9% to 11.5%.

Additionally, a correlation analysis investigated the influence of the corpus size and temporal disposition of included publications on the final graph's composition. While scatter plots and correlation coefficients indicated slight impacts of these conditions on the network graphs structure, there was no evidence of statistical significance.

This thesis provides a benchmark and allows further insights into the bibliometric analysis of literature reviews at a network level. However, further research is required to aggregate more results and enhance the statistical certainty of the assumptions made in this thesis.

Contents

List of Figures

List of Tables

1. Introduction

Systematic literature reviews are crucial in aggregating past research findings and guide following research advances [19]. However, due to their qualitative approach that provides room for biases [41, 68], quantitative methods are needed to introduce consistent, transparent, and replicable processes that help to mitigate subjective influences [83].

Throughout the last decade, there has been a significant increase in systematic reviews applying bibliometric analysis as a complementary quantitative method [22]. Bibliometric analysis helps summarize a field's intellectual structure by analyzing social and structural relationships among research constituents by employing network graphs, in which publications represent nodes and edges define their citing behavior [19]. However, there is a lack of research analyzing the bibliometric data of systematic reviews on a macro level [81]. Recent studies employed bibliometric methods in literature reviews to examine the importance and role of singular research publications and the thematic affiliation of groups they are building without concluding the composition of the whole corpus [81].

This thesis addresses this gap and pursues the question of how dense, connected, and clustered one would expect bibliometric network graphs to be on a macro level. In doing so, this research aims to provide a benchmark reference for future researchers to classify their bibliometric graph results computed from literature review corpora on the network level.

Sample graphs are created using the bibliometric data of 14 literature reviews published in the AIS College Senior Scholar's basket of eight journals. The extracted data for each corpus consists of a review's directly included publications and their references. Subsequently, this quantitative study analyzes the structural tendencies of network graphs by summarizing employed metrics with suitable statistical methods.

This thesis is composed of the following chapters:

First, relevant concepts for this work are explained in the second chapter. This includes the integration of bibliometric analysis into systematic literature reviews, the history, and the science behind methods and techniques employed to generate graph representations of scholarly knowledge referred to as science mapping. The theoretical chapter additionally introduces current technologies and databases relevant to bibliometric data retrieval.

Subsequently, the empirical part of this thesis starts with the formulations of hypotheses regarding the influence of the literature reviews size and temporal dispersion among included publications on the network graph's structure based on assumptions drawn from the theoretical chapter.

The following methodology chapter explains the different steps of this quantitative research in detail, including the initial selection of literature reviews, bibliometric data extraction, network graphs construction, graph visualization, explanation of employed metrics, and summarizing statistical analysis.

After that, results are presented, including tables and visualizations that summarize the computed graph metrics of each literature review. At the same time, conspicuous values are highlighted, and a correlation analysis provides insights about the validity of the formulated hypotheses.

The final discussion section makes inferences about the central tendency of the density, the connectedness, and the thematic clustering of bibliometric network graphs computed from literature reviews. It also attempts to explain the different variances among metrics while critically assessing the implications of the correlation results. Lastly, this chapter describes the limitations of this thesis and evaluates the validity of the presented statistics.

2. Theoretical Framework

2.1. Systematic Literature Reviews and Bibliometric Analysis

Systematic literature reviews and meta-analysis have traditionally been used to aggregate past research findings, highlight key developments, describe the current research front, and guide the following research advances [19]. The qualitative approach of systematic reviews allows for the examination of various studies and methodological approaches and provides an in-depth analysis of the literature and contextual issues [56, 83]. On the other hand, meta-analyses aim to synthesize empirical findings from quantitative studies [3]. Given their disjoint insights, these methods are often used complementary to one another [19].

A third method, referred to as science mapping, has gained increasing relevance as it enables summarizing a field's bibliometric and intellectual structure by analyzing the social and structural relationships between different research constituents [45, 83].

The qualitative nature of systematic literature reviews is prone to interpretational biases due to scholars having different academic backgrounds [41] and potential lack of rigor [68]. Therefore, the complementary use of quantitative methods, meta- and bibliometric analysis, provides a consistent, transparent, and reproducible process that can help mitigate these biases [83].

In addition, academic research databases such as Scopus and Web of Science are expanding their coverage of available bibliometric data, while the selection of accessible programs that can preprocess, visualize, and provide comprehensive insights into this data is growing [15]. As a result, there was near exponential growth in the number of bibliometric articles in multidisciplinary journals covering all kinds of scientific areas [33]. And naturally, systematic reviews drawing upon those methods have played a major role in this progression [9].

At what stages of a systematic review can a bibliometric analysis be practical?

Several studies propose using different bibliometric methods, even before the article selection. For example, authors can use performance analysis to identify the most significant publications for a research topic [16, 29]. Alternatively, methods have been developed that start with a small set of seed articles and propagate to other related articles based on citation relationships [9].

However, this thesis focuses on bibliometric appliances after the initial article selection. In this case, bibliometric analysis can be divided into two subcategories, both having their disjoint usefulness at different stages of systematic reviews [15].

Performance analysis can be employed in the evaluation stage to rank selected research constituents through various metrics and evaluate their importance and quality in the analyzed research field [19, 55].

On the other hand, science mapping enables the construction of complex network graphs to identify significant patterns and trends within the development of a research front [13]. Each node represents a research constituent in these graphs, while the connecting edges define their citation relationships. Science mapping assumes that citation networks do not merely describe amorphous relationships between articles but reflect clear structural relationships allowing the abstraction of scientific evolution [13].

The three mapping techniques, direct citation, co-citation, and bibliographic coupling, are mainly used to construct bibliometric network graphs and are explained in more detail in section 2.3. Prior research analyzed the resulting graphs at the micro and meso level while neglecting the macro level [81].

On the micro-level, singular publications are characterized by role and importance within a field with the help of various node-level metrics, such as the in-and-out-degree, betweenness centrality, and closeness centrality [81].

The meso-level represents the grouping of publications within a corpus. These groups are created by employing clustering techniques on the network graph, such as the frequently used Louvain method, which aims to maximize the modularity of a graph's division [73]. Subsequently, the created groups are mainly categorized by topic to gain a better overview of the analyzed research field [81].

However, the macro-level allowing a characterization of a literature corpus as a whole remains understudied [45, 71, 81].

This thesis addresses this gap and provides macro-level metrics for citation networks constructed from the corpora of 14 different literature reviews. Each corpus consists of directly included publications and their references. The methodology section describes utilized network-level metrics in more detail.

It is important to note that bibliometric analyses are primarily used with systematic literature reviews, but only seven of the fourteen sample reviews in this thesis are classified as systematic. Followingly, no differentiation is made between them, and the conclusions drawn in this thesis are applied to both. However, the discussion section will shed light on the possible distortions this can cause.

2.2. History of Science Mapping

Science mapping deals with the collection of methods and techniques employed to generate visual representations of scholarly knowledge. Its practice is deeply anchored in bibliometrics and scientometrics. Although there were earlier applications related to science mapping, the actual emergence of networks mapping human knowledge started after the 1960s with multidisciplinary databases indexing vast amounts of scientific publications [51].

The Science Citation Index (SCI), created by Eugene Garfield at the Institute for Scientific Information in the 1960s, started to store citation data of scholarly articles [40]. The contained citation data allowed the positioning of an article within its network of related publications and the creation of network graphs of entire scientific areas with the help of simultaneously emerging citation-based linkage methods [51].

In 1964, Garfield et al. proposed the method of direct citation to investigate the individual historical structures of scientific knowledge that lead to scientific breakthroughs [23]. One year earlier, Kessler (1963) proposed the method of bibliographic coupling, which according to Price (1965), allows for insights into a temporal state of a scientific composition of knowledge and shows what scientists are researching at a specific point in time [34, 54].

Co-citation analysis was introduced nearly a decade later [43, 61] and quickly prevailed as the standard method in bibliometric studies [12]. According to Small (1973), the relationships mapped through co-citation are strongly related to relationships emerging from direct citation relationships [61].

While the co-citation method remained the de facto standard at least until recently [12, 83], in the last decade there was a considerable increase in popularity regarding the use of direct citation and bibliographic coupling [20, 36]. Section 2.3 provides a detailed description of how these methods create different network graphs from bibliometric data.

These methods can be applied on the document level, describing the mapping of singular research publications. However, the publications can also be aggregated into higher research units to create networks of authors, institutions, countries, and journals [81].

Author citation networks can reveal the performance and collaboration of investigated authors [25, 50]. On the other hand, journal citation networks can be used to assess the impact, quality, and usefulness of journals [15], but also “the knowledge structure and publication patterns of scientific fields in prior research can be studied by analyzing citations among journals’ [47].

In sum, all mapping techniques and assessment levels have their specific use case. The principal applications of science mapping include performance evaluation of constituents, mapping the historical development of research fields, and information search and retrieval [21, 42].

Besides methodical aspects, an essential factor of a bibliometric analysis is the scope of selection. In the literature, the terms’ localism and globalism have been discussed [35]. Global maps are created to comprehensively represent all scientific disciplines and their relationships and thus provide insights that are not retrievable from smaller local maps of science [35]. However, comprehensive global maps require vast amounts of bibliometric resources that are not available to most research [14]. Conclusively, there is a limited number of studies focusing on global maps [36].

Local maps typically display a specific scientific area. The most popular approach is mapping scholarly output regarding a concrete research question [14]. In this case, a complex query is generally applied to academic research databases to select a limited number of articles among the vast possible ones in a scientific area [9]. Their formulation is a non-trivial task and requires extensive expertise in the specific field [14].

The far more frequent implementation of local maps is rooted in their broader use case, easily accessible resources through scholarly databases alongside emerging software tools allowing for the data's preprocessing and the construction of network graphs [14]. For this reason, due to their local scope, systematic reviews drawing on bibliometric analysis are a fast-growing trend [15].

2.3. Direct Citation, Co-Citation and Bibliographic Coupling

As already indicated, three citation-based methods, namely direct-citation, bibliographic coupling, and co-citation analysis, are mainly used in science mapping [12]. This section aims to provide insight into how these methods create different networks, each possessing its specific use cases and extractable information about the literature corpus to be analyzed. It is important to note that this thesis will only focus on document-level appliances and will not cover networks of aggregated units such as authors, institutions, journals, et cetera.

Direct citation is the most straightforward approach and reflects literature relationships as depicted in reality. That is to say, publications will only be linked together when there is a direct citation between these, resulting in a directed edge connecting the two research entities (nodes) in a directed network [61].

Direct citation networks are used to map the historical development and structure of research fields while enabling research performance analysis, such as tracking citation counts of publications [74]. However, direct citation networks are very sparse compared to the similarity-based networks bibliographic coupling and co-citation. A reason for this is that direct citation only maps explicit citation relationships between publications [20] and includes missing links resulting because of various reasons, "such as information overload, search failure, and journals non-use policy" [79] as cited in [82].

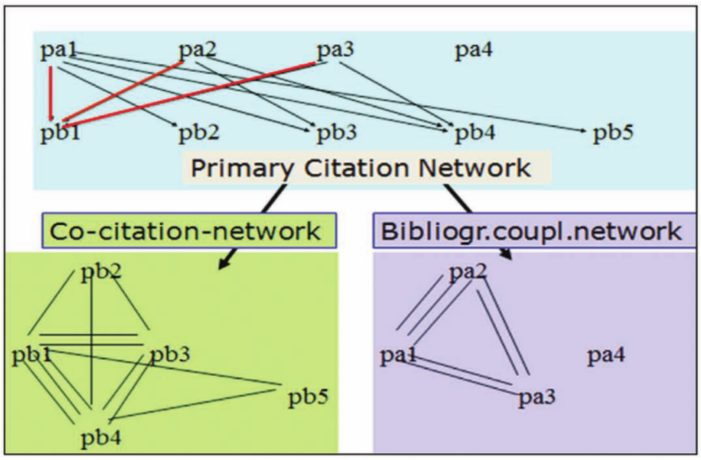
Bibliographic coupling links publications together if both cite one or more of the same publications, based on the assumption that they have similarities in their content if they share common references [19]. The number of the shared references both papers have in common determines the strength of the link [47]. Given that bibliographic coupling only creates relationships based on references and not citations, it fails to identify the importance of papers. However, this also enables bibliographic coupling to map current research fronts that lack citations or take niche publications into account, thus exposing a broader range of themes of analyzed research fields [83].

On the other hand, co-citation assumes that publications have cognitive similarities if they are being frequently cited together by other publications [82]. Thus, this mapping technique creates links between publications if they appear in the reference list of a third paper.

In comparison to bibliographic coupling, co-citation analysis discovers thematic clusters while it finds the most influential publications these clusters are built upon [19]. However, this comes with the price that co-citation analysis is biased towards these more influential papers resulting in an unreliable position of new and niche documents within their clusters [83].

In addition, compared to bibliographic coupling and direct citation, a co-citation network is not static. It will vary in time as future papers create new indirect co-citation relations [69].

Figure 2.1 shows examples of graphs computed from the different citation methods. In the primary direct citation network, pa1, pa2, pa3, pa4 are the citing publications and pb1, pb2, pb3, pb4, pb5 are the publications being cited.



P(r)

	pb1	pb2	pb3	pb4	pb5
pa1	1	1	1	1	1
pa2	1	0	1	1	0
pa3	1	0	0	1	0
pa4	0	0	0	0	0

$C = P^T(r) \times P(r)$

	pb1	pb2	pb3	pb4	pb5
pb1	3	1	2	3	1
pb2	1	1	1	1	1
pb3	2	1	2	2	1
pb4	3	1	2	3	1
pb5	1	1	1	1	1

Co-citation

$B = P(r) \times P^T(r)$

	pa1	pa2	pa3	pa4
pa1	5	3	2	0
pa2	3	3	2	0
pa3	2	2	2	0
pa4	0	0	0	0

Bibliographic coupling

Figure 2.1.: Different citation networks reprinted from Van Raan [74]

The derived co-citation network analyzes the relations of the cited publications. For example, pb1 and pb4 are connected with an edge of the weight three, symbolized by three singular edges, because they are both cited by the same three publications pa1, pa2, and pa3. It is to note, that this process is called co-citation clustering, while the prior mentioned co-citation analysis “takes the result of co-citation clustering, and then assigns current papers (or papers from the research front) to the co-citation clusters” [12].

Lastly, the bibliographic coupling network analyzes the relations of the citing papers based on their common share of references. For example, pa1 and pa2 reference the same three publications and are thus connected by an edge of the weight three. In contrast, pa4 shares no references with the remaining citing publications and is, therefore, not connected to the subgraph.

The figure 2.1 also indicates how the similarity-based citation networks are derived from the direct citation network. The adjacency matrices of both networks are computed with the shown corresponding formulas, where P(r) is the adjacency matrix of the direct citation network.

Several scholars tried to compare these methods and provide insights as to which of them is the most accurate in revealing the intellectual structure and the fundamental themes of research fields. While some studies have shown a better performance with the methods based on indirect relations [4, 67], other studies have shown the opposite and reported that direct citation yields better clustering results [12, 75, 76]. Then again, other publications emphasize that the accuracy ranking of these methods depends on multiple factors like the purpose of the analysis itself [82] and the composition of the data. Thus, comparing these methods with absolute accuracy will not lead to definite results [27].

It is also to note that all these methods are subject to several limitations. The creation of networks based solely on bibliographic data does allow the visualization of significant aspects of the data. However, this simplification also leads to a substantial loss of information about the content of the publications [20]. While citations are the only deciding factor that creates the network structure, it is impossible to determine why a particular citation was established initially [20]. Citations can confirm and help build upon previous literature, but they can also refute it [83]. In addition, authors and collaborators could be biased in citing their own literature [83]. Finally, the impact of prior technical decisions and parameters in data retrieval and preprocessing on the structure of the graph network is difficult to estimate due to an unknown network sensitivities [20].

These uncertainties suggest that a bibliometric network should be interpreted carefully and complementary to in-depth knowledge of the research fields to be analyzed [20].

This thesis draws on the direct citation and bibliographic coupling method to create the bibliometric network graphs. Direct citation allows insights into the development and the connectedness of the whole underlying corpus of a literature review as it maps directly included publications and their references. Whereas bibliographic coupling allows analyzing the thematic clustering and relatedness of only directly included publications based on their reference share.

It is important to note that a limited timeframe of incorporated publications is typically required when employing bibliographic coupling [26]. The problem with more extended periods is the effect of citation genealogy. “Citations form kind of branching process such that the probability that the reference lists of two related articles would still strongly overlap is rather low if 10, 20 years or more have elapsed between their publication" [26]. However, the temporal distribution of publications included in a review is typically broader. Thus, this thesis also analyzes the impact of the temporal distribution on the resulting graph structures concerning this effect.

Additionally, this thesis draws on a variation of bibliographic coupling referred to as fractional-counting to achieve better thematic clustering results. The advantages of fractional-counting over full-counting are described in [49]. The adapted formula for the adjacency matrix can be found in the methodology section 4.4.

2.4. Academic Research Databases

As bibliometric analysis is heavily dependent on data, its emergence is accompanied by the proliferation of accessible online databases collecting citation data [83].

In the 1960s, Eugene Garfield founded the first bibliographic database named Institute for Scientific Information (ISI), which since 1992 has been known as the Web of Science [53]. The Web of Science (WoS) profoundly affected the evolution of bibliometric analysis and established itself as the "gold-standard" data source. "As per the latest data of 2020, the Web of Science Core Collection covers more than 74.8 million scholarly data and datasets, 1.5 billion cited references (dating back to 1900)" [59]. The WoS database aggregated data of approximately 14400 journals and grew significantly in recent years while still being selective with respect to indexing decisions [59].

In 2004, Elsevier created the Scopus database, which since then expanded in coverage and emerged as the most prominent contender of WoS. It launched with 27 million publication records and has grown to over 77.8 million publications across 25,100 journals in 2020 [58]. In Scopus, indexing decisions are made by the Content Selection and Advisory Board reviewing all titles suggested for the database [58].

In 2018, Digital Science introduced the Dimensions database, which started with over 90 million publications and grew to 109 million publications in September 2019, covering over 74.000 journals by May 2020 [59]. Dimensions is promising for future bibliometric usage and a severe contender to Scopus and WoS [63]. However, only a few database comparison studies include Dimensions due to its new release [63]. Additionally, it frequently misses the reference lists of its indexed publication, thus making it unsuitable for this thesis's planned reference-based data aggregation [63].

Apart from that, other data sources exist, such as Microsoft Academic, CrossRef, ResearchGate, OpenCitations, PubMed, Medline, et cetera [53]. These can all be advantageous for specific purposes, given their more specialized coverage, but this makes them unsuitable for the analysis of multidisciplinary research units [53].

"Despite the fact that WoS and Scopus DBs have been extensively compared for more than 15 years, the scientometric community still have not reached the verdict of which one is better.” - [53]

However, the studies indicate a broader and more inclusive coverage for Scopus, while WoS is more selective regarding its coverage [53, 59]. Additionally, studies indicate that WoS has a higher rate of omitted references while having more duplicates and inaccurate references. [53]

Conclusively, this thesis uses Scopus for data retrieval assuming that it will have more articles indexed where the reference data can directly be extracted compared to WoS, thus diminishing the complementary manual data addition of not indexed publications.

2.5. Preprocessing, Visualization and Analysis

Researchers can use various preprocessing- and analyzing tools to utilize the extracted bibliometric data from scholarly databases.

For example, preprocessing methods include duplicate detection or grouping identical bibliometric items together if they were spelled differently with the help of string distance measures such as the Levenshtein distance [15].

Subsequently, several tools can create different types of bibliometric networks graphs and include various methods for further analysis and visualization. A detailed overview of the tools and their different use-cases is described in [15].

This thesis uses manually retrieved reference data from various sources in addition to the data extracted from Scopus, which was necessary for publications that were not indexed on Scopus. However, most science mapping tools expect standardized data formats offered by the academic research databases and are thus not suitable for the extracted bipartite data of this thesis.

As a result, the preprocessing, graph creation, and analysis steps were completed manually using the graph package “networkx” in Python. Subsequently, the visualization tool Gephi was used to cluster and visualize the constructed graphs, as it includes an editable and user-friendly visualization environment.

The methodology section 4 includes a more detailed description of the different steps conducted in this quantitative study.

2.6. Summary

In the past decades, there has been an emergence of conducting systematic literature reviews with the quantitative assistance of bibliometric analysis, reducing bias and allowing more significant insights in summarizing a research field’s intellectual structure.

Most of the reviews employed bibliometric methods on a micro-and meso-level to highlight key publications within the corpus or cluster the corpus thematically and analyze the distinct groups individually.

Thus, analyzing these networks on the macro level to make assumptions on the underlying research corpus as a whole has been neglected.

This thesis aims to fill this gap by analyzing bibliometric data of 14 different literature reviews published in the leading information science journals in the AIS College of Senior Scholars' basket of eight. The extracted data of each corpus consists of a review's directly included publications and their references.

The data was mainly retrieved through the Scopus database, while the data of not indexed publications and their references was manually integrated from various other sources.

Direct citation and bibliographic coupling graphs were created for each review, and several graph-related network-level metrics were computed and statistically summarized. The graphs were created and analyzed in Python and visualized using the Gephi visualization tool.

This work provides central tendencies of how connected, dense, and clustered these graphs are while allowing a better assessment and classification of the underlying knowledge structure for future literature reviews that include macro-level bibliometric analyses.

Important: To better explain the following contents of this thesis, publications that have been directly integrated into a literature review will be referred to as "core" publications.

3. Hypotheses

3.1. Descriptive Statistics

Overall, there seems to be a lack of macro level analyses for bibliometric network graphs computed from bibliometric data of literature reviews. Due to the absence of available data it is impossible to narrow down and predict the density, clustering, and connectivity that the employed metrics will indicate.

For these reasons, there are no concrete hypotheses for the descriptive statistical part of this thesis, and it mainly serves as a benchmark for future analyses.

3.2. Inferential Statistics

Regarding inferential statistics, a correlation among metrics was suspected based on the preliminary theoretical work. In detail, the author of this thesis expected the following independent variables to have a decisive impact on the resulting network graphs: The number of core publications, their dispersion of publication years, and their average count of references.

For the purpose of our research, only the metrics describing the share of direct reciprocal citations between publications included in a literature review (core publications) and their average common share of references are treated as dependent variables since they are considered as the most important in describing a corpus. The following assumptions were made for these variables.

First, a higher number of core nodes could indicate a broader or a more clustered composition of thematic influences in the review, which could diminish the mutual citation behavior of core publications and their average share of references. A reason for that might be that authors of literature reviews with a higher number of core publications are more likely to:

a) go after a broader or interdisciplinary research question that requires highlighting distinct additional topics, which again could require a greater number of included publications.

b) dive deeper into a monodisciplinary research question, leading to greater focus on distinct niche topics at the edge of the corresponding research area, which in turn leads to more included publications.

Thus, the following two Hypotheses were formulated.

Hypothesis 1: “A higher number of core publications included in a literature review negatively correlates with the share of direct reciprocal citations between them.”

Hypothesis 2: "A higher number of core publications included in a literature review negatively correlates with the common share of references among them."

Second, this study suspects that core papers published consecutively within a small-time frame have a lower chance of citing each other. One reason might be that the preliminary theoretical work of the later published publication could already be finished before the prior study is released. Additionally, the relevance and visibility of a recently released publication could take some time to build up, leading to the following hypothesis.

Hypothesis 3: "A larger dispersion of the publication dates of core papers included in a literature review positively correlates with the share of direct reciprocal citations among core papers."

Third, the following hypothesis was made regarding the effect of citation genealogy described in section 2.3.

Hypothesis 4: "A larger dispersion of the publication dates of core papers included in a literature review negatively correlates with the common share of references among them."

