DISCOVERING FREQUENT KEYWORD PAIRS ADDRESSED IN OPERATIONS MANAGEMENT RELATED ARTICLES PUBLISHED BETWEEN 2000 AND 2016 WITH DATA MINING[1]

2000-2016 ARASI İŞLEMLER YÖNETİMİ ALANIYLA İLGİLİ MAKALELERDE YAYGIN ANAHTAR KELİME ÇİFTLERİNİN VERİ MADENCİLİĞİ İLE BELİRLENMESİ

Arş. Gör. İnanç KABASAKAL
Ege Üniversitesi, İktisadi ve İdari Bilimler Fakültesi İşletme Bölüm,
inanc.kabasakal@ege.edu.tr

Yrd. Doç. Dr. Aydın KOÇAK
Ege Üniversitesi, İktisadi ve İdari Bilimler Fakültesi İşletme Bölümü
aydin.kocak@ege.edu.tr

Abstract

Keywords provide a summarization and abstraction of document content. Several studies utilized keyword data in academic publications to find out which concepts had been widely addressed. Focusing on Operations Management (OM) field, this study reveals and introduces the popular keywords frequently used together in articles published between 2000 and 2016. For this purpose, keyword data of articles published in 39 OM-related journals were collected into a database. After elimination of redundant data and standardization of keywords, the most common keywords were identified. Among unsupervised rule mining techniques, association rule mining was used to discover frequent keyword pairs in meta-data. As a result, the frequent keyword pairs are extracted that correspond to the topics and concepts frequently addressed together in articles. Since the data analyzed covers a large number of recent publications, the findings of the study might prove useful for researchers to notice the connections among topics and concepts in contemporary OM research.

Keywords: *Association Rule Mining, Keywords, Operations Management Research*

Öz

[1] This paper is a substantially revised and extended version of a paper entitled "Discovery and Trend Analysis of Keyword Associations in Operations Management Articles with Data mining" presented in the 16th Symposium on Production Research held between the dates 12-14 October 2016.

Anahtar kelimeler, belgelerin içeriğine dair özetleme ve soyutlama işlevleri görmektedir. Geçmiş çalışmalarda, İşlemler Yönetimi alanıyla ilgili akademik yayınları konularına ve araştırma yöntemlerine göre sınıflandıran çalışmalar yer almaktadır. Bununla birlikte, makalelerde anahtar kelime kullanımının ele alındığı çalışmalara da rastlanmaktadır. Bu çalışmada, İşlemler Yönetimi (İY) alanına odaklanılarak 2000-2016 arası yayımlanmış makalelerde yaygın biçimde kullanılmış anahtar kelimelerin ortaya çıkarılması hedeflenmiştir. Bu amaçla, İY alanıyla ilişkili 39 akademik dergiden makale ve anahtar kelime verisi toplanarak bir veri tabanı oluşturulmuştur. Gereksiz kayıtların ayıklanması ve kelimelerin standartlaştırılması sonrasında makalelerde en sık kullanılan anahtar kelimeler bulunmuştur. Sıkça birlikte kullanılan anahtar kelimelerin keşfi amacıyla, denetimsiz veri madenciliği yöntemleri arasında yer alan birliktelik kuralları madenciliği kullanılmıştır. Ortaya çıkarılan sık anahtar kelime çiftleri, İY alanında birlikte sıkça değinilen konu ve kavramlara karşılık gelmektedir. Çalışmada incelenen verinin güncel ve sayıca fazla makaleden oluşması nedeniyle, bulguların çağdaş İY çalışmalarında ele alınan konu va kavramlar arası bağlantıları araştırmacıların dikkatine sunarak yarar sağlayabileceği düşünülmektedir.

Anahtar Kelimeler: *Birliktelik Kuralları Madenciliği, Anahtar Kelimeler, İşlemler Yönetimi Araştırmaları*

1. INTRODUCTION

Since 1980's, there have been various studies in Operations Management (OM) discipline to investigate the new concepts, topics, research methods addressed in the recent literature. Amoako-Gyampah & Meredith (1989), Pannirselvam et al. (1999), Pilkington & Fitzgerald (2006) and Taylor & Taylor (2009) have focused on identifying the topics and corresponding methods in OM studies. Schniederjans et al. (2009) suggested that keywords indicate the article content fairly since the articles are evaluated through a review process by journal editors and peers. However, it has been noticed that there was no study where the progress in topics and methods are studied through the discovery and analysis of keyword associations. Based on this remark, we believe that analysis to reveal the emerging concepts addressed together in recent publications might prove useful for researchers.

The first objective of our study is to find the most popular keywords in OM related journals between 2000 and 2016. The other objective of the study is to find the most popular keyword pairs to find out the concepts addressed together in recent literature. The most frequent keywords and keyword pairs were revealed and presented in the results. Additionally, the results were projected with annual charts.

The second section of the study includes the literature review. The research methodology has been presented in the third section. Data collection for the study was elaborated, including the selection of journals and data pre-processing. Afterwards, data analysis that consists of

mining and discovery keyword pairs was remarked. In the next sections, the frequent keywords and keyword pairs were presented with annual charts. Finally, the conclusion comes up with the comments on the findings of the study.

2. LITERATURE REVIEW

Previous research includes various studies on the topics and methods in Operations Management field. Mostly, the objective is limited with identifying the themes and methods. Amoako-Gyampah and Meredith (1989) examined the recent agenda in Operations Management research. Filippini (1997) focused on topic areas and the approaches utilized in OM research. In both studies, the progress in OM was addressed with a focus on the topics covered and the methods employed.

With an analysis of 1754 articles published in seven Operations Management journals published between 1992 and 1997, Pannirselvam et al. (1999) examined the popular topics and research methods in OM research. They identified and classified the studies that address multiple topics in OM research, and presented the frequency of topics in previous research. As a result, the study revealed 44 pairs, five triples, and two quadruples of topics frequently addressed in articles. In the results, the study revealed the emerging topics in OM research in 90's. Moreover, scheduling and inventory were referred as the most popular topics.

