Unravelling the public procurement networks of architectural services in Finland as pathways of transformative innovations and tacit knowledge

Vincent Kuo
Aalto University, Department of Civil Engineering, Finland
vincent.kuo@aalto.fi

Günther H. Filz
University of Innsbruck, Faculty of Architecture, Austria
guenther.filz@uibk.ac.at

Abstract

Progressive transformation of the built environment is ultimately realized by industry practitioners and designers from companies, coupled with targets set by contracting authorities. Thus, insights on real-world collaborative networks in procurement can inform how progressive targets and innovative designs can be incentivized and proliferated through the ecosystem, as indicators of the responsiveness and transformational potential of the whole sector. However, deep insights on procurement networks of architectural services are currently not available for assessing the pathways of transformative knowledge. The aim of our research is to collect, clean and analyze open data about Finnish public procurements of architectural and related services, for insights on the landscapes companies, as key players for transformation in the building sector. We map the collaborative ecosystem country-wide, across more than 500 projects between 2018 and 2021, using data-mining, computational network modelling and graphing algorithms. In this way, communities and densities of companies can be detected and characterized based on their connectedness and influence based on explicit metrics of degree centrality, total link strengths, and distribution of market share. We create data visualizations as networks and graphs, aiding intuitive understanding of how transformative knowledge and innovations could spread through the network. Our principal findings reveal the gross inequality in social capital, as well as market share distribution across the architectural ecosystem, with various dichotomies of implications for transformation of the whole industry in terms of innovation adoption, creative thinking, and resilience. On this basis, we provide some practical recommends for government, industry practitioners, as well as researchers, and encourage deeper considerations on the utilization of network analytics metrics to characterization of complex problems in the architectural and related markets.

Keywords: Finnish public procurement, architectural services, company networks, data mining, data processing and visualization, clustering, innovation and knowledge transfer, sector transformation.

PEER-REVIEWED ARTICLE • REVALUE & REFORM

Introduction

The term “procurement” can be defined as the purchasing of well-specified items, or the acquisition of project resources to realize a constructed facility (Rowlinson and McDermott, 1999). Topics of procurement networks may seem outlandish among architecture practitioners and researchers, as it may not be self-evident how it relates to the traditional backdrop of architectural research and related topics. Using data mining, processing and network analysis, we investigate role of procurement of architectural services in Finland and how companies collaborate across the industry. This has direct implications on the exchange of knowledge and beliefs through the network of architectural and related companies, as enablers for progressive transformation of the architecture, engineering and construction (AEC) industry at large. To the authors’ knowledge, this research is the first to quantitatively characterize the architectural procurement networks in Finland utilizing open data, with data mining methods, network analysis and visualizations, as insights to promote the transfer of innovative design knowledge and practices. In doing so, we uncover the concentration of collaborative relationships within the Finnish architectural and related services industry, indicating the way knowledge can values are spread, and revealing potential inequalities and fragmentations and their relationships to market-wide creativity and innovation adoption.

Public procurement of innovation

Our research focusses on the quantitative modelling of collaborative networks of the market, and, as such, does not discriminate between various types of procurement. Nevertheless, it is useful for the reader to be aware of the conceptual background and distinction of public procurement, as well as that of innovation.

While “procurement” refers to the function of purchasing goods of services from an outside body (Arrowsmith, 2005), “public procurement” occurs when this function is performed by a public agency, governments or state-owned enterprises, all of which need to comply with public procurement laws and regulations (Rolfstam, 2013; OECD, 2023). Typically, a bidding or tendering process is utilized to facilitate the procurement, starting with the publication of a call for proposals, whereupon interested suppliers or consortia could submit bids, which are assessed and shortlisted, and eventually awarded. Herein lies a host of different variations, contexts and values underpinned by public procurement, which will not be discussed in detail in our paper. Our main concern is to analyse the resulting collaborative networks of suppliers in awarded architectural and related services in Finland.

With regards to the public procurement of innovation, some ambiguity prevails among available definitions. There is also a variation in what terms are used or defined. For instance, “public procurement of innovation”, “innovation procurement”, “public technology procurement”, “innovative procurement” and “pre-commercial procurement” are all some of the terms used in the literature (Rolfstam, 2013) with substantial semantic overlap and often used interchangeably. For the purposes of this paper, it suffices to acknowledge the distinction of “innovation”, as opposed to merely “production”, as referred to by Joseph Schumpeter, one of the most influential economists of the early 20th century. According to him, production concerns the utilization of “materials and forces within our reach”, while innovations are “new combinations of manifested as the introduction of a new good, a new method of production, the opening up of a new market, or the use of a new source of supply of raw materials or new ways of organizing industries” (Schumpeter, 1934). Public procurement of innovation could then be understood as a purchasing activity that delivers any of such aspects. Naturally, the line between production and innovation in the real-world market of architectural and related services could be subjective.

For our research, it is not of particular significance whether any specific awarded project of architectural and related services would be considered as "production" or "innovation". Our focus is rather on the characterization of the existing collaborative networks of the whole market, as an indicator of how potentially new ideas, values, and tacit knowledge leading to innovation and positive societal change, could spread through the network.

Role of procurement in AEC transformation

Over the last decades, the procurement field has evolved beyond mere descriptive, clerical purchasing activities, towards more appreciation of its strategic role and impact as an executive capability (Nissen, 2009), involving other features such as organizational partnering and networking, culture, management, economics, environment and political issues (Rahmani, Maqsood and Khalfan, 2017). These strategic aspects of procurement are closely intertwined with innovation and long-term sustainability of products/services, that require constant change in organizational knowledge and practice. Likewise in the AEC field, the pursuit of transformative targets, such as improving resilience, sustainability, and diversity, require change in both thinking and practice, i.e. both planning and implementation. Beside the importance of developing creative and innovative solution proposals, real societal impact is made only when new solutions are ultimately realized, that is, procured and implemented in building projects in the real world. Therefore, insightful understanding our procurement networks and knowledge ecosystems, based on real-world data of actualized projects and companies, can enable informed strategic planning towards transformation of the AEC industry for public, private and other sectors alike. Reshaping the procurement networks can have a pivotal role in reshaping the values of the AEC industry for the better.

Real societal impact is made only when new solutions are ultimately realized, that is, procured and implemented in building projects in the real world.

Successful procurement of building assets is the first step towards a successful project (Cartlidge, 2003) and naturally has high impact on the lifecycle towards the desired targets of progressive transformation (Figure 1). It is well known that the earlier actions are applied, the more impactful they are in the long run and at lower costs (Kohler and Moffatt, 2003). Procurement interventions can play a particularly impactful role on driving change, as procurement naturally precede all other phases of the project where actual work is done, be it design, construction, operation, decommissioning or beyond. Even in cases where architectural design concepts may have already been formulated before the actual procurement process, it is still the procurement criteria that dictates how the design solutions would need to be revised and finalized, and, in turn, what sort of companies would be awarded the project to implement or construct the solution.

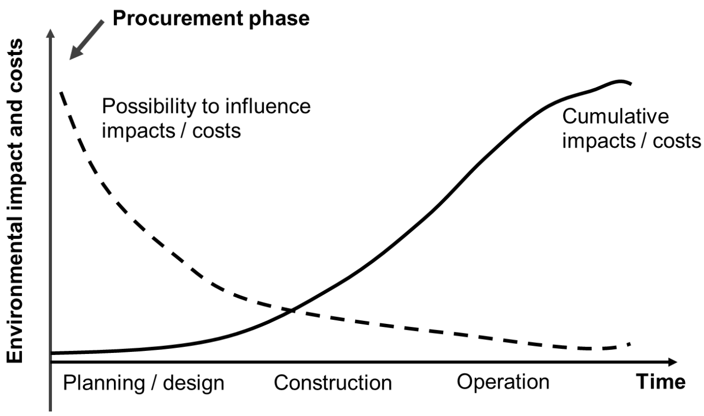

Figure 1. High impact of procurement, adapted from Kohler and Moffatt (2003).

Architectural research lacks procurement topics

Procurement topics seem to have received little attention in Finnish architectural research. As preliminary indication, we searched and compared the number of articles in various architectural journals containing “sustainability” and “procurement” over the last 10 years (Figure 2).

For instance, the journal of Architecture Research in Finland (ARF) contains only 4 articles in which the word “procurement” occurs, from its inception in 2017 until 2022. Similar results are seen in international examples, for instance, considering architecture journals with the highest qualitative rating of 3 accordingly to JUFO, the Publication Forum of the Finnish scientific community. In the Journal of Architecture (J. Archit.), “procurement” appears in only 36 articles in the 10-year period between 2012 and 2022, while in the Journal of the Society of Architectural Historians “procurement” appears in only 25 articles between 2012 and 2022. Furthermore, in all cases, “procurement” does not appear in the title, abstract, nor as a keyword, implying only a peripheral or incidental reference to procurement topics in those articles. For comparison, a common topic like “sustainability” appears in 21 articles of ARF between 2017 and 2022, in 214 articles of the Journal of Architecture between 2012 and 2022; and 96 articles in Journal of the Society of Architectural Historians between 2012 and 2022, often occurring in the title, abstract or as a keyword.

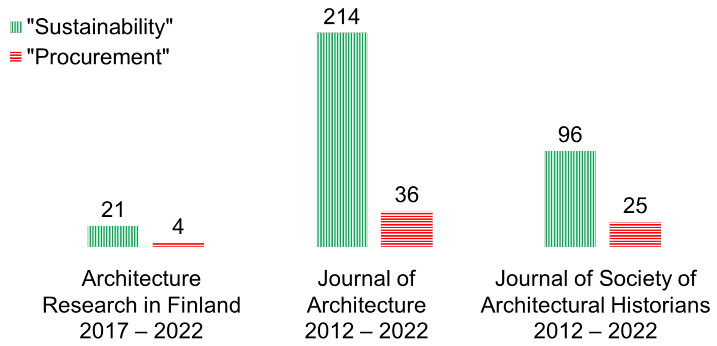

Figure 2. Number of articles in which “sustainability” and “procurement” appear in architectural journals.

The preliminary results give a clear relative indication about the meagre coverage of procurement topics in the context of architectural literature, both in Finland and internationally. Indeed, this may confirm the reader’s own anecdotal impression that procurement topics have been rarely covered in their own experiences in architectural research. For the purposes of our study, it is out of scope to further carry out a similar cross-comparison of journals from procurement and management traditions, that deals with architecture or related topics.

Procurement networks as pathways of knowledge and change

A key driver of innovation in the AEC industry is knowledge sharing through collaboration (Poirier, Forgues and Staub-French, 2016), which are important enablers of impactful transformation. Procurement networks are made up of buyers and suppliers, as well as the collaborative relationships between them, thus offering important indicators about how knowledge and values are transferred or shared across the industry (Senaratne, Jin and Denham, 2021).

Clearly this has implications on the drivers of transformative goals and the way in which new progressive values may be adopted through the knowledge pathways across the network.

Actions of organisations can best be explained in terms of their position within networks of relationships, similar to a social network.

Pryke (2005) and others have demonstrated the importance of adopting a network view of AEC organizational relationships and coalitions. The view is that AEC organisations operate in environments comprising networks of other organisations. Actions of organisations can best be explained in terms of their position within networks of relationships, similar to a social network. The value inherent in social connections is a fundamental concept in sociology, and has been referred to as *social capital* in related literature such as that of management, economics and business. Social capital is the fabric of social relations essential for knowledge sharing that can be mobilized to facilitate action (Nahapiet and Ghoshal, 1998; Portes, 1998; Leenders and Gabbay, 1999). The AEC industry comprises both the network of strong relationships and the knowledge resources that may be mobilised through that network (Senaratne, Jin and Denham, 2021).

Network analysis of people or organizations in AEC are not uncommon. Since the 1990s network analysis has been used frequently to investigate information and knowledge flow in construction, as the foundation for collaborative working and subsequently improving overall performance (Ruan, Ochieng and Price, 2011). Others have identified knowledge sharing processes within social networks in construction organisations using network analysis (Senaratne, Jin and Denham, 2021). However, few have considered characterizing the public procurement networks. Leiva *et al*. (2020) investigated the public procurement market in Chile using network analysis, and claimed to have found no other studies that characterize the public procurement market in terms of its main communities using network analysis.

To study the potential of transformation and dissemination of knowledge and values of the Finnish AEC sector, there is a need to see overall patterns of the complex, overlapping webs of organizational relationships, that are readily captured within open data on public procurements in Finland. Therefore, it is pertinent to analyse AEC organizations in terms of their networks of relationships using appropriate tools and methods. This enables us to investigate how knowledge and action can propagate through the industry and devise ways to promote transformative change, be it through improved governance or partnerships in design and construction.