Lastly, this paper hypothesizes that a higher average number of core publication references will lead to a greater likelihood of mutual citations among core publications in a review.

Hypothesis 5: "A higher count of average core paper references in a literature review correlates positively with the share of direct reciprocal citations among core papers."

4. Methodology

This section describes the different stages of this quantitative analysis in the order of conduction, starting from the selection of literature reviews to the quantitative analysis of the resulting graph metrics.

The methodology builds upon standard practices synthesized from relevant studies mentioned in the theoretical framework while introducing new elements that allow these methods to fit the research question.

Most processes were completed in Python version 3.7.4 and can be traced in more detail in the Github repository below:

`https://github.com/Eddyvmgit/bib_thesis`

4.1. Selection of Literature Reviews

The literature reviews from which this thesis draws its different bibliometric corpora stem from the Senior Scholars' Basket of Journals, which consists of the following journals:

European Journal of Information Systems
`https://www.tandfonline.com/toc/tjis20/current`

Information Systems Journal
`https://onlinelibrary.wiley.com/journal/13652575`

Information Systems Research
`https://pubsonline.informs.org/journal/isre`

Journal of AIS
`https://aisel.aisnet.org/jais/`

Journal of Information Technology
`https://journals.sagepub.com/home/jin`

Journal of MIS
`https://www.tandfonline.com/loi/mmis20`

Journal of Strategic Information Systems
`https://www.sciencedirect.com/journal/the-journal-of-strategic-information-systems`

MIS Quarterly
`https://misq.umn.edu/`

In September 2021, the above websites, each having the specific journal registered, were all searched for literature reviews using the search term: ("Title: Literature"). 43 publications were found, which were subsequently screened to determine if they match the research design.

Publications that were not addressing a specific research question, such as research guides for future literature reviews and reviews, including publications from the entire IS literature, were excluded. Additionally, only articles published after the year 2000 were investigated to ensure a comparable time structure of the different corpora and narrow down the scope of this thesis. Subsequently, only 22 literature reviews were left.

Each of these 22 reviews was examined in more detail to assess the content and design. Studies that included more than 150 publications were excluded. Two reviews shared the same topic and authors, which led to a significant similarity of their included publications. Consequently, only the newer review was integrated. Lastly, another article was excluded as it did not state a list of included publications.

Finally, the following 14 literature reviews were included in the bibliographic analysis:

D'arcy and Herath [17], Jiang et al. [30], Siponen and Vartiainen [60], Moeini and Rivard [44], Teubner and Stockhinger [65], Wiener et al. [78], Xiao et al. [80], Pereira and Serrano [48], Schneider and Sunyaev [57], Tsai et al. [70], Piccoli and Ives [52], Baghizadeh et al. [8], Günther et al. [28], Oehlhorn et al. [46]

4.2. Data Retrieval

As already indicated, the bibliometric data on which the several network graphs were built consists of publications being embedded into literature reviews (core publications) and their references.

It is important to note that the bibliometric data only includes scientific core publications, meaning journal or conference articles.

For each review, all the included publications were searched on Scopus. Articles that were indexed and had non-empty reference lists were extracted in a collective CSV format. In contrast, publications that were not indexed or had empty reference lists were manually retrieved from various external web sources and later integrated into a new combined CSV File.

Scopus only allows retrieval of articles and their corresponding reference lists where each reference is summarized into a complete string, thus not enabling a distinct extraction of the title, year, and author names.

However, it was planned for this thesis to create a unique identifier for each bibliometric entity consisting of the title, the year, and the first author name. Based on this identifier, the entities will be mapped together in the network graphs.

A solution was found by retrieving the complete list of references without the relational data of their citing articles, in which the bibliometric attributes are divided into different columns. Subsequently, the references were mapped back to their citing articles with the relational file, where a distinction between attributes was not possible in the first place.

This way, it was possible to create a combined CSV file where the bibliometric characteristics such as year, title, and author were divided into separate columns allowing the creation of manual identifiers. Otherwise, matching the whole bibliographic string would have been necessary, requiring approximate string-matching methods that this thesis does not cover.

The manually retrieved bibliometric data was copied into a text file with one record string per line. Afterwards, the author names and the titles were marked with tags, while the date was localized with regex pattern matching. Subsequently, a Python script created a combined CSV file of the adapted Scopus CSV file and the manually added text file for each distinct literature corpus.

The resulting CSV files contained distinct bibliometric attributes of each publication relevant in a literature review and their citing relations to construct the network graphs. The whole process can be tracked in the Github repository.

4.3. Sequence Matching

The bibliometric entities within a graph are mapped together by matching complete string sequences.

Two identifier formats were selected for the analysis:

1. Title_Year
2. LastNameFirstAuthor_Title_Year
    - For bibliometric records with titles consisting of less than five words

1. Example
    - Douglas G Bonett. Confidence interval for a coefficient of quartile variation. Computational statistics & data analysis, 2006.
        - Identifier: confidenceintervalforacoefficentofquartilevariation_2006
2. Example
    - Eugenio Petrovich. Science mapping. ISKO Encyclopedia of Knowledge Organization, 2020.
        - Identifier: petrovich_sciencemapping_2020

The rationale for these formats is grounded on the assumption that these identifiers provide a balanced combination of the following attributes:

First, the bibliometric characteristics title, year, and first author name usually have high applicability as all components will rarely be null [32].

Second, the combination of title and year allows a high selectivity describing the uniqueness of the identifier for each distinct bibliometric record, which is further enhanced by adding the first author's name for entities with short titles [32].

Third, the characteristic combination allows for moderate accuracy, which describes the average correctness of the values in an identifier. Here a limitation is given by using the complete title, as these have a greater chance of including different formats, spelling mistakes, and incomplete words. A solution to this problem would have been approximate string matching, but this also increases the likelihood of incorrectly matching unique bibliometric records [32]. Therefore, no approximate matching was performed since it is beyond this thesis's scope to assess related methods and their parameters accurately.

After creating the identifier for a bibliometric record, all special characters were deleted, and letters were converted to lowercase. Finally, bibliometric records were mapped together through exact string matching.

4.4. Graph Construction

Subsequently, after collecting the data, a python script created the graphs for each distinct corpus from the final CSV files by performing sequence matching.

First, the direct citation graphs were created with the help of the Python plugin “networkx”. All bibliometric entities of a corpus were added as distinct nodes, while directional edges were added to reflect their citing behavior.

Figure 4.1 shows a visualized example of one of the constructed direct citation graphs. The larger nodes indicate the core publications of a review, and the smaller nodes display their cited references with colors denoting the different cluster groups that the bibliometric entities were computed into.

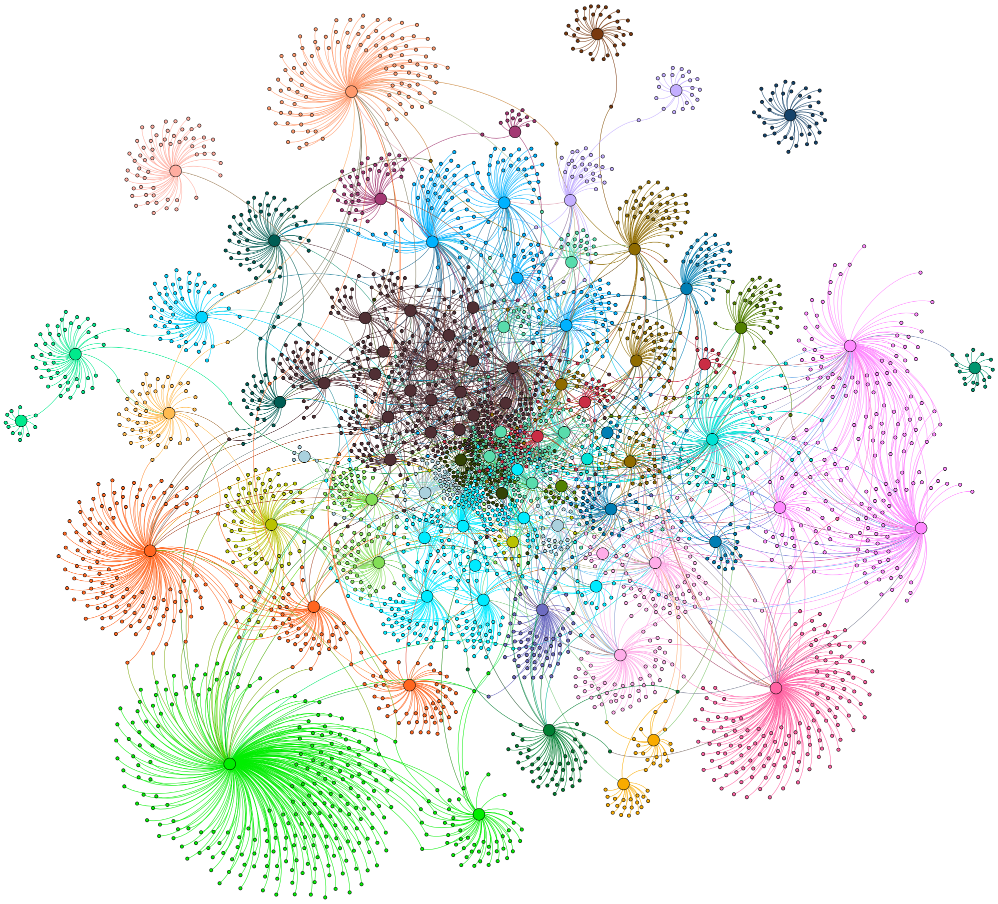


Figure 4.1.: Direct citation graph of Baghizadeh et al. [8]

Afterwards, adjacency matrices of the direct citation graphs were extracted to compute the adjacency matrices of the bibliographic coupling graphs with fractional counting using the following formula retrieved from [49]:

$$B = C\,\mathrm{diag}(C^T 1 - 1)^{-1}\, C^T,$$

where C symbolizes the adjacency matrix of the direct citation graph and “diag(v) denotes a diagonal matrix with the elements of the vector v on the main diagonal and where 1 denotes a column vector of length N with all elements equal to 1” [49]. The main diagonal elements of the bibliographic coupling matrix B are set to 0.

Figure 4.2 shows a visualized example of a constructed bibliographic coupling graph. Here only the core publications are displayed as bibliographic coupling graphs show the relations between the citing publications given their commonly shared references.

All constructed graphs were exported as GEFX Files. GEFX stands for Graph Exchange XML Format, a language describing complex network structures compatible with the visualization and clustering tool Gephi [1].

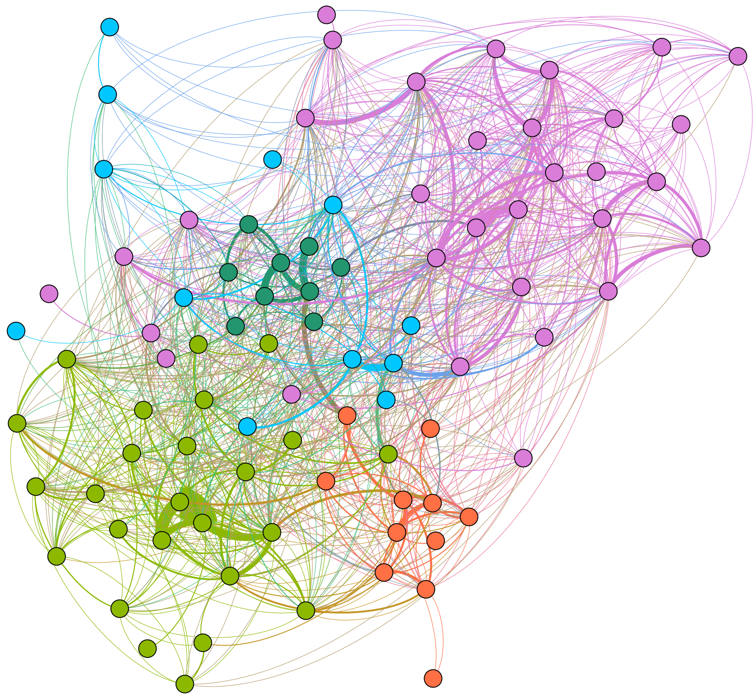


Figure 4.2.: Bibliographic coupling graph of Baghizadeh et al. [8]

4.5. Clustering and Visualization

All GEFX files were imported into Gephi to compute clusters with the inbuilt Louvain Algorithm while creating visualizations for each graph. The Louvain Algorithm is a greedy optimization method and attempts to optimize the modularity of a graphs partition [10]

After the algorithm has computed an optimal partition, Gephi colored the nodes and edges according to their cluster group. This thesis used different layout settings for direct citation and bibliographic coupling graphs, given their distinct nature. Figure A.1 and A.2 in the appendix demonstrate the employed layout settings for the visualizations in Gephi.

Finally, cluster results and visualizations were extracted in PDF formats. The visualizations are not crucial in this quantitative analysis. However, they are still provided for every graph in the appendix section A.4 with their distribution of cluster sizes providing more clarity for each corpus. On the other hand, the computed number of clusters and resulting modularity values were manually integrated in the subsequent statistical analysis.

4.6. Employed Metrics

After successfully creating and visualizing the network graphs of the literature review samples, a python script iterated over the exported GEFX files and computed the corresponding metrics for each graph. The modularity values and the number of clusters were manually integrated from the prior computations with the Gephi tool.

Below are the lists of the collected outer graph and inner graph metrics for each corpus:

4.6.1. Outer Graph Metrics

1. Number of Core Publications
2. Average Number of References of Core Publications
3. Interquartile Range of Core Publications's Release years
    - The interquartile range gives a robust measure of dispersion for unknown distributions. Section 4.7 provides a more detailed rationale behind selecting the interquartile range over the standard deviation.
4. Number of all Bibliographic Entities
    - Core Publications and their References

4.6.2. Graph Metrics

1. Density Measures
    - This thesis employs two different measures of density, each providing distinct insights. In general, density measures the ratio between all existing edges and all possible edges within a graph [7].
    - Core Density
        - Core Density is employed on the direct citation graph to measure the share of mutual direct citations among core publications in a review, which is done by dividing the existing citations by all possible citations between these. The corresponding formula is:

$$\frac{\#citations}{n(n-1)/2},$$

where n denotes the number of core publications. The formula was retrieved from [66].
    - Bibliographic Coupling Graph Density

   - On the other hand, the standard graph density formula is employed to compute the density of the bibliographic coupling graph. While the bibliographic coupling graph is weighted, it is essential to note that the density formula does not consider edge weights. No matter how many references two core publications share, every connection counts once. Thus, the bibliographic coupling graph density indicates the percentage of direct relationships among core publications in a corpus over one or more common references.

2. Average Citation Overlap
   - The average citation overlap is the common share of references between all core publications in a corpus. High values indicate a greater thematic coherence among publications, while low values indicate greater thematic distinctness. The corresponding formula is:

   $$\frac{2}{n(n-1)} \sum_{ij} \frac{C_i \cap C_j}{C_i \cup C_j},$$

   where n denotes the number of core publications and C denotes the citations of a core publication. The formula was retrieved from [66].

3. Number of Unconnected Core Publications
   - Describes the number of core publications that do not share any references with the connected main component of the corpus.

4. Number of Direct Citation Graph Components
   - This measure is similar to the number of unconnected core publications but gives additional insights into the grouping of the disjunct core publications, which means if the unconnected core nodes share reference relations among each other.

5. Clusters
   - The number of clusters could denote the number of thematic components in a literature corpus [77]. The Louvain clustering algorithm was employed on both graph types, the direct citation and the bibliographic coupling graph, to allow further insights and compare differences among these results. The Louvain algorithm creates the clusters by optimizing the modularity and was employed with the following settings in Gephi:
     - Randomize: Yes
     - Use weights: Yes
     - Resolution: 1.0

6. Modularity
   - Modularity describes the strength of the network's division into modules by comparing the density within and between computed clusters [2].

7. Average Shortest Path

- This measure describes the shortest average path between nodes [64]. It thus provides insights into the overall direct or indirect relatedness of core publications over at least one common reference in the bibliographic coupling graph. The average shortest path is only applicable for the bibliographic coupling graph because it requires that the graph consists of one connected component [64].

4.7. Statistical Analysis

This quantitative study draws on a descriptive and inferential statistical part to allow insights into central tendency, variation, confidence intervals, and correlation measures.

4.7.1. Descriptive Statistics

Due to the small sample size, normality tests have little statistical power [24]. Thus, the ongoing analysis was carried out without making assumptions about the underlying distributions of the metrics' populations.

Furthermore, the applicability of indicators for measuring central tendency and dispersion of the data depends on the data's symmetry measured by skewness [39]. The Pearson-Fisher skewness coefficients for each metric were computed and summarized in figure 4.3. Values far from 0 in a range from -3 to +3 indicate a non-normal (skewed) population of the metric at hand [18], and it is evident that the skewness coefficients vary significantly depending on the metric. On the other hand, skewness is hard to judge for small samples, and thus, the illustrated coefficients should be interpreted with caution [18].

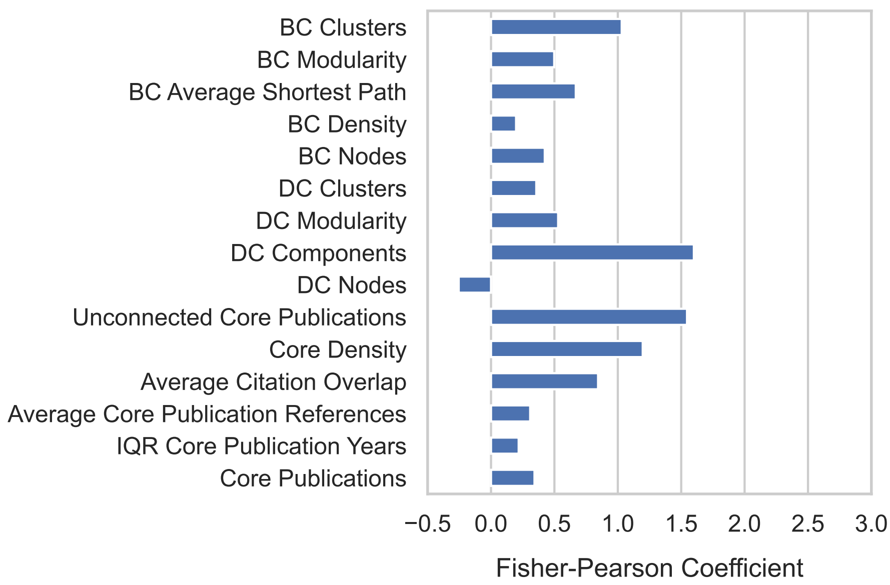


Figure 4.3.: Barplot - Fisher-Pearson skewness coefficient values of metrics employed on the sample data

Given these statistical uncertainties, it was decided to focus on the median and the interquartile range to summarize the central tendency and dispersion as these indicators are more robust compared to the mean and the standard deviation when the sample data signals substantial skewness and inherits outliers [39].

The same statistical methods were employed on all metrics to reduce complexity and enhance uniformity in this thesis. The tables in the appendix section A.2 provide the mean and standard deviation for completeness. However, given the above-described reasons, these statistical measures will not be regarded in the subsequent analysis.

Furthermore, this thesis employs the unitless coefficient of quartile variation to compare the dispersion and measure the homogeneity among the metrics [5, 6, 11] and boxplots to provide a graphical overview of the sample data.

4.7.2. Inferential Statistics

In the inferential part, a confidence interval for the median was computed to make further assumptions about the populations of each analyzed graph metric. As this thesis makes no assumptions about the distribution of the populations, the intervals were computed using order statistics as described in [62].

Additionally, this study employs scatter plots to determine if the number of core papers within a corpus, the interquartile range (dispersion) of the publication dates, or the number of average references affects the core density and the average citation overlap.

Subsequently, if a linear correlation between these variables is evident in the graphical display, this paper plans to compute the Spearman correlation coefficient and employ a two-tailed significance test for the correlation at hand.

The non-parametric Spearman correlation was chosen over the Pearson Correlation, as it provides more robustness when the distributions are unknown or non-normal [38].

5. Results

As noted in section 4.7, this thesis uses uniform statistical parameters to summarize all metrics. The sample size for all of the following metrics is n = 14 as 14 different literature corpora were analyzed.

This thesis employed the median and the interquartile range as central tendency and variation measures and the coefficient of quartile variation (CQV) to compare homogeneity across different metrics.

Additionally, for each metric, a confidence interval was computed for the median using order statistics allowing a probability of 94.26% that the population's median falls within the constrained intervals.

The raw data listing the metrics for each distinct literature corpus can be found in tables in the appendix section A.1.

5.1. Core Publications

First, table 5.1 summarizes the statistical information about metrics focusing on core publications of the literature reviews. Basic information such as the number of core publications for each corpus and their average references is statistically presented.

The release date dispersion of core publications in a review is reflected by the interquartile range (IQR) of publication years. The minimum value of 1 signals that the middle 50% of the included core publication years of one literature corpus have a spread of only one year. The maximum value of 14, on the other hand, indicates a literature review with a scattered distribution of core publication years, where the middle 50% of publication years have a spread of 14 years.

It becomes evident that the values for the core density and average citation overlap vary significantly across the different literature reviews, as both coefficients of quartile variation reveal comparatively high values of 0.52 and 0.69.

This variation is also reflected in higher interquartile ranges compared to the same metric's median and looser confidence intervals. A looser confidence interval is especially notable for the core density ranging from 0.039 to 0.115, implying that a share of 3.9% to 11.5% of all possible citations among core articles reflects the central tendency for the population with a probability of 94.26%. The confidence interval for the average citation overlap indicates an average reference share between 1.1% to 2.6% among publications within a literature review as the central tendency, with the same certainty.

The boxplots in the figures 5.1 to 5.6 provide a graphical overview of the metrics at hand. It becomes evident that most of these boxplots show a relatively symmetrical allocation of values. The whiskers seem to be about the same length on both sides, while only the boxplots of the core nodes and the average reference overlap indicate skewed distributions.

Additionally, outliers are evident for the average core publication references of the publication Teubner 2020, with an average of 60 being one of the more recently published literature reviews. The publication Siponen 2004 has an average count of 20, being the oldest analyzed literature review.

Another outlier was present in the core density distribution. One literature corpus has a core density value of 0.267, displaying a significantly higher share of direct reciprocal citations among core publications, while it only includes 15 of them. Below, the correlation section 5.4 analyzes the relation between the number of core publications and the core density.