Pilkington and Fitzgerald (2006) investigated the major themes and emerging topics of Operations Management. Taylor and Taylor (2009) focused on main research themes and methods addressed in 310 articles that had been published in two journals: *International Journal of Operations* and *Production Management Journal*. Schniederjans, Schniederjans, and Schniederjans (2009) investigated the research methodologies in Operations Research field to explore the progress in Operations Research methodologies.

Among the studies examined during the literature review, none had utilized data mining techniques on a considerably large number of articles. In a study proposed to generate keywords from abstracts, HaCohen-Kerner (2003) argued that keywords provide an additional kind of document abstraction. Regarding the importance of keywords in publications, an analysis of the keywords with data mining might prove useful to discover frequent patterns within the keyword data. With such motive, this study is aimed to reveal the recent and popular pairs of concepts addressed in OM-related publications. Furthermore, with an intention to provide insights about the recent progress in the OM field, the annual frequency of keywords and keyword pairs were presented in results.

3. RESEARCH METHODOLOGY

In accordance with the objective of the study, a list of OM-related journals is necessary for data collection. Initially, a number of Operations Management related journals were selected according to their scope. Accordingly, previous studies that list and rank the prominent journals in OM field were taken into consideration. The previous studies by Goh et al. (1997), Soteriou, Hadjinicola, and Patsia (1999), Barman, Hanna, and LaForge (2001), Theoharakis et al. (2007), Fry and Donohue (2013) and Shang et al. (2015) were referred for journal selection. Besides, the journals were taken into consideration according to their impact factors announced regarding the last two years and five years. The list of journals chosen for data collection is listed in Table 1.

Table 1: List of journals chosen for data collection

	Impact Factor			Impact Factor	
Journal	**2014/15**	**5-yr**	**Journal**	**2014/15**	**5-yr**
4OR-A Quarterly Journal of Operations Research	1	0,88	Int. Journal of Technology Management	0.625	0,70
Annals of Operations Research	1.217	1,50	Int. Transactions in Operational Research	0.977	1,10
Applied Stochastic Models in Business and Industry	0.725	0,81	Journal of Global Optimization	1.287	1,45
Asia-Pacific Journal of Operational Research	0.346	0,58	Journal of Manufacturing Systems	1.682	2,08
Central European Journal of Operations Research	0.832	0,93	Journal of Operations Management	3.818	7,69
Computational Optimization and Applications	1.317	1,47	Journal of Optimization Theory and Applications	1.509	1,63
Computers & Operations Research	1.861	2,45	Journal of Scheduling	1.028	1,79
Engineering Economist	0.844	0,80	Journal of the Operational Research Society	0.953	1,25
Engineering Optimization	1.076	1,29	Management Science	2.482	3,40
European Journal of Operational Research	2.358	2,91	Manufacturing & Service Operations Management	1.462	2,39
Flexible Services and Manufacturing Journal	1.872	1,77	Mathematical Methods of Operations Research	0.625	0,89
Fuzzy Optimization and Decision Making	2.163	2,61	Mathematics of Operations Research	1.307	1,59
IIE Transactions	1.371	1,58	Operations Management	3.818	N/A
INFOR	0.171	0,50	Operations Research Letters	0.617	0,86
Informs Journal on Computing	1.077	1,54	OR Spectrum	0.987	1,96
Interfaces	0,42	0,91	Production and Operations Management	1.439	2,12

Int. Journal of Computer Integrated Manufacturing	1.012	1,14	Production Engineering	0	NA
Int. Journal of Operations & Production Management	1.736	1,99	Production Planning & Control	1.466	1,73
Int. Journal of Production Economics	2.752	3,07	Queueing Systems	0.839	0,83
Int. Journal of Production Research	1.477	1,77			

In order to find out the popular concepts addressed in articles published in journals listed, article data was collected through online academic databases. The data was based on the keywords provided by the databases. In this regard, we assume that the keywords were appropriately selected and all are coherent with their context.

3.1. Data collection and pre-processing

To collect the keyword data; a computer program, which had been developed by the lead author, was used. To access online databases and parse article data, Selenium Browser Automation Tool (http://www.seleniumhq.org/) was used. Moreover, a relational database was created in Microsoft SQL Server to collect the article data with corresponding keywords. The software has been run for each of the journals individually. The raw data extracted from the results involves 397,657 instances of 71,322 keywords in 64,600 articles.

After the collection of raw data, the database was checked for redundancy. It was noticed that data collected included editorials, promulgations, announcements, etc. Consequently, the pre-processing stage required the elimination of irrelevant content. A blacklist was prepared including the titles of articles to be excluded. Specifically 565 titles such as 'A special issue on', 'Call for Papers: Special Issue on Some Subject' were involved in the blacklist. Afterwards, the articles without any corresponding keywords were excluded. The pre-processing step for the elimination of redundant articles is demonstrated in Table 2.

Table 2. Count of articles evaluated through data pre-processing phase

Pre-Processing Step	**Article Count**
Raw data	64,600
Articles after elimination of special issues, notes, editorial, etc.	61,433
Article after elimination of articles without keywords	60,199
Articles having at least one frequent (≥ 5) keyword	56,173

The final data included in the study consists of 56,173 articles. The count of articles in selected journals is presented in Table 3.

After the pre-processing of articles, keyword data was examined in a similar manner. Since the concepts might be expressed with several variations of keywords, the keyword data required unification. Moreover, the lack of standardization in abbreviations also required attention. In accordance, the pre-processing phase required carefully unification of keywords. However, the size of keyword data required excessive time to analyze all keywords. Since the analysis aims to discover frequent keywords, analysis on rare keywords would not contribute to the objective of the study. In this regard, it was justifiable to exclude rare keywords. As an assumption, the keywords that were addressed in less than five articles were presumed rare. Accordingly, 63,383 of 71,322 keywords (88.86%) were rare. As a side effect, 4026 articles without any frequent keywords were also ignored.

After the elimination of rare keywords, it was feasible to review 7939 keywords manually. The keywords that differ because of abbreviations or differences in punctuation required standardization. For example, two keywords such as ‘agile production.’ and ‘agile production’ were unified. Afterwards, it was remarked that keyword data contains NAICS codes on articles. North American Industry Classification System (NAICS) is a standard used to classify business establishments for statistical analysis (http://www.census.gov/eos/www/naics). It was convenient to ignore NAICS codes; since they were not specific about neither the topics covered nor the methodology employed in articles. The pre-processing steps of keyword data were summarized in Table 4.