Finland is an ideal context for studying procurement networks

According to some recent sources, Finland could be considered a posterchild of leadership in progressive procurement. In an elaborate study by the European Commission published in March 2021, more than 30 EU countries have been benchmarked for their progressiveness in national procurement policy and execution, and Finland has been ranked the top country by a huge margin above the average metrics (European Commission, 2021). This highlights the need to address the lack of procurement topics within architectural and related research, given that the built environment is such an important part of the national agenda.

On national level, Finland continues to reform procurement practices through numerous programmes and initiatives. For instance, the national Hankinta-Suomi (Procurement-Finland) programme, set up by the Ministry of Finance, is revolutionizing procurement strategies and approaches country-wide, to improve economic, social and ecological sustainability, as well as supporting the uptake of innovations in procurement practices (Valtiovarainministeriö, 2021). Such efforts are likewise evident in the Finnish construction domain. For instance, Helsingin Asuntotuotanto (residential development unit of the City of Helsinki) has been involved in Hankinta-Suomi activities together with Hansel (central procurement unit of the Finnish government) to improve their procurement practices.

In terms of digitalization and data management in building and infrastructure, bodies like Väylä (Finnish Transport Infrastructure Agency) are actively leading large alliance projects such as Ihku (Ihku Alliance, 2021), to improve industry-wide project cost data aggregation and estimation standards. There are also substantial efforts to improve transparency and publish open data. For instance, the Finnish Ministry of Finance and Hansel (central procurement unit of the Finnish government) provide a service called Tutki Hankintoja, which publishes data on public spending and invoices of state organs and municipalities. Additionally, there is a digital service for notices on public procurement in Finland called Hilma (Hankintailmoitukset.fi), where buyers publish notices on upcoming and ongoing tendering procedures and also on the results of procedures. Hilma also makes the related data available via their open API (application programming interface).

In Finland, the broad appreciation of the role, practices and impacts of procurement are in constant development. The availability of open data on procurement makes comprehensive in-depth quantitative analyses feasible. Given the gap in academic AEC research contributions in these efforts, and the mature context of Finland's activities in advancing procurement, the Finnish context is very suitable for our analysis of the procurement networks of architectural and related services.

In light of the overwhelmingly positive characterization of public procurement in Finland, it is important also to acknowledge that such progressive developments in are only fairly recent, compared to the last twenty odd years of innovative procurement in Finland, for instance, as assessed in a comprehensive European Commission study by Fraunhofer in 2005 (Edler *et al.*, 2005). Furthermore, collaborative public procurement philosophies are novel in Finland and broad decentralization is still known to be a great challenge. Even the meaningful developments with Hilma have taken place only in the last few years, and the end targets of Hankinta-Soumi have yet to come to fruition to be assessed rigorously. The need is clear to examine the current sentiments and ambitious vision of public procurement through rigorous exploration in quantitative terms. In our study, we focus on the scope of architectural and related services as the core context of our analysis.

Purpose and methodology

Research questions

This research characterizes the collaborative networks of the architectural and related companies in Finland, based on open data of actualized projects from public procurement, which are published by Hansel (central procurement unit of the Finnish government) through a service called Hilma (Hankintailmoitukset.fi). This investigation is underpinned by the following research questions:

1. How usable is the open data on procurement for network analysis?
2. How are knowledge networks distributed across the industry?
3. How can architectural network insights be utilized to stimulate AEC industry transformation and innovation adoption?

The subsections below describe the steps of the analysis in more detail addressing the research questions.

Data mining and segmentation

The primary source of data is from Hansel (central procurement unit of the Finnish government) via an open API (https://www.hankintailmoitukset.fi/fi/info/rajapinta-lisaarvopalveluille). We acquire and utilize data specifically on the public procurement award notice (jälki-ilmoitus), between September 2018 and March 2021. An example excerpt of the raw data is shown in Figure 3. Each row of data consists of a procurement project and related metadata:

- *Procurement ID*: Unique numerical identifier of the project.
- *Procurement title*: Name of the project as a text string.
- *CPV-code* and *-name*: Common Procurement Vocabulary (CPV) is the standard hierarchical classification system used in the EU to describe the subjects of public procurement contract i.e. products and services.
- *Date*: The date of the award notice.
- *Organization of buyer*: The public organization procuring the project.
- *Value*: The monetary size of the awarded procurement in Euros
- *Awarded supplier/s*: List of companies that are awarded the project.

Procuren	Procurement title	CPV code	CPV name	Date	Organizati	Value	All contractors
24016	Projektinjohtopalvelu ja valvonta erilaisiin perusk	71530000	Rakentamiseen liittyvät konsulttipalvelu	11/06/2019 20:15	Kiinteistö Oy Nikkarinkruun		['Raksystems Insinööritoimisto Oy', 'Rakennuttajatoimisto Ca
28417	Rakentamislogistiikka 2019	71530000	Rakentamiseen liittyvät konsulttipalvelu	11/10/2019 19:30	Kaupunkiympä	2380000	['Sitowise Oy']
39284	Rakennuttamispalvelut 2020-26, Osa-alue: RAP 1	71530000	Rakentamiseen liittyvät konsulttipalvelu	19/02/2020 09:00	Senaatti-kiinte	10000000	['Granlund Oy', 'Ramboll CM Oy', 'DUCO Rakennuttaja Oy', 'In
39306	Rakennuttamispalvelut 2020-26, Osa-alue: RAP 2	71530000	Rakentamiseen liittyvät konsulttipalvelu	19/02/2020 09:00	Senaatti-kiinte	65000000	['Genpro Solutions Oy', 'Ramboll CM Oy', 'DUCO Rakennuttaj
39513	Rakennuttamispalvelut, RAP 3 - Kunnossapito- ja	71530000	Rakentamiseen liittyvät konsulttipalvelu	21/02/2020 09:00	Senaatti-kiinte	20000000	['Granlund Oy', 'Ahma Insinöörit Oy', 'Sweco PM Oy', 'Rakenn
23855	Katujen ja yleisten alueiden rakennuttamisen puit	71530000	Rakentamiseen liittyvät konsulttipalvelu	31/05/2019 18:30	Oulun Kaupunk	500000	['CC Control Oy', 'Carement Oy', 'Plaana Oy', 'Welado Oy']
11434	Luksia, rakennuttamisen ja valvonnan puitejärjest	71530000	Rakentamiseen liittyvät konsulttipalvelu	08/10/2018 19:45	Länsi-Uudenm	800000	['Ramboll CM Oy', 'Pöyry Finland Oy', 'Indepro Oy', 'Rakennut
39120	Havainnollistamispalvelut 1.10.2018-30.9.2020	71530000	Rakentamiseen liittyvät konsulttipalvelu	16/02/2020 15:45	Senaatti-kiinte	1000000	['3D Talo Finland Oy', 'WSP Finland Oy', 'FCG Suunnittelu ja te
39284	Rakennuttamispalvelut 2020-26, Osa-alue: RAP 1	71530000	Rakentamiseen liittyvät konsulttipalvelu	19/02/2020 09:00	Senaatti-kiinte	10000000	['Granlund Oy', 'Ramboll CM Oy', 'DUCO Rakennuttaja Oy', 'In
39306	Rakennuttamispalvelut 2020-26, Osa-alue: RAP 2	71530000	Rakentamiseen liittyvät konsulttipalvelu	19/02/2020 09:00	Senaatti-kiinte	65000000	['Genpro Solutions Oy', 'Ramboll CM Oy', 'DUCO Rakennuttaj
39513	Rakennuttamispalvelut, RAP 3 - Kunnossapito- ja	71530000	Rakentamiseen liittyvät konsulttipalvelu	21/02/2020 09:00	Senaatti-kiinte	20000000	['Granlund Oy', 'Ahma Insinöörit Oy', 'Sweco PM Oy', 'Rakenn
56575	Kosteus- ja ja sisäilmatekniset kuntotutkimuspalv	71530000	Rakentamiseen liittyvät konsulttipalvelu	23/10/2020 14:40	Hämeenlinnan	800000	['Ramboll Finland Oy', 'Ideastructura Oy', 'WSP Finland Oy', 'S
17465	Tulevaisuuden maankäyttöpäätökset hankinta	73000000	Tutkimus- ja kehityspalvelut ja niihin liitt	18/02/2019 18:30	Ympäristömini	368900	['Ramboll Oy']
17787	Suunnittelupalveluiden puitesopimuksen hankinna	71240000	Arkkitehti-, insinööri- ja suunnittelupalve	27/02/2019 18:45	Väylävirasto	30000000	['Destia Oy', 'NRC Group Finland Oy', 'Ramboll Finland Oy', 'Pö
17859	Taitorakenteiden suunnitteluohjeiden laatiminen,	71000000	Arkkitehti-, rakennus-, insinööri- ja tarka	04/03/2019 18:45	Väylävirasto	1800000	['Sweco Rakennetekniikka Oy', 'Ramboll Finland Oy', 'Inststo I
18833	Valtatien 3 parantaminen välillä Lempäälä-Pirkka	71000000	Arkkitehti-, rakennus-, insinööri- ja tarka	15/03/2019 18:45	Pirkanmaan eli	521390	['Sitowise', 'Ramboll Finland Oy']
19285	YVA ja yleissuunnitelma vt 3 välillä Helsingby – Lai	71000000	Arkkitehti-, rakennus-, insinööri- ja tarka	22/03/2019 18:45	Etelä-Pohjanm	360000	['Ramboll Finland Oy']

ID	Procurement title	CPV code	CPV name	Date	Buyer	Value	Awarded company / supplier
17787	Suunnittelupalveluiden puitesopimuksen hankinnat. Osatehtävä 1) Radanpidon ja rataverkon käytön suunnittelu- ja asiantuntijatehtävät 28.9.2018	71240000	Arkkitehti-, insinööri- ja suunnittelupalvelut	27/02/2019 18:45	Väylävirasto	30000000	['Destia Oy', 'NRC Group Finland Oy', 'Ramboll Finland Oy', 'Pöyry Finland Oy', 'WSP Finland Oy', 'Sitowise Oy', 'Proxion Plan Oy', 'Finnmap Infra Oy']

Figure 3. Example of raw data of procurement award notices.

Using the CPV classification, we segment the data and filter it down to the most relevant subjects for our analysis, by selecting which CPVs are related to architectural services. The data of projects used in our analysis has the following CPV codes and names:

71000000	Architectural, construction, engineering and inspection services
71200000	Architectural and related services
71220000	Architectural design services
71240000	Architectural, engineering and planning services
71250000	Architectural, engineering and surveying services
71400000	Urban planning and landscape architectural services

Subjects that are related to architecture, but that are focussed on engineering or construction related services are not included. We utilize mainly the list of awarded suppliers for each project for network analysis, described in the next section, to investigate the links between companies as knowledge pathways distributed across the industry. The CPV-filtered data of suppliers then undergo parsing and cleaning, to remove character inconsistencies such as spaces or tabs within the strings. This separates the text string into discrete company entities that can be appropriately indexed and modelled as network nodes.

Network analysis

Network analysis is based upon Graph Theory (Scott, 1992) and entails the modelling of relationships (links or edges) between discrete entities (nodes or vertices). Network theory and analysis methods are used across many disciplines e.g. mathematics, engineering, biology, sociology and economics. A network represents groupings as systems of nodes linked to each other by a specified relationship type (Pryke, 2005). Therefore, a network of companies is an indicator of their collaborative knowledge pathways, and how ideas, beliefs or values can spread through the system. For our analysis, we characterize each company as a node with links to other companies, based on how they collaborate in procurement projects together.

The outcome of the network analysis of companies enables us to:

- Reveal centralities of companies and clusters, indicating the distribution of influence.
- Aid informed decisions of strategic partnerships for impact or maximizing interventions.
- Identify the nature of potential knowledge fragmentation vs cohesion and inequalities.
- Explicitly represent the knowledge pathways between companies, and how ideas beliefs or values can spread through the ecosystem.

These insights can help architects, contractors, researchers and policy makers to gain better understanding of the real-world market environment, in which solutions are implemented. Such knowledge not only helps companies make strategic decisions about their most impactful partnerships, but also to become more aware and active in advising policy makers towards more desirable collaborative industry networks that promote innovation through inclusive and responsive knowledge sharing.