Metric	Min	Max	Median	CI Median$_{p=0.9426}$	IQR	CQV
Core Publications	15	131	61.50	(49, 92)	37.25	0.28
IQR Core Publication Years	1.0	14.0	6.75	(5.5, 10.0)	3.70	0.26
Average Core Publication References	20.4	65.1	40.90	(38.7, 46.3)	7.33	0.09
Average Citation Overlap	0.002	0.045	0.014	(0.011, 0.026)	0.017	0.52
Core Density	0.006	0.267	0.070	(0.039, 0.115)	0.092	0.69
Unconnected Core Publications	0	12	1.50	(1, 6)	1.75	0.47

Table 5.1.: Statistical summary of metrics regarding the core publications

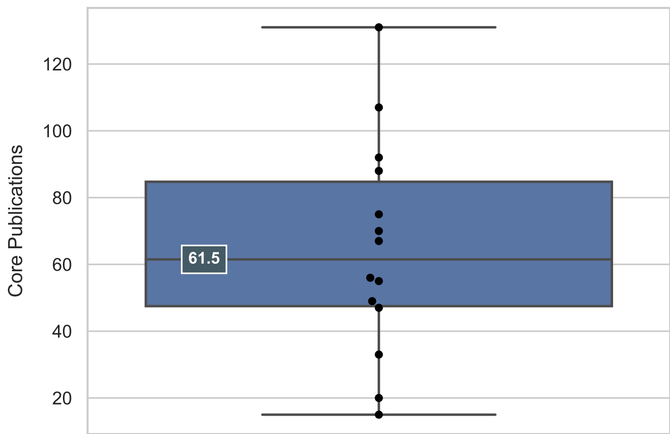


Figure 5.1.: Boxplot - Number of core publications

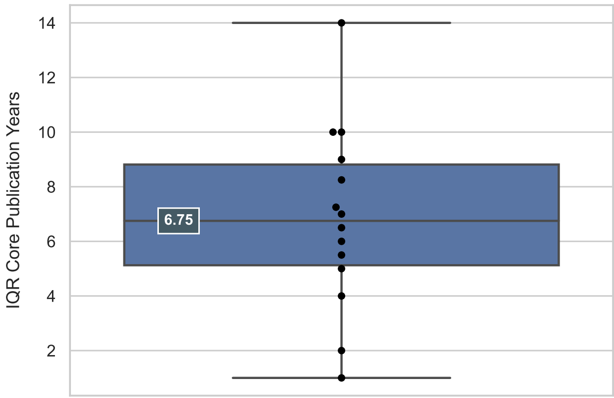


Figure 5.2.: Boxplot - IQR of core publication release years

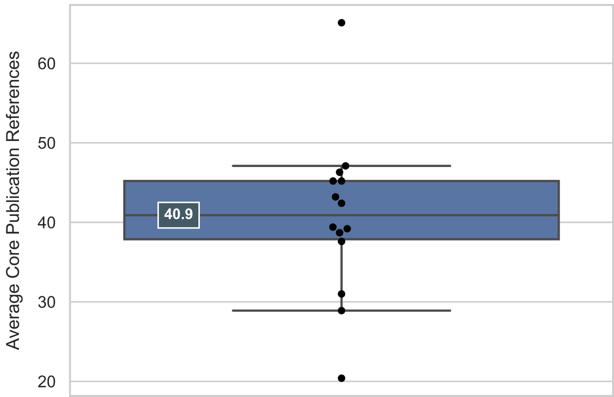


Figure 5.3.: Boxplot - Average reference count of core publications

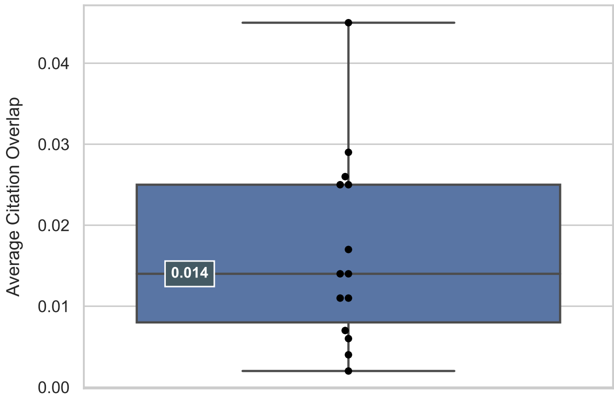


Figure 5.4.: Boxplot - Average citation overlap

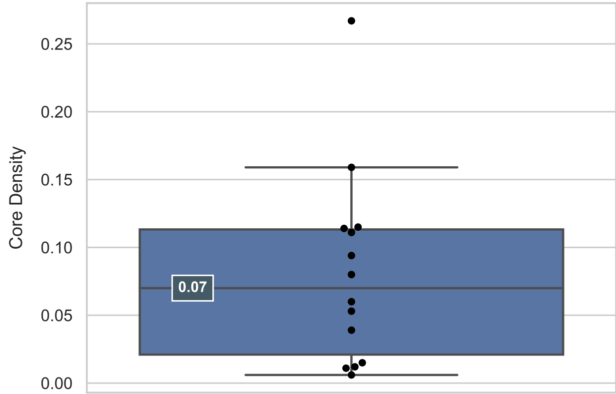


Figure 5.5.: Boxplot - Core density

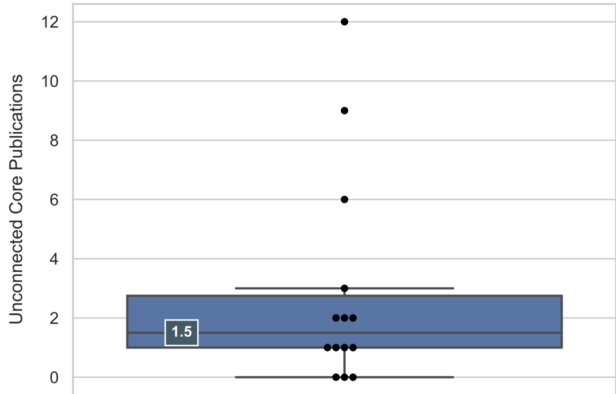


Figure 5.6.: Boxplot - Number of unconnected core nodes

5.2. Direct Citation Graphs

Table 5.2 summarizes the descriptive statistical measures of the 14 direct citation graphs.

Regarding the size of the different literature corpora, the smallest literature corpus consisted of only 332 bibliometric records, while the largest corpus consisted of 3,317 records. However, most corpora sizes were within the range of approximately 1,500 to 2,400 records, which is evident in the boxplot in figure 5.7.

The number of direct citation graph components indicates how many distinct groups of publications within a review corpus share no joint references at all with each other.

For the created 14 direct citation graphs, most smaller components apart from the main component consist of only one core publication and its references. Only the direct citation graph of Xiao et al. [80] had one minor component composed of two interconnected core papers. Thus, the number of unconnected core publications inherits similar values to the component number of the direct citation graph.

Both measures inherit a significant dispersion among the literature corpora, indicated by moderately high CQV values of 0.46 for the unconnected core publications and 0.30 for the graph component count.

Additionally, it becomes evident that direct citation graphs with a higher number of components tend to have more clusters when comparing the different corpora in the appendix table A.2. This makes natural sense considering that each distinct component will count automatically as one cluster.

The cluster counts are significantly high, and the substantial modularity scores exhibit a remarkably low dispersion between graphs indicated by a low CQV value and a narrow-computed confidence interval of (0.68, 0.88).

These values indicate that the Louvain clustering method created a similarly high ratio of the connectedness within the clusters compared to the connectedness among the clusters, showing that the clustering of publications underlies a stable pattern among the graphs.

The majority of boxplots in the figures 5.7 to 5.10 denote a somewhat skewed distribution regarding the metrics of the direct citation graphs as the whiskers are uneven. This holds true for all, except for the node count in figure 5.7, indicating a relatively symmetric distribution and the component count is subject to several outliers.

Metric	Min	Max	Median	CI Median$_{p=0.9426}$	IQR	CQV
DC Nodes	332	3317	1945	(1462, 2697)	870	0.23
DC Components	1	13	2.50	(2, 7)	1.75	0.30
DC Clusters	11	43	23.50	(20, 33)	11	0.22
DC Modularity	0.61	0.91	0.71	(0.68, 0.88)	0.13	0.08

Table 5.2.: Statistical summary of metrics regarding the direct citation graphs

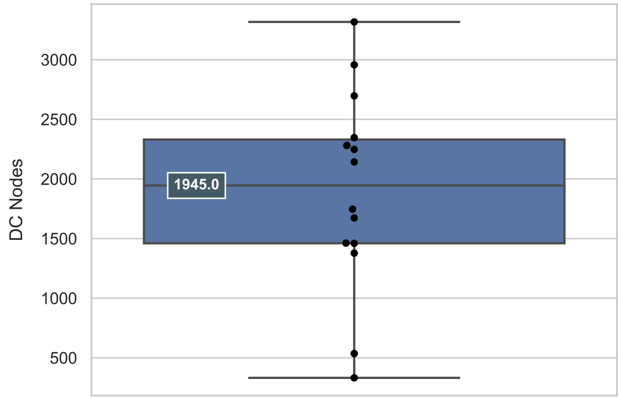


Figure 5.7.: Boxplot - Number of nodes in the direct citation graphs

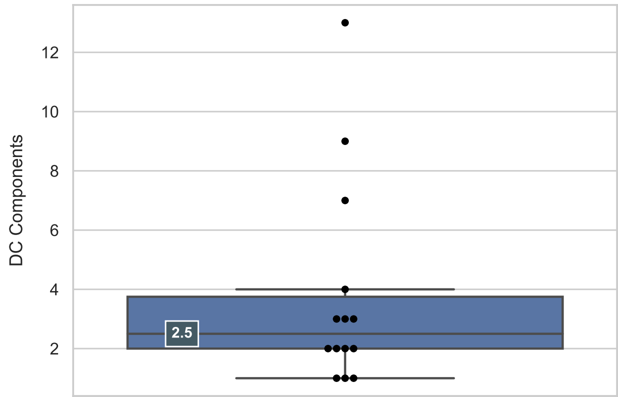


Figure 5.8.: Boxplot - Number of components in the direct citation graphs

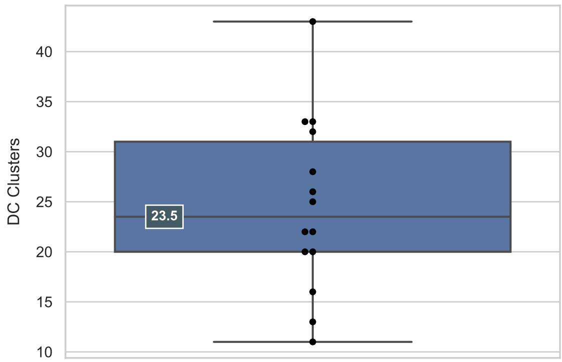


Figure 5.9.: Boxplot - Number of clusters in the direct citation graphs

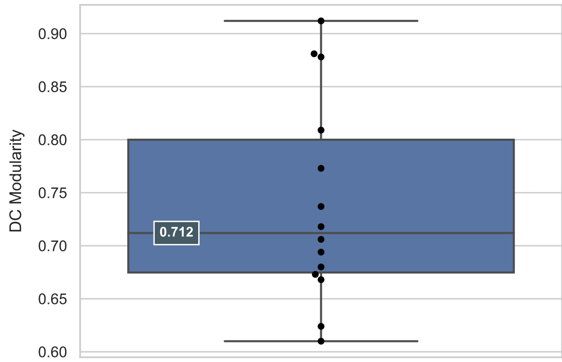


Figure 5.10.: Boxplot - Modularity values for the cluster computations in the direct citation graphs

5.3. Bibliographic Coupling Graphs

Lastly, table 5.3 summarizes the metrics resulting from the 14 bibliographic coupling graphs.

These graphs allow additional insights about the relation of the core papers due to their common share in references.

For this graph, the number of nodes and the density exhibit the highest dispersion among the metrics given CQV values of 0.33 and 0.29.

Note that the node numbers differ slightly from the above-summarized core publication numbers because core publications that share no references with any other core article of the review were not integrated as nodes in the bibliographic coupling graphs.

Regarding the density, the confidence interval of the median indicates that approximately 21-61% of core publications share at least one common reference in a literature review corpora and are thus connected. Conversely, the confidence interval denotes that 39-79% of core papers are unconnected as these do not share any references.

Nearly all average shortest path values are below the threshold level of 2, showing that most of the core publications share a direct or indirect relation over one other article. The average shortest path values correspond to the density values of a bibliographic coupling graph which is indicated by the table A.3 summarizing information for each distinct corpus.

The cluster counts are consistent among the corpora considering that the number of publications differs significantly and that the clustering of the direct citation graphs varies to a greater extent.

Regarding the boxplot in 5.14, most bibliographic coupling graphs have 5 to 6 computed clusters, indicating the main range for possible thematic distinctive groups among the analyzed literature reviews.

In contrast, the modularity values are relatively low among the clustering results of the graphs. In the corresponding boxplot in figure 5.15, all values except one outlier are within a range of approximately 0.25 to 0.43, signaling that the computed cluster groups were not very disjunct to each other, as the density ratio within and among the clusters was relatively low.

Regarding the boxplots in the figures 5.11 to 5.15, all metrics apart from the node count of the bibliographic coupling graph show a somewhat skewed sample distribution, and outliers are present for the average shortest path values, the modularity values, and the number of clusters.

Metric	Min	Max	Median	CI Median$_{p=0.9426}$	IQR	CQV
BC Nodes	14	131	59.50	(49, 89)	41.50	0.33
BC Density	0.14	0.88	0.46	(0.39, 0.79)	0.29	0.29
BC Average Shortest Path	1.12	2.37	1.59	(1.46, 1.88)	0.37	0.12
BC Clusters	3	10	5	(5, 6)	1	0.09
BC Modularity	0.13	0.50	0.26	(0.25, 0.43)	0.10	0.17

Table 5.3.: Statistical summary of metrics regarding the bibliographic coupling graph

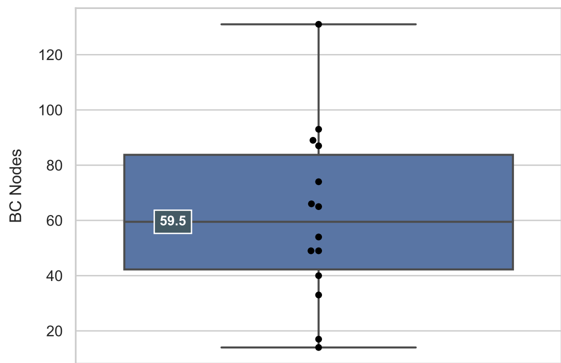


Figure 5.11.: Boxplot - Number of nodes in the bibliographic coupling graphs

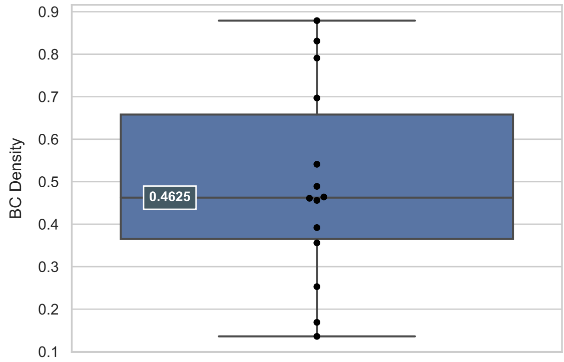


Figure 5.12.: Boxplot - Density values of the bibliographic coupling graphs

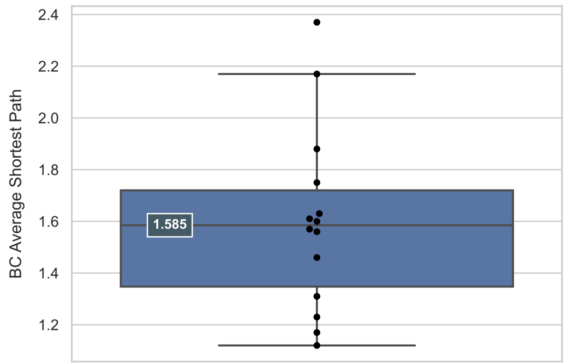


Figure 5.13.: Boxplot - Average shortest path values in the bibliographic coupling graphs

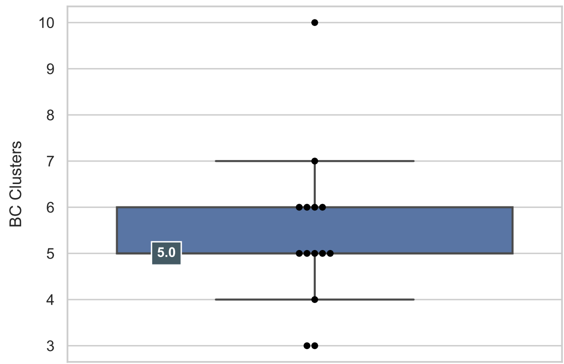


Figure 5.14.: Boxplot - Number of clusters in the bibliographic coupling graphs

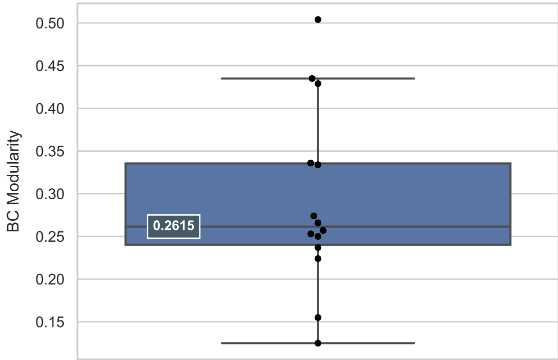


Figure 5.15.: Boxplot - Modularity values for the cluster computations in the bibliographic coupling graphs

5.4. Correlation

Subsequently, the correlations among the independent and dependent variables described in the hypothesis section were analyzed to support or reject the formulated hypotheses.

If a relationship becomes evident through a scatter plot, the spearman correlation coefficient was computed along with its p-value from a two-tailed significance test with the "spearmanr()" function of the "Scipy" package in Python.

The following hypotheses were supported by a linear relation being evident from the orange line in the scatter plots figures 5.16 to 5.18 and weak to slightly moderate computed Pearson correlation coefficients:

Hypothesis 1: "A higher number of core publications included in a literature review negatively correlates with the share of direct reciprocal citations between them." - Figure 5.16

- Pearson correlation coefficient = -0.22
- p-value = 0.45

Hypothesis 2: "A higher number of core publications included in a literature review negatively correlates with the common share of references among them." - Figure 5.17

- Pearson correlation coefficient = -0.34
- p-value = 0.24

Hypothesis 3: "A larger dispersion of the publication dates of core papers included in a literature review positively correlates with a share of direct reciprocal citations among core papers." - Figure 5.18

- Pearson correlation coefficient = 0.30

- p-value = 0.30

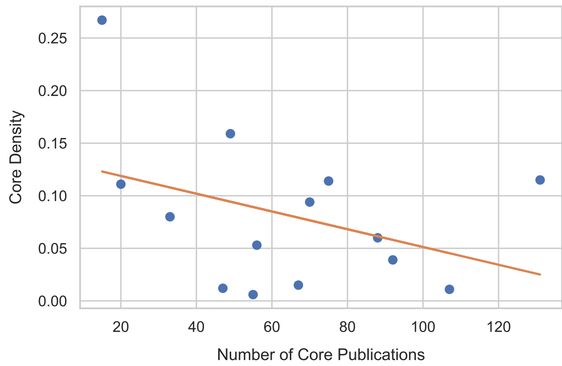


Figure 5.16.: Scatterplot indicating the relationship between the number of core publications and the core density

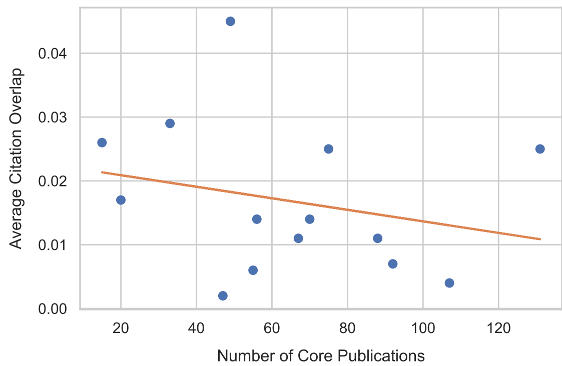


Figure 5.17.: Scatterplot indicating the relationship between the number of core publications and the average citation overlap

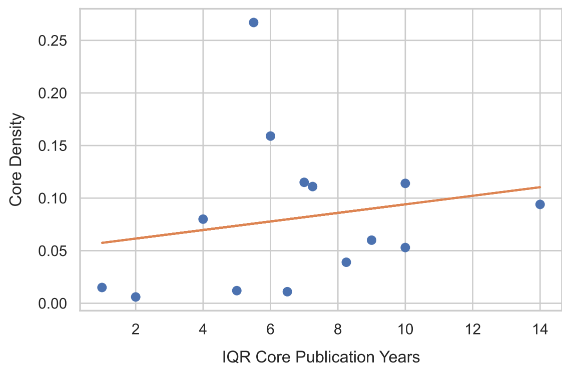


Figure 5.18.: Scatterplot indicating the relationship between the interquartile range of core publication release years and the core density

Apart from that, the graphical display resulted in no relation being evident for the following hypotheses given a horizontal line:

Hypothesis 4: "A larger dispersion of the publication dates of core papers included in a literature review negatively correlates with the common share of references among them." - Figure 5.19

Hypothesis 5: "A higher count of average core publication references in a literature review positively correlates with the share of direct mutual citations among core publications." - Figure 5.20

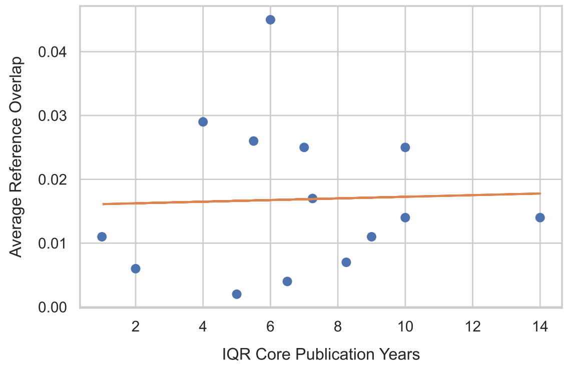


Figure 5.19.: Scatterplot indicating the relationship between the interquartile range of core publication release years and the average citation overlap

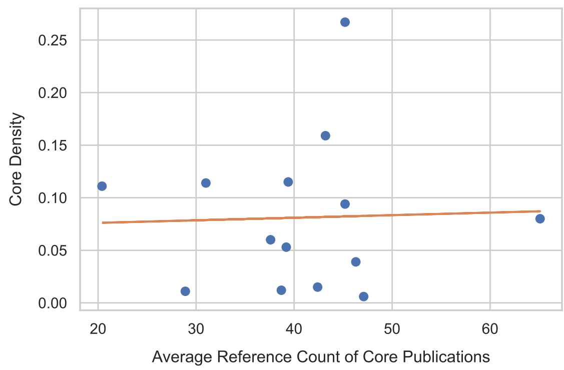


Figure 5.20.: Scatterplot indicating the relationship between the average reference count of core publications and the core density

However, none of the computed p-values for the initially supported hypotheses showed statistical significance given their high values. The p-values ranging from 0.24 to 0.48 indicate the probability that an uncorrelated system will produce data sets with a Spearman correlation at least as extreme as those analyzed [37]. Therefore, the significance tests could not verify any of the formulated hypotheses.

6. Discussion

This thesis employed various density, connectivity, and clustering metrics to analyze network graphs created from the bibliometric data of 14 literature reviews in the information systems field to provide a benchmark for future analysis.

The findings in the result section indicate diverse population distributions among the various metrics.

Summarizing the outer graph measures, the number of publications included in the literature reviews ranged from 15 to 131, with most reviews containing approximately 50 to 90 publications. The dispersion of release dates varied considerably among the corpora, while the number of average references stayed relatively consistent.

These external measures show that this quantitative analysis and its inferences cover a variety of literature review types with different scopes and periods.

6.1. Variance

The core density and average citation overlap vary the most among the sample literature corpora compared to the other metrics. These measures are considered the most important in describing a literature review corpus, as they summarize the direct citations among all directly included publications and their indirect connectedness based on their common share of references.

The computed confidence intervals for the median with a certainty of 94.26% denote that most literature reviews within the information systems field have a density of core publications citing each other reciprocally of 3.9% to 11.5%. Furthermore, the average share of common references between the core publications is likely to be in the range between 1.1% to 2.6%.

A reason for the more significant variance in these values is likely to be their multilateral dependency. This means that the network graph's compositions could, on the one hand, be dependent on general properties like the size and temporal structure of a corpus, but more significantly on superordinate aspects such as the review's research question and the included articles' selection process.

Regarding the article selection, different academic backgrounds and search methods will lead to a subjective collection of integrated core papers in a review [41, 68], which could lead to a more significant variance of network graph metrics computed from their bibliometric corpora.

For example, systematic literature reviews conducting rigorous and precise literature selection processes could have a greater chance to come up with a final list of included publications that precisely reflect the actual bibliometric development of the research field they are trying to summarize.

On the other hand, a literature review without a systematic selection process could miss important articles or include articles that were not decisive in the bibliometric development leading to a more distorted reflection of the actual research field.

These examples need to be combined with the assumption that bibliometric data and its citation structures could inherit reoccurring patterns disregarding the field of research due to the underlying standardized processes of modern scientific research in aggregating knowledge. Yet, no prior literature that supports this assumption could be found.

Conclusively, systematic literature reviews that reflect the bibliometric development more precisely could lead to network graphs including these reoccurring patterns and thus, have lesser variance across computed metrics.

On the contrary, literature reviews including subjectively selected articles could contain a deformed reflection of the research field where these exact reoccurring patterns are distorted, leading to more significant variance among their graph metrics.

However, this argument is very hypothetical. Each field of research is subject to its disjunct evolution, which likely emerges vastly different relationships among included bibliographic entities, depending on the historical depth and scattering of subfields. Thus, the main reason for the diverse composition of the bibliographic network graphs is most likely the scope of the research question and the thematic influences in the review.

With that said, literature reviews investigating interdisciplinary questions could show a more clustered and less connected inner citational behavior compared to monothematic literature reviews.

Still, given the above-stated arguments, this paper assumes that the network graph metrics stemming from the corpora of systematic literature reviews with a rigorous and algorithmic article selection process could inherit a lower variance than literature reviews with a somewhat subjective selection process.

Even though this thesis includes literature reviews and systematic literature reviews, a comparison of variances among these groups was not conducted as the sample size of both groups was considered too small.

Therefore, this paper draws inferences from systematic reviews and literature reviews without distinguishing them.

In summary, different literature corpora will most likely always exhibit some degree of variance in certain network graph metrics that cannot be attributed to general characteristics of the review, such as the temporal dispersion of the included articles and the size of the review itself. The graph's structure also depends significantly on the research field, its interdisciplinarity, the specificity of the research question, and the selection process of the included articles.

6.2. Time Periods

Regarding average references of core publications, it is noticeable that the outliers, which indicate the minimum and maximum, stem from literature reviews with the oldest and one of the most recent publication dates.

It would make sense for recently published reviews to have a higher chance of including more recent articles.

The bibliometric databases provide a better reference list coverage for newer articles. For example, in the case of Scopus, a study showed that older publications more frequently had fewer references indexed than these publications reference in reality [72].

Thus, the author of this paper suggests that future studies should consider the time tendency in which the core publications of a literature review were released because different coverage periods could lead to significantly different quality of the extracted data. And thus more accurate or inaccurate graph results.

6.3. Components and Clusters

The expected number of components for a direct citation graph stemming from a literature review in the information systems ranges from 2 to 7, as indicated by the median's confidence interval. When comparing the number of the components to the unconnected core publications, it becomes evident that smaller components consist of only one core publication except for one, which consists of two. Thus, larger separate components within a bibliometric corpus of a literature review are expected to be unlikely.

The cluster groups could, in general, denote different thematic groups among the publications, as the connectedness of publications addressing dissimilar research topics is likely to be low in network graphs, as these are less likely to have a direct or indirect citation relationship [77].

However, the median number of computed clusters for the direct citation graphs is around 24, a remarkably high thematic division for a literature review addressing a specific research question. Thus, this thesis suggests that the bibliographic coupling graph should be used for thematic clustering when employing the Louvain method, as the median of 5 denotes a more realistic thematic division. However, this should be done with caution, as the cluster computation of the bibliographic coupling graph resulted in low modularity values showing a less significant distinction of the cluster groups. It is also questionable whether the number of thematic subdivisions in the sample literature reviews is as consistent as the distribution indicates.

It is important to note that the difference in the number of clusters among these graphs can partly be explained by the bibliographic coupling graph not integrating unconnected core articles that form singular clusters in the direct citation graph. Still, the remaining difference among the cluster numbers indicates that the Louvain Algorithm computes more strictly related clusters for the direct citation graph. In contrast, it creates a broader cluster division for the bibliographic coupling graph.

The general nature of the different graph types could explain this, as the direct citation graph only integrates natural relations among the publications within, leading to a sparsely connected network with more distinct groups. At the same time, the bibliographic coupling method creating connections based on indirect similarities allows a greater cohesion among bibliographic entities and fewer separate groups among them [20].

6.4. General Thematic Cohesiveness

The bibliographic coupling graph showed considerably high density values among all core publications, reflecting a literature review's general thematic cohesiveness. This is also confirmed by the average shortest path results, which indicate a direct or indirect relation over one other article for most core publications.

6.5. Correlation

A linear relation evident in the scatter plots and weak to slightly moderate Pearson correlation coefficients supported the following hypotheses:

Hypothesis 1: "A higher number of core publications included in a literature review negatively correlates with the share of direct reciprocal citations between them."

Hypothesis 2: "A higher number of core publications included in a literature review negatively correlates with the common share of references among them."

Hypothesis 4: "A larger dispersion of the publication dates of core papers included in a literature review positively correlates with a share of direct reciprocal citations among core papers."

The supported hypotheses 1 and 2 reinforce the thesis' claim that a higher number of publications integrated into a literature review could be related to a lesser thematic cohesiveness among these.

The reasoning for this claim is the assumption that authors including more articles in their study were more likely to go after an interdisciplinary research question requiring more references. Or dive deeper into their monodisciplinary research question, leading to the inclusion of niche and interdisciplinary sub-articles.

The supported hypothesis 4 backs the assumption that a mutual citation relationship among publications is more likely when both papers have a sufficient temporal distance in their publication years. This makes sense as the preliminary theoretical work of the later published publication could already be finished before the prior study is released. Additionally, the relevance and visibility of a newly released publication could take some time to build up.

On the other hand, graphical display resulted in no correlation relation being evident the following hypotheses:

Hypothesis 3: "A larger dispersion of the publication dates of core papers included in a literature review negatively correlates with the common share of references among them."

Hypothesis 5: "A higher count of average core paper references in a literature review positively correlates with the share of direct reciprocal citations among core papers."

The rejected third hypothesis contradicts the author's assumption that a more significant difference in release dates of thematic related publications will diminish the probability of overlaps within their reference list. This effect was called citation genealogy and was initially fathomed by [26].

Lastly, the rejected fifth hypothesis negates the assumption of the author that the reciprocal citation behavior among core publications in a literature review is biased to their average reference count.

Nevertheless, it is essential to note that the computed p-values for the initially supported hypotheses could not show any statistical significance. Additionally, the small sample sizes lead to low reliability of the p-values and correlational analysis in general. [37]. Thereby diminishing any certainty of hypotheses 1, 2, and 4 being valid and the falsification of hypotheses 3 and 5.

6.6. Limitations

6.6.1. Sample Size

The literature reviews that resulted from the systematic search in the Senior Scholars' Basket of Journals proved to be a good fit for this thesis, as extracting bibliometric data of the corpora and the subsequent creation of graphs requires a lot of work. Thus, the resulting 14 reviews provided an applicable frame of effort for this thesis. However, in light of this small sample size, the empirical analysis lacks reliability.

It was impossible to assess the different distributions of the metric populations with certainty. Thus, the small sample limited the applicability of many statistical methods not used in this thesis while diminishing the outcome certainty of applied statistical methods used in this thesis.

More robust measures were utilized, such as the median, the interquartile range, a confidence interval with order statistics, and the Spearman correlation. Other statistical measures could have provided further insights but were avoided given the ambiguous data situation.

Thus, future research should aggregate more samples and employ the same analytical measures to make more assumptions about how connected, dense, and clustered one would expect these graphs to be based on empirical data.

6.6.2. Lack of Preprocessing

In bibliometric analysis, It is essential to preprosses the retrieved data before mapping and analyzing it, as algorithms employed on flawed incoherent input data will result in output data providing room for misleading interpretations [31]. Bibliometric data could have spelling errors, duplicates, or different names or abbreviations for the same bibliographic entity [15].

This is done by employing preprocessing methods such as duplicate detection or grouping different spelled items together with the help of string distance measures such as the Levenshtein distance [15].

However, given the limited scope of this thesis, no such preprocessing methods were employed as they also inherit the chance of creating false positives when mapping two bibliographic items together based on similar identifiers. Especially in this case, as the author of this thesis has no prior experience with these methods and their required threshold settings. For these reasons, the inferences made in this thesis about the corpora of literature reviews are to some extent based on data that could contain a significant amount of noise.

In this context, the author of this paper proposes that future researchers in bibliometric analysis should develop a guide to allow standardized extraction and preprocessing of bibliometric data with clear defined thresholds. Therefore, improving the quality of inferences being drawn from a future bibliometric network analysis of literature reviews.

7. Conclusion

The present study analyzed network graphs created from the bibliometric corpora of 14 different literature reviews given suitable metrics and statistically addressed how dense, clustered, and connected one would expect these types of network graphs to be.

Two different types of bibliometric network graphs were created from each corpus. A direct citation graph mapping the entire corpus, including directly included publications of a review and their references, and a bibliometric coupling graph solely mapping the relation among core publications, based on their common share of references. The literature corpora varied in size and temporal composition of release dates of included publications.

The results show that the metrics exhibit a degree of variability that cannot be explained solely by the influence of general external characteristics such as the size of the corpus and the temporal disposition of the included publications but requires further examination of the research question and the author's selection process. Especially the core density and the average citation overlap were varying to a great extent.

Regarding these main measures, the computed confidence intervals denote that most literature reviews within the information systems field have a density of directly included publications (core publications) citing each other reciprocally of 3.9% to 11.5%. Furthermore, the average share of common references between core publications is likely to be around 1.1% and 2.6%.

The outliers in the average amount of references and prior research led to the assumption that the extraction of bibliometric data from literature reviews that include publications of distinct time periods could lead to a different quality of the extracted data, thus impacting the resulting network graph composition. However, this assumption was not further analyzed and needs further confirmation.

It is evident that the literature reviews mostly contained a small number of core publications that were entirely disjunct from the main corpus, given no common share in references. Also, these disjunct publications were not forming groups among each other, indicating no significantly large disjunct components, only disconnected singular core publications from the main corpus.

The clustering analysis of the corpora led to significantly different results for the two graph types. The median number of computed clusters for the bibliographic graphs is 5, which denotes a more realistic thematic division of a literature review than the median of around 24 for the direct citation graph. Therefore, this paper suggests the usage of the bibliographic coupling graph instead of the direct citation graph for thematic clustering when applying the Louvain method.

Correlation analysis revealed a weak to slightly negative relationship between the number of included publications in a review and their thematic coherence. Additionally, a higher temporal dispersion among the release dates of included publications in a review indicated a weak positive relation to the density of their reciprocal citations.

On the other hand, the sample data could not verify the expected negative correlation between the temporal dispersion of release dates and the common share of references among core publications in a review. Also, no correlation bias was found between the average number of references and the mutual citation among core publications. Additionally, it is to note that statistical tests could not prove significance for any of the correlations initially supported through the graphical display in the scatter plots. Generally, the reliability of the correlation analysis is to be put into question, given the small sample size.

This quantitative research thesis provides a baseline of expectancy for the resulting density, connectedness, and clustering of network graphs created from the bibliometric data of literature reviews and insights into their dependency relationships.

However, further research is required to aggregate more graph samples and provide additional statistical certainty to the computed values and the inferences drawn from them in this thesis.

This thesis stresses the importance of a guideline in extracting and preprocessing bibliometric data to streamline the process of the network-level analysis of literature reviews, thus allowing for comparable macro-level graph results in future research.

Bibliography

[1] Gexf file format. `http://gexf.net/`. Accessed: 2022-02-15.

[2] Gaurav Agarwal and David Kempe. Modularity-maximizing graph communities via mathematical programming. The European Physical Journal B, 66(3):409–418, 2008.

[3] Herman Aguinis, Charles A Pierce, Frank A Bosco, Dan R Dalton, and Catherine M Dalton. Debunking myths and urban legends about meta-analysis. Organizational Research Methods, 14(2):306–331, 2011.

[4] Per Ahlgren, Cristian Colliander, and Olle Persson. Field normalized citation rates, field normalized journal impact and norwegian weights for allocation of university research funds. Scientometrics, 92(3):767–780, 2012.

[5] Ahmed N Albatineh, BM Golam Kibria, Meredith L Wilcox, and Bashar Zogheib. Confidence interval estimation for the population coefficient of variation using ranked set sampling: a simulation study. Journal of Applied Statistics, 41(4):733–751, 2014.

[6] Bulent Altunkaynak and Hamza Gamgam. Bootstrap confidence intervals for the coefficient of quartile variation. Communications in Statistics-Simulation and Computation, 48(7):2138–2146, 2019.

[7] Tasleem Arif. The mathematics of social network analysis: metrics for academic social networks. International Journal of Computer Applications Technology and Research, 4 (12):889–893, 2015.

[8] Zeinab Baghizadeh, Dubravka Cecez-Kecmanovic, and Daniel Schlagwein. Review and critique of the information systems development project failure literature: An argument for exploring information systems development project distress. Journal of Information Technology, 35(2):123–142, 2020.

[9] Christopher W Belter. Citation analysis as a literature search method for systematic reviews. Journal of the Association for Information Science and Technology, 67(11): 2766–2777, 2016.

[10] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment, 2008(10):P10008, 2008.

[11] Douglas G Bonett. Confidence interval for a coefficient of quartile variation. Computational statistics & data analysis, 50(11):2953–2957, 2006.

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

[13] Chaomei Chen. Science mapping: a systematic review of the literature. Journal of data and information science, 2(2), 2017.

[14] Chaomei Chen and Min Song. Visualizing a field of research: A methodology of systematic scientometric reviews. PloS one, 14(10):e0223994, 2019.

[15] Manuel J Cobo, Antonio Gabriel López-Herrera, Enrique Herrera-Viedma, and Francisco Herrera. Science mapping software tools: Review, analysis, and cooperative study among tools. Journal of the American Society for information Science and Technology, 62(7): 1382–1402, 2011.

[16] Thomas Cowhitt, Timothy Butler, and Elaine Wilson. Using social network analysis to complete literature reviews: a new systematic approach for independent researchers to detect and interpret prominent research programs within large collections of relevant literature. International Journal of Social Research Methodology, 23(5):483–496, 2020.

[17] John D'arcy and Tejaswini Herath. A review and analysis of deterrence theory in the is security literature: making sense of the disparate findings. European journal of information systems, 20(6):643–658, 2011.

[18] David P Doane and Lori E Seward. Measuring skewness: a forgotten statistic? Journal of statistics education, 19(2), 2011.

[19] Naveen Donthu, Satish Kumar, Debmalya Mukherjee, Nitesh Pandey, and Weng Marc Lim. How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133:285–296, 2021. ISSN 0148-2963. doi: https://doi.org/10.1016/j.jbusres.2021.04.070. URL `https://www.sciencedirect.com/science/article/pii/S0148296321003155`.

[20] Nees Jan Van Eck and Ludo Waltman. Visualizing bibliometric networks. In Measuring scholarly impact, pages 285–320. Springer, 2014.

[21] Leo Egghe and Ronald Rousseau. Introduction to informetrics: Quantitative methods in library, documentation and information science. Elsevier Science Publishers, 1990.

[22] Ole Ellegaard and Johan A Wallin. The bibliometric analysis of scholarly production: How great is the impact? Scientometrics, 105(3):1809–1831, 2015.

[23] Eugene Garfield, Irving H Sher, and Richard J Torpie. The use of citation data in writing the history of science. Technical report, Institute for Scientific Information Inc Philadelphia PA, 1964.

[24] Asghar Ghasemi and Saleh Zahediasl. Normality tests for statistical analysis: a guide for non-statisticians. International journal of endocrinology and metabolism, 10(2):486, 2012.

[25] Wolfgang Glänzel. National characteristics in international scientific co-authorship relations. Scientometrics, 51(1):69–115, 2001.

[26] Wolfgang Glänzel and Bart Thijs. Using 'core documents' for detecting and labelling new emerging topics. Scientometrics, 91(2):399–416, 2012.

[27] Jochen Gläser, Wolfgang Glänzel, and Andrea Scharnhorst. Same data, different results? towards a comparative approach to the identification of thematic structures in science. Scientometrics, 111(2):981–998, 2017.

[28] Wendy Arianne Günther, Mohammad H Rezazade Mehrizi, Marleen Huysman, and Frans Feldberg. Debating big data: A literature review on realizing value from big data. The Journal of Strategic Information Systems, 26(3):191–209, 2017.

[29] ACJW Janssens, Marta Gwinn, J Elaine Brockman, Kimberley Powell, and Michael Goodman. Novel citation-based search method for scientific literature: a validation study. BMC medical research methodology, 20(1):1–11, 2020.

[30] Hemin Jiang, Mikko Siponen, and Aggeliki Tsohou. Personal use of technology at work: a literature review and a theoretical model for understanding how it affects employee job performance. European Journal of Information Systems, pages 1–15, 2021.

[31] LI Jie1 GUO Xiao-hong JIANG and Kang1 Lü Peng-hui. Preliminary study of knowledge map of safety science: Base on data of safety science. China Safety Science Journal, 2013.

[32] Yu Jiang, Can Lin, Weiyi Meng, Clement Yu, Aaron M Cohen, and Neil R Smalheiser. Rule-based deduplication of article records from bibliographic databases. Database, 2014, 2014.

[33] Koen Jonkers and GE Derrick. The bibliometric bandwagon: Characteristics of bibliometric articles outside the field literature. Journal of the American Society for Information Science and Technology, 63(4):829–836, 2012.

[34] Maxwell Mirton Kessler. Bibliographic coupling between scientific papers. American documentation, 14(1):10–25, 1963.

[35] Richard Klavans and Kevin W Boyack. Using global mapping to create more accurate document-level maps of research fields. Journal of the American Society for information Science and Technology, 62(1):1–18, 2011.

[36] Richard Klavans and Kevin W Boyack. Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? Journal of the Association for Information Science and Technology, 68(4):984–998, 2017.

[37] Stephen Kokoska and Daniel Zwillinger. CRC standard probability and statistics tables and formulae. Crc Press, 2000.

[38] Charles J Kowalski. On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. Journal of the Royal Statistical Society: Series C (Applied Statistics), 21(1):1–12, 1972.

[39] Martin G Larson. Descriptive statistics and graphical displays. Circulation, 114(1): 76–81, 2006.

[40] Jie Li, Judy F Burnham, Trey Lemley, and Robert M Britton. Citation analysis: Comparison of web of science®, scopus, scifinder®, and google scholar. Journal of electronic resources in medical libraries, 7(3):196–217, 2010.

[41] Robert J. MacCoun. Biases in the interpretation and use of research results. Annual Review of Psychology, 49(1):259–287, 1998. doi: 10.1146/annurev.psych.49.1.259. URL `https://doi.org/10.1146/annurev.psych.49.1.259`. PMID: 15012470.

[42] Irina Marshakova. Citation networks in information science. Scientometrics, 3(1):13–25, 1981.

[43] Irina V Marshakova. System of document connections based on references. Nauchno-Tekhnicheskaya Informatsiya Seriya 2-Informatsionnye Protsessy I Sistemy, (6):3–8, 1973.

[44] Mohammad Moeini and Suzanne Rivard. Sublating tensions in the it project risk management literature: a model of the relative performance of intuition and deliberate analysis for risk assessment. Journal of the Association for Information Systems, 20(3):1, 2019.

[45] Pachayappan Murugaiyan and Venkatesakumar Ramakrishnan. A methodology for structured literature network meta-analysis. Journal of Modelling in Management, 2020.

[46] Caroline E Oehlhorn, Christian Maier, Sven Laumer, and Tim Weitzel. Human resource management and its impact on strategic business-it alignment: A literature review and avenues for future research. The Journal of Strategic Information Systems, 29(4):101641, 2020.

[47] Farideh Osareh. Bibliometrics, citation analysis and co-citation analysis: A review of literature i. 1996.

[48] Rúben Pereira and João Serrano. A review of methods used on it maturity models development: A systematic literature review and a critical analysis. Journal of information technology, 35(2):161–178, 2020.

[49] Antonio Perianes-Rodriguez, Ludo Waltman, and Nees Jan Van Eck. Constructing bibliometric networks: A comparison between full and fractional counting. Journal of Informetrics, 10(4):1178–1195, 2016.

[50] H Peters and A Van Raan. Structuring scientific activities by co-author analysis: An expercise on a university faculty level. Scientometrics, 20(1):235–255, 1991.

[51] Eugenio Petrovich. Science mapping. ISKO Encyclopedia of Knowledge Organization, 2020.

[52] Gabriele Piccoli and Blake Ives. It-dependent strategic initiatives and sustained competitive advantage: a review and synthesis of the literature. MIS quarterly, pages 747–776, 2005.

[53] Raminta Pranckutė. Web of science (wos) and scopus: The titans of bibliographic information in today's academic world. Publications, 9(1):12, 2021.

[54] Derek J De Solla Price. Networks of scientific papers: The pattern of bibliographic references indicates the nature of the scientific research front. Science, 149(3683):510–515, 1965.

[55] Nonthapat Pulsiri and Ronald Vatananan-Thesenvitz. Improving systematic literature review with automation and bibliometrics. In 2018 Portland International Conference on Management of Engineering and Technology (PICMET), pages 1–8. IEEE, 2018.

[56] Sumita Raghuram, Philipp Tuertscher, and Raghu Garud. Research note: mapping the field of virtual work: A cocitation analysis. Information Systems Research, 21(4): 983–999, 2010.

[57] Stephan Schneider and Ali Sunyaev. Determinant factors of cloud-sourcing decisions: reflecting on the it outsourcing literature in the era of cloud computing. Journal of Information Technology, 31(1):1–31, 2016.

[58] Scopus. Scopus: Content coverage guide, 2020.

[59] Vivek Kumar Singh, Prashasti Singh, Mousumi Karmakar, Jacqueline Leta, and Philipp Mayr. The journal coverage of web of science, scopus and dimensions: A comparative analysis. Scientometrics, 126(6):5113–5142, 2021.

[60] Mikko Siponen and Tero Vartiainen. Unauthorized copying of software and levels of moral development: A literature analysis and its implications for research and practice. Information Systems Journal, 14(4):387–407, 2004.

[61] Henry Small. 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, 1973.

[62] Hyebin Song. Stat 415 introduction to mathematical statistics, 2021. URL `https://online.stat.psu.edu/stat415/book/export/html/835`.

[63] Stephan Stahlschmidt and Dimity Stephen. Comparison of web of science, scopus and dimensions databases. KB Forschungspoolprojekt; DZHW: Hannover, Germany, 2020.

[64] Shazia Tabassum, Fabiola SF Pereira, Sofia Fernandes, and João Gama. Social network analysis: An overview. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(5):e1256, 2018.

[65] R Alexander Teubner and Jan Stockhinger. Literature review: Understanding information systems strategy in the digital age. The Journal of Strategic Information Systems, 29(4):101642, 2020.

[66] Timm Teubner. E-mail correspondence with thesis supervisor prof. dr. timm teubner regarding the metrics employed in this thesis. Received: 2022-01-25.

[67] Bart Thijs, Edgar Schiebel, and Wolfgang Glänzel. Do second-order similarities provide added-value in a hybrid approach? Scientometrics, 96(3):667–677, 2013.

[68] David Tranfield, David Denyer, and Palminder Smart. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management, 14(3):207–222, 2003. doi: https://doi.org/10.1111/1467-8551.00375. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-8551.00375`.

[69] Caleb M Trujillo and Tammy M Long. Document co-citation analysis to enhance transdisciplinary research. Science advances, 4(1):e1701130, 2018.

[70] HsingYi Tsai, Deborah Compeau, and Darren Meister. Voluntary use of information technology: an analysis and synthesis of the literature. Journal of Information Technology, 32(2):147–162, 2017.

[71] Lidwien Van De Wijngaert, Harry Bouwman, and Noshir Contractor. A network approach toward literature review. Quality & Quantity, 48(2):623–643, 2014.

[72] Nees Jan van Eck and Ludo Waltman. Accuracy of citation data in web of science and scopus. arXiv preprint arXiv:1906.07011, 2019.

[73] Nees Jan Van Eck, Ludo Waltman, Rommert Dekker, and Jan Van Den Berg. A comparison of two techniques for bibliometric mapping: Multidimensional scaling and vos. Journal of the American Society for Information Science and Technology, 61(12):2405–2416, 2010.

[74] AF Van Raan. Advances in bibliometric analysis: research performance assessment and science mapping. Bibliometrics Use and Abuse in the Review of Research Performance, 87:17–28, 2014.

[75] Ludo Waltman and Nees Jan Van Eck. A new methodology for constructing a publication-level classification system of science. Journal of the American Society for Information Science and Technology, 63(12):2378–2392, 2012.

[76] Ludo Waltman, Kevin W Boyack, Giovanni Colavizza, and Nees Jan van Eck. A principled methodology for comparing relatedness measures for clustering publications. Quantitative Science Studies, 1(2):691–713, 2020.

[77] Shenghui Wang and Rob Koopman. Clustering articles based on semantic similarity. Scientometrics, 111(2):1017–1031, 2017.

[78] Martin Wiener, Carol Saunders, and Marco Marabelli. Big-data business models: A critical literature review and multiperspective research framework. Journal of Information Technology, 35(1):66–91, 2020.

[79] Patrick Wilson. Unused relevant information in research and development. Journal of the American society for Information Science, 46(1):45–51, 1995.

[80] Xiao Xiao, Christopher B Califf, Saonee Sarker, and Suprateek Sarker. Ict innovation in emerging economies: a review of the existing literature and a framework for future research. Journal of Information Technology, 28(4):264–278, 2013.

[81] Erjia Yan and Ying Ding. Scholarly network similarities: How bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and coword networks relate to each other. Journal of the American Society for Information Science and Technology, 63(7):1313–1326, 2012.

[82] Jinhyuk Yun, Sejung Ahn, and June Young Lee. Return to basics: Clustering of scientific literature using structural information. Journal of Informetrics, 14(4):101099, 2020.

[83] Ivan Zupic and Tomaz Cater. Bibliometric methods in management and organization. Organizational Research Methods, 18(3):429–472, 2015. doi: 10.1177/1094428114562629. URL https://doi.org/10.1177/1094428114562629.

A. Appendix

A.1. Metrics for each Literature Review Corpus

Publications	Core Publications	IQR Publication Years	Average References	Average Citation Overlap	Core Density
Baghizadeh et al. [8]	92	8.25	46.3	0.007	0.039
D'arcy and Herath [17]	15	5.50	45.2	0.026	0.267
Günther et al. [28]	67	1.00	42.4	0.011	0.015
Moeini and Rivard [44]	75	10.00	31.0	0.025	0.114
Oehlhorn et al. [46]	70	14.00	45.2	0.014	0.094
Pereira and Serrano [48]	107	6.50	28.9	0.004	0.011
Piccoli and Ives [52]	56	10.00	39.2	0.014	0.053
Schneider and Sunyaev [57]	88	9.00	37.6	0.011	0.060
Siponen and Vartiainen [60]	20	7.25	20.4	0.017	0.111
Jiang et al. [30]	131	7.00	39.4	0.025	0.115
Teubner and Stockhinger [65]	33	4.00	65.1	0.029	0.080
Tsai et al. [70]	49	6.00	43.2	0.045	0.159
Wiener et al. [78]	55	2.00	47.1	0.006	0.006
Xiao et al. [80]	47	5.00	38.7	0.002	0.012

Table A.1.: Metrics regarding the core publications of each literature review

Publication	DC Nodes	DC Components	DC Modularity	DC Clusters
Baghizadeh et al. [8]	3317	2	0.809	32
D'arcy and Herath [17]	536	2	0.706	11
Günther et al. [28]	2248	3	0.773	28
Moeini and Rivard [44]	1462	2	0.673	20
Oehlhorn et al. [46]	2143	4	0.680	25
Pereira and Serrano [48]	2697	13	0.878	43
Piccoli and Ives [52]	1673	3	0.737	22
Schneider and Sunyaev [57]	2281	2	0.694	26
Siponen and Vartiainen [60]	332	3	0.718	13
Jiang et al. [30]	2957	1	0.610	22
Teubner and Stockhinger [65]	1460	1	0.668	16
Tsai et al. [70]	1378	1	0.624	20
Wiener et al. [78]	2347	7	0.881	33
Xiao et al. [80]	1747	9	0.912	33

Table A.2.: Metrics regarding the direct citation graphs of each literature review

Publication	BC Nodes	BC Density	BC Average Shortest Path	BC Diameter	BC Modularity	BC Clusters
Baghizadeh et al. [8]	89	0.253	1.88	4	0.429	5
D'arcy and Herath [17]	14	0.791	1.23	3	0.237	3
Günther et al. [28]	65	0.464	1.60	4	0.266	6
Moeini and Rivard [44]	74	0.489	1.56	4	0.334	7
Oehlhorn et al. [46]	66	0.541	1.46	3	0.250	5
Pereira and Serrano [48]	93	0.169	2.17	5	0.504	10
Piccoli and Ives [52]	54	0.461	1.57	3	0.253	6
Schneider and Sunyaev [57]	87	0.392	1.63	4	0.257	5
Siponen and Vartiainen [60]	17	0.456	1.61	3	0.125	3
Jiang et al. [30]	131	0.697	1.31	3	0.274	5
Teubner and Stockhinger [65]	33	0.831	1.17	2	0.224	5
Tsai et al. [70]	49	0.879	1.12	2	0.155	4
Wiener et al. [78]	49	0.356	1.75	4	0.336	6
Xiao et al. [80]	40	0.136	2.37	5	0.435	6

Table A.3.: Metrics regarding the bibliographic coupling graphs of each literature review

A.2. Additional Statistics

Metric	Mean	Standard Deviation	Fisher-Pearson Skew Coefficient
Core Publications	64.643	31.413	0.343
IQR Core Publication Years	6.821	3.270	0.218
Average Core Publication References	40.693	9.929	0.309
Average Citation Overlap	0.017	0.011	0.844
Core Density	0.081	0.069	1.196
Unconnected Core Publications	2.857	3.502	1.546

Table A.4.: Additional statistical measures of metrics regarding the core publications

Metric	Mean	Standard Deviation	Fisher-Pearson Skew Coefficient
DC Nodes	1898.429	816.675	-0.255
DC Components	3.786	3.384	1.600
DC Modularity	0.740	0.093	0.529
DC Clusters	24.571	8.441	0.357

Table A.5.: Additional statistical measures of metrics regarding the direct citation graphs

Metric	Mean	Standard Deviation	Fisher-Pearson Skew Coefficient
BC Nodes	61.500	30.819	0.424
BC Density	0.494	0.226	0.198
BC Average Shortest Path	1.602	0.346	0.668
BC Modularity	0.291	0.103	0.498
BC Clusters	5.429	1.678	1.031

Table A.6.: Additional statistical measures of metrics regarding the bibliographic coupling graphs

A.3. Gephi Visualization Settings

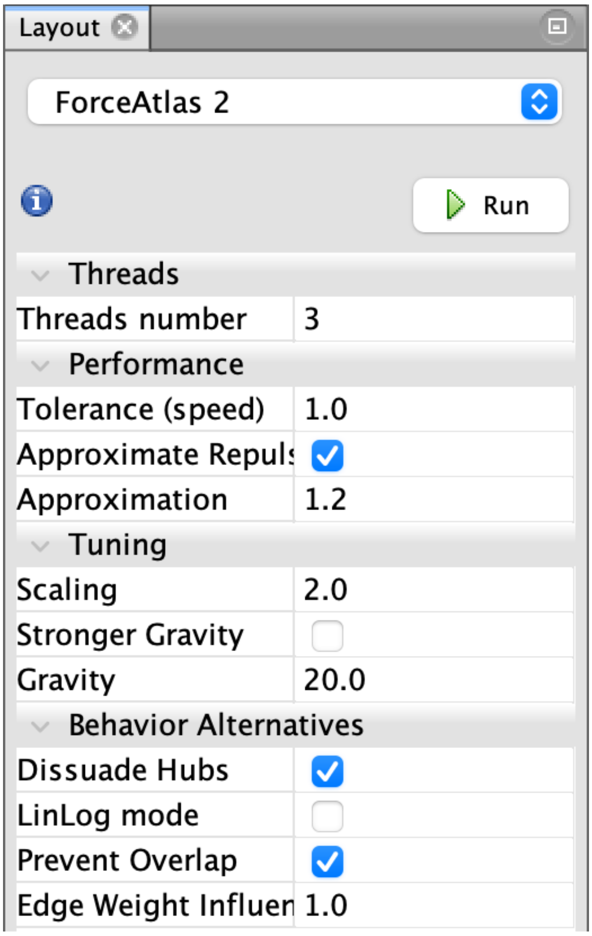


Figure A.1.: Gephi visualization settings for the direct citation graphs

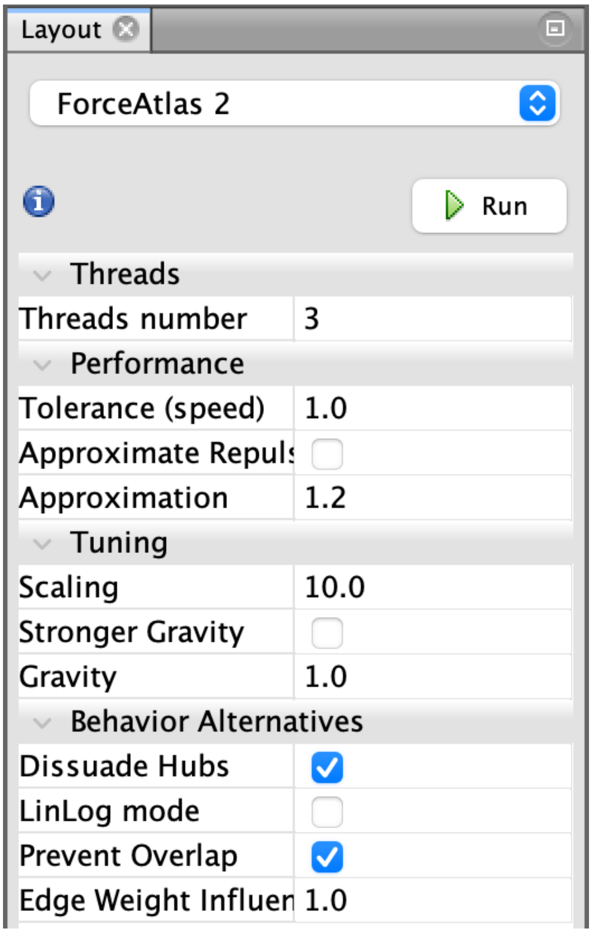


Figure A.2.: Gephi visualization settings for the bibliographic coupling graphs

A.4. Citation Network Graphs

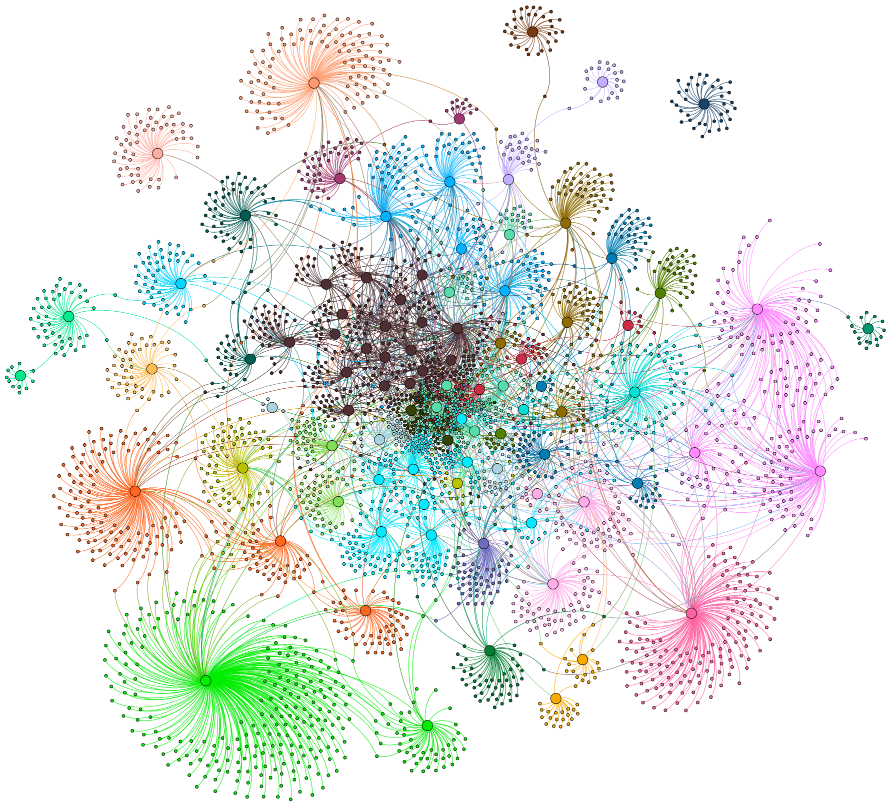


Figure A.3.: Direct citation graph of Baghizadeh et al. [8]

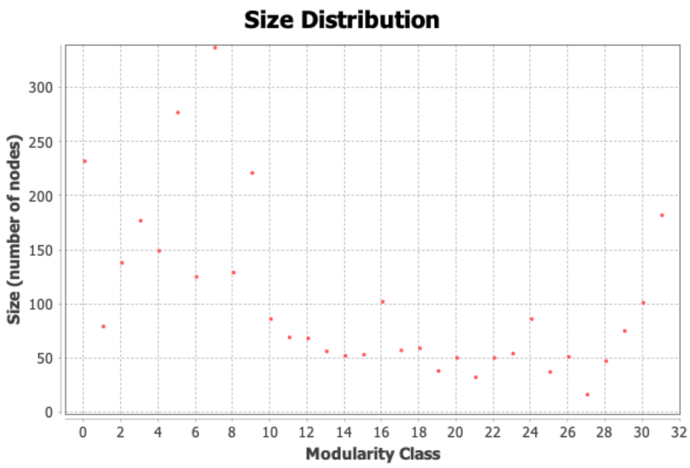


Figure A.4.: Cluster size distribution of direct citation graph of Baghizadeh et al. [8]

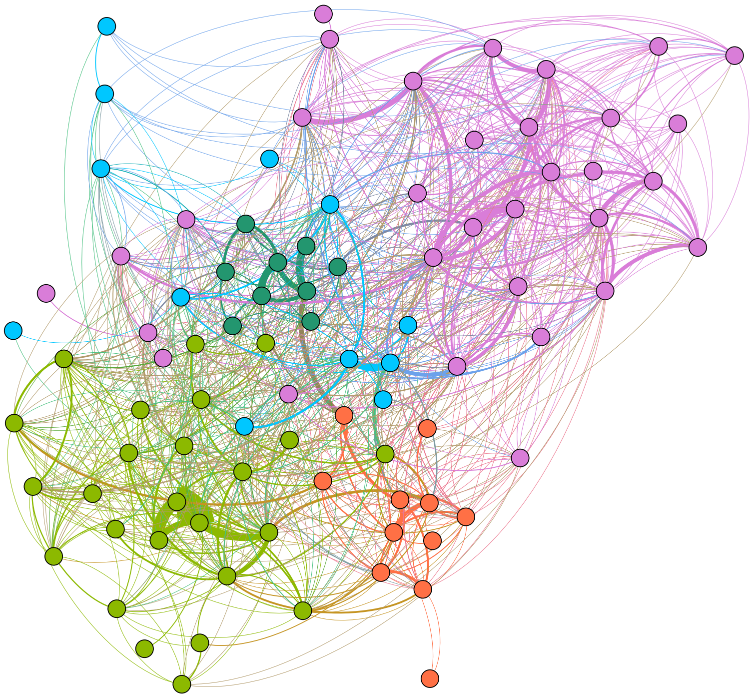


Figure A.5.: Bibliographic coupling graph of Baghizadeh et al. [8]

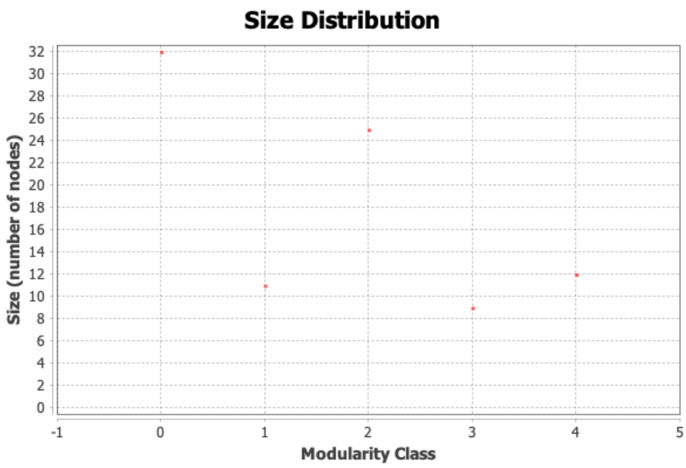


Figure A.6.: Cluster size distribution of bibliographic coupling graph of Baghizadeh et al. [8]

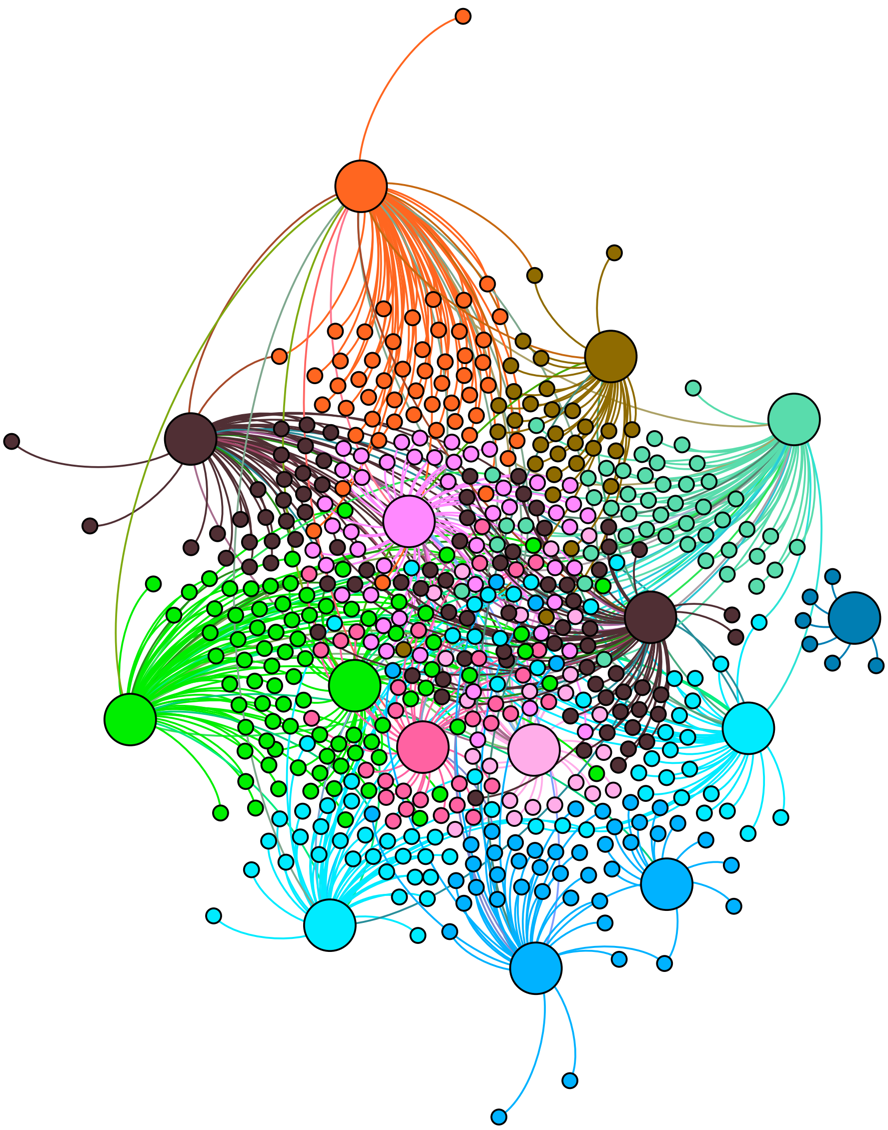


Figure A.7.: Direct citation graph of D'arcy and Herath [17]

Figure A.8.: Cluster size distribution of direct citation graph of D'arcy and Herath [17]

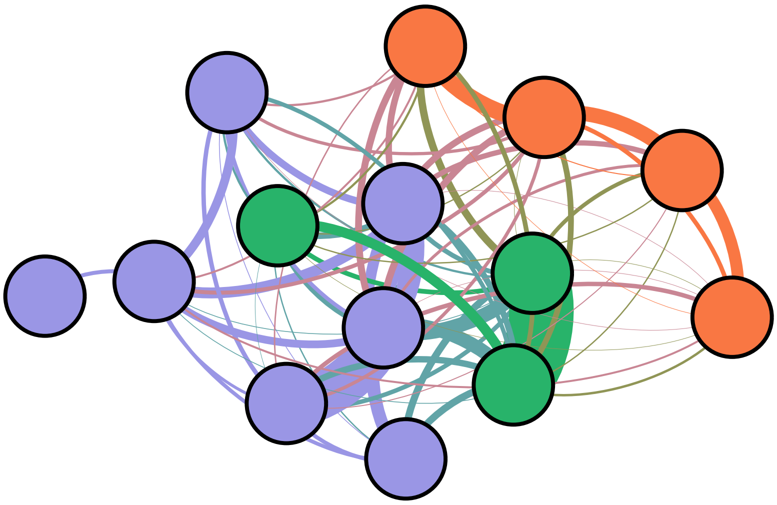


Figure A.9.: Bibliographic coupling graph of D’arcy and Herath [17]

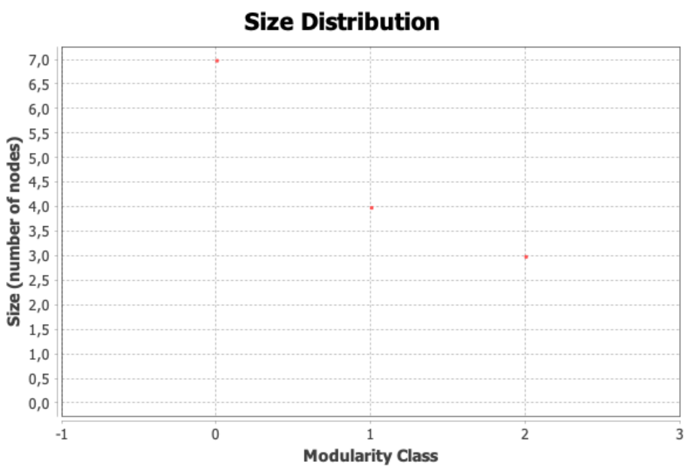


Figure A.10.: Cluster size distribution of bibliographic coupling graph of D’arcy and Herath [17]

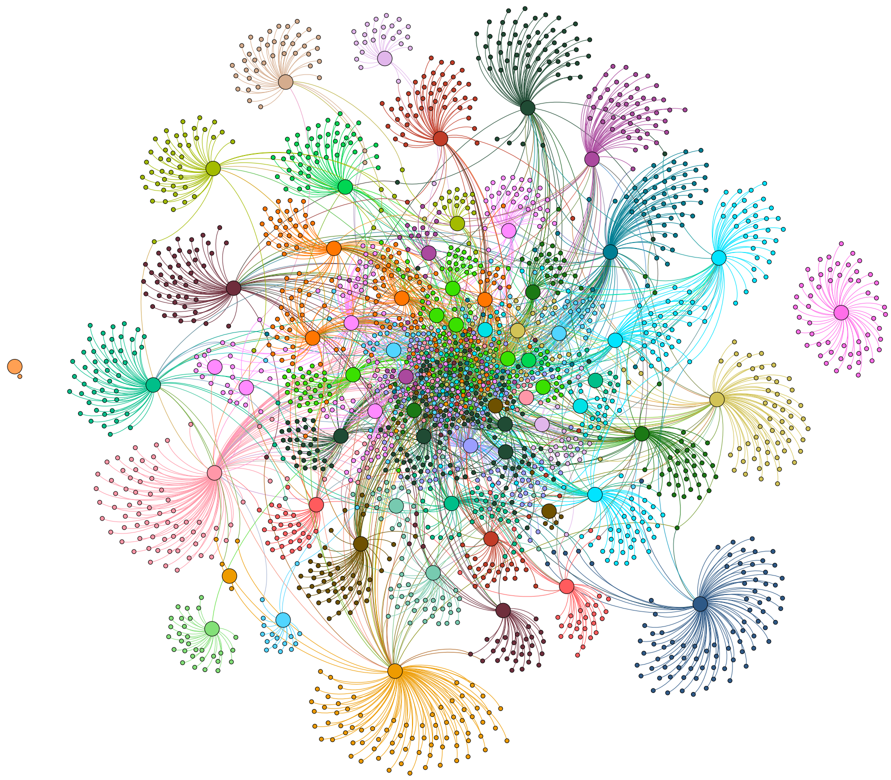


Figure A.11.: Direct citation graph of Günther et al. [28]

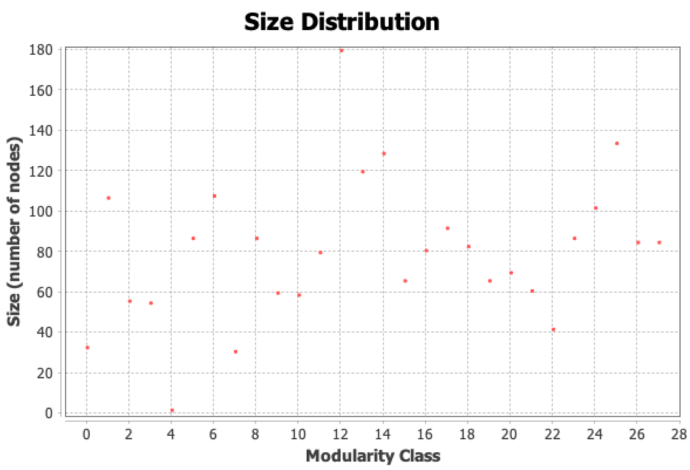


Figure A.12.: Cluster size distribution of direct citation graph of Günther et al. [28]

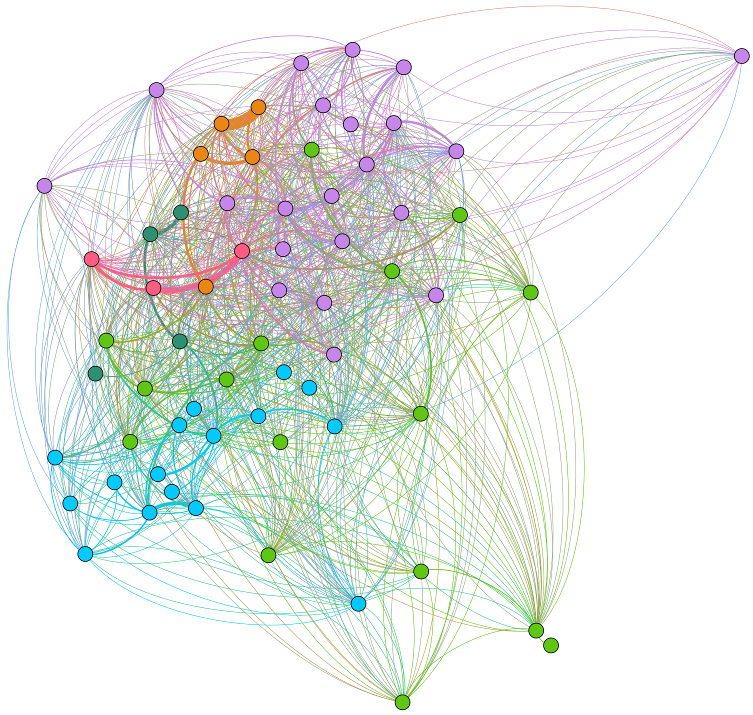


Figure A.13.: Bibliographic coupling graph of Günther et al. [28]

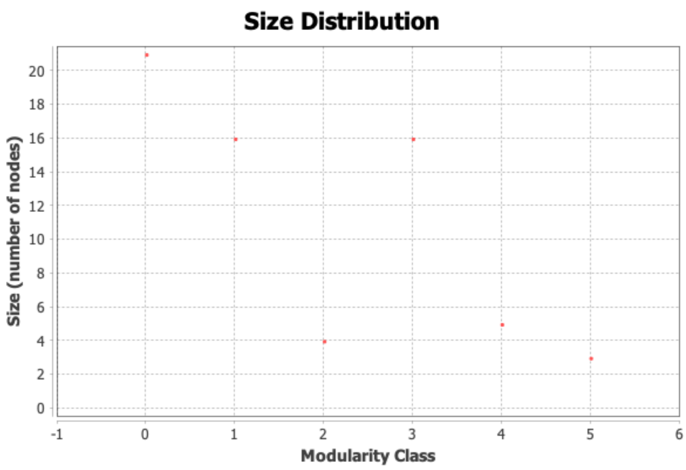


Figure A.14.: Cluster size distribution of bibliographic coupling graph of Günther et al. [28]

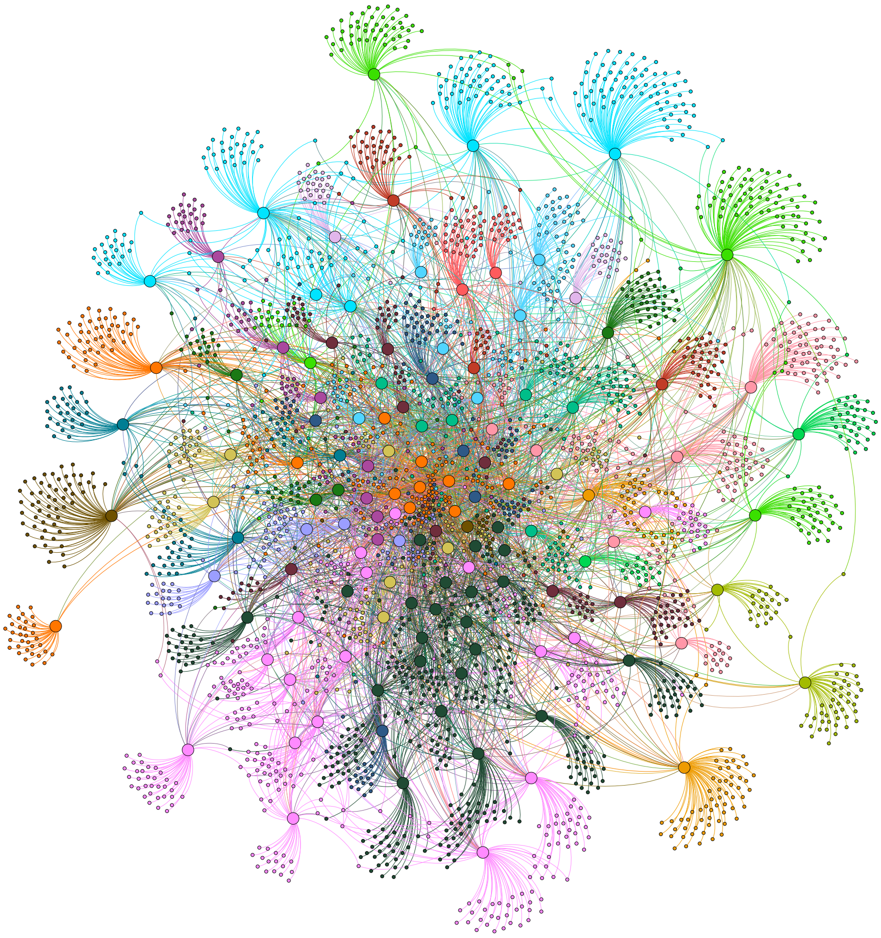


Figure A.15.: Direct citation graph of Jiang et al. [30]

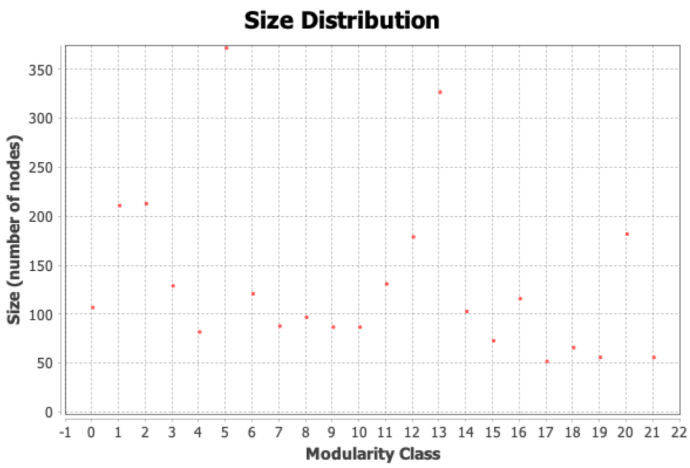


Figure A.16.: Cluster size distribution of direct citation graph of Jiang et al. [30]

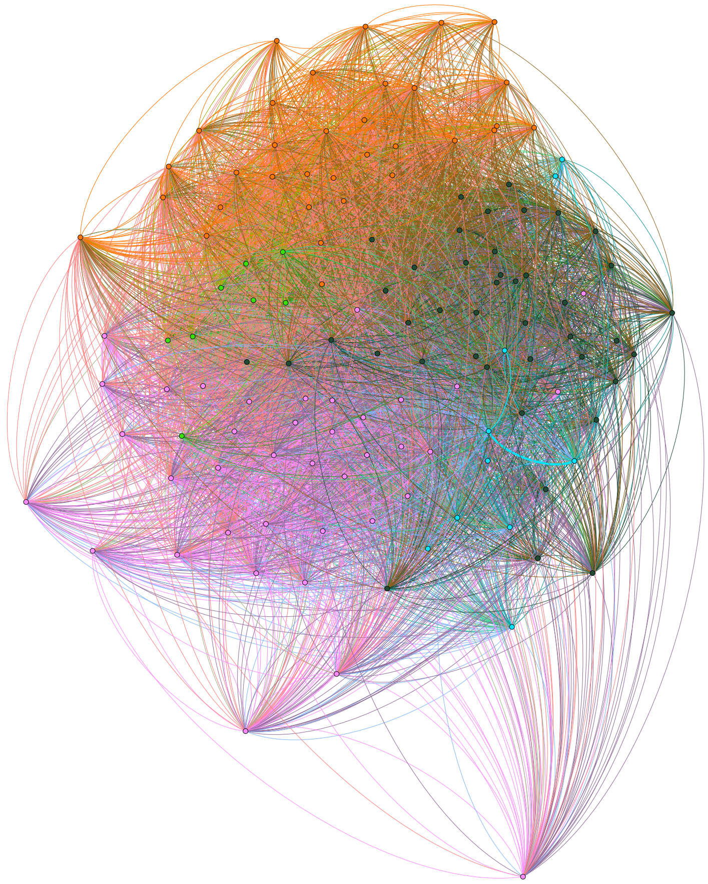


Figure A.17.: Bibliographic coupling graph of Jiang et al. [30]

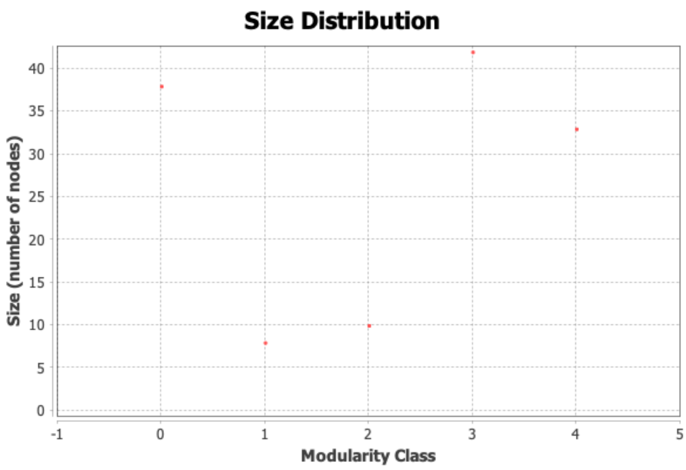


Figure A.18.: Cluster size distribution of bibliographic coupling graph of Jiang et al. [30]

Figure A.19.: Direct citation graph of Moeini and Rivard [44]

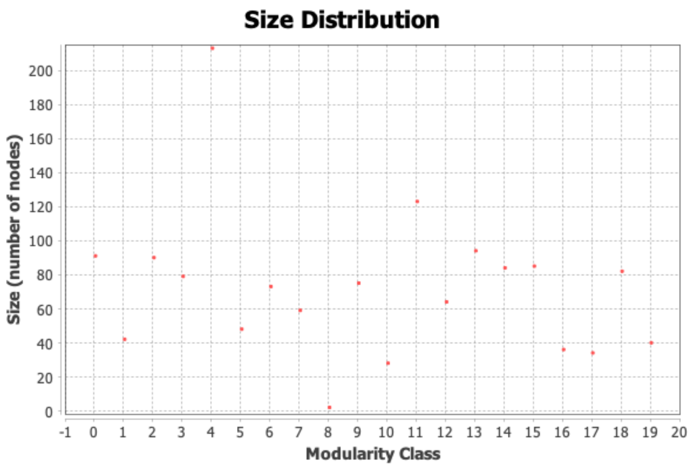


Figure A.20.: Cluster size distribution of direct citation graph of Moeini and Rivard [44]

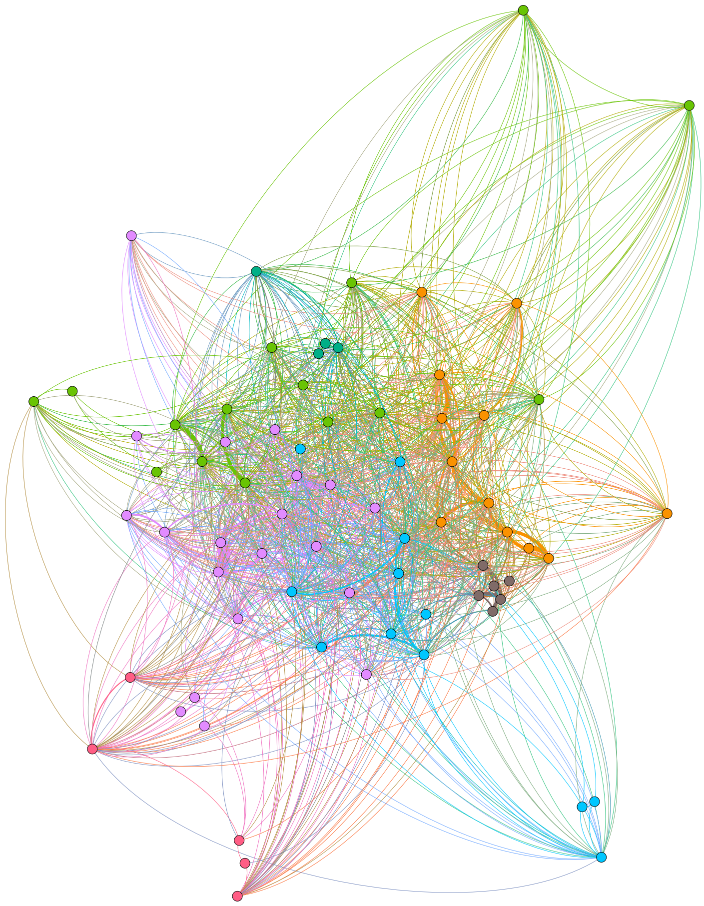


Figure A.21.: Bibliographic coupling graph of Moeini and Rivard [44]

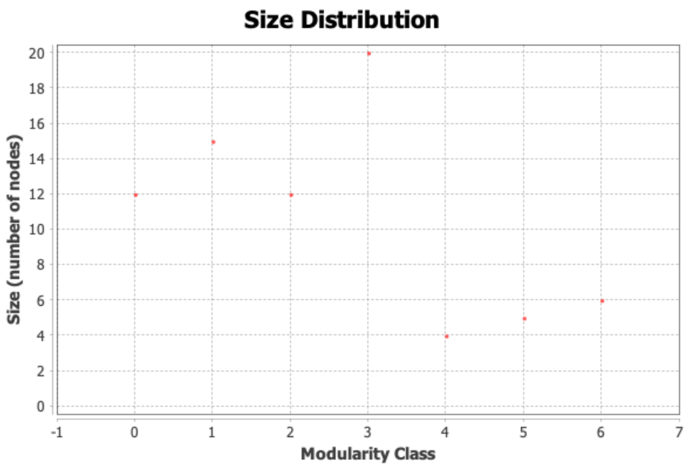


Figure A.22.: Cluster size distribution of bibliographic coupling graph of Moeini and Rivard [44]

Figure A.23.: Direct citation graph of Oehlhorn et al. [46]

Figure A.24.: Cluster size distribution of direct citation graph of Oehlhorn et al. [46]

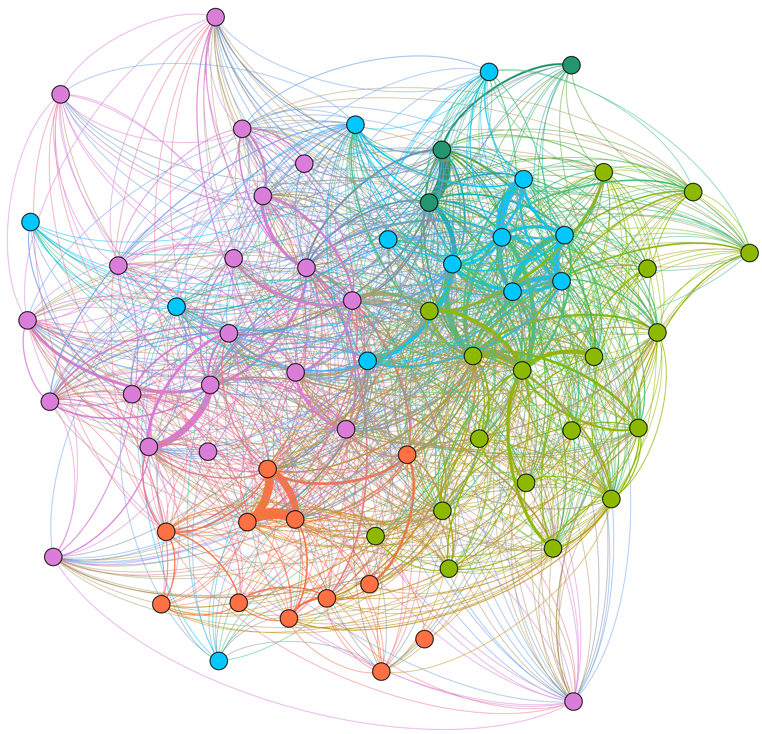


Figure A.25.: Bibliographic coupling graph of Oehlhorn et al. [46]

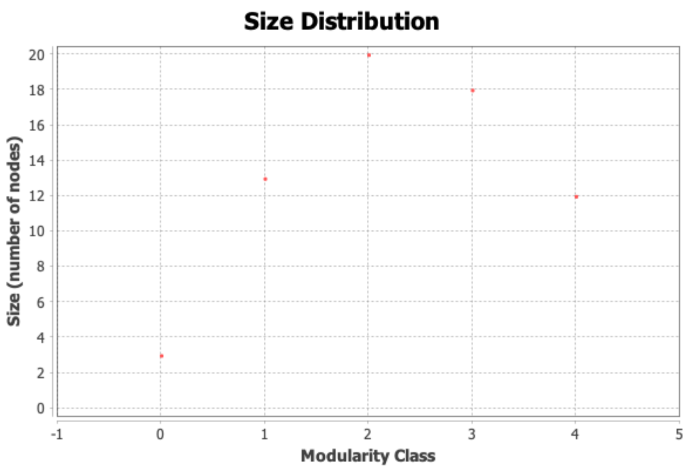


Figure A.26.: Cluster size distribution of bibliographic coupling graph of Oehlhorn et al. [46]

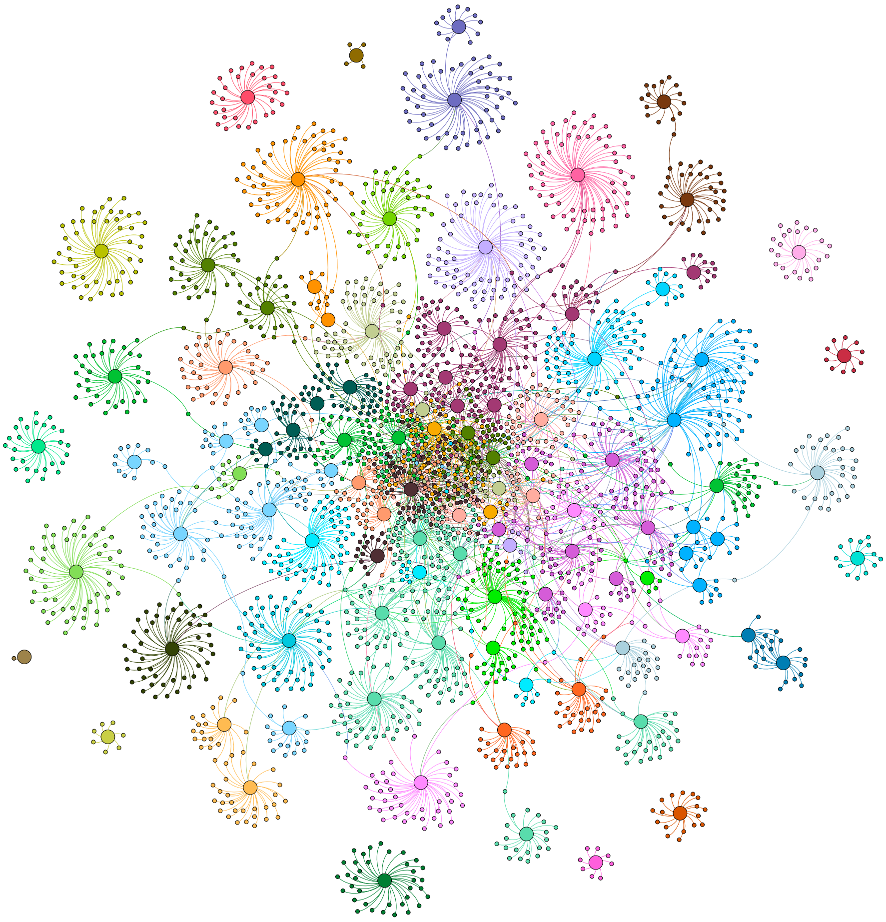


Figure A.27.: Direct citation graph of Pereira and Serrano [48]

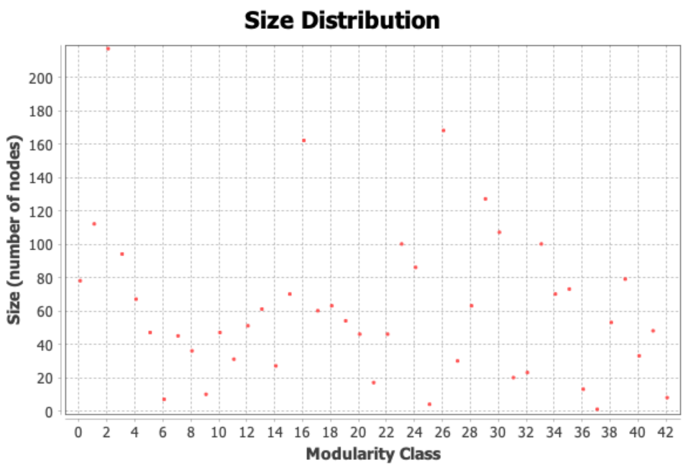


Figure A.28.: Cluster size distribution of direct citation graph of Pereira and Serrano [48]

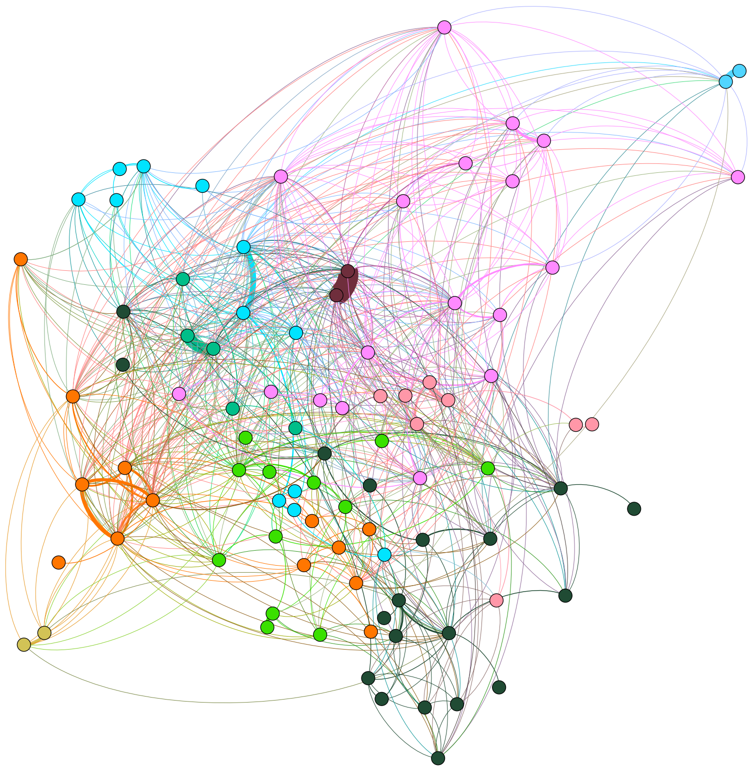


Figure A.29.: Bibliographic coupling graph of Pereira and Serrano [48]

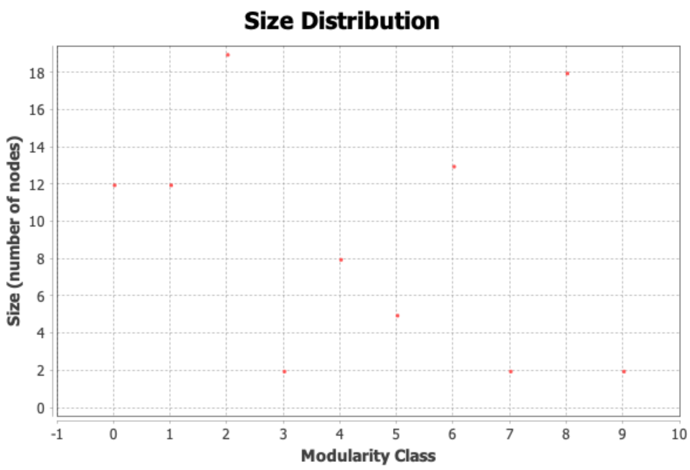


Figure A.30.: Cluster size distribution of bibliographic coupling graph of Pereira and Serrano [48]

Figure A.31.: Direct citation graph of Piccoli and Ives [52]

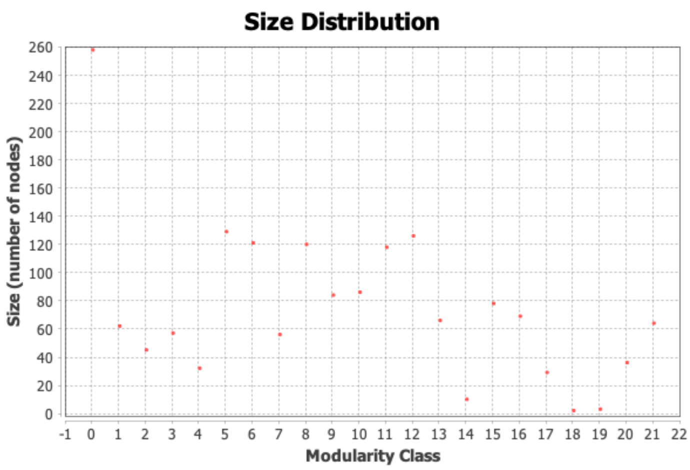


Figure A.32.: Cluster size distribution of direct citation graph of Piccoli and Ives [52]

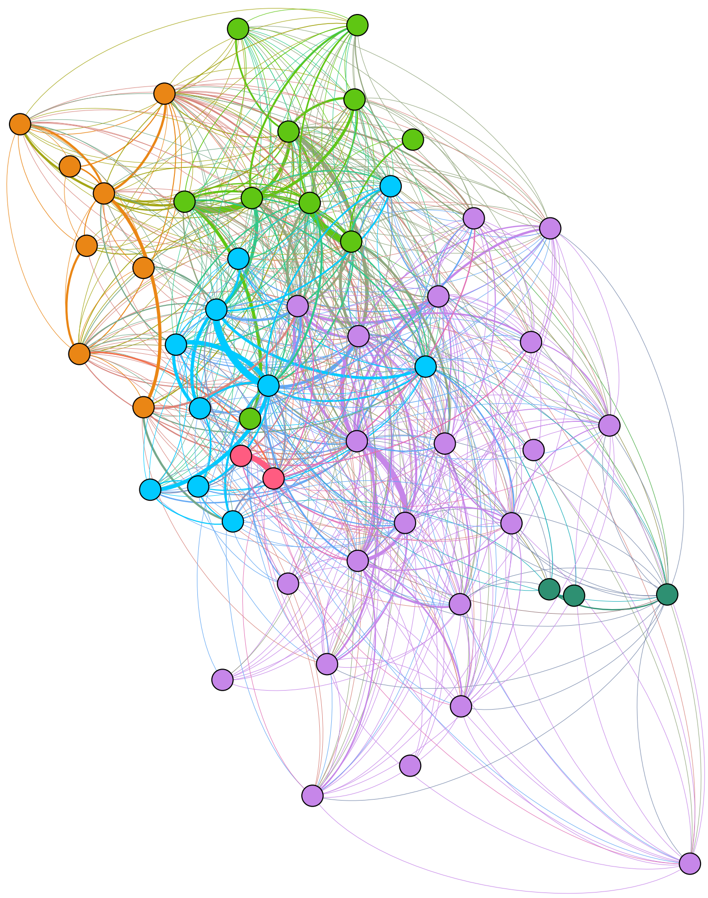


Figure A.33.: Bibliographic coupling graph of Piccoli and Ives [52]

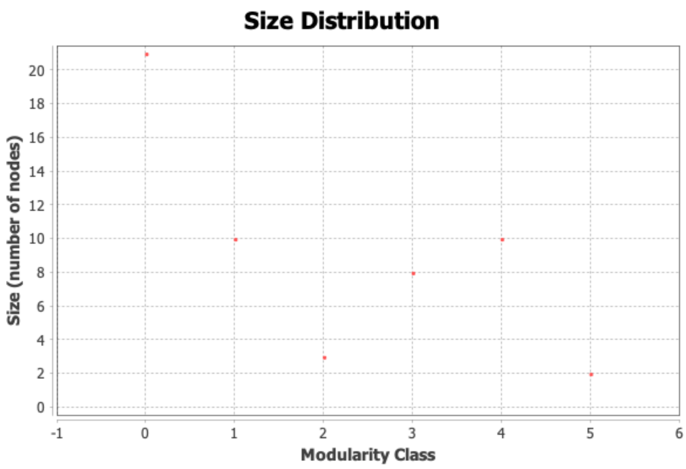


Figure A.34.: Cluster size distribution of bibliographic coupling graph of Piccoli and Ives [52]

Figure A.35.: Direct citation graph of Schneider and Sunyaev [57]

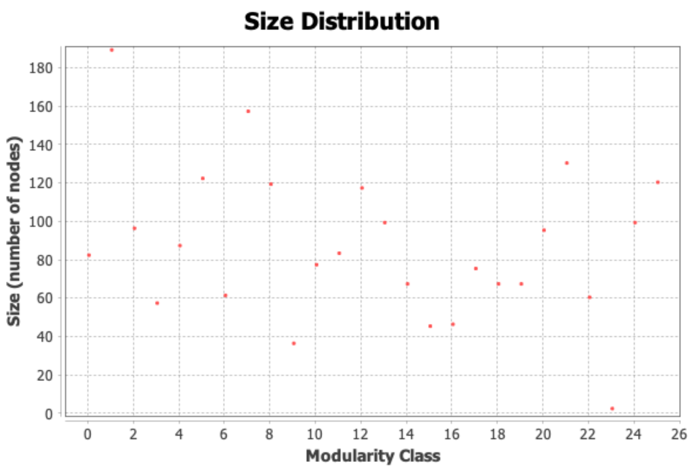


Figure A.36.: Cluster size distribution of direct citation graph of Schneider and Sunyaev [57]

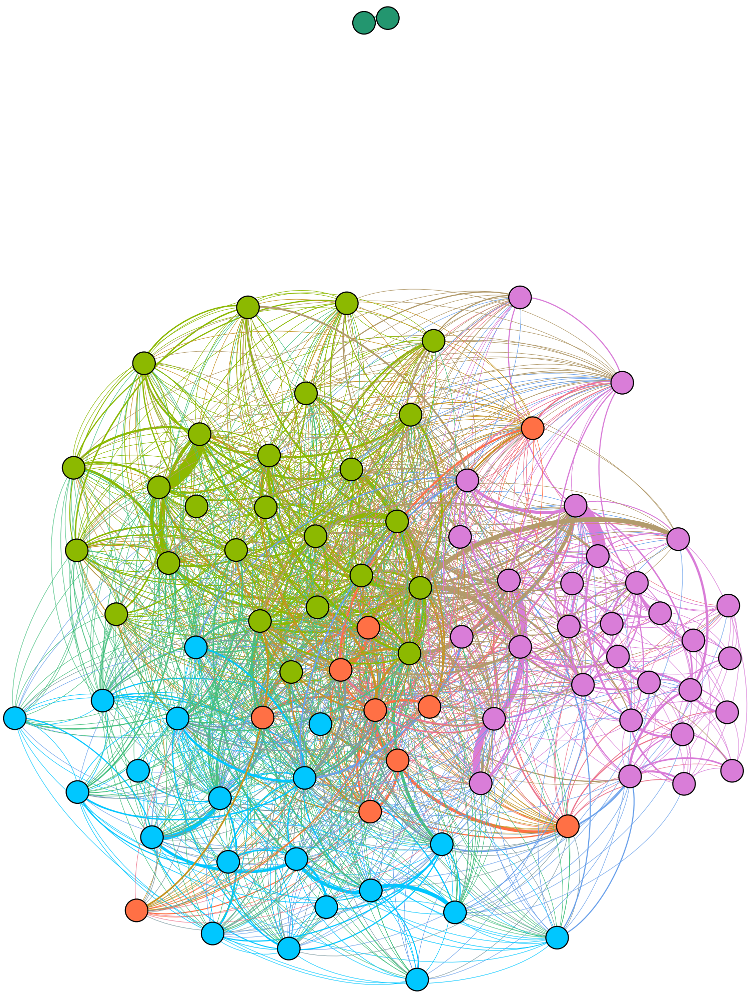


Figure A.37.: Bibliographic coupling graph of Schneider and Sunyaev [57]

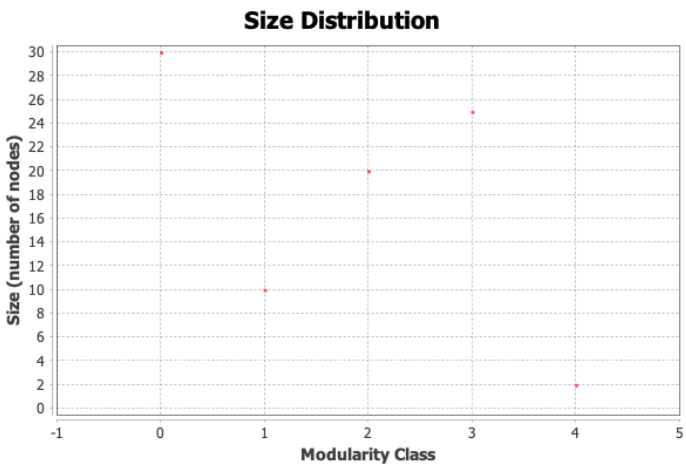


Figure A.38.: Cluster size distribution of bibliographic coupling graph of Schneider and Sunyaev [57]

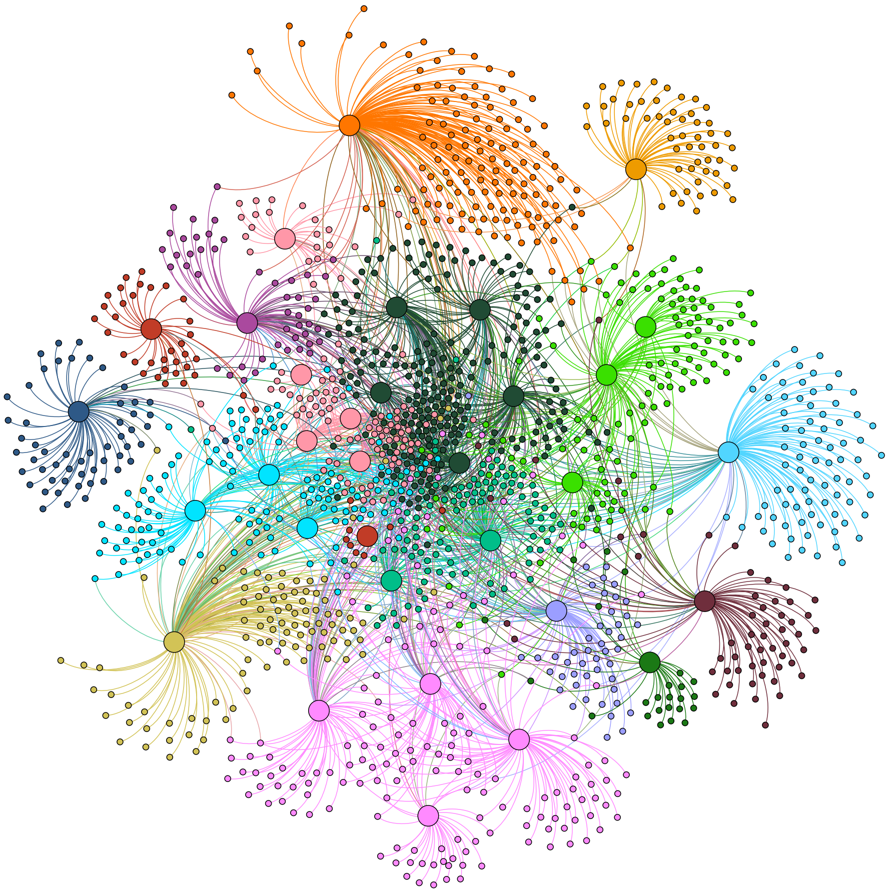


Figure A.39.: Direct citation graph of Teubner and Stockhinger [65]

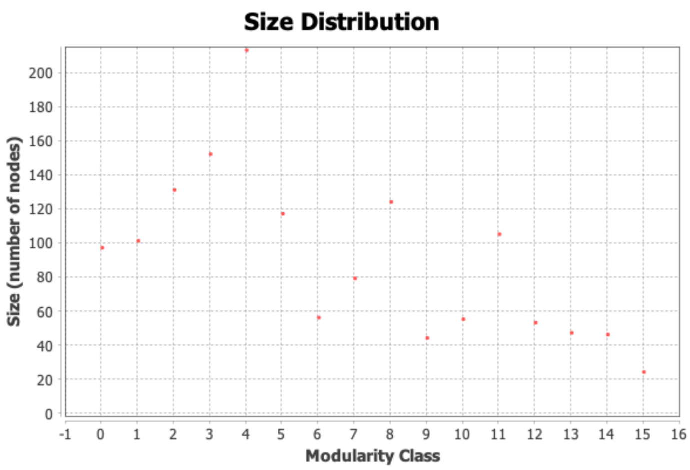


Figure A.40.: Cluster size distribution of direct citation graph of Teubner and Stockhinger [65]

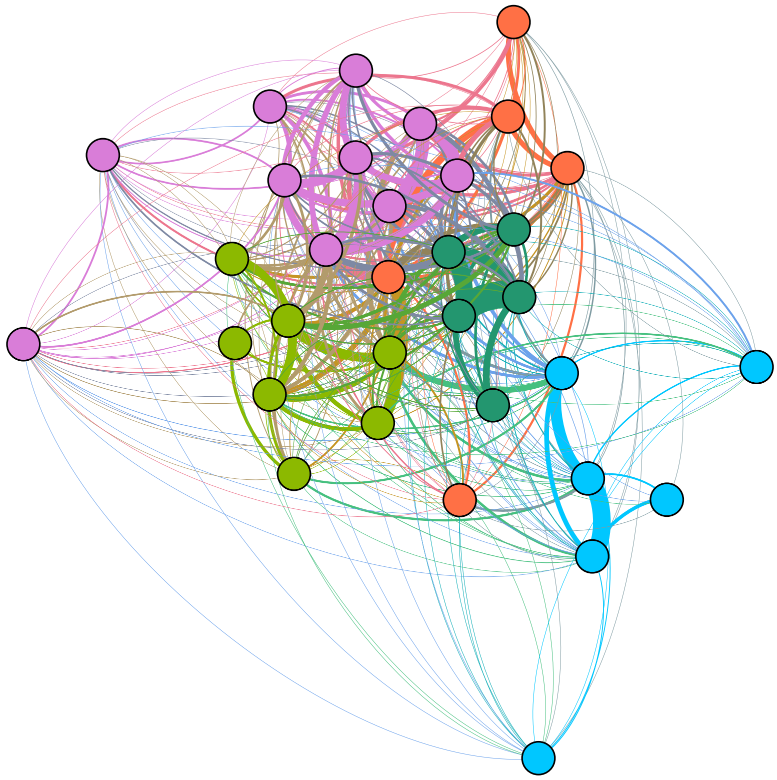


Figure A.41.: Bibliographic coupling graph of Teubner and Stockhinger [65]

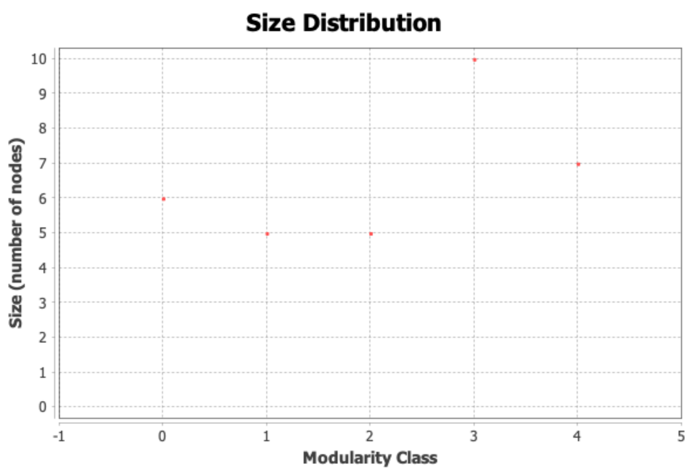


Figure A.42.: Cluster size distribution of bibliographic coupling graph of Teubner and Stockhinger [65]

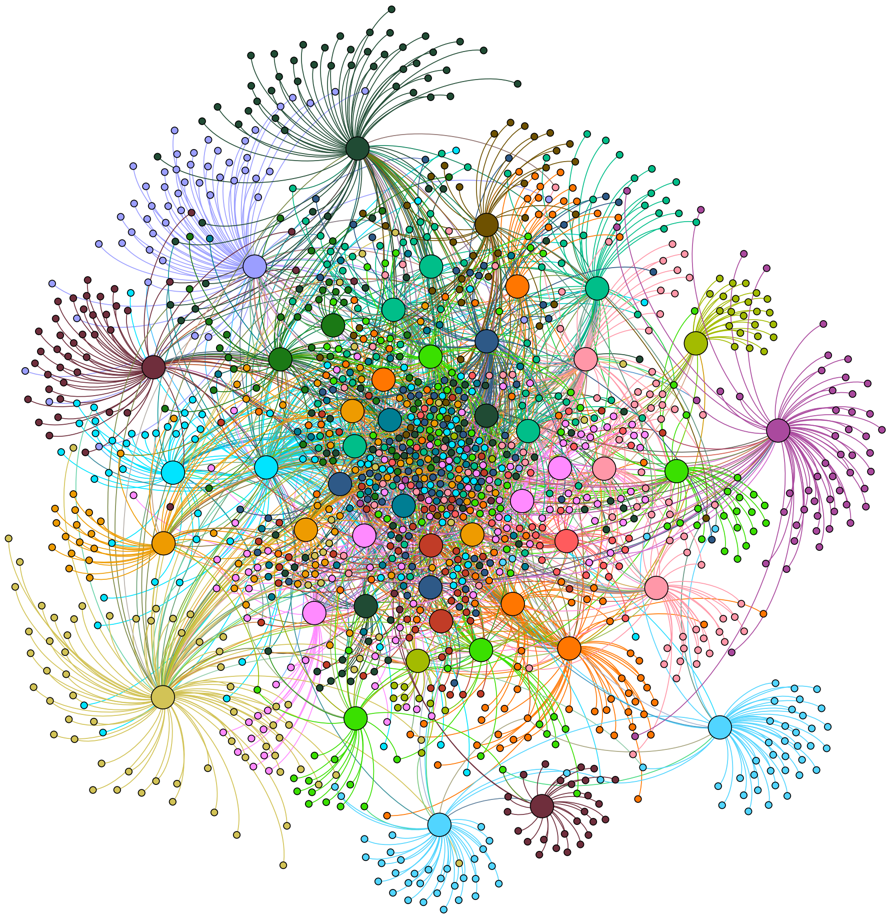


Figure A.43.: Direct citation graph of Tsai et al. [70]

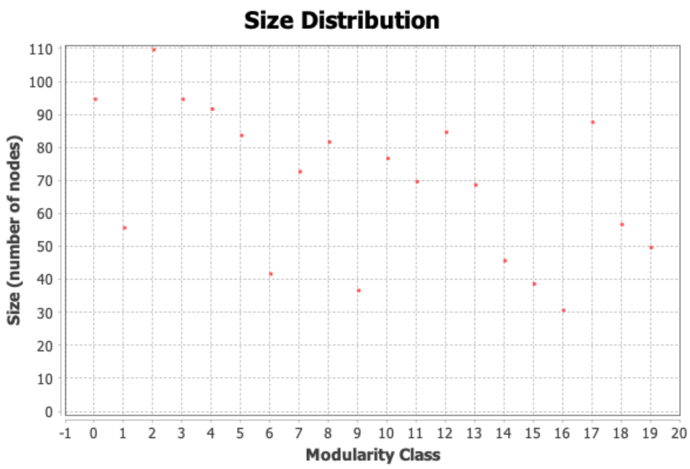


Figure A.44.: Cluster size distribution of direct citation graph of Tsai et al. [70]

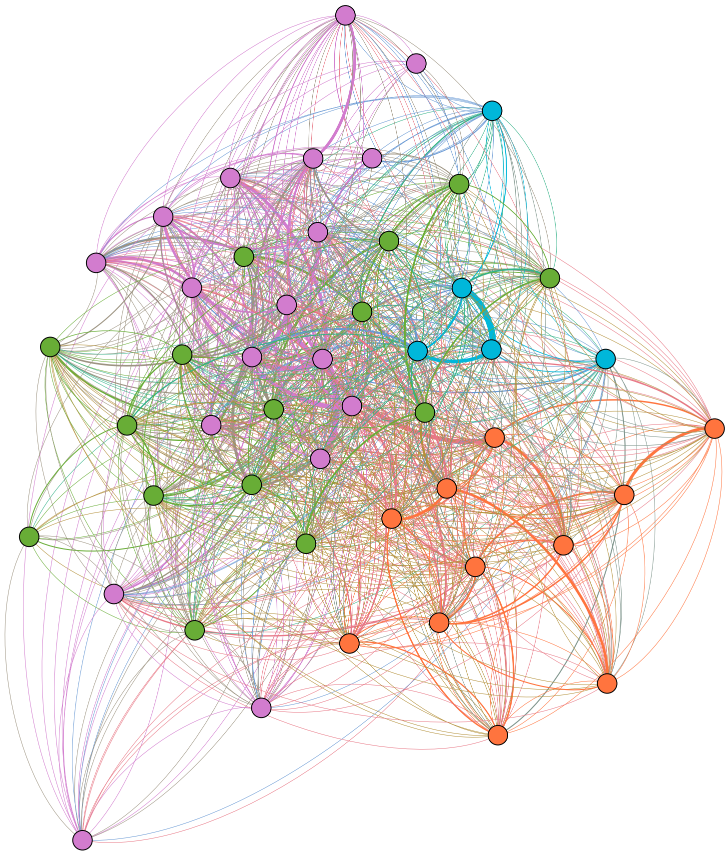


Figure A.45.: Bibliographic coupling graph of Tsai et al. [70]

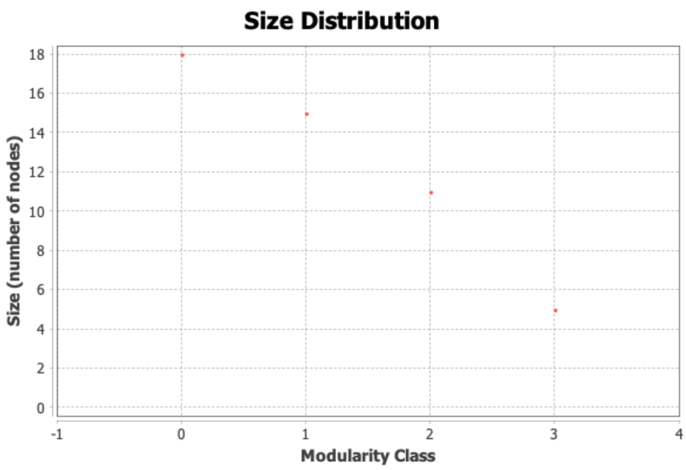


Figure A.46.: Cluster size distribution of bibliographic coupling graph of Tsai et al. [70]

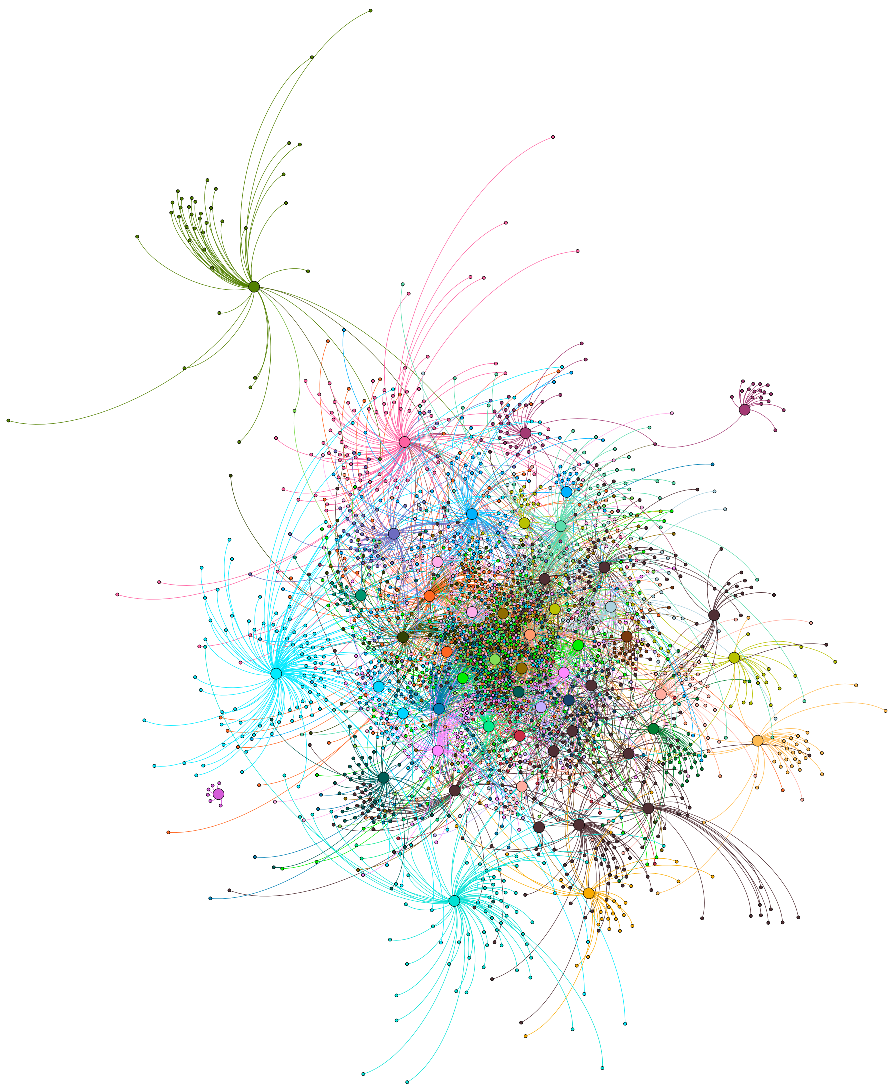


Figure A.47.: Direct citation graph of Wiener et al. [78]

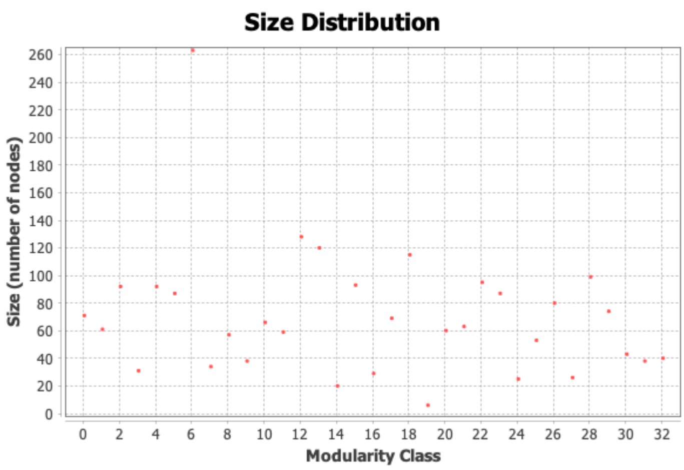


Figure A.48.: Cluster size distribution of direct citation graph of Wiener et al. [78]

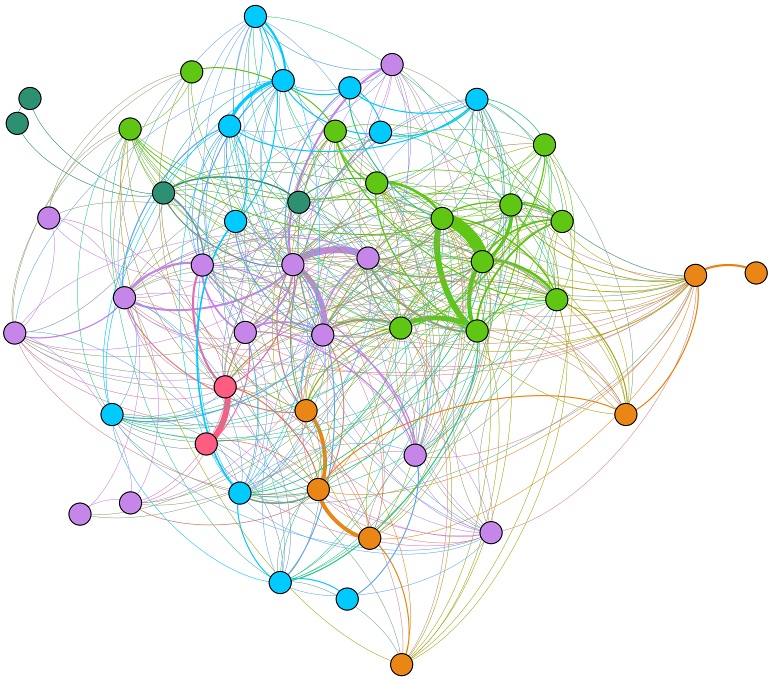


Figure A.49.: Bibliographic coupling graph of Wiener et al. [78]

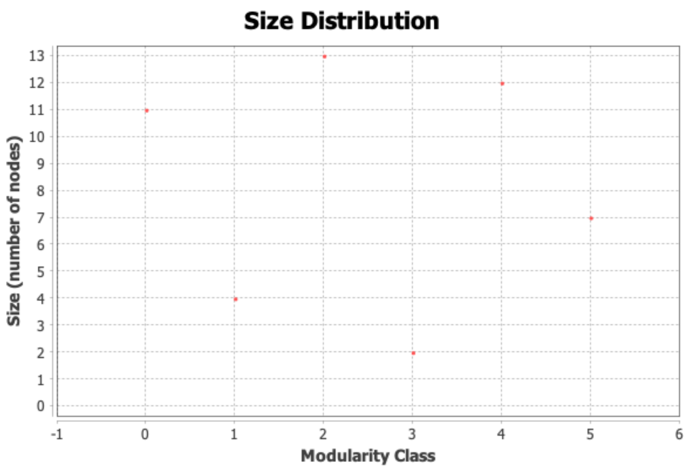


Figure A.50.: Cluster size distribution of bibliographic coupling graph of Wiener et al. [78]

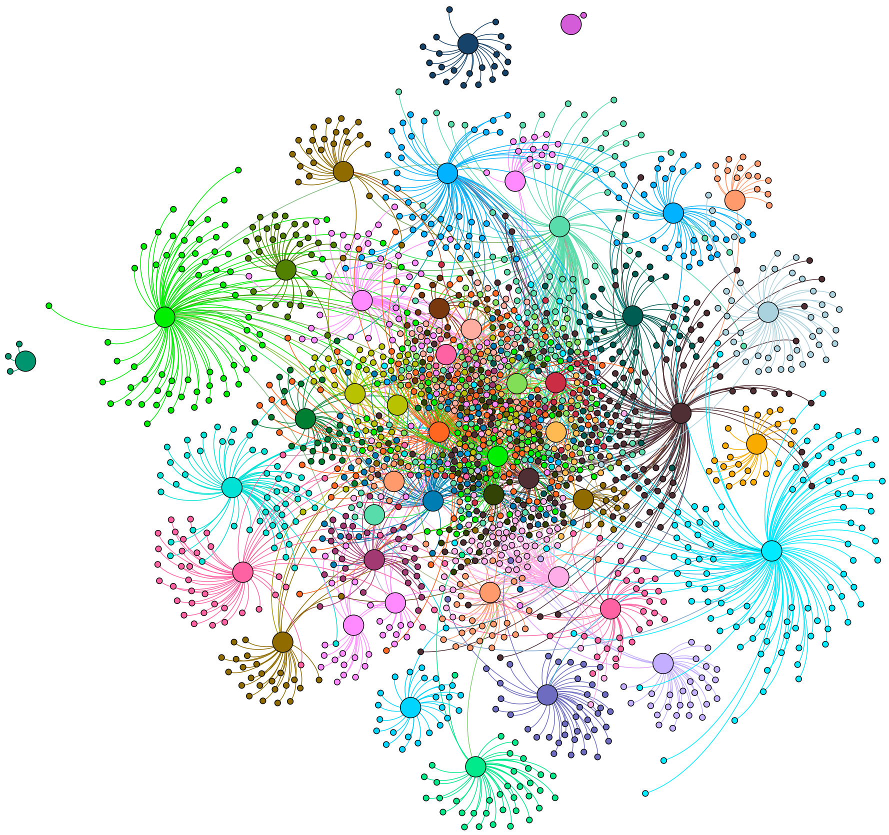


Figure A.51.: Direct citation graph of Xiao et al. [80]

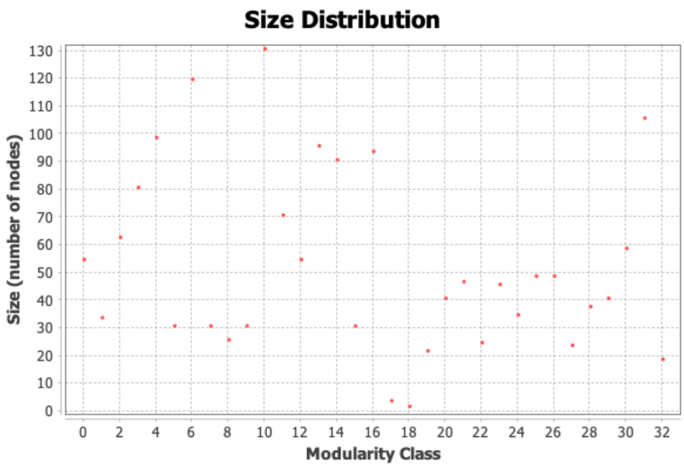


Figure A.52.: Cluster size distribution of direct citation graph of Xiao et al. [80]

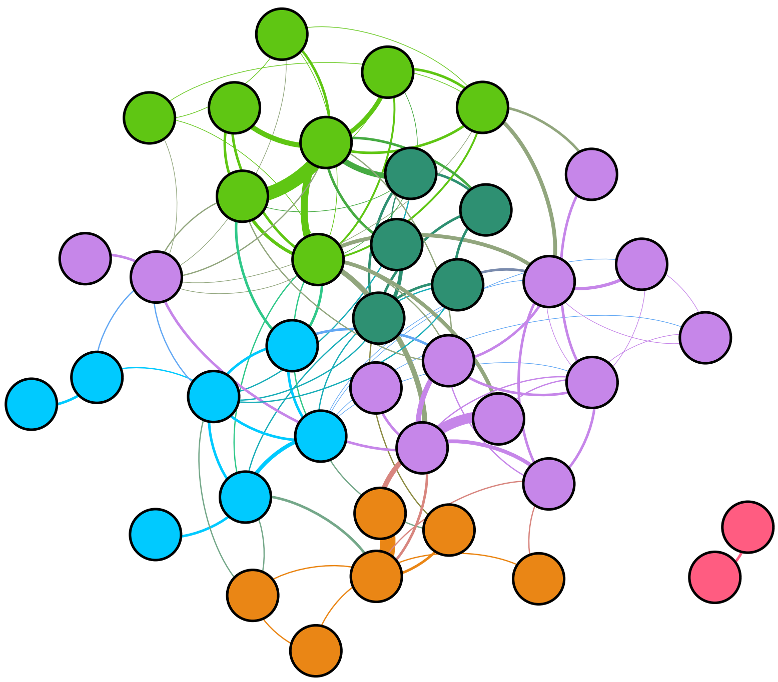


Figure A.53.: Bibliographic coupling graph of Xiao et al. [80]

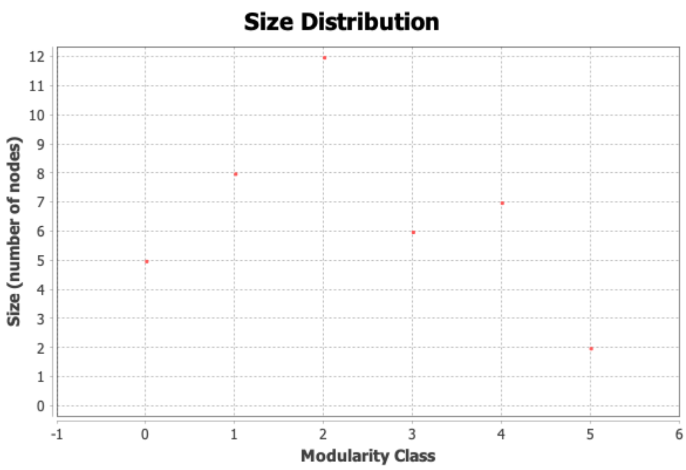


Figure A.54.: Cluster size distribution of bibliographic coupling graph of Xiao et al. [80]

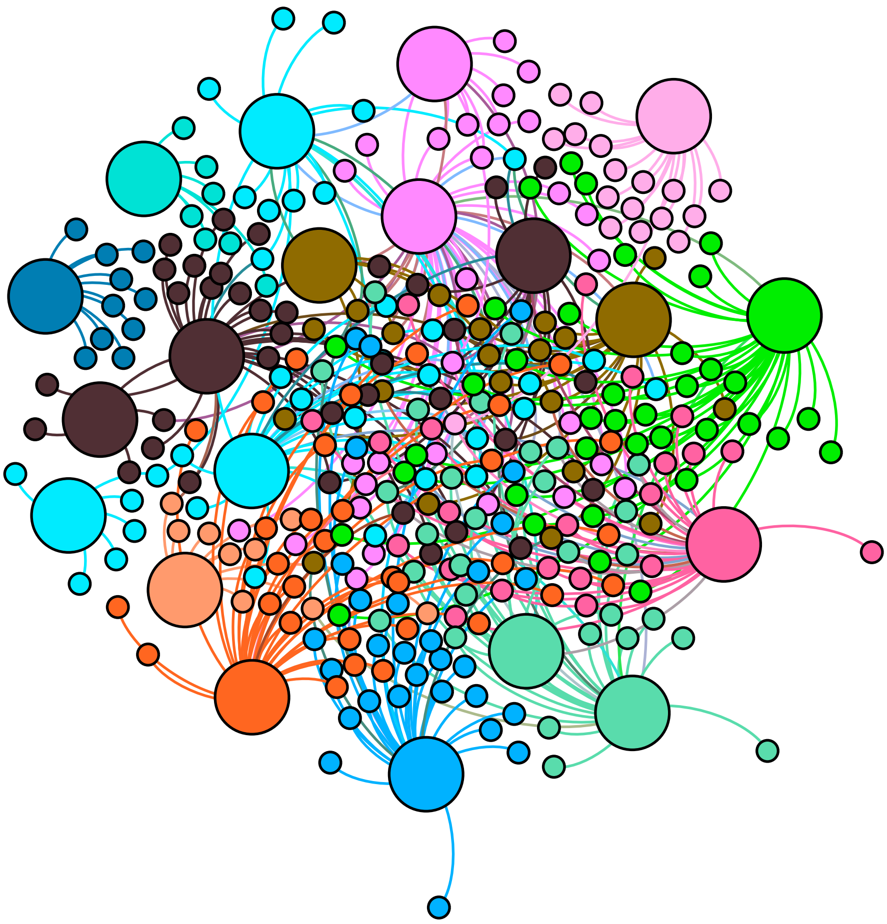


Figure A.55.: Direct citation graph of Siponen and Vartiainen [60]

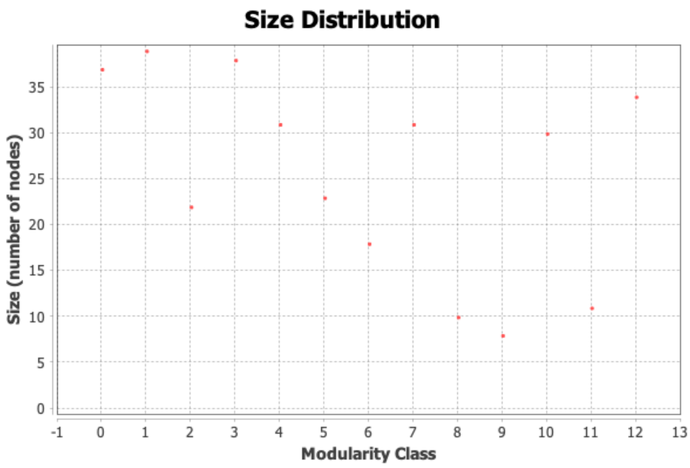


Figure A.56.: Cluster size distribution of direct citation graph of Siponen and Vartiainen [60]

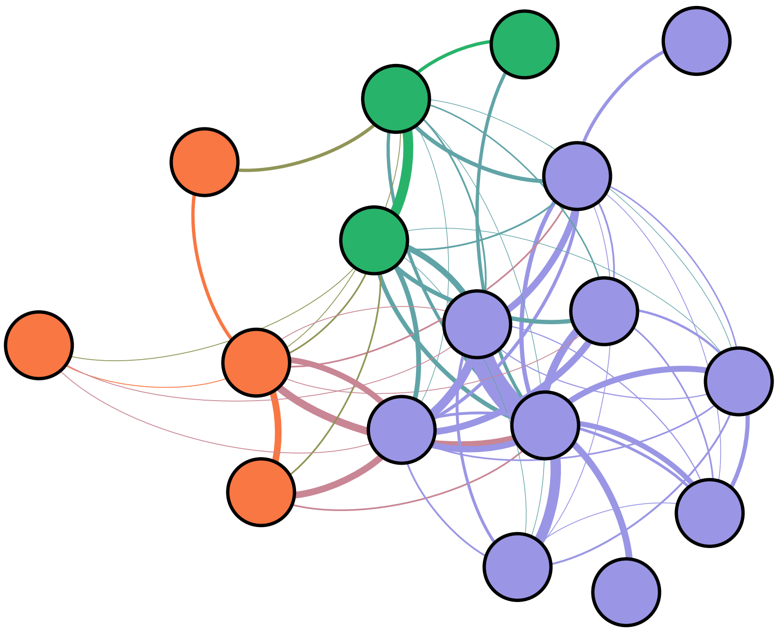


Figure A.57.: Bibliographic coupling graph of Siponen and Vartiainen [60]

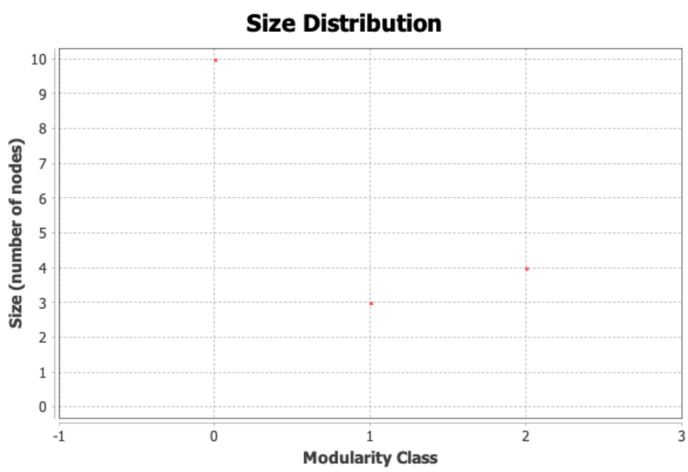


Figure A.58.: Cluster size distribution of bibliographic coupling graph of Siponen and Vartiainen [60]