Table 3. Final count of articles by journal

Journal	Articles	Journal	Articles
European Journal of Operational Research	8,977	Applied Stochastic Models in Business and Industry	784
International Journal of Production Research	5,726	Interfaces	771
International journal of production economics	3,810	Queueing Systems	767
Computers & Operations Research	3,033	International Transactions in Operational Research	766
Journal of the Operational Research Society	2,740	Journal of Manufacturing Systems	717
Annals of Operations Research	2,678	Informs Journal on Computing	690
Journal of Optimization Theory and Applications	2,353	Asia-Pacific Journal of Operational Research	674
Management Science	2,349	Manufacturing & Service Operations Management	590
Journal of Global Optimization	1,684	Journal of Operations Management	583
IIE Transactions	1,468	Central European Journal of Operations Research	544
Operations Research Letters	1,386	Operations Management	536
International Journal of Technology Management	1,268	Production Engineering	532
Production Planning & Control	1,261	Journal of Scheduling	530

International Journal of Computer Integrated Manufacturing	1,191	OR Spectrum	512
Production and Operations Management	1,071	INFOR	368
Computational Optimization and Applications	1,038	Engineering Economist	320
International Journal of Operations & Production Management	1,029	4OR-A Quarterly Journal of Operations Research	320
Engineering Optimization	924	Fuzzy Optimization and Decision Making	295
Mathematical Methods of Operations Research	869	Flexible Services and Manufacturing Journal	177
Mathematics of Operations Research	842		

Table 4. Count of keywords evaluated through pre-processing phase

Pre-Processing Step	Keyword Count
The initial count of keywords	71,322
Frequent keywords (≥ 5)	7,939
Keywords, after standardization	5,543
Keywords, after elimination of NAICS codes	4,905

The final data set for the analysis includes 244,511 instances of 4,905 distinct keywords in 56,173 articles.

3.2. Discovery of frequent keyword pairs with Association Rule Mining

The keywords frequently used together in articles were analyzed through data mining. In particular, Apriori algorithm was used for association analysis. As a result, significant keyword pairs that were frequently used together in articles were discovered.

Association rule mining can be addressed among the unsupervised models in data mining (Tsiptsis and Chorianopoulos, 2009). Apriori is regarded as one of the most popular algorithms in association modeling techniques. The algorithm was proposed by Agrawal (1993) to discover the products bought together by the customers. More generally, the objective of the algorithm is to discover the patterns of item groups and generate association rules to model the associations among items.

The algorithm consists of two steps (Agrawal and Srikant, 1994): finding all sets of frequent items that have transactions over a minimum support and using itemsets to generate association rules. Since the objective of the study is limited to finding the frequent keyword pairs, only the steps required will be addressed in this context.

The problem can be formulated as the following (Tan et al, 2005): let $I = \{ i_1, i_2, .. , i_n \}$ be the set of all items in a market basket and $T = \{ t_1, t_2, \dots , t_n \}$ be the set of all transactions. Each transaction t_i consists of items in *I*. The support of an item i_j is the count of transactions t_i that contains i_j in its sub-items. The algorithm uses a minimum support criterion (*min-sup*) to check the frequency of items: $support(i_j) < min\text{-}sup$ and prune those below the threshold. In this way, pruning of non-frequent items helps to avoid redundant computation (Tan et.al, 2005).

The pruning based on support threshold was justified on a mathematical basis. If a set of items, namely *{Y}*, contains a subset *{X}* of items, the following statement holds true by definition: $Support(X) \geq Support(Y)$. In such case, if the subset *{X}* is not frequent, the item-set *{Y}* cannot be frequent either. This observation guarantees that the algorithm follows a solid strategy by extending frequent itemsets to find larger frequent itemsets (Maimon and Rokach, 2005).

The first phase of Apriori algorithm generates frequent item-sets. The discovery of frequent itemsets in Apriori algorithm is given in Figure 1. Additionally, there is another phase of the algorithm where the item-sets are used to create association rules. In this phase, association rules *{X}* → *{Y}* are checked against the minimum confidence criteria and only the rules that satisfy the criteria are selected. A critique on the confidence criteria (Silverstein et al, 1998) argues that the confidence in the evaluation of association rules might lead to false inferences. Moreover, several algorithms and indicators have been proposed on interestingness of rules such as lift, conviction, and gain (Bayardo Jr. & Agrawal, 1999). However, the discovery of frequent keywords was sufficient for our analysis. The rule generation step of the algorithm was out of the scope, thus was not covered in this study.

1) L_1 = {large 1-itemsets}
2) for ($k = 2$; $L_{k-1} \neq \emptyset$; k++)
3) C_k = apriori_gen(L_{k-1}) // Generate new candidates
4) forall transactions $t \in D$ do begin
5) Ct = subset(C_k, t); // Candidates contained in t
6) forall candidates $c \in C_t$ do
7) c.count++; // Increment the support for candidate
8) end
9) $L_k = \{ c \in C_k \mid c.count \leq minsup \}$ // Select the candidates above support criteria
10) end
11) FrequentItems = $\cup L_k$;

Figure 1: Frequent itemset discovery in Apriori Algorithm (Agrawal & Srikant, 1994:5)

4. RESULTS

The keyword data collected in research database was analyzed with SQL queries to find the most popular keywords. The most frequent keywords in data were presented with the annual frequencies of corresponding articles in Table 5.

Table 5: Annual frequencies of most frequent keywords in articles

Rank	Keyword	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
1	supply chain management	29	65	65	106	165	159	189	252	257	258	288	276	353	367	377	363	339	3908
2	algorithms	155	150	141	145	204	209	212	227	257	242	214	240	195	200	169	97	128	3185
3	research	31	32	29	14	28	25	161	173	133	155	135	105	319	441	410	505	218	2914
4	manufacturing processes	167	111	167	139	203	192	155	156	194	200	232	154	191	185	102	136	147	2831
5	mathematical optimization	81	119	84	94	164	177	161	197	190	162	174	209	198	227	201	157	186	2781
6	operations research	137	110	120	199	272	193	178	190	225	168	157	95	131	141	61	80	63	2520
7	production/scheduling	152	126	104	126	167	119	104	109	147	127	125	150	193	204	147	144	106	2350
8	mathematical model	108	124	70	60	88	112	52	70	97	137	123	281	300	305	135	101	137	2300
9	decision making	76	67	64	71	92	123	125	115	129	161	167	182	167	180	140	156	121	2136
10	heuristic	90	89	83	93	126	104	157	158	169	155	101	129	133	110	91	80	105	1973
11	scheduling	48	54	63	56	84	103	103	162	148	132	148	114	149	129	135	121	135	1884
12	simulation methods	74	66	55	62	71	77	91	107	126	154	112	149	141	117	113	93	77	1685
13	inventory control	48	73	63	84	108	74	94	98	132	106	117	106	111	143	107	90	96	1650
14	linear programming	67	70	65	73	117	99	74	105	112	109	108	111	135	104	110	93	95	1647
15	production planning	93	73	58	82	110	81	93	76	122	130	84	105	182	131	87	64	64	1635
16	integer linear programming	52	51	65	45	87	75	101	133	123	114	119	118	142	96	94	115	103	1633
17	genetic algorithm	52	50	51	77	71	80	101	112	107	127	100	131	146	113	105	109	77	1609
18	mathematical programming	52	87	60	77	130	149	98	122	106	74	114	70	126	99	82	78	70	1594
19	inventory	60	67	48	75	73	92	109	118	131	98	77	78	90	96	108	73	55	1448
20	production management	62	73	96	91	117	80	97	67	131	101	51	50	130	93	41	30	42	1352
21	combinatorial analysis	35	29	43	46	71	82	101	96	104	109	89	77	128	79	73	58	69	1289
22	production economics	93	66	138	109	136	106	61	63	101	85	79	66	91	41	13	13	19	1280
23	production control	70	48	45	47	100	57	74	69	91	79	75	76	142	85	52	63	43	1216
24	industrial management	75	70	38	83	143	99	94	80	140	96	78	40	48	34	42	19	33	1212
25	management	66	39	45	54	42	62	68	61	102	63	84	50	135	137	58	66	57	1189
26	dynamic programming	46	55	35	47	42	54	85	70	81	87	82	93	89	91	69	71	72	1169
27	data envelopment analysis	35	38	40	31	62	41	58	63	50	99	102	85	99	75	93	89	88	1148
28	stochastic process	51	53	60	56	68	90	74	69	69	57	50	50	63	75	62	54	34	1035
29	problem solving	32	29	16	25	32	21	44	54	47	62	73	54	67	131	128	144	66	1025
30	management science	23	81	81	49	127	76	82	102	128	103	26	49	38	32	3	11	9	1020
31	markov decision process	39	47	44	38	47	54	49	71	57	52	67	64	72	103	73	56	76	1009
32	game theory	22	32	32	42	45	46	47	49	44	46	77	66	98	99	83	86	70	984
33	queueing theory	31	32	46	27	37	34	37	70	74	84	64	79	72	68	62	54	48	919
34	probability theory	25	40	43	30	42	64	55	47	58	69	56	59	52	77	60	72	64	913
35	pricing	22	30	26	25	27	56	56	57	73	68	73	53	65	60	65	69	81	906
36	industrial engineering	46	35	41	69	104	86	69	59	79	45	65	33	64	40	17	14	17	883
37	product management	40	70	54	41	103	86	83	71	85	58	44	23	34	41	21	12	14	880

38	production	40	25	31	19	30	44	64	73	96	152	26	47	34	38	48	43	62	872
39	manufacturing industries	17	16	21	16	18	31	35	67	43	79	49	57	94	62	80	93	52	830
40	quality control	30	28	27	33	46	34	37	45	61	76	51	62	65	64	38	38	48	783
41	manufactures	36	45	30	37	40	38	56	69	87	30	30	36	54	52	46	40	49	775
42	industrial efficiency	5	6	5	5	20	22	22	23	92	99	121	85	68	63	38	33	40	747
43	strategic planning	22	35	33	21	48	25	50	47	60	65	41	48	69	51	28	43	54	740
44	supply & demand	20	23	20	20	45	51	47	58	68	50	66	44	38	35	43	42	46	716
45	flexible manufacturing system	44	49	49	27	40	33	28	41	58	46	55	50	73	54	25	14	21	707
46	numerical analysis	12	26	14	21	24	41	33	35	40	59	30	62	70	79	58	44	42	690
47	decision support systems	44	34	26	29	37	33	51	43	38	55	43	47	51	40	40	43	27	681
48	new product development	18	35	37	35	30	33	45	42	40	49	31	28	44	66	55	52	38	678
49	methodology	30	46	22	17	29	41	80	44	51	32	54	66	51	63	17	17	11	671
50	nonlinear programming	23	24	26	21	31	34	44	48	51	42	45	32	53	48	48	66	32	668

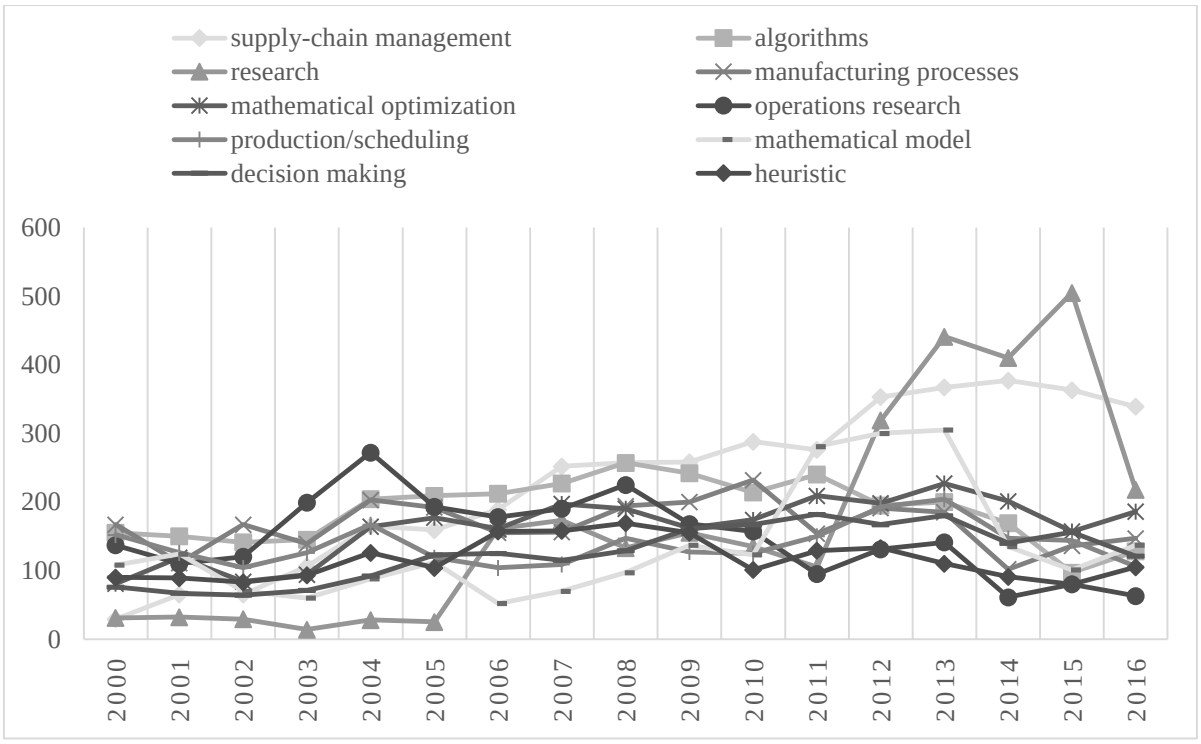

Figure 2: Annual frequency of most frequent keywords

Figure 2 indicates that the keywords “research” and “supply chain management” have been substantially used in more articles between 2010 and 2016. Moreover, the count of articles that have the keyword “mathematical model” have increased until 2013, sooner the keyword has shown a decline. Additionally, the graph indicates that the attention for the keyword “operations research” have peaked in 2004; however, the keyword has been used in fewer articles since then. The keyword “algorithms” also was in a decline for the last 5 years. The annual frequency of keywords in articles was presented in Table 5.

4.1. Frequent keyword pairs discovered

In our study, association analysis was performed to find keywords frequently addressed in articles. Apriori algorithm was used in order to find frequent keywords that were used together in articles. The algorithm had already been implemented by the corresponding author; no further software was necessary for the study. Since the discovery of frequent keyword pairs was sufficient to accomplish the objective of the study, rule generation step of the algorithm was neglected.

The minimum support criteria provided for the analysis was 0.1%. The percentage settled implies that a keyword is qualified as frequent if it had been addressed in at least 57 articles (0.1% of 56,173). The same condition holds true for the keyword pairs as well. Among the candidates generated by the algorithm, 832 frequent keyword groups were discovered. The top 50 of most frequent keyword pairs were listed in Table 6, with the count of corresponding articles denoted by the support.

The keyword pairs in the results mostly consist of research methods, topics, specific concepts on topics, etc. It can be argued that the lack of taxonomy in keyword data restricts deductions over the broad topics in OM-related studies. However, the findings include interesting pairs of concepts that require attention.

Table 6 demonstrates that “manufacturing processes” and “production economics” have been the most frequently addressed keyword pair in the results. Both keywords have been used in 608 articles together. The secondly most popular couple was “Algorithms” and “mathematical optimization” in 589 articles. Afterwards, “supply chain management” and “inventory control” have been both addressed in 522 articles. The results imply that “algorithms” keyword is frequent in articles where “mathematical optimization” or “production/scheduling” keywords were also addressed.

The annual statistics of most frequent keywords pairs addressed together have also been fetched from research data via SQL queries. The yearly statistics of corresponding keyword pairs frequently used together in articles have been demonstrated in the following figures.

Table 6: The most frequent keyword pairs addressed together in articles

Rank	Keyword Pair		Support
1	manufacturing processes	production economics	622
2	algorithms	mathematical optimization	613
3	supply chain management	inventory control	554
4	production/scheduling	production control	493
5	research	supply chain management	437
6	manufacturing processes	research	435
7	algorithms	production/scheduling	433
8	operations research	production/scheduling	412
9	manufacturing processes	production/scheduling	403
10	production planning	production management	396
11	manufacturing processes	production planning	375
12	production/scheduling	production planning	370
13	algorithms	operations research	369
14	production planning	production control	364
15	manufacturing processes	production management	361
16	research	management	361
17	mathematical optimization	mathematical programming	360
18	algorithms	mathematical programming	352
19	operations research	mathematical optimization	344
20	supply chain management	supply & demand	334
21	supply chain management	industrial procurement	330
22	manufacturing processes	production control	329
23	research	production/scheduling	327
24	manufacturing processes	industrial engineering	325
25	manufacturing processes	process control	325
26	research	mathematical model	324
27	operations research	research	307
28	research	production planning	303
29	mathematical optimization	mathematical analysis	293
30	production/scheduling	mathematical model	288
31	algorithms	mathematical model	288
32	supply chain management	suppliers	284
33	production planning	production economics	284
34	linear programming	mathematical programming	281
35	production/scheduling	production management	276
36	mathematical optimization	mathematical model	273
37	research	production control	272
38	supply chain management	business logistics	272
39	production management	production economics	260
40	operations research	heuristic	258
41	mathematical programming	integer linear programming	256
42	research	production management	255
43	manufacturing processes	mathematical model	252
44	algorithms	heuristic	251
45	operations research	mathematical programming	250
46	decision making	supply chain management	248
47	production/scheduling	heuristic	247
48	production/scheduling	genetic algorithm	244
49	algorithms	manufacturing processes	244
50	manufacturing processes	assembly line methods	241

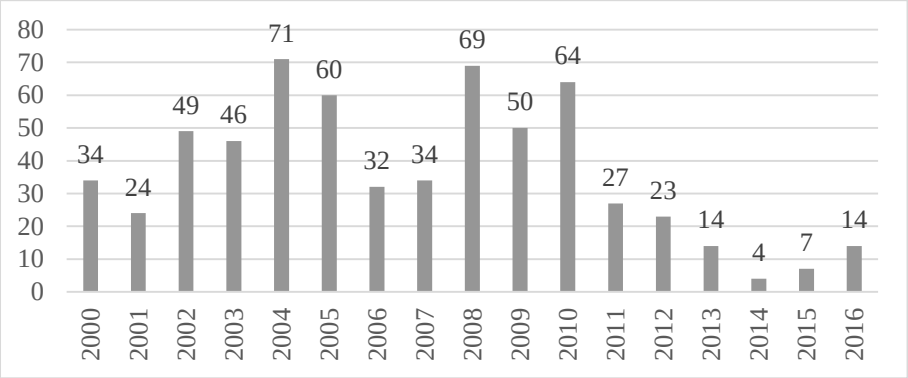

Figure 3: Frequency of articles with keywords: "manufacturing processes" & "production economics"

The average count of articles where the keywords "manufacturing processes" and "production economics" have been cited together was 36.59. According to the graph, both keywords were mostly used together until 2010. However, the average count of articles with both keywords has dropped from 48.45 to 14.83 since 2010. Although the keyword pair has been the most common pair that was cited together, Table 5 indicates a substantial decline in the keyword 'production economics'. Accordingly, the drop in the frequency articles with the keyword pair is remarkable.

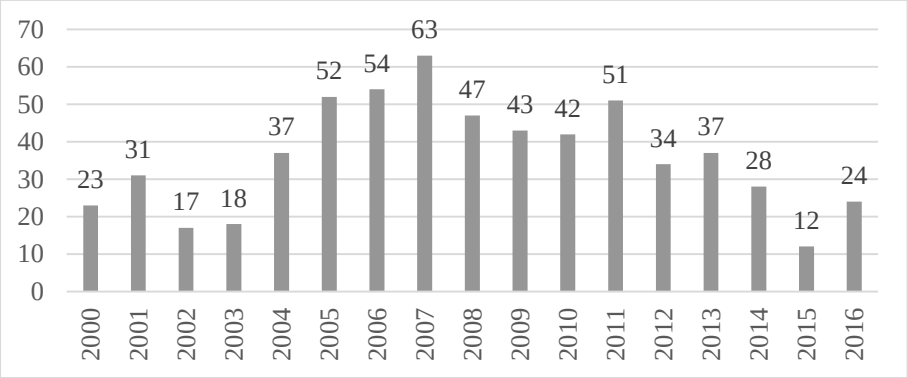

Figure 4: Frequency of articles with keywords: "algorithms" & "mathematical optimization"

The average count of articles where the keywords "algorithms" and "mathematical optimization" have been cited together was 36.06. The frequency of those articles has mostly increased until 2007. However, the frequency has been in decline for the last decade. The two keywords are among the fundamental methods used in problems. The results in Table 5

indicate no decline in the popularity of both keywords individually; however, Figure 4 reveals that the articles that address both topics together are becoming less popular.

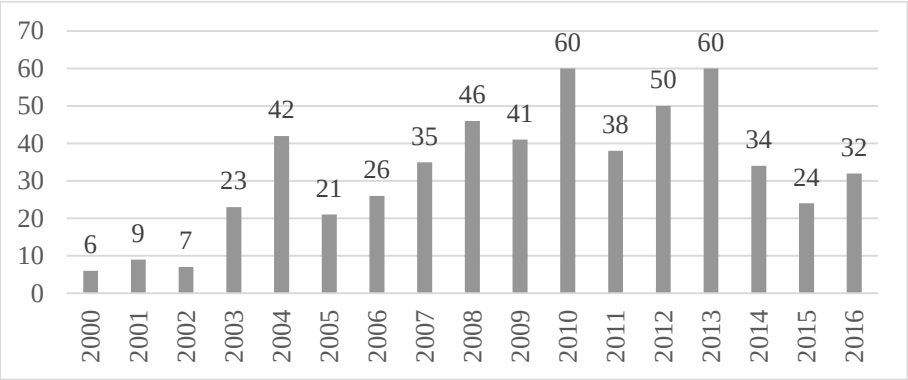

Figure 5: Frequency of articles with keywords: "supply chain management" & "inventory control"

From 2000 to 2016, the average number of articles where "supply chain management" & "inventory control" were both addressed was 32.59. According to the graph, the two topics cited in keywords together mostly increased from 2003 to 2013. Since the mentioned keywords are the third most frequent among the results, it can be suggested that inventory control is the most popular topic within the articles related to supply chain management. However, that decline of the keyword pair for the last few years is noticeable in the results.

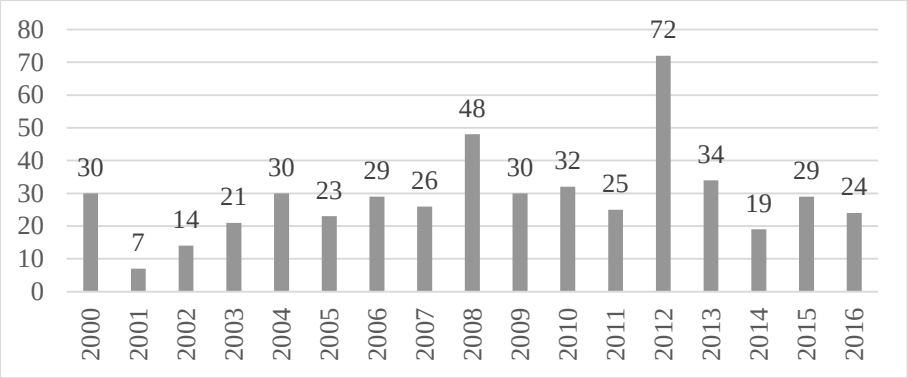

Figure 6: Frequency of articles with keywords: "production/scheduling" & "production control"

The average count of articles where the keywords "production/scheduling" and "production control" have been cited in keywords together was 29. The graph of annual statistics in Figure 6 indicates fluctuations from 2000 to 2016. However, it cannot be interpreted as a continuing trend of increase or decline. The most remarkable year in data was 2012. In that year, two keywords were addressed together in 72 articles.

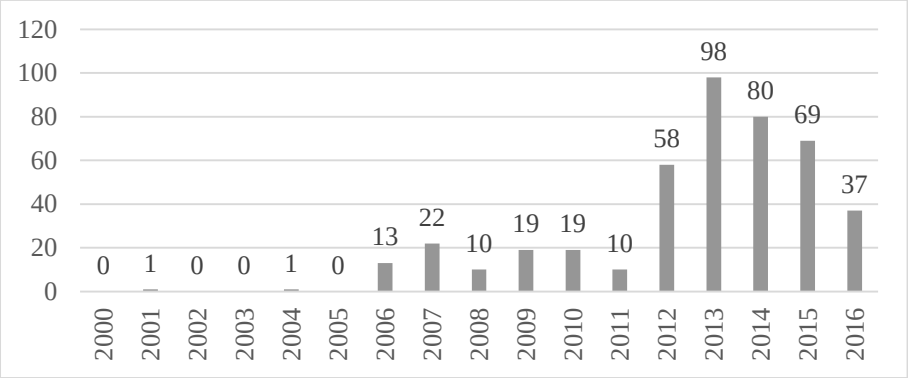

Figure 7: Frequency of articles with keywords: "research" & "supply chain management"

The average count of articles where "research" and "supply chain management" were addressed was 25.71. The graph indicates an overall increase in years. In the first five years, virtually no articles have addressed both topics together; however, the next two 5-year periods represent an increase compared to the previous period. Moreover, the keywords "supply-chain management" is the most frequent keyword in article data. The keyword "research" is also prevalent in articles; it is the third most frequent keyword. As visualized in Figure 1, the frequency of both keywords was much fewer articles at the beginning of 2000's. However, each of both keywords was more popular in recent publications. The frequency of the citations of the keyword pair "research" and "supply chain management" is no different.

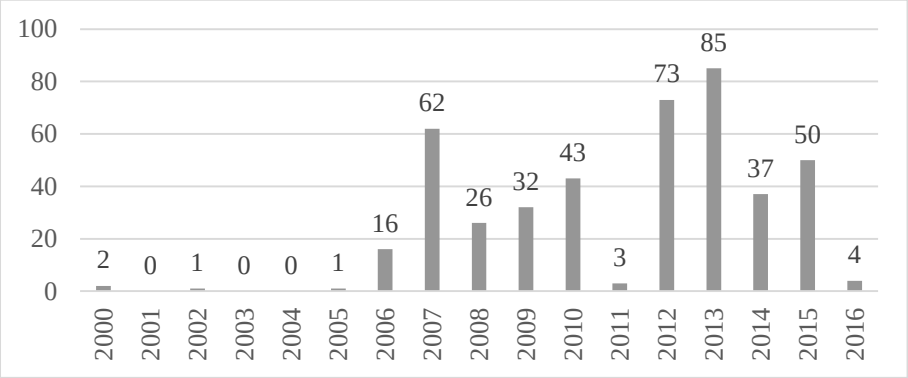

Figure 8: Frequency of articles with keywords: "manufacturing processes" & "research"

According to the results, "manufacturing processes" and "research" have been much more frequently addressed in articles since 2006. The average count of articles was 25.59. Although the year 2011 represents an exception, it can be suggested that two keywords were often used together in articles last decade. Since "research" is the most dramatically

increasing keyword within all keyword data, it is perceptible that the pairs of keywords that contain “research” have been addressed in more articles within the last decade.

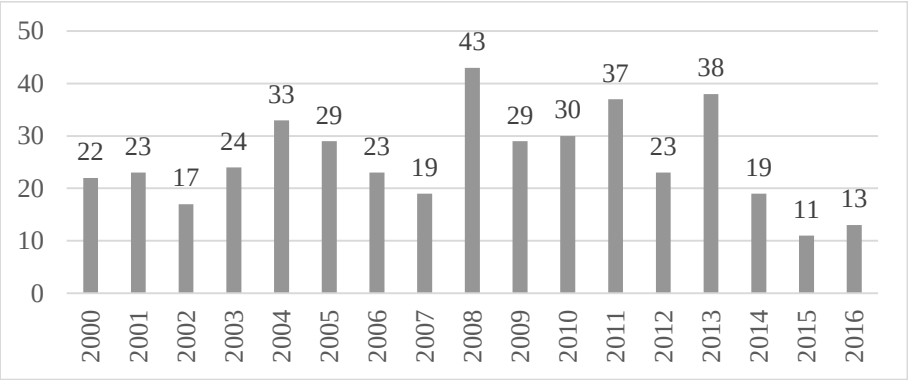

Figure 9: Frequency of articles with keywords: “algorithms” & “production/scheduling”

The average count of articles where keywords “algorithms” and “production/scheduling” have been addressed together was 25.47. Although the graph demonstrates noticeable fluctuations over the years, it is hard to interpret the graph as a noticeable trend. As Figure 9 demonstrates, the usage of both keywords in articles has dropped since 2014. However, the keyword pair remains the seventh most frequent pair in results.

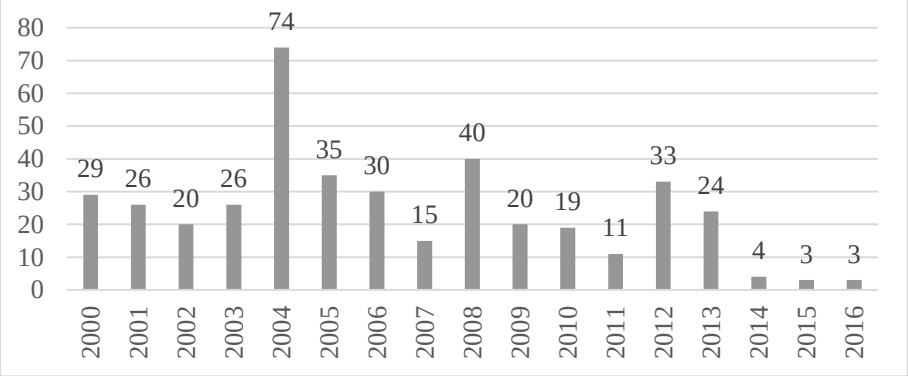

Figure 10: Frequency of articles with keywords: “operations research” & “production/scheduling”

The average count of articles where the keywords “operations research” and “production/scheduling” have been cited in keywords together was 24.24. The graph demonstrates a significant decline since 2014. The average count of articles that address both keywords has declined from 28.71 to 3.33 after 2013. In Table 5, it is noticeable that the frequency of “operations research” has dropped from 165.42 to 68 after 2013. However, a direct conclusion of decline in “operations research” studies might be misleading. Since the keyword corresponds to a field of research in taxonomy of disciplines, it can be argued that

the decline of articles with the keyword 'operations research' requires a thorough explanation in further studies.

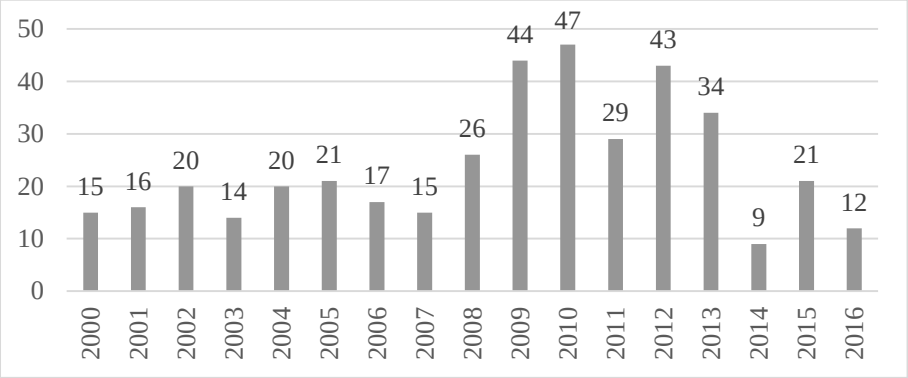

Figure 11: Frequency of articles with keywords: "manufacturing processes" & "production/scheduling"

The keyword pair "manufacturing processes" and "production/scheduling" were addressed in 23.71 articles on average between 2000 and 2016. From the graph, it is noticeable that the frequency of articles that have both keywords had been on the top in the period between 2009 and 2013.

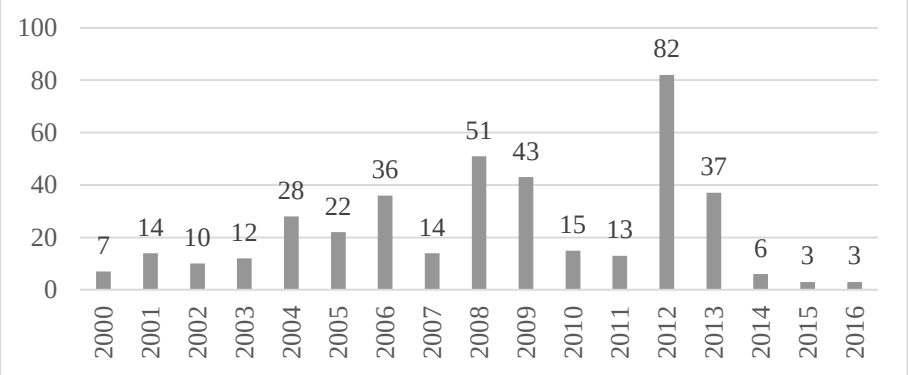

Figure 12: Frequency of articles with keywords: "production planning" & "production management"

The final pair of keywords among the most frequent 10 pairs involves the keywords "production planning" and "production management". Both keywords were addressed together in 23.29 articles in average. Although the graph indicates significant fluctuations, it is hard to summarize the graph in a trend of decrease or increase. The year 2012 was significantly different in the count of articles in Figure 12. However, the last two years demonstrate a decline in the citation of both keywords together.

5. CONCLUSION

In the study, the articles Operations Management related journals have been analyzed to identify the most popular keywords. Moreover, keywords that were frequently used together have been discovered. In the literature review, several studies that focused on the Operations Management topics and methods were referred. Amoako-Gyampah & Meredith (1989), Pannirselvam et al. (1999), Pilkington & Fitzgerald (2006) and Taylor & Taylor (2009) focused on the topics and methods in Operations Management research. Among the studies reviewed, the study by Pannirselvam et al. (1999) was the only one that presented the topics frequently addressed together. The data involved 1754 articles from 7 journals published in a period of 5 years. However, our study covers a much larger number of articles; thus the discovery of frequent keywords required utilization of data mining.

The methodology contributes to the originality of the study in two aspects: the automation process to collect a large amount of meta-data on articles, and the data mining process that include pre-processing phase and discovery of topics pairs frequently address together. The data enabled to evaluate the popularity of keywords in articles, thus provided a basis for further discussion on popular concepts on Operations Management discipline.

The results provided might provide useful for researchers to identify which concepts are getting more interest in recent years. Additionally, the frequent pairs might assist to diagnose which concepts were mentioned together over the past studies. However, the lack of taxonomy in topics or concepts in keyword data is a limitation. For instance, one might argue that the keywords ‘Management’ and ‘Research’ were used in a broader sense than the others. Moreover, more recent studies inherently involve new terms coined to express new concepts; often resulting in the decline of relevant keywords. Although the discussion section includes an assessment of keyword pairs based on the magnitudes in annual frequency, the remarks are subjective and open to dispute due to the limitations mentioned.

In further studies, the pre-processing of keyword data might be supported with more advanced techniques. Semantic similarity techniques and tools are used to identify topically similar words (Sridhara et al., 2008). Such techniques could assist to identify similar keywords to unify and consolidate data.

In conclusion, we believe that our results might provide interpretation and summarization of recent progress in OM-related studies from a broader perspective. The frequency of keyword couples in articles might help to identify the popular concepts and the focus on the recent research. