Journal of Purchasing & Supply Management 30 (2024) 100892

Contents lists available at ScienceDirect

Journal of Purchasing and Supply Management

journal homepage: www.elsevier.com/locate/pursup

Note and Debate

Text mining and network analytics for literature reviews: Exploring the landscape of purchasing and supply management research

Robert Suurmond [a,*], Finn Wynstra [b], André Vermeij [c], Erick Johan Haag [b]

[a] Maastricht University, School of Business and Economics, Netherlands
[b] Erasmus University, Rotterdam School of Management, Netherlands
[c] Kenedict Innovation Analytics, Netherlands

ARTICLE INFO

Keywords:
Purchasing and supply management
Literature review
Scientometric study

ABSTRACT

In this paper, we state and debate the use and usefulness of text similarity and network analytics using natural language processing for our field. While previous reviews of Purchasing and Supply Management have relied on manual coding and classification, the large scale and variety of the field calls for new approaches. In this Notes and Debates article, we therefore review different approaches from bibliometric and scientometric studies to explore literature using (semi)automated approaches. We exemplify one approach, leveraging text similarity and network visualization, to complement earlier analysis. Along the way, we discuss the researcher's role at critical vantage points in reviews that are augmented by natural language processing. We compare and contrast the results of this exploration to previous manual reviews and sketch opportunities and provide recommendations for future use.

1. Introduction

The field of Purchasing and Supply Management (PSM) has become increasingly larger and more mature, but also more diverse and varied. Previous reviews of this literature have been instrumental in the reflection on and further development of the identity of PSM. Most of these reviews have relied on manual coding and analysis. For example, coding of topics based on article contents (Carter et al., 2014; Chicksand et al., 2012; Hult and Chabowski, 2008; Spina et al., 2013; Wynstra et al., 2019). Other reviews have used social network analysis, based on citation ties (Carter et al., 2007) or co-citation ties (Hult and Chabowski, 2008). Generically, these studies can be called bibliometric analyses–using mathematical and statistical methods to study the influence of particular authors and publications.

Relatedly, scientometric studies analyse the evolution of scientific practice over time and over domains – the subject of analysis being science itself. Such studies typically take the shape of extremely large systematic reviews (e.g., Boyack et al., 2011 covering two million biomedical publications), which drives the attention towards the structure, nature, and network of the research field overall rather than the contributions of any individual study per se (Meredith and Pilkington, 2018; van Eck et al., 2010). While previous attempts to represent the PSM research field have thus relied on manual coding and top-down classification approaches, such methods can now be complemented by scientometric analysis such as semi-automated text-mining approaches using natural language processing. The aim of text-mining is to identify the nature of PSM research by analyzing the corpus of text of scientific publications in a typically exploratory fashion. These methods are urgently required to help deal with the large scale and variety of publications that are associated with our topical domain. Furthermore, semi-automated methods can reduce biases in the search, appraisal, synthesis, and analysis (Grant and Booth, 2009) of the underlying papers, even though there is a continued role for the researcher in classification processes augmented by natural language processing as we will explore.

Our aim in this Notes and Debates article is to illustrate the use and usefulness of text similarity analysis using natural language processing, by applying it to the body of research in PSM. In partial contrast and partial complement to existing reviews, this methodology provides a bottom-up classification of the PSM phenomena in clusters, rather than top-down, and can be used to investigate the strength of content linkages between clusters of research output (i.e., flows of knowledge; Wynstra et al., 2019). These approaches therefore help to build scholarly understanding of *what* our field studies (Chicksand et al., 2012; Larson and

* Corresponding author.
E-mail address: r.suurmond@maastrichtuniversity.nl (R. Suurmond).

https://doi.org/10.1016/j.pursup.2023.100892
Received 12 January 2023; Received in revised form 10 November 2023; Accepted 29 November 2023
Available online 12 December 2023

Halldorsson, 2002) and *how* we can strengthen our field further by identifying varying connections in a growing and heterogeneous research landscape.

By using (largely) the same corpus of PSM research as a previous review that used manual coding (Wynstra et al., 2019), we can compare and contrast the two approaches and their outcomes. This corpus consists of 3477 journal publications from the period 1995–2019. Like Wynstra et al. (2019), we see PSM research as a multidisciplinary research field. Therefore, we also analyse the disciplinary backgrounds of groups of PSM publications in the current mapping of the landscape of PSM research, to illustrate the differences with the process and results of the manual coding approach.

2. Techniques for large-scale literature reviews

We can summarize the different approaches for large-scale bibliometric and scientometric reviews as in Table 1. In the top row, various approaches are listed that leverage citation data. Applications within our research field of this type of approach include Carter et al. (2007) and Hult and Chabowski (2008). The second row includes methods to leverage the characteristics of structured text. Structured text refers to textual data that is organized and formatted in a consistent and predictable manner, making it amenable to automated processing and analysis. Unlike unstructured text, which lacks a defined format and may include free-form language, structured text follows a predefined pattern or layout that facilitates easier extraction of specific information and meaningful insights. For journal articles, structured text elements typically include author names, affiliations and keywords; e.g., MEDLINE (MeSH) or JEL standardized keywords. Examples of structured text analysis within the wider Operations and Supply Chain Management (OSCM) discipline include a series of examinations by Behara and colleagues that have looked in particular at authors and institutions in various subfields and geographies (Babbar et al., 2019, 2020; Behara et al., 2014; Koufteros et al., 2021).

Table 1
Analysis techniques for bibliometric and scientometric reviews.

	Analysis Type	Input Data	Output	Examples (Methods/ Tools)
Citation Analysis	Citations	Documents, Backward Citations, Forward Citations	Citation Networks: - Cite each other - Share references - Cited by the same papers	VOSviewer (Van Eck and Waltman, 2011) CiteSpace (Chen, 2006) Connected Papers (Tarnavsky et al., 2023) Inciteful (Weishuhn, 2023) Litmaps
	Bibliographic Coupling	Documents, Backward Citations,		
	Co-citation	Documents, Forward Citations		
Structured Text Analysis	Collaboration	Authors, Institutions	Author/ Institution Collaboration Networks	VOSviewer (van Eck and Waltman, 2011)
	Keywords	Author Keywords	Keyword Networks	Bibliometrix (Aria and Cuccurullo, 2017) Kenelyze
Unstructured Text Analysis	Topic Modelling	Titles, Abstracts, or Full Texts	Topic Lists & Probability Distributions	Latent Dirichlet Allocation (Blei et al., 2003) BERTopic (Grootendorst, 2022) Top2Vec (Angelov, 2020)
	Term Co-Occurrence		Term Co-Occurrence Networks	VOSviewer (van Eck and Waltman, 2011) Kenelyze
	Text Similarity		Text Similarity Networks	Netbase Quid Kenelyze

As illustrated on the bottom row, analysis of unstructured text allows research to explore titles, abstracts, or even the full text of scientific papers using natural language processing. These methods focus more directly on the content of the papers than citation or collaboration analyses. In topic modelling, the analysis is conducted as many-to-one: many topics can be represented by one paper. Topic modelling has—to the best of our knowledge, only once been applied in context of PSM using supplier documents (Bodendorf et al., 2021), but there are also applications of topic modelling to study literature fields: Retail Operations (Rooderkerk et al., 2022), Healthcare Operations (Ali and Kannan, 2022), and Public Procurement (Rejeb et al., 2023).

In contrast to topic modelling, term co-occurrence and text similarity analysis rely on one-to-one matching of each document to one cluster, similar to cluster analysis in empirical research. Within our field, analyses of this type have been conducted on specific 'practical' data sets (Bodendorf et al., 2022), primarily leveraging open data in public procurement (El Haddadi et al., 2021; Torres-Berru et al., 2023; Zhu et al., 2023). However, in contrast to topic modelling, text mining has not been leveraged in the context of a literature review in PSM. The next section explains this text similarity analysis technique in more detail by describing how it was applied in our study.

3. Methods

We describe our approach to perform cluster analysis based on text similarity here. Along the way, important vantage points are discussed in detail, to both transparently communicate about the underlying methods here, as well as to inform future scholars on the intricacies of the methodology. Our approach uses the following process steps: 1) define a corpus of documents; 2) extract and pre-process textual information; 3) compute pairwise document-document cosine similarity using term frequency-inverse document frequency; 4) perform cluster analysis and; 5) label each cluster.

3.1. Study corpus

We set out to collect a corpus of academic publications on the topics of Purchasing and Supply Management. Starting from the full dataset of Wynstra et al. (2019) for the period of 1994–2014, we additionally collected publications for the years 2015–2019. Using the same keywords[1] and bibliometric database (Scopus) to search in the same set of journals (see Table 2), we expanded the database to include more recent publications. After manually filtering false positives based on inspection of the titles and abstracts—as for the original dataset (Wynstra et al., 2019), the database now contains 3477 publications from 18 journals in the field of management. This includes 955 recent articles that were not included in the topic analysis of Wynstra et al. (2019).[2] This addition should not be seen as a substantive contribution; our intention here is not to make a comparison between publications from the 1994–2014

[1] Keywords: Purchas*, Buy*, Suppl*, Sourc*, and Contract*. See footnote 4, Wynstra et al. (2019).

[2] In an appendix, Wynstra et al. (2019) reports simple descriptives on 612 articles published in 2015–2017 but do not present any content analysis nor show what topics those papers cover.

Table 2
Journals per disciplinary group.

Operations Management (OM)	Marketing (MA)
IEEE Transactions on Engineering Management (IEEE-TEM) International Journal of Operations & Production Management (IJOPM) International Journal of Production Economics (IJPE) Journal of Operations Management (JOM) Production and Operations Management (POM)	Industrial Marketing Management (IMM) Journal of Marketing (JM) Journal of Marketing Research (JMR) Journal of the Academy of Marketing Science (JAMS) Marketing Science (MS)
Strategy & Organization (SO)	**Purchasing and Supply Management (PSM)**
Academy of Management Journal (AMJ) Academy of Management Review (AMR) Administrative Science Quarterly (ASQ) Journal of Management (JoM) Organization Science (OS) Strategic Management Journal (SMJ)	Journal of Purchasing and Supply Management (JPSM) Journal of Supply Chain Management (JSCM)

period and those from the 2015–2019 lustrum. Yet, adding this most recent lustrum provides a more comprehensive view of the entire scope of PSM research.

3.2. (Pre-)processing textual information

The text in abstracts of the publications was pre-processed using common techniques in text analysis: tokenisation,[3] stopword filtering, part-of-speech tagging (Brill, 1995), and lemmatization of nouns, verbs, and adjectives. We also eliminated nouns related exclusively to methodology (e.g., *systematic literature review* and *case study*). This process ensures that the noun phrases are as comparable as possible, in order to later compute the similarity amongst noun phrases in each pair of articles.

3.3. Computing textual similarity

Processed abstracts were converted into document vectors using a TF-IDF (Term Frequency-Inverse Document Frequency) BM25 (Best Matching 25) approach (Sparck Jones et al., 2000; Whissell and Clarke, 2011). In particular, tokens in each article abstract were scored based on the frequency of that token in that abstract—Term Frequency, and punished for how common such a token is across all abstracts in the full dataset—Inverse Document Frequency. For example, *supply* is a common token in abstracts in this PSM database, and hence receives relatively high Term Frequency scores in many documents, when it appears multiple times in a single abstract. At the same time, because the token appears so often across all abstracts, it receives a very low Inverse Document Frequency score—close to 0 if it appears in almost all abstracts. The final multiplication of Term Frequency and Inverse Document Frequency for this token then leads to a relatively low score, reflecting the generic nature of this token in this dataset. Hence, 'supply' is not a keyword that accurately discriminates between the articles. The BM25 adaptation of TF-IDF controls for saturation of particular tokens to ensure that extreme values do not affect the results. We then construct a vector for each abstract, which contains the TF-IDF scores for all the tokens in the abstract. See Seitz (2020) for details on TF-IDF.

A link between any two abstracts is said to occur when they have a significant combined overlap in tokens between two abstracts: i.e., when there is a substantial similarity between the vectors with token scores. We therefore calculated the pairwise cosine similarities between TF-IDF-based document vectors and established a link between any two documents when the cosine similarity score was equal to or higher than 0.15 (Erkan and Radev, 2004), which is also a default threshold for large datasets. We chose this specific threshold because it led to a network with the clearest clustering compared to other thresholds which were tested iteratively. For example, a threshold of 0.10 led to significantly more links and less clear clustering. Setting this threshold requires manual intervention, interpretation, and transparent reporting.

3.4. Constructing topic clusters

We employ an approach based on a combination of natural language processing and network analysis, using text analysis of abstracts to detect clusters of closely related publications (Boyack et al., 2011). As a form of content analysis, our text-based network analysis develops categories inductively from the material under investigation and is therefore less vulnerable to authors' own biases in assessing the field. Rather than topic modelling, which focuses on deriving a predetermined number of topics and associated keywords from a given set of documents (e.g., Rooderkerk et al., 2022), our approach entails the creation of a network with links between documents, where each document belongs to exactly one cluster.

We established the network of publications using the 'Leiden community detection algorithm' (Traag et al., 2019). The quality function behind this optimization is a Constant Potts Model (Traag et al., 2011, 2013, 2019) that maximizes the number of links within each cluster, while keeping clusters relatively small. This iterative method starts with each node in its own cluster. It then assigns—'moves'—nodes to clusters with which it shares similarity. The algorithm continues to operate on the aggregated network level to refine clusters by assigning nodes to different clusters, or aggregating clusters, until no further improvements can be made. This means that no communities can be further merged or split to reach a higher quality solution and that no nodes can be moved to improve the solution. See Traag et al. (2019, Fig. 3) for details.

These clusters each represent a coherent research domain at its core, while articles on the periphery of a cluster are more marginally related to a cluster and are in addition related to papers in other clusters as well—hence, pulled outside the centre of a cluster. The clusters therefore cannot be considered to provide hard boundaries between one article and the next, since a particular paper on the periphery may be (almost) as related (in terms of cosine similarity) to its cluster neighbours as it is to its non-cluster neighbours. Yet, those articles in turn will then be more related to the high-scoring noun phrases of another cluster.

After establishing the clusters, we conducted network analysis to create visualizations. To create network maps, nodes were positioned using the ForceAtlas2 algorithm (Jacomy et al., 2014) and adapted to only picture the links with 10% highest pairwise cosine similarity between nodes to provide a better readable layout with clearly delineated clusters. Nodes that are connected are pulled together (gravity), whereas nodes that are unconnected are pushed apart (repulsion), in an iterative process until a good layout emerges. The position of the nodes on the map and the distance between any two nodes are all relative, and dependent on the specific body of publications used for this analysis.[4]

3.5. Labelling topic clusters

Each identified cluster was labelled with a unique identifier based on high-scoring noun phrases (e.g., supplier selection) across all abstracts

[3] Tokenisation involves separating a piece of text into smaller units (sequences of characters) called tokens, which are grouped as semantic unit for processing. Tokens can be either words, characters, or subwords.

[4] For instance, consider a set of publications where 2 and 3 have strong links with 1 and are thus pulled together. If 1 is removed, 2 and 3 could be pulled into opposing directions on the map by 4 and 5 respectively.

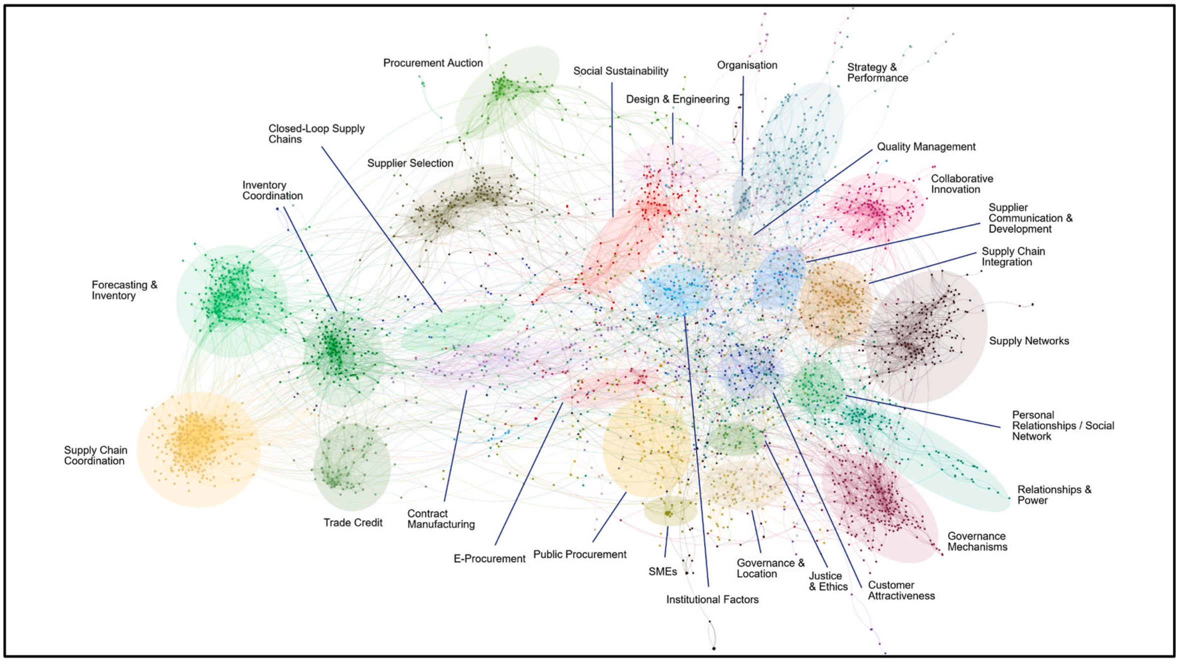

Fig. 1. Network of PSM publications with 35 clusters.

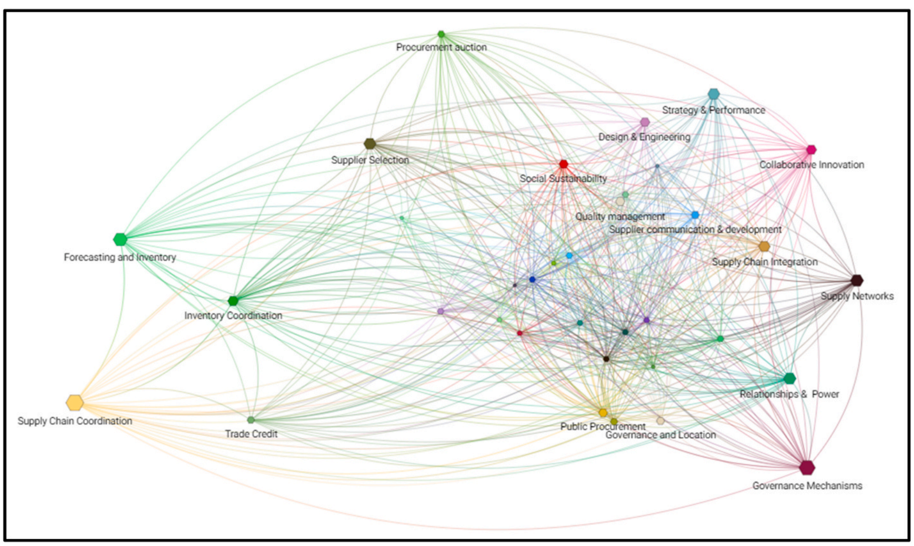

Fig. 2. Network of PSM research with linkages between clusters.

within that cluster, as well as the authors' substantive and expert knowledge of the field. Specifically, some clusters were relatively straightforward to label based on both the common noun phrases and the commonality between the papers, e.g., a cluster of papers focused on *Governance Mechanisms*, specifically contractual and relational governance and their sub-dimensions. In other cases, the authors had to dive deeper into the titles and abstracts of all articles in a cluster to identify the proper cluster label. For example, the cluster on *Institutional Factors* includes keywords such as institutional context and institutional theory. Finally, some clusters still include a mix of papers, in which case a more generic and abstract keyword was used, such as in the last cluster of *Contracts*, which includes papers on property rights, bargaining power, and optimal contracts, as well as others. This requires specific expert knowledge to make the cluster labels as insightful as possible, in this case, that of the authors. Note that the (post-hoc) naming of the clusters does not affect their computation, in contrast to for instance topic modelling.

3.6. Term maps

The results of these analyses include term maps of the academic field of Purchasing and Supply Management from 1995 to 2019. The network analyses and visualizations are produced using Kenelyze (www.kenelyze.com).

All figures in this paper are available in full colour online via the Publisher's website and on Open Science Framework, which furthermore includes an interactive network map with article-level information (Suurmond et al., 2023; www.osf.io/WT8F3).

4. Results

The descriptive statistics are provided in Table 3. A total of 3477

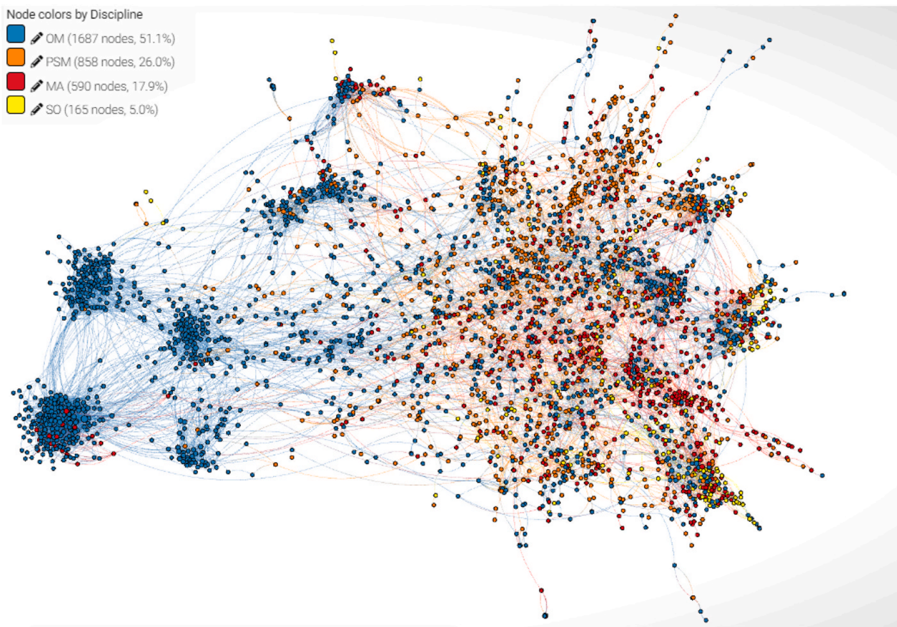

Fig. 3. Nodes clustered by disciplinary origin.

Table 3
Descriptive statistics of PSM research.

	1995–1999	2000–2004	2005–2009	2010–2014	2015–2019	1995–2019
Number of nodes	334	478	680	1030	955	**3477**
Number of nodes in the LCC	308	451	649	982	910	**3300**
Share of nodes part of the LCC	92,22%	94,35%	95,44%	95,34%	95,29%	**94,91%**

articles (nodes) are included in our main analysis, of which 3300 (95%) are represented in the main network—labelled as the largest connected community (LCC). The remaining five percent of the papers are not connected to any of the other papers because of high levels of text dissimilarity (cosine similarity <0.15). Within this main network, the Leiden community detection algorithm has determined 35 clusters of publications. Our analysis enables us to explore three aspects of the PSM research landscape: the clustering of nodes, the network of clusters, and further subgrouping analysis on the disciplinary background of nodes and clusters.

4.1. Clustering of nodes

For the entire set of publications (1995–2019), the network of clusters is visualized in Fig. 1. Table 4 provides additional information on the clusters, their 'top' (most prevalent) disciplinary backgrounds, 'top' journals, prolific authors, and highest scoring noun phrases. Each node represents an article and the node colour represents the cluster. Relationships between nodes are sized by the weight of the similarity score and if nodes are very similar, they tend to be co-located on the map. For ease of interpretation, each cluster is marked (manually) with a background in the form of a similarly-coloured ellipse. Eight clusters are not marked in this way, because they are too stretched out over the network space and less populated. These clusters still exist, but cannot accurately be plotted onto a two-dimensional plane. As similarity between two abstracts is based on the vectors of all keywords, some nodes are pulled into different directions by their commonality with more than one other node, for instance in terms of content and methodology.

The largest cluster, with 222 articles associated with it, is labelled as *Supply Chain Coordination*, which has gained particular prominence in our data set in the years after 2009 and the majority of which is published in IJPE and then other operations management journals. In contrast, the second largest cluster is *Governance Mechanisms*, which covers a much broader area of the field with diverse publication outlets, including SMJ, IMM, and JOM.

Compared to other review methodologies, a number of features of our analysis stand out. First, manual, top-down approaches to classify and cluster academic research depend on a limited number of dimensions for topic classification (e.g., one dimension in Spina et al., 2013, four dimensions in Wynstra et al., 2019). Our analysis instead leads to clusters based on many dimensions. For instance, the cluster of *Buyer-Supplier Relationships*, which is spread out over the network space and not encircled in Fig. 1, connects topically, phenomenologically, and epistemologically to various other clusters. Similarly, the clusters on the left of the space, related to *Supply Chain Coordination*, Forecasting, and Inventory, share also methodological linkages based on Operations Research analysis.

Second, boundaries of our network space are to some extent artificial. Our analysis provides insights into our specific field, yet we can easily identify that at least some clusters are in turn related to, and perhaps represented in, networks of other domains of study. For example, the cluster on Procurement Auction (e.g., Jap, 2002; Wagner and Schwab, 2004) relates phenomenologically to our field (e.g., Supplier Selection), but relates mathematically to other domains as well. We hence observe that several clusters, specifically those towards the periphery of the network space, might need to be separately considered in relation to their other connections (outside our scope) as well.

4.2. Network of clusters

To further analyse the PSM research landscape in terms of clusters, rather than individual papers, Fig. 2 displays the clusters as nodes, with node size matching article count. The weight (thickness) of the link between any two clusters is determined based on the cumulative cosine

Table 4
Clusters in PSM research.

Cluster	No.	HHI	Top Discipline	Top Journal	Top Authors	Noun Phrases
Supply Chain Coordination	222	86%	OM	IJPE	Chen J. (8); Wang Y. (6); Gan X. (4); Sethi S.P. (4); Yan H. (4)	wholesale price/retail price/order quantity/channel coordination/wholesale price contract
Governance Mechanisms	204	27%	OM	IMM	Heide J.B. (8); Wathne K.H. (6); Lumineau F. (6); Zhao X. (5); Ghosh M. (4)	transaction cost economics/relational governance/governance mechanism/opportunistic behavior/asset specificity
Forecasting and Inventory	173	98%	OM	IJPE	Cheng T.C.E. (6); Kouvelis P. (5); Dolgui A. (4); Sethi S.P. (4); Zhang J. (4)	optimal policy/demand uncertainty/demand forecast/option contract/selling season
Supply Networks	158	30%	OM	IMM	Lawson B. (6); Choi T.Y. (5); Cousins P.D. (5); Revilla E. (5); Handfield R.B. (4)	social capital/absorptive capacity/relational capital/innovation performance/organizational learning
Relationships & Power	150	47%	MA	IMM	Katsikeas C.S. (5); Skarmeas D.A. (3); Van der Valk W. (3); Johnston W.J. (3);	buyer-seller relationship/coercive power/service-dominant logic/relationship quality/interpersonal relationship
Supplier Selection	148	62%	OM	IJPE	Ghodsypour S.H. (7); O'Brien C. (7); Talluri S. (7); Sarkis J. (5); Ho W. (4)	analytic hierarchy process/total cost/decision maker/multiple criterion/order allocation
Strategy & Performance	144	49%	PSM	JPSM	Handfield R.B. (5); Gelderman C.J. (5); Van Weele A.J. (5); Ellram L.M. (4)	competitive advantage/strategic role/business performance/global marketplace/other function
Supply Chain Integration	138	44%	OM	IJPE	Vickery S.K. (5); Hartley J.L. (3); Siguaw J.A. (3); Simpson P.M. (3);	operational performance/financial performance/positive effect/customer integration/external integration
Inventory Coordination	126	91%	OM	IJPE	Kelle P. (4); Bylka S. (4); Glock C.H. (4); Miller P.A. (3); Disney S.M. (3)	vendor-managed inventory/total cost/inventory policy/inventory level/order quantity
Collaborative Innovation	122	34%	OM	JPSM; JSCM	Wynstra F. (8); Van Weele A.J. (6); Ellram L.M. (5); Hartley J.L. (5);	new product development/product development/npd project/concurrent engineering/new product
Design & Engineering	111	35%	PSM	JPSM	Giunipero L.C. (3); Roehrich J.K. (3); Newnes L.B. (3); Parry G. (3);	current practice/value creation/different type/product life cycle/recent year
Quality management	109	32%	OM	JPSM	Wisner J.D. (5); Dyer J.H. (4); Stanley L. L. (4); Cousins P.D. (3)	service quality/internal service quality/product quality/make-or-buy decision/vehicle manufacturer
Social Sustainability	106	44%	OM	IJPE	Foerstl K. (7); Reuter C. (5); Hartmann E. (4); Blome C. (4)	corporate social responsibility/social sustainability/social performance/csr activity/triple bottom line
Public Procurement	104	35%	PSM	JPSM	Choi T.Y. (5); Telgen J. (4); Wu Z. (4); Harland C. (4); Virolainen V.-M. (3)	service triad/public sector/public procurement/dyadic relationship/managerial implication
Supplier communication & development	93	36%	OM	JSCM	Carr A.S. (5); Fynes B. (4); Paulraj A. (4); Krause D.R. (3); Smeltzer L.R. (3)	electronics sector/chain management/critical factor/information technology/performance outcome
Governance and Location	91	31%	PSM	JPSM; JSCM	McIvor R. (4); Ellram L.M. (3); Carter C.R. (3); Tate W.L. (3);	transaction cost economics/resource-based view/governance mode/managerial implication/make-buy decision
Trade Credit	82	93%	OM	IJPE	Teng J.-T. (6); Cárdenas-Barrón L.E. (5); Chang C.-T. (4); Chung K.-J. (4)	trade credit/permissible delay/economic order quantity/theoretical result/optimal solution
Small and Medium-sized enterprises/SMEs	81	35%	OM	IJPE; JPSM	Bocconcelli R. (3); Lehtinen U. (2); Ellram L.M. (2); Esposito E. (2)	medium-sized enterprise/medium enterprise/edi adoption/information technology/industrial district
Procurement auction	78	35%	OM	POM	Jap S.D. (6); Carter C.R. (4); Schoenherr T. (4); Haruvy E. (4)	electronic reverse auction/online reverse auction/procurement auction/reverse auction/total cost
Personal relationships/ social network	72	36%	OM	IMM	Cheng T.C.E. (6); Yeung A.C.L. (5); Bao Y. (5); Humphreys P. (4);	boundary spanner/u-shaped relationship/hong kong/bullwhip effect/positive effect
Contract Manufacturing	69	53%	OM	POM	Nellore R. (4); Niu B. (4); Huang G.Q. (3); Zhang F. (3); Howard M. (2)	original equipment manufacturer/contract manufacturer/perfect equilibrium/modular assembly/component procurement
Institutional factors	69	41%	OM	JPSM	Lai K.-H. (4); Sarkis J. (3); Min H. (2); Galle W.P. (2); Cox A. (2)	gscm practice/institutional theory/environmental performance/institutional pressure/institutional environment
Buyer-Supplier Relationships	67	37%	MA	IMM	Kaufmann L. (6); Walter A. (3); Ritter T. (3); Kleinaltenkamp M. (3);	buyer-seller relationship/business relationship/cross-functional sourcing team/customer retention/sourcing team
Customer Attractiveness	65	33%	PSM	IMM	Kaufmann L. (4); Carter C.R. (4); Ellegaard C. (3); Thomas R.W. (3);	social exchange theory/preferred customer/customer attractiveness/preferred customer status/other party
Interdependence & Power	63	31%	MA	JPSM	Kumar N. (4); Scheer L.K. (4); Steenkamp J.-B.E.M. (3); Ramsay J. (3);	dependence asymmetry/interdependence asymmetry/automobile dealer/power position/complex adaptive system
Performance-based contracting	62	39%	OM	IMM	Van der Valk W. (4); Caldwell N.D. (3); Essig M. (3); Sumo R. (3);	performance-based contracting/performance-based contract/enterprise resource planning/erp system/horizontal collaboration
Global sourcing	61	33%	PSM	JPSM; IMM	Jia F. (4); Sartor M. (4); Orzes G. (4); Nassimbeni G. (4); Romano P. (3)	global sourcing/financial performance/apparel company/global sourcing context/multinational corporation
Management consulting	60	39%	OM	JPSM	González-Benito J. (4); Nollet J. (3); Germain R. (2); Droge C. (2);	consulting service/cultural difference/management consulting service/quality management practice/acquisition process
E-procurement	55	39%	OM	IJPE	Croom S.R. (3); Sundarraj R.P. (3); Min H. (2); Galle W.P. (2);	electronic commerce/e-procurement adoption/mro item/new technology/e-procurement system
Psychological contract	50	46%	OM	IJPE	Weber L. (3); Eckerd S. (3); Shao X.-F. (2); Sheu J.-B. (2); Mir S. (2)	reactive strategy/psychological contract breach/contract frame/psychological contract/order allocation
Lean	45	37%	PSM	JPSM; JSCM	Smeltzer L.R. (3); Carter J.R. (3); Narasimhan R. (3); Cooper R.W. (2);	total quality management/job satisfaction/network picture/various factor/such practice
Justice & Ethics	33	35%	OM	IMM	Luo Y. (3); Liu Y. (3); Huang Y. (3); Carter C.R. (2); Griffith D.A. (2)	distributive fairness/unethical behavior/distributive justice/procedural justice/interactional justice
Closed-loop supply chains	33	47%	OM	IJPE	Carter C.R. (3); Hendrick T.E. (2); Wee H.-M. (2); Wang X. (2); Liang L. (2)	raw material/acquisition cost/local authority/high uncertainty/reverse logistics

(*continued on next page*)

Table 4 (*continued*)

Cluster	No.	HHI	Top Discipline	Top Journal	Top Authors	Noun Phrases
Organisation	30	66%	PSM	JSCM	Leenders M.R. (11); Johnson P.F. (6); Fraser Johnson P. (5); Fearon H.E. (4);	major corporate activity/organizational structure/ organizational change/major change/free trade agreement
Contracts	26	61%	OM	IJPE	Lippman S.A. (3); Tang C.S. (3); Dharma Kwon H. (2); Perrone G. (2);	property right/governance form/optimal contract/ bargaining power/private information

Note. No.: number of publications. HHI: Herfindahl-Hirschman Index (HHI) of disciplinary concentration. Top Discipline (Journal): discipline (journal) with highest share of publications. Authors: authors with the largest share of publications (co-authorships are counted equally).

similarity between articles belonging to the two clusters. This analysis reveals some further interesting findings. First, even though clusters appear to occupy different corners of the graph, there are many interconnections between almost all the clusters. Second, some clusters appear to draw more heavily on the network than others. For example, the clusters of *Governance Mechanisms* and *Supply Networks* have tighter and stronger connections with other clusters.

Second, our analysis allows us to visually observe how closely related topic clusters are. Clusters which are co-located share higher degrees of text similarity than those farther apart. Previous manual reviews typically stop at analysing these clusters independently, while this text similarity analysis and specifically the network visualization reveals the intricate relations between nodes, both within and across clusters. For example, papers in the cluster of *Social Sustainability* (near the top) relate to both *Institutional Factors* (to its right) and *Supplier Selection* (to its left).

A close inspection of the landscape reveals which topics are closely placed and closely linked together. For example, it is well established that *Collaborative Innovation* can be studied in relation to *Supply Chain Integration* and *Supplier Communication & Development*—yet, our analysis reveals further important linkages with *Social Sustainability* and *Institutional Factors*. It thereby uncovers latent connections that might be pursued for further grounding of any piece of research in the PSM field.

4.3. Subgroup analysis of disciplinary origin

Many further analyses can be conducted on the basis of this methodology. For example, it is possible to combine top-down classifications with bottom-up text similarity analysis. Here, we illustrate this approach by comparing the disciplinary origin of nodes and clusters – based on the distinct journal groups as identified in Table 2 and following the manual analyses by Wynstra et al. (2019).

As a measure for the concentration of the cluster publications across the four disciplines—as in Table 2, we calculate a Herfindahl-Hirschman Index (HHI), see Table 4.[5] The higher the index, the more concentrated the cluster publications are within one or several particular disciplines (or, more precisely: within the set of journals associated with that discipline). The most concentrated ones are *Forecasting and Inventory*, *Inventory Coordination*, and *Trade Credit*. Together with nine other relatively concentrated clusters, these three are dominated by publications in Operations Management with just a few contributions each from other disciplines. Thus, there appears to be a separate branch of PSM research belonging to the Operations Management discipline. This is further visualized in Fig. 3, where each node is given the colour of the disciplinary origin of the publication rather than of the cluster. In this figure, we can observe a community of clusters in dark blue on the left side of the plot, in which contributions are relatively more concentrated in the OM discipline than in the rest of the PSM research landscape. In contrast, clusters with an HHI below 0.39 (the median value) are relatively spread out over different disciplines, and therefore these clusters in particular could involve multidisciplinary or even transdisciplinary research (Choi and Pak, 2006). Some topics are clearly grounded in more multidisciplinary perspectives, for example *Governance Mechanisms*, *Supply Networks*, and *Collaborative Innovation*. This also includes, amongst others, *Buyer-Supplier Relationships*, and *Public Procurement*, both of which are multidisciplinary in terms of breadth of concepts, methodologies, and applied theories. Hence, our investigation provides further evidence that PSM research displays characteristics of multidisciplinarity, with some clusters developing as separate strands of investigations with its own set of authors and journals to publish in (Chicksand et al., 2012; Harland et al., 2006; Zsidisin et al., 2019).

The manual analysis in Wynstra et al. (2019) highlights the most popular topics for each respective discipline over time (pp. 7–9). A recalculation of the manually coded data leads to the index for concentration (HHI) per topic as provided in Appendix B – Table A1. From the results, we can observe a similar distribution of topics over disciplines, albeit with a number of important differences. First, we still observe a stronger level of multidisciplinarity (lower concentration overall) of the field than the manual analysis has revealed (median = 0.39(a) vs 0.46(m)). This is surprising because automated text analysis can be considered to be more sensitive to semantic divides between disciplines, and could therefore result in more mono-disciplinary clusters. Second, the automated text analysis produces more extreme concentration values (high and low); it produces a small number of topic clusters that are extremely concentrated and thus appear highly mono-disciplinary (e.g., *Forecasting and Inventory*).

4.4. Further analyses

Sub-analyses were performed for each five-year period (lustrum) within the timeframe to investigate the evolution of the field over time Figures B1–5, in Appendix B, display the evolution of this network of PSM research over time, per five-year period (1995–1999; 1995–2004; 1995–2009; 1995–2014; 2015–2019, respectively). Note that we describe these consecutive networks by looking back through the 'lens' of the overall 1995–2019 network. In other words, no new network of clusters are constructed for each lustrum, but the existing clusters are visualized as they appeared.

Some clusters appear with a multitude of nodes from an early date, for example *Governance Mechanisms*, *Quality Management*, and *Strategy & Performance*. Others, however, mainly emerged later, for example *Procurement Auction*, *Supply Chain Coordination* (only 3 publications in first lustrum) and *Social Sustainability*. Wynstra et al. (2019) observes that over time, an increasing share of publications is focusing on tactical and operational PSM processes. This observation is consistent with what we find through the current text-similarity analysis, but the latter produces more fine-grained results in terms of the topics covered and their evolution. For example, the relatively quick evolution post-2005 (Appendix B) of the clusters Supply Chain Coordination, Forecasting & Inventory, and Inventory Coordination, shows a comparable increase in the number of publications with such topical focus.

It is also interesting to contrast this bottom-up analysis with perceptions of the evolution of topics. For instance, we increasingly understand (environmental) sustainability in terms of circular and regenerative supply chains, but our analysis reveals that there were already studies in the cluster on *Closed Loop Supply Chains* – as this was

[5] The Herfindahl–Hirschman Index is a commonly accepted measure of market concentration and is calculated by squaring the 'market' share of each 'actor' competing in the market and then summing the resulting numbers. Here, it is computed as the disciplinary share in each cluster.

once called – in the period 1994–1999.

5. Conclusion and recommendations

This article provides an exploratory scientometric analysis of PSM research, in order to illustrate the use and usefulness of text similarity analysis for our field. Based on the network of 3477 PSM publications, the cluster maps provide interesting and fresh perspectives on the structure and nature of the research field, as illustrated in the previous section. In this final section, we provide a number of considerations in using this type of semi-automated text analysis and network analytics in literature reviews. Clearly, these tools can also be applied in PSM research to other types of documents, but a review of such applications is outside the scope of this article.

It may be obvious, but it is important to note that text similarity and network analysis techniques are tools to be used once the corpus of literature has been defined. Thus, in terms of the typical literature review steps – Search, Appraisal, Synthesis and Analysis (Grant and Booth, 2009) – it is a tool that can primarily assist in synthesis and analysis. As a primarily exploratory technique, it seems logical to use text similarity and network analysis at the start of the review process. However, in some situations, as in the case of the corpus of PSM publications used here, the manual review and coding may precede the text similarity analysis, with the latter being used as a follow-up to create alternative perspectives of the same 'landscape'. Still, the outcome of a text similarity and network analysis may help to identify a particular subset of literature (i.e., one or more clusters) as a starting point for a more targeted review, and it that way also serve as a tool for 'search'. For example, a researcher interested in research on 'Supplier Selection' could take the 148 papers identified in that cluster as a basis for search, appraisal, synthesis, and analysis, potentially updating this subset with papers from other journals and other timeframes. While the corpus identified for this paper serves its purpose for the network of PSM research, researchers interested in more specific topics should be careful that the identified clusters depend on the composition of this initial corpus.

As we have illustrated, text similarity and network analysis techniques are helpful for reviewing literature on two different levels. First, it can identify specific clusters or groups of similar studies and second, it can identify the relations between such clusters. These features make it particularly suitable for so-called "mapping reviews"; literature reviews that map out and categorize existing literature, based on which more specific reviews or new primary research can be initiated (Grant and Booth, 2009). The usefulness of the cluster identification is not a given. Subsequent manual inspection and qualitative review of the nodes within a cluster is needed to be able to verify any underlying logic of the similarities, and to enhance the interpretability of cluster. Such a manual review – of the full publication – is especially important when the text similarity analysis is based on abstracts and not the full text of the publication. Manual review of a cluster may identify particular heterogeneities, for instance in terms of the theories being addressed or research methods being used in a study, if the text similarity analysis has resulted in a clustering based on the empirical phenomena being discussed. Which publication elements – conceptualisations, theory, methods, etcetera – will drive the text similarity analysis is not determined up front. In the illustrational application in this article, it turns out that the clusters are more based on conceptualisations or conceptualised phenomena, but this may not always be the result of the text similarity and network analysis, compared to more process-oriented approaches (Spina et al., 2013; Wynstra et al., 2019). For instance, restricted formatting of abstracts or full text (if that is being used in the analysis) of articles in a particular corpus may be emphasising research method descriptions: then this is likely to drive the cluster identification. Researchers should be aware of such potential effects when interpreting the clusters. In our illustrative application, such effects do not seem to occur as our corpus covers a wide variety of journals with different, relatively flexible abstract formats (cf. 'Structured Abstracts' in Emerald and other journals). Clearly, also the usefulness of the entire 'landscape overview' consisting of the relative sizes and positions of clusters is also not a given and is fundamentally dependent on the interpretability of the clusters themselves. These cluster maps can be particularly relevant for junior researchers, such as PhD candidates, who, otherwise, start to uncover the landscape of a new research field from a more narrow vantage point.

In concluding, the most important recommendation would be that in conducting literature reviews, text similarity and network analysis is used to augment decision and analysis, not to automate: manual review remains an essential part of the process. In this paper, we have illustrated how we dealt with several of these decisions at critical vantage points.

Our research is not without limitations. First, the text-based analysis used the abstracts as input and not the full text of each article. While abstracts typically present the most important topics and structure of an article, it is possible that a full-text analysis would reveal different links altogether. Still, abstract-based text analysis is common in other network and bibliometric studies and we are confident that sufficiently reliable insights can be derived from this approach. Second, our analysis is limited to the (extended) database of Wynstra et al.'s (2019) investigation of PSM research. This database includes the top 18 highest (citation) impact journals in each journal group (Table 2), but PSM is potentially represented more generically in other (management) journals as well. As noted before, however, our intended contribution consists of an illustration of the use and usefulness of text similarity plus network analysis, focusing on the processes and the *type* of findings they produce. This study does not pretend to present a substantive and exhaustive review of PSM research, but it has hopefully provided some useful insights in how text similarity and network analysis could be used for such reviews.

Author statement

We ensure that we have written entirely original works, and if the authors have used the work and/or words of others, that this has been appropriately cited or quoted and permission has been obtained where necessary.

Declaration of competing interest

There are no conflicts of interest.

Data availability

Data is available on https://osf.io/wt8f3/ (Suurmond et al., 2023)

Appendix A. Disciplinary origins and concentration for original set of manually coded publications

For the comparison in concentration between the original analysis in Wynstra et al. (2019) and the current paper, we needed to calculate the Herfindahl-Hirschman Index (HHI) based on the original topics as coded before. We therefore conducted further analysis on the original dataset (based on 2522 articles from 1995 to 2014 only). The Herfindahl–Hirschman Index is a commonly accepted measure of market concentration and is calculated by squaring the 'market' share of each 'actor' competing in the market and then summing the resulting numbers. Here, it is computed as the disciplinary share in each topic.

Table A1
Disciplinary origins and and Herfindahl-Hirschman Index (HHI) per topic

Strategic Processes	OM	MA	SO	PSM	HHI
Corporate & PSM Strategy	25%	9%	0%	**62%**	0.46
Make-or-Buy/Outsourcing	**42%**	10%	22%	26%	0.30
Category Sourcing Strategy	36%	17%	0%	**44%**	0.36
Global Sourcing	**41%**	16%	2%	38%	0.34
Supplier Relationship Management	28%	**34%**	8%	27%	0.28
Supplier Integration in NPD	**38%**	20%	14%	28%	0.28
Supplier Integration in Order Fulfillment	**72%**	6%	1%	16%	0.55
Supplier Development & Quality Management	**60%**	9%	3%	29%	0.45
Strategic Cost Management	38%	3%	0%	**56%**	0.46
Tactical Processes	OM	MA	SO	PSM	HHI
Specify	**81%**	0%	0%	19%	0.70
Select	**49%**	19%	2%	30%	0.37
Contract	**69%**	16%	7%	9%	0.51
Order	**92%**	2%	0%	6%	0.85
Receive	**88%**	4%	0%	8%	0.77
Evaluate	**67%**	11%	0%	22%	0.51
Pay	**86%**	7%	0%	7%	0.74
Various	44%	9%	0%	**47%**	0.42

Notes: Dominant discipline printed bold. Median HHI: 0.46.

Appendix B. Term maps for each lustrum

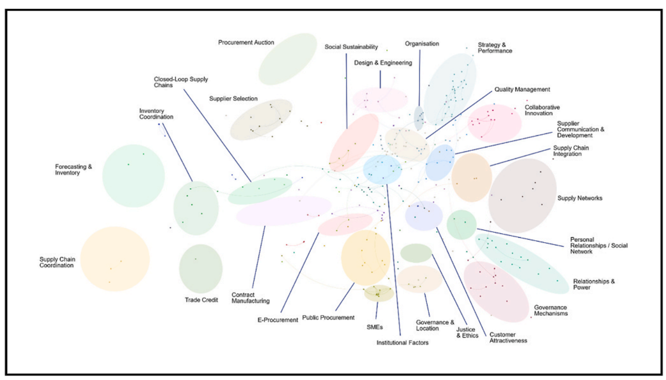

Figure B1. Term map of PSM research published from 1995 to 1999.

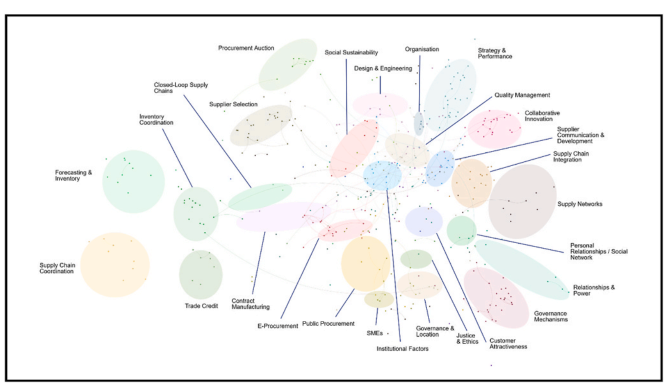

Figure B2. Term map of PSM research published from 2000 to 2004.

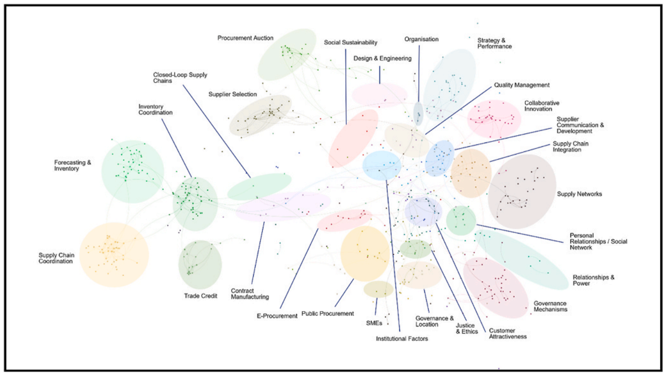

Figure B3. Term map of PSM research published from 2005 to 2009.

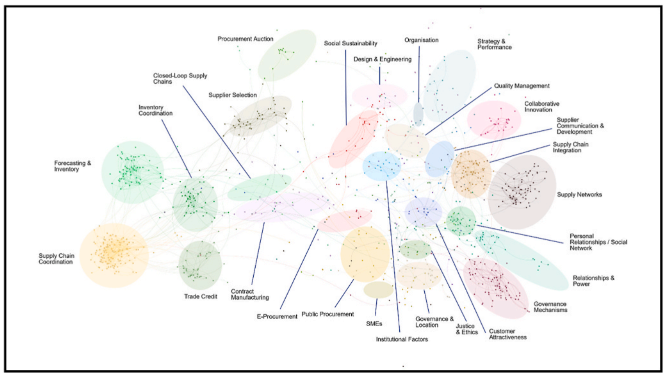

Figure B4. Term map of PSM research published from 2010 to 2014.

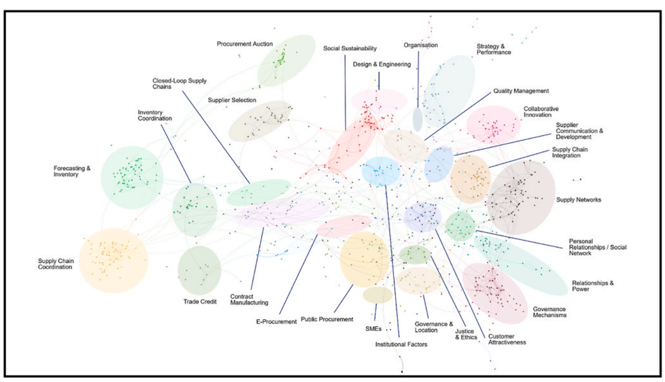

Figure B5. Term map of PSM research published from 2015 to 2019.

References

Ali, I., Kannan, D., 2022. Mapping research on healthcare operations and supply chain management: a topic modelling-based literature review. Ann. Oper. Res. 315, 29–55. https://doi.org/10.1007/s10479-022-04596-5.

Angelov, Dimo. "Top2Vec: Distributed Representations of Topics." arXiv, August 19, 2020. https://doi.org/10.48550/arXiv.2008.09470.

Aria, Massimo, and Corrado Cuccurullo. "Bibliometrix : An R-Tool for Comprehensive Science Mapping Analysis." *Journal of Informetrics* 11, no. 4 (November 2017): 959–75. https://doi.org/10.1016/j.joi.2017.08.007.

Blei, David M, Andrew Y Ng, and Michael Jordan. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (2003): 993–1022.

Babbar, S., Koufteros, X., Behara, R.S., Wong, C.W.Y., 2019. SCM research leadership: the ranked agents and their networks. SCM 24, 821–854. https://doi.org/10.1108/SCM-11-2018-0386.

Babbar, S., Koufteros, X., Bendoly, E., Behara, R., Metters, R., Boyer, K., 2020. Looking at ourselves: lessons about the operations management field learned from our top journals. J. Ops. Manag. 66, 349–364. https://doi.org/10.1002/joom.1081.

Behara, R.S., Babbar, S., Andrew Smart, P., 2014. Leadership in OM research: a social network analysis of European researchers. Int. J. Oper. Prod. Manag. 34, 1537–1563. https://doi.org/10.1108/IJOPM-08-2013-0390.

Bodendorf, F., Dimitrov, G., Franke, J., 2022. Analyzing and evaluating supplier carbon footprints in supply networks. J. Clean. Prod. 372, 133601 https://doi.org/10.1016/j.jclepro.2022.133601.

Bodendorf, F., Wytopil, B., Franke, J., 2021. Business analytics in strategic purchasing: identifying and evaluating similarities in supplier documents. Appl. Artif. Intell. 35, 857–875. https://doi.org/10.1080/08839514.2021.1936423.

Boyack, K.W., Newman, D., Duhon, R.J., Klavans, R., Patek, M., Biberstine, J.R., Schijvenaars, B., Skupin, A., Ma, N., Börner, K., 2011. Clustering more than two million biomedical publications: comparing the accuracies of nine text-based similarity approaches. PLoS One 6. https://doi.org/10.1371/journal.pone.0018029.

Brill, E., 1995. Transformation-based error-driven learning and natural language processing: a case study in part-of-speech tagging. Comput. Ling. 21, 543–565.

Carter, C.R., Ellram, L.M., Kaufmann, L., Autry, C.W., Zhao, X., Callarman, T.E., 2014. Looking back and moving forward: 50 years of the journal of supply chain management. J. Supply Chain Manag. 50, 1–7. https://doi.org/10.1111/jscm.12045.

Carter, C.R., Leuschner, R., Rogers, D.S., 2007. A social network analysis of the Journal of Supply Chain Management: knowledge generation, knowledge diffusion and thought leadership. J. Supply Chain Manag. 43, 15–28. https://doi.org/10.1111/j.1745-493X.2007.00028.x.

Chen, 2006. Chaomei. "CiteSpace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature. Journal of the American Society for Information Science and Technology 57 (3), 359–377.

Chicksand, D., Watson, G., Walker, H., Radnor, Z., Johnston, R., 2012. Theoretical perspectives in purchasing and supply chain management: an analysis of the literature. Supply Chain Manag. 17, 454–472. https://doi.org/10.1108/13598541211246611.

Choi, B.C., Pak, A.W., 2006. Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. Clin. Invest. Med. 29, 351.

El Haddadi, T., El Haddadi, O., Mourabit, T., El Allaoui, A., Ben Ahmed, M., 2021. Automatic analysis of the sustainability of public procurement based on Text Mining: the case of the Moroccan ICT markets. Clean Respons. Consum. 3, 100037 https://doi.org/10.1016/j.clrc.2021.100037.

Erkan, G., Radev, D.R., 2004. LexRank : graph-based centrality as salience in text summarization. J. Artif. Intell. Res. 22, 457–479.

Grant, M.J., Booth, A., 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Inf. Libr. J. 26, 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x.

Grootendorst, Maarten. "BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure." arXiv, March 11, 2022. https://doi.org/10.48550/arXiv.2203.05794.

Harland, C.M., Lamming, R.C., Walker, H., Phillips, W.E., Caldwell, N.D., Johnsen, T.E., Knight, L.A., Zheng, J., 2006. Supply management: is it a discipline? Int. J. Oper. Prod. Manag. 26, 730–753. https://doi.org/10.1108/01443570610672211.

Hult, G.T.M., Chabowski, B.R., 2008. Sourcing research as an intellectual network of ideas. Decis. Sci. J. 39, 323–335. https://doi.org/10.1111/j.1540-5915.2008.00194.x.

Jacomy, M., Venturini, T., Heymann, S., Bastian, M., 2014. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. PLoS One 9, e98679. https://doi.org/10.1371/journal.pone.0098679.

Jap, S.D., 2002. Online reverse auctions: issues, themes, and prospects for the future. J. Acad. Market. Sci. 30, 506–525. https://doi.org/10.1177/009207002236925.

Koufteros, X.A., Babbar, S., Behara, R.S., Baghersad, M., 2021. OM research: leading authors and institutions. Decis. Sci. J. 52, 8–77. https://doi.org/10.1111/deci.12452.

Larson, P.D., Halldorsson, A., 2002. What is SCM? And, where is it? J. Supply Chain Manag. 38, 36–44. https://doi.org/10.1111/j.1745-493X.2002.tb00141.x.

Meredith, J.R., Pilkington, A., 2018. Assessing the exchange of knowledge between operations management and other fields: some challenges and opportunities. J. Oper. Manag. 60, 47–53. https://doi.org/10.1016/j.jom.2018.05.004.

Rejeb, A., Rejeb, K., Appolloni, A., Kayikci, Y., Iranmanesh, M., 2023. The landscape of public procurement research: a bibliometric analysis and topic modelling based on Scopus. J. Public Procure. 23, 145–178. https://doi.org/10.1108/JOPP-06-2022-0031.

Rooderkerk, R.P., DeHoratius, N., Musalem, A., 2022. The past, present, and future of retail analytics: insights from a survey of academic research and interviews with practitioners. Prod. Oper. Manag. 31, 3727–3748. https://doi.org/10.1111/poms.13811.

Seitz, R., 2020. Understanding TF-IDF and BM-25 [WWW Document]. URL. https://kmwllc.com/index.php/2020/03/20/understanding-tf-idf-and-bm-25/. accessed 9.3.23.

Tarnavsky, Eitan, Eddie Smolyansky, Itay Knaan Harpaz, and Sahar Perets. "Connected Papers," 2023. https://www.connectedpapers.com/.

Sparck Jones, K., Walker, S., Robertson, S.E., 2000. Probabilistic model of information retrieval: development and comparative experiments. Part 2. Inf. Process. Manag. 36, 809–840. https://doi.org/10.1016/S0306-4573(00)00016-9.

Spina, G., Caniato, F., Luzzini, D., Ronchi, S., 2013. Past, present and future trends of purchasing and supply management: an extensive literature review. Ind. Market. Manag. 42, 1202–1212. https://doi.org/10.1016/j.indmarman.2013.04.001.

Suurmond, R., Wynstra, F., Vermeij, A., Haag, E.J., 2023. PSM Research. https://doi.org/10.17605/OSF.IO/WT8F3.

Torres-Berru, Y., Lopez-Batista, V.F., Zhingre, L.C., 2023. A data mining approach to detecting bias and favoritism in public procurement. Intell. Autom. Soft Comput. 36, 3501–3516. https://doi.org/10.32604/iasc.2023.035367.

Traag, V.A., Krings, G., Van Dooren, P., 2013. Significant scales in community structure. Sci. Rep. 3, 2930. https://doi.org/10.1038/srep02930.

Traag, V.A., Van Dooren, P., Nesterov, Y., 2011. Narrow scope for resolution-limit-free community detection. Phys. Rev. E 84, 016114. https://doi.org/10.1103/PhysRevE.84.016114.

Traag, V.A., Waltman, L., van Eck, N.J., 2019. From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. 9, 5233. https://doi.org/10.1038/s41598-019-41695-z.

van Eck, N.J., Waltman, L., Noyons, E.C.M., Buter, R.K., 2010. Automatic term identification for bibliometric mapping. Scientometrics 82, 581–596. https://doi.org/10.1007/s11192-010-0173-0.

Eck, Nees Jan van, and Ludo Waltman. "Text Mining and Visualization Using VOSviewer." arXiv, September 9, 2011. http://arxiv.org/abs/1109.2058.

Wagner, S.M., Schwab, A.P., 2004. Setting the stage for successful electronic reverse auctions. In: Journal of Purchasing and Supply Management, the 12th Annual IPSERA Conference 2003, vol. 10, pp. 11–26. https://doi.org/10.1016/j.pursup.2003.11.001.

Whissell, J.S., Clarke, C.L.A., 2011. Improving document clustering using Okapi BM25 feature weighting. Inf. Retr. 14, 466–487. https://doi.org/10.1007/s10791-011-9163-y.

Wynstra, F., Suurmond, R., Nullmeier, F., 2019. Purchasing and supply management as a multidisciplinary research field: unity in diversity? J. Purch. Supply Manag. 25, 100578 https://doi.org/10.1016/j.pursup.2019.100578.

Zhu, D., Zhang, Y., Lu, Z., 2023. Policy evolution of government procurement of public services in China: a text-mining perspective. Complexity 2023, 3139117. https://doi.org/10.1155/2023/3139117.

Zsidisin, G.A., Lamming, R., Harland, C., Wynstra, F., Ancarani, A., Tate, W.L., Knight, L., 2019. Reflecting on the past 25 years of the journal of purchasing and supply management: the editors' perspectives. J. Purch. Supply Manag. 25 https://doi.org/10.1016/j.pursup.2019.100559.