We primarily use the software VOSviewer (Leiden University, 2022) to facilitate the network analysis and visualizations, explained in the sub-headings below. A later section is an illustrative example utilizing these steps on a simplified hypothetical "mini network" consisting of just 8 companies for clarity.

Convert bipartite to monopartite data

The raw data on awarded tenders/projects can be described as bipartite in nature, each tender consisting of 1 buyer (public organization) and 1 or more companies. The network modelling in our investigation is enabled by converting the bipartite data into monopartite data, that is, consisting only of links between companies. This is done simply by assigning a link between companies that occur together within the same project. Monopartite network analysis enables straightforward representation of the relationships between companies, by virtue of how they have worked together in the same projects. The conversion of bipartite to monopartite data for procurement network analysis is extensively discussed and motivated in Leiva *et al.* (2020) and is a common step in preparing data for network analysis. Each link between companies thus indicates the extent to which those 2 companies have collaborated in projects.

Calculate degree centrality

The concept of *centrality* is fundamental in graph theory and consists of various methods to characterize the importance of nodes within a network. We determine the basic *degree centrality* for each company, which is simply defined as a count of how many unique direct links a company has to other companies, as an indication of its connectedness within the network. (Golbeck, 2015)

Counting of total link strengths

There are 2 methods to count the total link strength of a particular link, namely, *full* and *fractional* counting.

Full counting of the link strength of a specific link between any 2 nodes is self-explanatory and entails merely adding up the number of times that 2 nodes co-occurred within the same project. On the other hand, *fractional* counting of links means that when there are many companies involved in the same project, then the weights of those links should be discounted proportionally. Depending on the assumptions about the network, and the purpose of the interpretations, fractional counting can be argued as a more accurate representation of link strengths across the network. Intuitively, for instance, where only 2 companies are involved in a project, it is reasonable to assume that they would generally have more or stronger collaboration, than 2 companies within a consortium of numerous companies. Of course, this is only a general assumption, given that the data does not indicate the exact nature or distribution of interactions between companies within a project consortium. Perianes-Rodriguez, Waltman, and Van Eck (2016) offers an in-depth discussion of the differences between full and fractional counting.

In our investigation we predominantly use the fractional counting method for the total link strength of each company for further visualizations. However, we also show some results of the full counting for comparison and analysis.

Clustering, layout and visualizations

In general, community detection and clustering form a broad field in computer and data science, and there are countless methods to do so. VOSviewer has default unified clustering and layout functions to assign each company to a discrete cluster, and calculate appropriate X and Y coordinates to plot the nodes of the network for visualization. In this investigation, we use the VOSviewer default functions without further detailed discussions and justifications on the intricacies of the clustering and layout algorithms, which can be found in detail in related literature (Van Eck *et al.*, 2010; Waltman, Van Eck and Noyons, 2010; Waltman and Van Eck, 2013; Van Eck and Waltman, 2014).
We utilize the clustering outputs for inspecting the distribution of the number of clusters, cluster sizes, the topography/structures of the visualized landscape of companies, and distributions of degree centralities within clusters. The layout visualization enables us to see the relationship between the network structures the quantitative calculations for each company and their connections, such as degree centrality and total link strengths.

Plot distribution of companies' share of projects
We plot the distribution of all the companies ranked by their involvement in projects (i.e. *rank–size plot*). See the illustrative example in the following sections for the explicit step-by-step processes.

We first compile a *i* x *j* company-to-project matrix, with company along the vertical *i* axis and projects along the horizontal *j* axis. The entries in the matrix are the number of times company *i* occurs in project *j*. Naturally, this yields a Boolean matrix of 1s and 0s, because companies do not occur more than once in each project. This matrix can also be regarded as a relational matrix between all the companies and projects.

Since many companies can occur within one project, we normalize each column vector *j* to weigh each project proportionally by the number of companies involved. For example, when there are 2 companies involved in a project, each company *i* will have an entry of 0.5 in the project *j*. Likewise, when there are 10 companies involved in a project, each company will have an entry of 0.1. In principle, this ensures a more accurate representation of the total involvement or share of the project per company, similar to the reasoning for using fractional counting of network link strengths. For instance, companies that often participate in projects with large consortia would be considered to have less share of the project and thus the involvement would be discounted proportionally by virtue of the normalization.

For each company, all the entries of the respective *i*-th row in the matrix are summed to yield a number that indicates the company's level of involvement with respect to all projects. Then, we sort (rank) the companies in decreasing order of the level of involvement (size), enabling a *rank–size plot* that explicitly describes the distribution by which companies are involved in projects in descending order. With the same data, we further plot the *cumulative percentages* of total projects represented by companies ranked in descending order. The steepness and curvature of these plots are indicators of how the level of involvement per company is distributed across the market. For instance, this enable clear insights on what percentage of all projects are represented by what percentage of top companies, showing the level of equality or disparity of companies in the market with respect to the distribution of their involvement in projects and the implied knowledge pathways between companies.

We also use the same rank–size plot to graph distributions of other network parameters mentioned earlier, such as degree centrality and total link strengths. For those parameters, no pre-processing is required, and the data can be plotted directly after sorting in descending order i.e. ranking.

Example of a hypothetical small dataset

For transparency and replicability, we provide a step-by-step example of the methods and outputs with a tiny dataset consisting of 4 projects and 8 companies. The parsing and cleaning of the raw open dataset yields Table 1, showing each project along with a comma separated text string of companies involved in each project.

Table 1. Example of small dataset

Project	Companies
P1	['Arcadia Oy arkkitehtitoimisto']
P2	['Ramboll Finland Oy', 'Gravicon Oy', 'FCG Suunnittelu ja tekniikka Oy', 'A-Insinöörit Rakennuttaminen Oy', 'Tietoa Finland Oy', 'Arcadia Oy Arkkitehtitoimisto']
P3	['Arcadia Oy Arkkitehtitoimisto']
P4	['Arcadia Oy Arkkitehtitoimisto', 'Architecture Workshop Finland Oy', 'Raami Arkkitehdit Oy']

By inspection, Table 1 shows that there are 8 unique companies, of which "Arcadia Oy Arkkitehtitoimisto" occurs in all 4 projects, co-occurring with other companies in 2 projects, i.e. projects P2, and P4. The project-to-company bipartite data is converted to company-to-company monopartite data, so that a link is formed when companies co-occur within the same project. Applying network analysis operations by calculating the degree centralities (number of unique links per node), fractional counting of total link strengths, and clusters in VOSviewer, we get the results shown in Table 2, in descending order of the total link strength. In this case, the link strengths from full counting (not shown) would be identical to the degree centrality, because all companies that have links with another company co-occur no more than once.

Table 2. Network analysis of hypothetical mini data.

Company	Total link strength	Degree centrality	Cluster
arcadia oy arkkitehtitoimisto	2	7	2 (green)
a-insinöörit rakennuttaminen oy	1	5	1 (red)
fcg suunnittelu ja tekniikka oy	1	5	1 (red)
gravicon oy	1	5	1 (red)
ramboll finland oy	1	5	1 (red)
tietoa finland oy	1	5	1 (red)
architecture workshop finland oy	1	2	2 (green)
raami arkkitehdit oy	1	2	2 (green)

It is worthy to note that the difference between fractional and full counting is only in the strength of the link, with no effect on whether the link exists or not. Using the same data, we use VOSviewer to visualize the corresponding network (Figure 4).

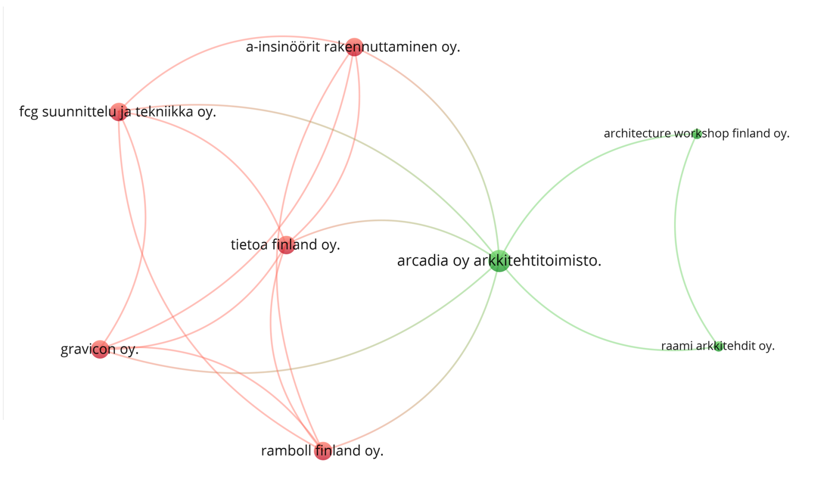

Figure 4. Example network visualization of mini data.

The 2 clusters are represented by colours red (Cluster 1) and Green (cluster 2), making the communities easy to discern visually. It is easily evident that "Arcadia Oy Arkkitehtitoimisto" is the most central, and acts as a bridge between companies in the 2 clusters. Companies of the red Cluster 1 has higher degree centralities than those of the green Cluster 2, simply because there are more interconnections, given that each company has more unique links to other companies. However, note that fractional counting has discounted those link strengths of the companies in the red cluster (Table 2), given that more companies are involved in the same project (P2) making up the red cluster, compared to the companies in project P4 (green cluster), where only 3 companies co-occur.

With the same raw data from Table 1, we can compile the *i* x *j* company-to-project matrix, showing where company *i* occurs in project *j* (Table 3). Then the columns are normalized, and the rows summed and sorted in descending order (Table 4).

Table 3. Example company-to-project matrix of 8 companies and 4 projects.

Example companies	P1	P2	P3	P4
Arcadia Oy arkkitehtitoimisto	1	1	1	1
Ramboll Finland Oy	1	0	0	0
Gravicon Oy	1	0	0	0
FCG Suunnittelu ja tekniikka Oy	1	0	0	0
A-Insinöörit Rakennuttaminen Oy	1	0	0	0
Tietoa Finland Oy	1	0	0	0
Architecture Workshop Finland Oy	0	1	0	0
Raami Arkkitehdit Oy	0	1	0	0

Table 4. Example company-to-project matrix with columns normalized and ranked by summed rows.

Example companies	P1	P2	P3	P4	Sum
Arcadia Oy arkkitehtitoimisto	0.17	0.33	1.00	1.00	**2.50**
Architecture Workshop Finland Oy	0.00	0.33	0.00	0.00	**0.33**
Raami Arkkitehdit Oy	0.00	0.33	0.00	0.00	**0.33**
Ramboll Finland Oy	0.17	0.00	0.00	0.00	**0.17**
Gravicon Oy	0.17	0.00	0.00	0.00	**0.17**
FCG Suunnittelu ja tekniikka Oy	0.17	0.00	0.00	0.00	**0.17**
A-Insinöörit Rakennuttaminen Oy	0.17	0.00	0.00	0.00	**0.17**
Tietoa Finland Oy	0.17	0.00	0.00	0.00	**0.17**

The values in the sum column in Table 4 are used to plot the *rank–size distribution* (Figure 5) and cumulative percentage distribution (Figure 6). To interpret Figure 5 as an example, it is explicitly clear that the company with the highest involvement in projects is much more prominent than the others, even compared to just the next ranked in the "mini market", as the line of the plot drops steeply.

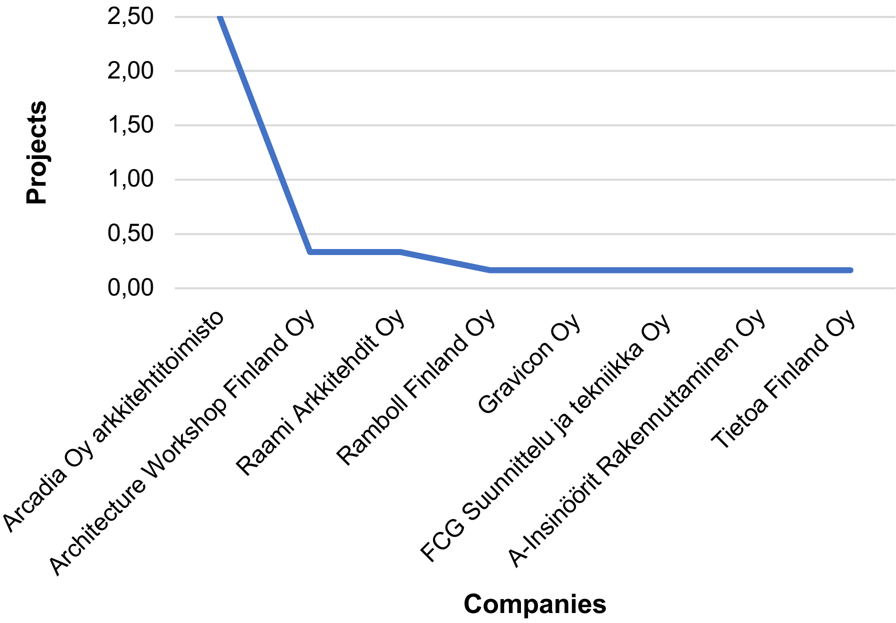

Figure 5. Number of projects (normalized) per company

Figure 6 is a more explicit representation of the percentage of projects represented by companies in descending order. For instance, we can see the first data point from the left that 63% of all projects in the mini market belong to the top 13% of the companies, and that 79% of the mini market belongs to the top 38% of companies (3rd data point from the left). Figures 5 and 6 generally show a rather skewed playing field dominated by 1 very prominent player, with a long tail of companies with little influence on the total market.

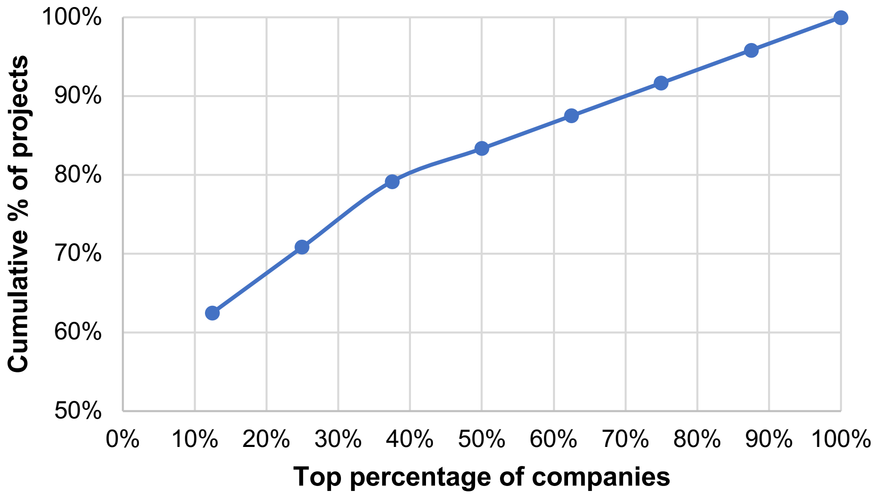

Figure 6. Percentage of projects per top percentage of companies

The network analysis, rank–size, and cumulative plots give additional dimensions of the public procurement networks. It is interesting to note, for instance, that companies "Architecture Workshop Finland Oy" and "Raami Arkkitehdit Oy" are equality ranked 2nd and 3rd in terms of their share in the market (Figure 5), however, they are ranked last in terms of their degree centrality (Table 2), that is, their connectedness in the network based on the number of unique links they have with other companies. Such representations of information and data of the public procurement provides insights and enable further questions to analyse how the nature of the network may or may not promote innovation within the AEC industry or encourage specific progressive transformation goals of the industry.

Analysis results and discussions of full-scale data

Dataset at a glance:

2.5 years (2018-2021)
509 projects
482 companies
4381 links

Overview of the sizes of project consortia

After filtering dataset of all the projects by CPV code, we arrive at a total 509 projects between September 2018 and March 2021 (2.5-year timeframe). The size of project consortia, that is, the number of companies involved in each project, ranges between 38 and 1. To give a more complete idea, Figure 7 shows the number of projects with certain sizes of consortia. For instance, out of the 509 projects, 314 projects (62% of all projects) have just 1 company participant, 26 projects have 2 participants, while 46 projects have 3 participants. One can see that the relationship between the number of projects and the size of project consortia are generally inversely proportional. Therefore, very few projects have very large consortia, and most projects have only 1 company.

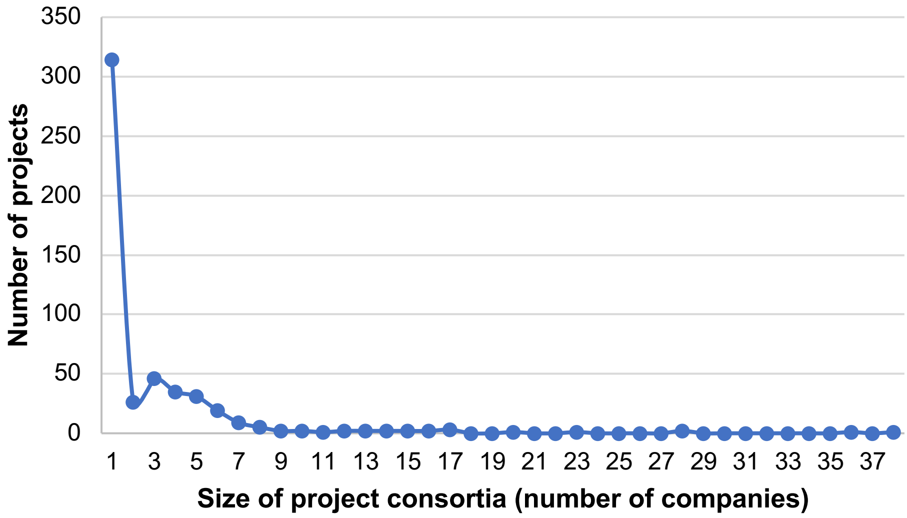

Figure 7. Distribution of consortia sizes (number of companies) of all 509 projects.

Text parsing steps, challenges, and data usability

The parsing of the company names from the raw data held some challenges. The company list of the raw data has the following format: ['Company_1', 'Company_2',...]. Therefore, the first step is to separate the strings by the comma, and then to strip the apostrophes and square brackets, leaving the unique company names. Applying this to all the company lists and removing duplicates yielded 529 unique companies in total.

However, by inspection of the 529 companies, there were still many apparent duplicates. It turned out that there were many double spaces between multi-word company names, as well as spaces in the beginning and end of company name strings, which caused these to be recognized as different companies than their correct counterparts without erroneously additional spaces. Furthermore, there were also invisible non-printable characters found, such as tabs and line breaks, within the strings. To address these, we used the TRIM function in Microsoft Excel to remove the extra spaces, resulting in 512 unique companies (17 duplicates removed). Then, we use the CLEAN function in Microsoft Excel to remove the non-printable characters (22 duplicates removed), leaving us with 490 unique companies.

Additionally, there were a few other issues found where the same company name would have had explicitly different strings, that cause duplicates:

- A few companies had name variants where suffixes of “Oy” (osakeyhtiö i.e. limited liability company) may or may not be included.
- There were spelling variations such as “&” vs “and”, or inconsistencies in capitalizations.

Such inconsistencies are not easy to identify in an automated way, because it would not have been apparent without manual inspection. Furthermore, it is uncertain what other types of problems may occur with another dataset. Nevertheless, in terms of the dataset for our investigation, these cases are very few and thus insignificant to the overall goal of the project. For example, through manual inspections, we identified only 8 companies out of the 490 (1.6%) that were duplicated due to company name spelling variants, and thus do not impact the aggregate results of the analysis in any meaningful way. Nevertheless, for the purposes of this investigation, we will regard 482 companies as the ground truth for the number of companies for further investigation.

Beyond the parsing challenges, there may also be logical inconsistencies in the data to keep in mind. For instance, some larger companies that operate in many regions in the country have been registered as separate entities, such as Granlund Oy, and its many regional entities: Granlund Oulu Oy, Granlund Pohjanmaa Oy, Granlund Jyväskylä Oy, Granlund Joensuu Oy etc. It is worth mentioning that such inconsistencies identified in the dataset are a minority of cases in the context of our analysis. Furthermore, the only implication of these logical inconsistencies is only that companies, that could have been considered as a single entity, would be separated. Therefore, the only potential impact it may have on the quantitative analysis is that there will appear to be a few more companies, of which the degree centralities (number of links) and total link strengths (either by fractional or full counting) would be smaller, compared to the situation when all the “sub-companies” would be combined. The potential difference is not observable on the market level aggregate, and even if it would theoretically be observable, it would only make the market appear slightly less unequal. That is, the total links in the market would appear slightly better distributed or slightly less concentrated. Therefore, it is not an essential consideration how to specifically handle these small inconsistencies caused by parsing, because combining the “sub-companies” would only further confirm the problem of inequality that we would observe in the data.

In summary, the following steps of actions were taken, along with the resulting number of companies remaining from each step after removing duplicates:

1. Separate by commas, removal of apostrophes and square brackets: 529
2. VOSviewer default parsing or TRIM function, removing extra spaces: 512
3. CLEAN function, removing non-printable characters: 490
4. Manual inspection and duplicate removal: 482

Alongside the manual parsing, we test the default functions of VOSviewer in parsing the company lists into individual discrete entities representing a node in the network, to evaluate its feasibility as a plug-and-play research tool for other researchers. VOSviewer’s output is 512 companies, which means that it considers the issue of additional spaces before, in between, and after the company names, but not the non-printable characters such as tabs and line breaks. Nevertheless, since the difference between 512 and the ground truth 482 is only 30 (6.2%), we consider the difference negligible for proceeding with the network analysis in VOSviewer. Therefore, we regard VOSviewer’s default functions as an effective tool for researchers without programming or data science backgrounds.

Researchers or analysts, who wish to utilize public procurement open data, should assume that no “structured datasets” are completely consistent.

Our findings show that the raw open data of project award notices are of reasonable usability, but has room for improvement for making them public, particularly regarding the removal of erroneous extra spaces, non-printable characters, and spelling variants. This has very important practical implications on promoting data-driven knowledge and insights, because poor quality of data may require so much pre-processing and cleaning that it discourages the analysis and utilization of open data, which holds a wealth of potential knowledge that can improve the society. It is even more important that public open data maintains a high standard of consistency to enable direct ingestion into third party software or applications that elevates the potential of the data, such as VOSviewer. On the other hand, we recommend that researchers or analysts, who wish to utilize public procurement open data, should assume that no “structured datasets” are completely consistent. Furthermore, an appropriate methodology and reasoning for interpreting the results should be devised to be able to consider the inherent inconsistencies, while at the same time also seek indication within the nature of the dataset that the level of variance or inconsistencies are intuitively within reasonable bounds for the investigation. Therefore, it is important to reserve ample resources for data discovery and pre-processing during planning of such investigations to ensure sufficient quality of research outputs. Unfortunately, there is no agreed standard to define or achieve a sufficient quality of outputs in such analyses, because so much depends on the nature of data being analysed. Nevertheless, additional care can be taken to, for instance, make lucid arguments about the inconsistencies of the data, or justify the selection of a subset of sufficiently structured data to analyse so that its outputs could be extrapolated to the whole dataset.

Procurement market fragmentation vs cohesion

Using VOSviewer’s default output of 512 companies as nodes of the network, we detect that some of them are not connected to each other. The largest set of connected items consists of 428 companies (~83.5%), meaning that 84 companies (~16.5%) are not connected to the main network and thus are fragmented. Almost all (79) of the 84 that are fragmented or detached from the main large network have no links at all, implying that they have only participated in projects on their own without any other collaborators. We can thus make the general deduction that a company would have a high likelihood to be connected to the large network of companies, even with just one other collaborator. In other words, there are almost never a group of collaborating companies that are fragmented from the large main network, thus indicating a very cohesive network.

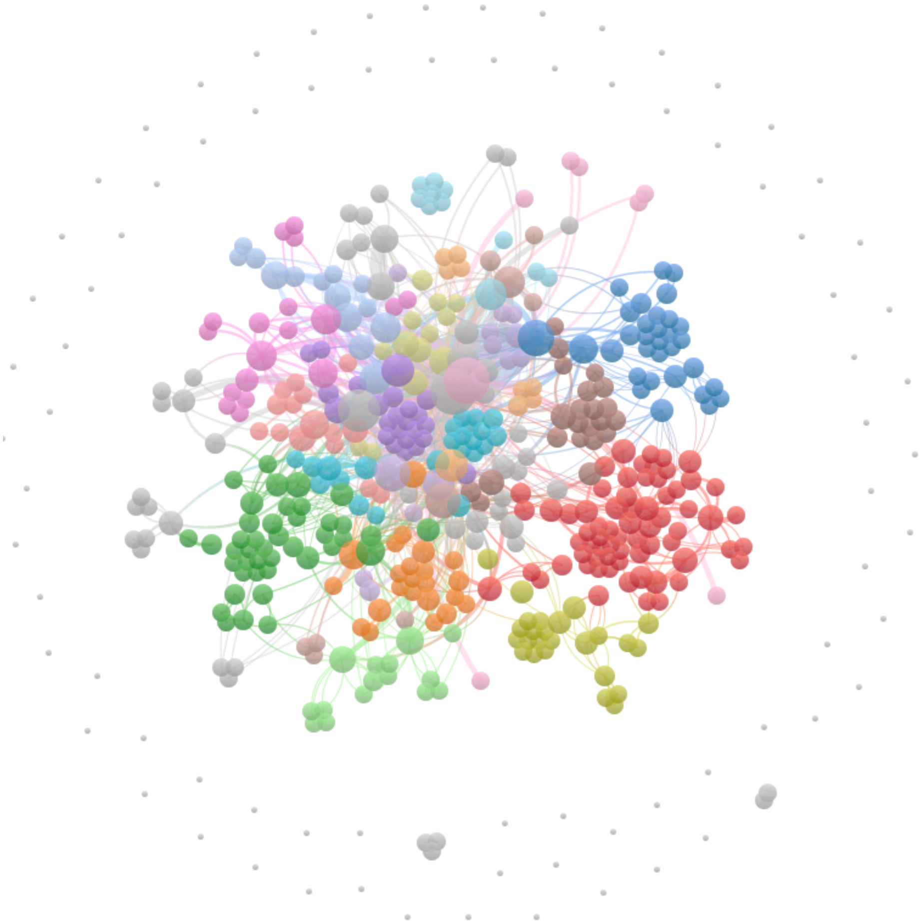

Figure 8. Overview of the network visualization, showing the large main connected network in the centre and fragmented companies on the periphery as grey dots.

Most companies are collaboratively connected to one another by some degree of separation, thus forming a highly cohesive network.

The same can be discerned visually in Figure 8 that most companies are connected to the main network, implying a cohesive procurement market where most companies are collaboratively connected to one another by some degree of separation. At the same time, it is not surprising that there are companies that are fragmented from the main network of interconnected companies. Nevertheless, it is difficult to intuitively make sense of the fragmentation ratio of companies quantitatively, without relative examples or results of other cases for comparative interpretation. There are no broad explicit definitions for fragmentation and there are no studies from literature that investigates how to quantify the fragmentation or cohesion of a public procurement supplier network, nor the implications on transformative or innovation adoption processes.

In financial literature, market fragmentation is typically considered detrimental (Stoll, 2001), which is also aligned with the general agreement that knowledge fragmentation in the AEC industry is undesirable (Kuo, 2019). Market fragmentation, as referred to in financial literature, may not be a direct analogy to the fragmented companies within the procurement network in our investigation, but can nevertheless provide some frame of reference. Accordingly, the fact that only 16.5% of companies are not connected to the main network may be considered a positive indicator for the cohesion, inclusiveness, and responsiveness of the market, that is, a market in which fairness can be maintained so that the most suitable companies are reachable for new project calls. Likewise, low network fragmentation also indicates good transferability of tacit knowledge within the collaborative networks of companies providing architectural services in Finland.

However, high network cohesion may intuitively also lead to excessive conformity of ideas and beliefs within the network, akin to the well-known anecdote that “birds of a feather tend to flock together”, compared to networks that would be more fragmented. One may assume that a more fragmented group of companies is also more likely to have varying beliefs and behaviours that go hand in hand with more diverse or creative solutions. This raises questions about whether there could be a quantitatively optimal level of fragmentation vs cohesion in a collaborative network that would be more conducive to the emergence of more creative innovations, but that still maintains enough connectedness to ensure that these same creative innovations would not be locked within its own closed fragmented islands, bubbles, or schools of thought.

The ratio of fragmented companies does not give a detailed indication on the distribution of the connections within the main connected network. The following sections delve deeper into the connected main network using network analysis methods.

Degree centrality and distributions

For the interconnected network consisting of 512 companies from VOSviewer’s outputs, we calculate the degree centrality for each company, that is, the number of unique links to other companies. The 20 highest ranked companies are shown in Table 5. Indeed, for readers familiar with the Finnish market, most of these companies should be commonly known and are active within architectural developments to various extents.

Table 5. Top 20 companies by degree centrality

Rank	Company	Degree centrality
1	sitowise oy.	252
2	ramboll finland oy.	208
3	wsp finland oy.	174
4	fcg suunnittelu ja tekniikka oy.	135
5	sweco talotekniikka oy.	102
6	sweco rakennetekniikka oy.	81
7	destia oy.	79
8	arkkitehtitoimisto a-konsultit oy.	77
9	a-insinöörit suunnittelu oy.	74
10	taratest oy.	71
11	hmt arkkitehdit oy.	71
12	rejlers finland oy.	68
13	geopalvelu oy.	67
14	profit interior oy.	67
15	afry finland oy.	66
16	vahanen suunnittelupalvelut oy.	66
17	a-insinöörit civil oy.	64
18	rakennuttajatoimisto htj oy.	63
19	sweco ympäristö oy.	62
20	welado oy.	59

Distribution of social capital in the network that is highly skewed in favour of the most central companies.

There are in total 4381 links between the 512 companies, of which the top 20 companies make up a total of 951 links. Therefore, 21.75% of all links in the network are held by only 3.9% of the top companies by degree centrality. This indicates a distribution of social capital in the network that is highly skewed in favour of the most central companies, as shown explicitly in Figure 9 throughout the whole market. We see that the degree centrality drops significantly into a steep elbow, followed by a disproportionately long tail of companies with very gradual decrease of connectedness down the ranks. This graph indicates a clear so-called power-law pattern, where very few companies are disproportionally better connected than the average. Interestingly, the power-law distribution is a common emergent pattern observed in many chaotic systems, regardless of whether human consciousness would be involved, such as in physics and the natural world, as well as economics and sociology.

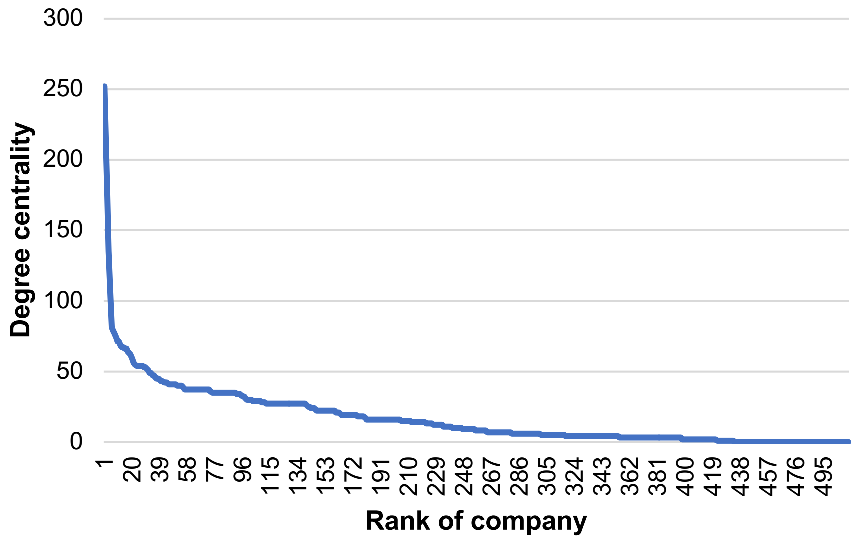

Figure 9. Distribution of degree centrality across the market.

Using VOSviewer, we can hone into any node to see the visual representation of the specific company. Figure 10 shows the specific network of the company with the highest degree centrality i.e. “Sitowise”, and its broad tentacles of connections across the market. This makes the complex landscape of collaborative procurement networks more intuitively explorable.

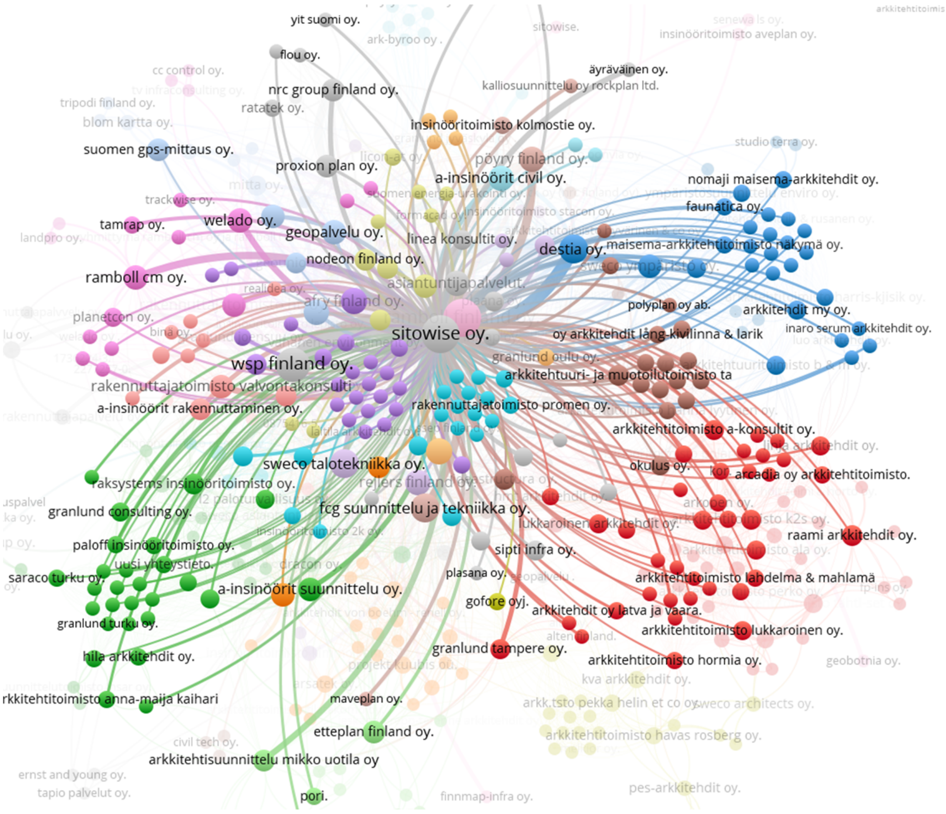

Figure 10. Network of the company “Sitowise” with highest degree centrality of 252.

We note that companies, or clusters of companies, with highest degree centralities are often also linked to one another. For instance, many of the companies ranked in the top 20 in Table 5 can be seen in Figure 10 explicitly connected to the highest ranked, namely, Sitowise.

To illustrate how other companies with a lower degree centrality are visualized, Figure 11 shows the immediate network of Arkkitehtitoimisto Erat Oy with a degree centrality of 15, which can be counted even manually. By contrast, Figure 12 shows a pair of companies with just degree centralities of 1, collaborating with each another in a project.

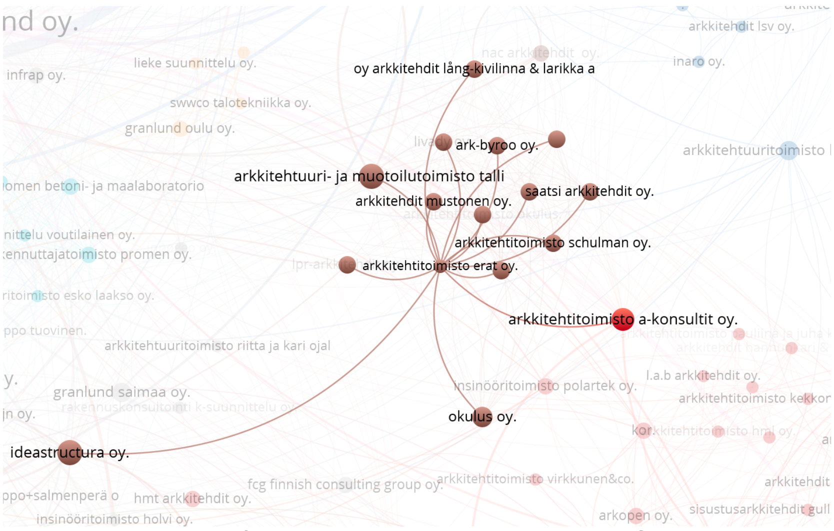

Figure 11. Network of company Arkkitehtitoimisto Erat Oy with degree centrality of 15.

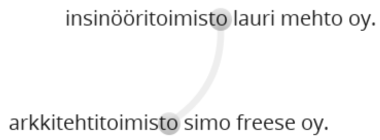

Figure 12. Network visualization of a small 2 company consortium.

Aggregate graphs of the degree centralities and interactive visualization serves as an additional way to explore the distribution of social capital in the market.

The degree centrality can be a good indicator of the capability for a particular company to share knowledge or values with its linked companies, which in turn can also be seen as a measure of influence. In addition to aggregate graphs of the degree centralities, the interactive visualization serves as an additional way to explore the distribution of social capital in the market. This enables inspection and identification of the main companies that are central knowledge hubs, as critical nodes for knowledge dissemination in the network. Such knowledge hubs can also be good targets for strategic partnerships for maximum impact when applying certain solutions in the market.

Full and fractional link strengths

While the degree centrality indicates the number of unique connections to other companies, it does not indicate the weight of those links. The counting of link strength is done to include the consideration of the weight of the links, which can indicate the proverbial "bandwidth" of knowledge pathways, since the more times companies collaborate, the higher their total link strengths. It is reasonable to assume that high frequency of collaboration between companies relates to smoother communication, higher trust, and deeper relationship between those companies. To compare the 2 counting methods, namely the full and fractional counting, the top 20 ranked companies of each are shown in Table 6.

Table 6. Top 20 ranked companies by full vs fractional counting of total links strengths.

Full counting		Factional counting	
sitowise oy.	516	sitowise oy.	77
ramboll finland oy.	451	ramboll finland oy.	77
wsp finland oy.	370	wsp finland oy.	49
fcg suunnittelu ja tekniikka oy.	200	destia oy.	24
destia oy.	194	sweco talotekniikka oy.	23
sweco talotekniikka oy.	149	fcg suunnittelu ja tekniikka oy.	22
afry finland oy.	126	afry finland oy.	15
sweco rakennetekniikka oy.	121	sweco rakennetekniikka oy.	14
taratest oy.	117	rejlers finland oy.	13
asiantuntijapalvelut.	110	a-insinöörit civil oy.	13
a-insinöörit civil oy.	109	asiantuntijapalvelut.	13
a-insinöörit suunnittelu oy.	106	taratest oy.	12
geopalvelu oy.	98	pöyry finland oy.	12
rakennuttajatoimisto htj oy.	95	geopalvelu oy.	11
sweco ympäristö oy.	91	welado oy.	10
welado oy.	90	ramboll cm oy.	10
rejlers finland oy.	87	a-insinöörit suunnittelu oy.	9

vahanen suunnittelupalvelut oy.	85	rakennuttajatoimisto htj oy.	9
arkkitehtitoimisto a-konsultit oy.	81	finnmap infra oy.	9
finnmap infra oy.	77	vahanen suunnittelupalvelut oy.	8

The top 20 ranked companies using full and fractional counting do not differ much. Therefore, in this investigation the effects of full vs fractional counting appear insignificant for ranking the companies by total link strengths. Furthermore, we can plot the total link strengths for all the companies of the network and compare the results of full vs fractional counting, as in Figure 13.

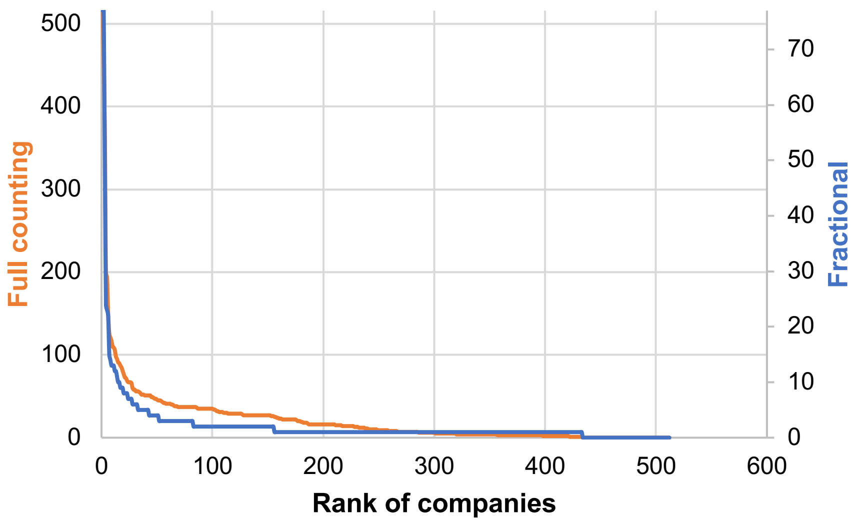

Figure 13. Distribution curve of total link strengths comparing full and fractional counting

Companies with high degree centralities occurs more often in large consortia.

The distribution curves only exhibit very minor differences between the full and fractional counting, when plotted on 2 separate y axis, showing how the total link strengths are distributed among all companies from the highest to the lowest. Like the degree centrality distribution graph (Figure 9), the total link strengths of both the full and fractional counting are distributed with extreme inequality in favour of those companies that are highly ranked. These curves clearly follow a power-law distribution, showing that there are extremely few companies with very high total link strengths compared to the whole market, and a long tail of companies with very low total link strengths.

Full counting (orange curve in Figure 13) appears to have a slightly more gradual or even distribution of total link strengths compared to that of fractional counting. Visually it seems that the difference is negligible. Nevertheless, this difference theoretically indicates that the companies with high degree centralities, i.e. many unique links to other companies, occurs more often in large consortia, to the extent that it impacts their relative total link strength per fractional counting, which discounts the link strengths of each pair of companies in a consortium proportionally to the size of the consortium. The larger the consortium in which 2 companies co-occur, the weaker their link strength by fractional counting.

Clustering layout and visualizations

The total of 512 companies are algorithmically clustered into 108 groups based on their relatedness using the VOSviewer default clustering algorithms. Out of the 108 clusters, 29 clusters consist of 1 or more companies, meaning that 79 clusters are of single companies that are not connected with another company, as was discussed in a previous section on cohesion. The top 5 clusters are shown in Figure 14 as the full visualization layout, and Table 7 lists the top 5 clusters and corresponding colours and sizes.

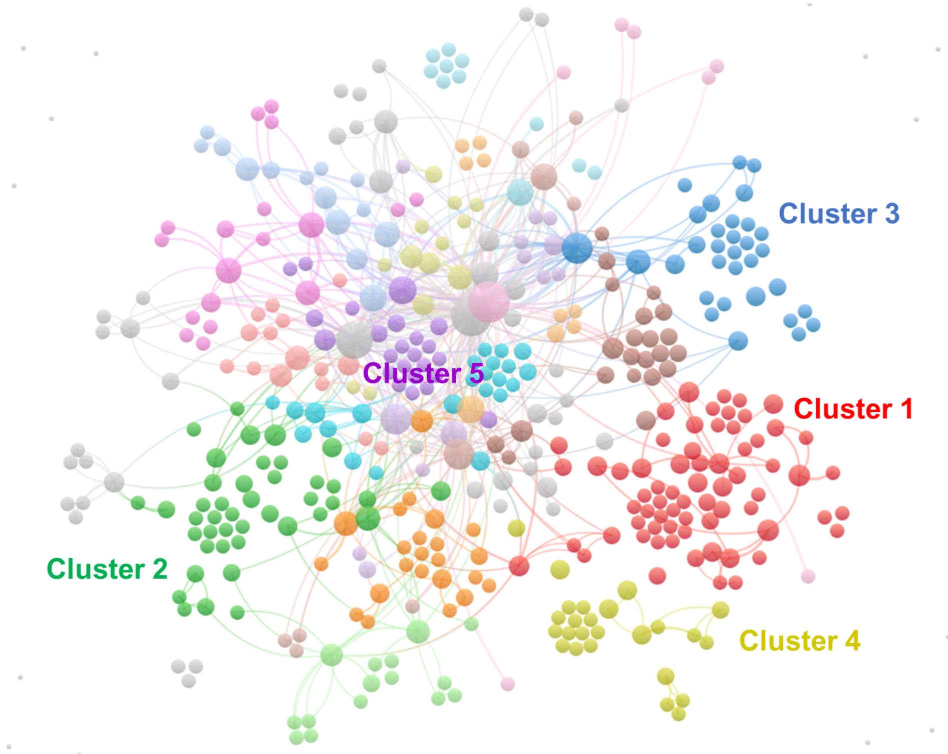

Figure 14. The whole network landscape with the top 5 clusters labelled.

Table 7. Sizes of top 5 clusters and corresponding colours codes

Cluster nr	Cluster size
1 (red)	68
2 (green)	42
3 (blue)	32
4 (yellow)	25
5 (purple)	24

We note that just the top 5 clusters out of 108 clusters in total consists of 191 companies, that is, 37.3% of all the companies, which also indicates a very steep distribution of communities in favour of the largest. This can be seen even more clearly through plotting the distribution of cluster sizes of all the companies in Figure 15.

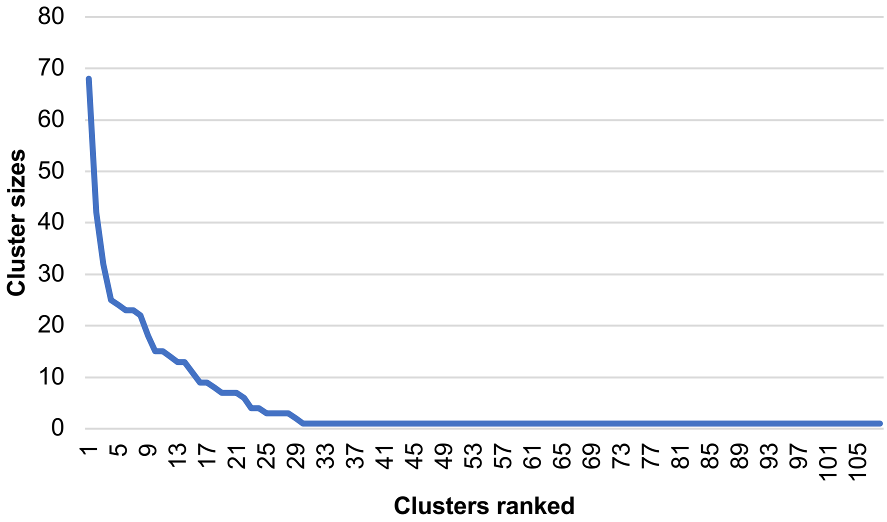

Figure 15. Distribution of cluster sizes.

Furthermore, we briefly investigate the distribution of degree centrality within the largest Cluster 1, to explore how social capital as per degree centrality varies within the largest cluster. Table 7 shows the top 10 companies by degree centrality within just Cluster 1, accompanied by the network layout in Figure 16. Even within the cluster, we see a similar trend of highest ranked companies having non-linearly higher degree centralities than others ranked lower down.

Table 7. Top 10 companies by degree centrality within in the largest Cluster 1.

Rank	Company	Degree centrality
1	arkkitehtitoimisto a-konsultit oy.	77
2	hmt arkkitehdit oy.	71
3	avario oy.	54
4	arkopen oy.	49
5	pook arkkitehtitoimisto oy.	49
6	arkkitehtitoimisto k2s oy.	47
7	arkkitehtitoimisto ala oy.	46
8	innovarch oy.	45
9	siren arkkitehdit oy.	43
10	verstas arkkitehdit oy.	42

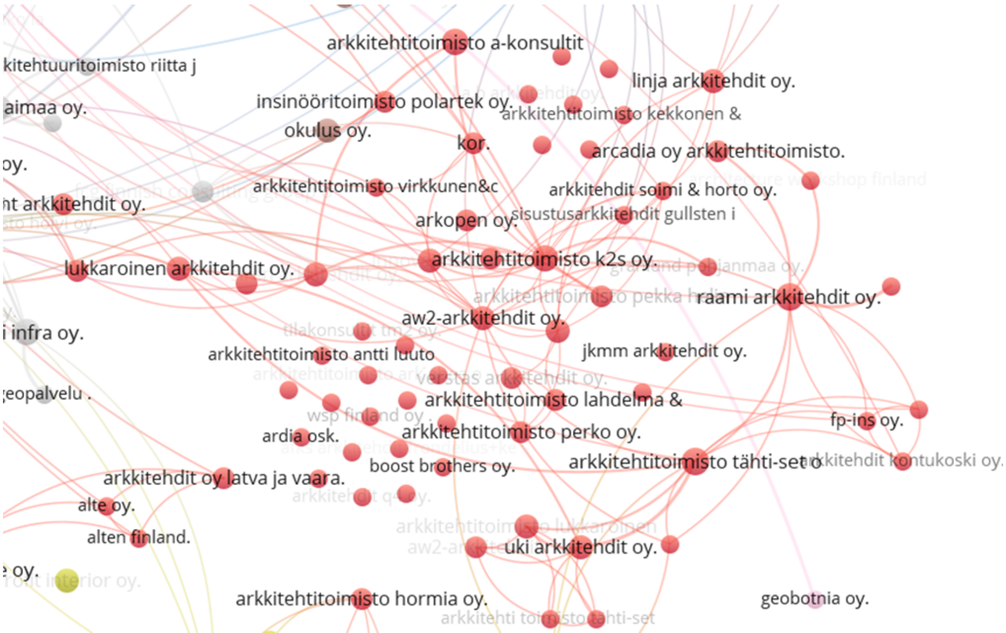

Figure 16. Network visualization of the largest Cluster 1.

We go a step further (Figure 17) to plot the distribution curves of all top 5 clusters, but instead of using the rank of the company as an absolute value on the x axis, we scale all the ranks to 100%, in order to compare the degree centrality distributions per the percentage of ranked companies across all 5 clusters with different sizes.

The power-law curve is revealed not only in the distribution of cluster sizes over the whole network, but there is a similar non-linear inequality even within each of the top 5 clusters, exhibiting traits of self-similarity.

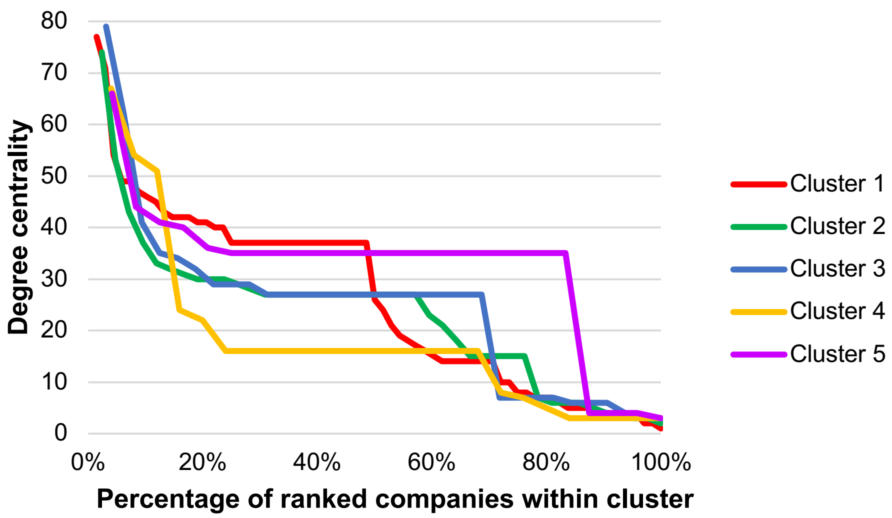

Figure 17. Comparison by distribution of degree centrality within the top 5 clusters

Clusters give an indication of the collaborative and knowledge pathways of the market on a lower level of abstraction. The power-law curve is revealed not only in the distribution of cluster sizes over the whole network, but there is a similar non-linear inequality even within each of the top 5 clusters, exhibiting traits of *self-similarity*, which is a pattern also common in many mathematical expressions describing natural elements, e.g. fractal structures, coastlines, leaf structures etc.

This leads us to deduce that the type of inequality distribution of social capital in procurement communities may be an emergent property that arises from complex interactions of the network. Furthermore, they take place on different levels of abstraction even within clusters consisting of just 10s of companies, thus suggesting some level of *scale invariance*, i.e. describing the pattern where there is a smaller piece of the object that is similar to the whole, regardless of magnification. We can see some aspects of these patterns on different levels of the company networks.

Companies' share of projects

Using the method of compiling the company-to-project matrix, normalizing the project columns, and adding up the totals per company, we can plot an approximate graph showing the number projects captured by companies in descending order of rank. Again, we see the power-law distribution clearly in Figure 18, indicating that very few highly influential companies capture a large share of the market.

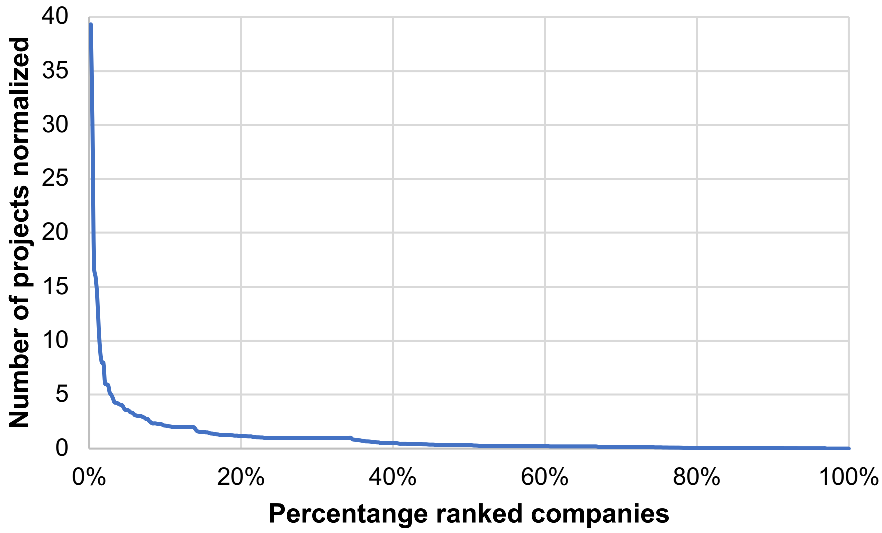

Figure 18. Rank-size plot of the number of projects captured by companies ranked.

80% of the market is captured by top 27% of companies.

With the cumulative plot of the same data (Figure 19) we can clearly inspect the cumulative share of projects in the total market by the percentage of top companies. For instance, on various points along the curve, we see that the top 10% of companies hold 58% share of the market, top 20% of companies hold 73% of the market, and so forth. The insights can be intuitively stated in a different way, for instance, 50% of the market is capture by just the top 7% of the companies, while 80% of the market is captured by top 27% of companies.

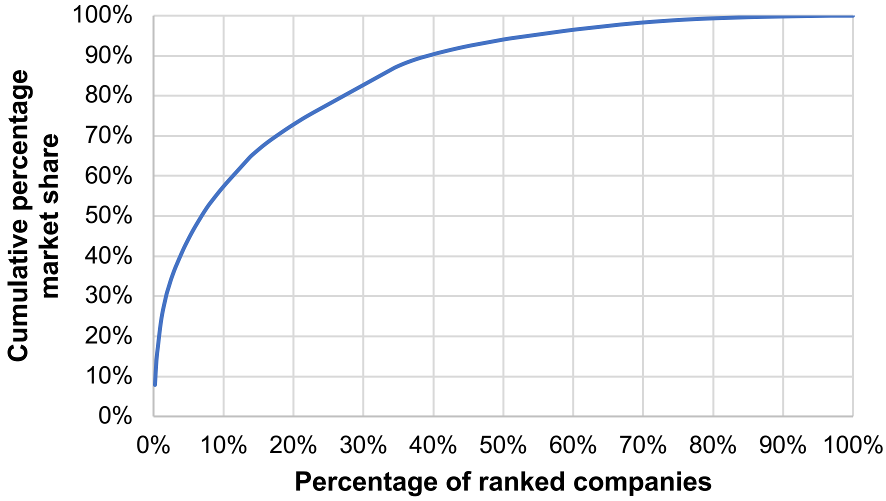

Figure 19. Cumulative plot of the share of the market by percentage of companies.

Implications on AEC transformation, resilience, and innovation

Analytics and insights on the public procurement networks of architectural and related services provide additional dimensions to assess the potential of transformation and innovative development of the market. Networks show how ideas, values, and tacit knowledge can be shared across the ecosystem and indicate how social capital, influence and power are distributed among companies that drive real impacts in architectural and related services sector.

We have found that there is high cohesiveness among companies, which indicates good efficiency of how change can be incentivised and proliferate through the industry. This is especially important in terms of applying progressive transformative goals such as carbon neutrality, anti-corruption, or innovative products and solutions. It is promising, and not surprising, to see the high cohesion of Finnish companies, forming a robust and focussed network of knowledge pathways to facilitate continuous improvements of industry values as a closely knit unit.

However, the high market cohesion also implies a potential for the lack of diversity. At the same time our analyses on the distributions of degree centralities, total link strengths, and companies' market share, show that social capital and knowledge within the whole network are dominated by very few powerful hubs of extremely influential companies. These reveal the susceptibility of "the rich getting richer" effect in commerce or economics terms as an emergent network phenomenon. The risk is that such emergent forces and may overwhelm the desired goals of meritocracy, that is, distributing power based more on intrinsic qualitative value of solution proposals and companies.

Distributions of degree centralities, total link strengths, and companies' market share, show that social capital and knowledge within the whole network are dominated by very few powerful hubs of extremely influential companies.

The extremely skewed advantage in favour of the highest ranked companies and communities raises questions about some potential fundamental problems. There are disproportionately high influence residing among very few companies, which is against the principles of diversity. Typically, the holders of the power are incentivized to defend the status quo as it favours them at least in the shorter-term pragmatic goals. This potentially results in a resistance to change and can stifle the creative and experimentation mindset, which are considered risky, in favour of consistency and security. Nevertheless, this type of power-law distributions actually occurs very often in apparently optimal non-deterministic systems such as in nature e.g. biology, geology, chemistry, astronomy, or human behavioural systems e.g. sociology, psychology etc. All cases are underpinned by some type of *preferential attachment* in statistical terms, which describes a situation where well-connected entities are more likely to get new connections compared to those entities with fewer or weaker connections. Seeing that it emerges from both living and non-living natural scenarios, this pattern or phenomenon may hold some other optimal properties, which deserves deeper exploration. For instance, there may be emergent aspects of how our markets work that are inherently out of our control, regardless of any public or private sector interventions or resources. On the other hand, there may be analogous dynamics or actions that we may learn from other domains or systems that may be effectively applied in the AEC ecosystem.

With the quantitative interpretations in our investigation, we realize that resilience and responsiveness of the Finnish architectural services market, hailed as unquestionably positive goals, start to become ambiguous concepts that are difficult to define and thus pursue. On the one hand, a cohesive network enables efficient knowledge flows. Such cohesive capabilities for the sharing of values thus indicates a market that is resilient to disruptions. However, the flip side of the same coin is that the inherent resistance to change, as a result of the skewed distribution of influence, may stifle pervasive radical innovations on the broad market level, which are essential for longer term sustainability or even survival. This dichotomy is reminiscent of a classical anecdote that the architecture mindsets tend to try new and innovative things. There is, after all, the Frank Lloyd Wright's well-known statement that, "if the roof doesn't leak, the architect hasn't been creative enough" (Donahue, 1989). However, novelty is also often untested, and thus holds inherent shortcoming in pragmatic utility or function, which are just as important particularly in the short to medium term.

Multi-faceted contributions

This paper places the emphasis on the algorithmic network analysis and related methods and contributes foremost to advancing the utilization of data science in the AEC field. The network analysis method can be scaled and applied to different datasets of the same problem context, that is, mapping and quantifying organizational relationships of awarded contracts, or otherwise, modelling the sociological nature of other relational or transactional phenomena, such as supply chains and contractual partnerships, as in Pryke (2005). Our study contributes to the currently meagre quantitative analyses of public procurement networks in architectural literature, building on similar methods employed by Leiva *et al*. (2020) that focussed on the overall Chilean market. Beyond public procurement, our methodology can be used to model a variety of other relational systems, not only of people and organizations (Senaratne, Jin and Denham, 2021), but also of technology, in the form of data exchange and information systems.

Whereas public procurement and related business research within architectural journals might not be common, applying studies especially from the field of public procurement for architectural innovation adoption have good potential. For instance, there are myriad possibilities to reflect the implications of our research in terms of public procurement for innovation, focussed on the procurer's commitment to buying innovative products and services to mitigate societal problems (Edquist and Zabala-Iturriagagoitia, 2012). Furthermore, network modelling approaches can complement existing research on assessing the value of procurement from the perspectives of the ecosystem and end-users (Torvinen and Haukipuro, 2017; Malacina *et al.*, 2022), which has employed more qualitative and interpretive approaches, such as interviews, participant observations, literature reviews and case studies.

Aside from academic discussions and generic policy viewpoints, there are many practice-orientated contributions attributed to this study. Industry practitioners and especially researchers need more holistic understanding of the business environment in which they operate. Our study contributes to the viewpoint that innovations are not only a matter of creativity and excellence of quality inherent in proposed solutions, but that is also a result of systemic collaborative networks between actors in the market, enabling the discovery and eventual adoption of innovation, proliferating through the society overtime. The results of our study emphasize the sociological nature of how companies collaborate and network, as a potential indicator of their collective influence on the architectural and related services on our society. To promote the realization of innovation and transformation as solutions to be procured, it does not suffice only to focus on the quality of the solutions. It is just as, if not more, important to consider the networks and collaborative relationships with other companies and procurers.

Practitioners should devote ample thought and effort to the discoverability of their ideas and tacit knowledge by other professionals and promote outreach and engagement towards different actors of the ecosystem. At the same time, it is worth acknowledging that focussing only on connections and neglecting the quality of the solutions offered is also detrimental, and thus a balanced strategy is appropriate. Afterall, the core purpose of architectural and related expertise is to provide innovative solutions, while the networks and connections merely facilitate the dissemination of said solutions.

In an analogous context, one study utilized similar data-driven network analysis methods to quantify the reputation and success in the art world (Fraiberger *et al.*, 2018). They reconstructed the exhibition history of half a million artists, mapping out the co-exhibition network that captures the movement of art between institutions. They found that centrality within this network captured institutional prestige, and an artist’s early access to prestigious central institutions offered life-long access to high-prestige venues and reduced dropout rate. On the other hand, an artist who started at the network periphery resulted in a high dropout rate. We may compare the sociological nature of success in the world of artists and exhibitions, with that of our study of the architectural and related market networks. For artists, however, it may be that the value of the artistic artefact succumbs to higher subjectivity, compared to that of architectural services. This may explain the seemingly higher significance of an artists’ connections to central hubs as a factor of their success, compared to that of architectural practitioners, which rely also significantly on the quality of their artefacts.

A similar study in the world of sports, Yucesoy and Barabasi (2016) showed that tennis players have high correlation between their popularity and objectively measurable performance. They concluded that the agreement between the performance and observed popularity suggests that in most areas of achievement, exceptional visibility may be rooted in detectable performance measures. The comparison between popularity or success in professional tennis players and artists is an interesting lens through which architectural practitioners may develop their own trajectory and influence. The “performance” of an artist’s artefact is less, if at all, intrinsically measurable, compared to that of a tennis players performance. Similarly for architects, where the performance, that is, quality of work outputs can be measured, it is also likely to attribute to popularity and success and connection to central hubs.

Architectural and related services embody both measurable outputs, akin to sports, as well as artistic, social, that is, less measurable performance outcomes, akin to an artist’s development in the path to success. Thus, for companies offering innovative architectural and related services, it would be a valuable endeavour to optimize their strategy accordingly, to balance the efforts devoted to quality of the solutions and proposals, as well as the fostering of skills and processes to promote networking with well-connected actors in the ecosystem.

Conclusions and recommendations

Our investigation is the first that characterizes the public procurement networks of architecture and related companies in Finland using data mining and quantitative network analysis. We motivate the importance of procurement topics in decision-making for optimizing impact and transformative innovations in the AEC industry. Our research provides a set of methods and tools to mine/process raw open data published by the procurement body of the Finnish government; assess the emergent collaborative procurement networks of companies; and explain how the results can be interpreted to generate helpful and actionable insights.

Firstly, we show that the raw open data of project award notices are of reasonable usability, but has room for improvement for making them public, particularly regarding the removal of erroneous extra spaces, non-printable characters and spelling variants. There are inconsistencies in the data during the parsing stage that are not easy to identify in an automated way that can generalize to all datasets, because many inconsistencies would not have been apparent without manual inspection and often are specific to certain datasets. Therefore, ample attention should be given to surveying and discovery of the data and its relation to the problem domain, as well as the analysis methods.

Arising from our network analyses, the graphs and visualizations reveal the distribution of collaborative connections of companies responsible for real solution implementation in the market, as well as explicit quantification of the share of the market belonging to top percentage of companies. These insights help researchers, practitioners and policy makers by providing a glimpse into the real market, and enable assessment of one’s own potential in promoting innovative design knowledge across the complex AEC market via partnerships.

Our principal findings show that the Finnish ecosystem of architecture and related services is very cohesive and has extreme preferences for the most connected companies. Companies that are ranked high based on their centralities, total ink strengths, and market share, have exponentially more influence than those ranked lower, as shown by the power-law curves of the distribution plots. Furthermore, there are evidence of self-similar patterns within the network, so that the same type of advantages held by the few can be observed on different levels of abstraction or magnification of the network. While we have demonstrated the feasibility of utilizing open data to generate insights using network analysis methods, we have also shown the challenges of processing the data and making intuitive interpretations that can inform appropriate and impactful actions, be it for policy makers or companies.

It would be pertinent in future that governments would take further considerations whether these distributions of social capital, value, work, wealth, or power amid the architectural and related services market is indeed optimal to achieving the national or municipal goals. Following the reasoning, incentives should be applied to guide the industry into a more desirable state with the help of indicators shown in our research. For instance, tax exemptions or deductions could be used to incentivise cross domain collaboration e.g. if specialist companies are encouraged to collaborate with non-specialists etc. or between architecture, engineering and contractor firms, to even out existing skewed favouritism of certain powerful companies. Perhaps first-time collaborations between companies could be incentivized more than collaborations with previous existing relationships, to distribute the total link strengths more evenly across the market. Such incentives may be encouraged through the setting of more appropriate procurement criteria by procurers. For instance, setting certain minimum procurement requirements in the procurement call for references, size, or financial status of companies can be adjusted to reduce biases and favouritism.

On the other hand, companies or research organizations can utilize the tools and methods, or simply the outputs of our research to identify concentrations of hubs or communities as targets for most impactful strategic partnerships or potential clients.

Our research paves the way towards further questions on the relationship of our collaborative networks of companies and its implications on other market phenomena in terms of transformation and innovation. For instance, a cohesive collaborative network of companies enables the focus and conformity for efficient planning and execution of impactful market-wide developments and change. However, such a cohesive network would have a strong mainstream of ideas or practices that tends to prevail over unique or different new ideas which are often marginalized, often for sake of reducing risk and/or maintaining the status quo. More investigations are required to better understand these complex forces and their observable effects on architectural innovations.

As engineers, architects, contractors, or researchers, we need better understanding of the environment in which our creative solutions need to be accepted, adopted, and eventually thrive, in addition to merely considering the solutions themselves. We also need more awareness about our potential role in strategic collaborations and influence towards government and policy makers that shape the market. In this way, we may have more influence on stimulating the company networks positively through expert knowledge, resulting in a good balance between diverse, creative solutions, as well as the effectiveness in the discoverability, dissemination and adoption of such ideas and solutions in the real-world market.

References

Arrowsmith, S. (2005) *The Law of Public and Utilities Procurement*. London: Sweet & Maxwell.

Cartlidge, D. (2003) *Procurement of Built Assets*. 1st edn. Routledge.

Donahue, J. (1989) 'Fixing Fallingwater's flaws: The Leaks and Deteriorating Concrete', *Architecture*, pp. 99–101.

Van Eck, N.J. *et al.* (2010) 'A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS', *Journal of the American Society for Information Science and Technology*, 16(12), pp. 2405–2416.

Van Eck, N.J. and Waltman, L. (2014) 'Visualizing Bibliometric Networks', in Y. Ding, R. Rousseau, and D. Wolfram (eds) *Measuring Scholarly Impact*. Springer, Cham, pp. 285–320.

Edler, J. *et al.* (2005) *Innovation and Public Procurement: Review of Issues at Stake.* Germany.

Edquist, C. and Zabala-Iturriagagoitia, J.M. (2012) 'Public Procurement for Innovation as mission-oriented innovation policy', *Research Policy*, 41(10), pp. 1757–1769.

European Commission (2021) *The strategic use of public procurement for innovation in the digital economy*. Brussels.

Fraiberger, S.P. *et al.* (2018) 'Quantifying reputation and success in art', *Science*, 362(6416), pp. 825–829.

Golbeck, J. (2015) 'Analyzing networks', in J. Golbeck (ed.) *Introduction to Social Media Investigation*. Syngress Elsivier, pp. 221–235.

Ihku Alliance (2021) *Infrahankkeiden kustannuslaskentajärjestelmä ja -palvelualliansi*. Available at: https://ihkuallianssi.fi/ (Accessed: 6 June 2021).

Kohler, N. and Moffatt, S. (2003) ‘Life-cycle analysis of the built environment’, *UNEP Industry and Environment*, 26(2), pp. 17–21.

Kuo, V. (2019) *Latent semantic analysis for knowledge management in construction*. Aalto University.

Leenders, R.T.A.J. and Gabbay, S.M. (1999) *Corporate Social Capital and Liability*. Dordrecht/Boston/London: Kluwer Academic Publishers.

Leiden University (2022) ‘VOSviewer’.

Leiva, F.H. *et al.* (2020) ‘Characterization of the Chilean Public Procurement Ecosystem Using Social Network Analysis’, *IEEE Access*, 8, pp. 138846–138858.

Malacina, I. *et al.* (2022) ‘Capturing the value creation in public procurement: A practice-based view’, *Journal of Purchasing and Supply Management*, 28(2), p. 100745.

Nahapiet, J. and Ghoshal, S. (1998) ‘Social Capital, Intellectual Capital, and the Organizational Advantage’, *The Academy of Management Review*, 23(2), pp. 242–266.

Nissen, M.E. (2009) ‘Procurement: process overview and emerging project management techniques’, in P. Morris and J.K. Pinto (eds) *The Wiley Guide to Project Technology, Supply chain, and Procurement Management*. Wiley, Hoboken, New Jersey.

OECD (2023) *Public procurement*. Available at: https://www.oecd.org/gov/public-procurement/.

Perianes-Rodriguez, A., Waltman, L. and Van Eck, N.J. (2016) ‘Constructing bibliometric networks: A comparison between full and fractional counting’, *Journal of Informetrics*, 10(4), pp. 1178–1195.

Poirier, E., Forgues, D. and Staub-French, S. (2016) ‘Collaboration through innovation: implications for expertise in the AEC sector’, *Construction Management and Economics*, 34(11), pp. 769–789.

Portes, A. (1998) ‘Social Capital: Its Origins and Applications in Modern Sociology’, *Annual Review of Sociology*, 24(1), pp. 1–24.

Pryke, S.D. (2005) ‘Towards a social network theory of project governance’, *Construction Management and Economics*, 23(9), pp. 927–939.

Rahmani, F., Maqsood, T. and Khalfan, M. (2017) ‘An overview of construction procurement methods in Australia’, *Engineering, Construction and Architectural Management*, 24(4), pp. 593–609.

Rolfstam, M. (2013) *Public Procurement and Innovation: The role of institutions*. Edward Elgar Publishing.

Rowlinson, S. and McDermott, P. (1999) *Procurement System: A Guide to Best Practice*. London, UK: Spon Press.

Ruan, X., Ochieng, E.G. and Price, A.D.F. (2011) ‘The evaluation of social network analysis application’s in the UK construction industry’, in C. Egbu and E.C.W. Lou (eds) *27th Annual ARCOM Conference*. Bristol, UK: Association of Researchers in Construction Management, pp. 423–432.

Schumpeter, J. (1934) *The Theory of Economic Development*. New York: Oxford University Press.

Scott, J. (1992) *Social Network Analysis: A Handbook*. London: Sage.

Senaratne, S., Jin, X. and Denham, K. (2021) ‘Knowledge sharing through social networks within construction organisations: case studies in Australia’, *International Journal of Construction Management* [Preprint].

Stoll, H.R. (2001) ‘Market Fragmentation’, *Financial Analysts Journal*, 57(4), pp. 16–20.

Torvinen, H. and Haukipuro, L. (2017) ‘New roles for end-users in innovative public procurement: case study on user engaging property procurement’, *Public Management Review*, 20(10), pp. 1444–1464.

Valtiovarainministeriö (2021) *Hankinta-Suomi – julkiset hankinnat vaikuttavammiksi*. Available at: https://vm.fi/hankinta-suomi (Accessed: 6 June 2021).

Waltman, L. and Van Eck, N.J. (2013) ‘A smart local moving algorithm for large-scale modularity-based community detection’, *The European Physical Journal B*, 86(471).

Waltman, L., Van Eck, N.J. and Noyons, E.C.M. (2010) ‘A unified approach to mapping and clustering of bibliometric networks’, *Journal of Informetrics*, 4(4), pp. 629–635.

Yucesoy, B. and Barabasi, A.-L. (2016) ‘Untangling performance from success’, *EPJ Data Science*, 5(17).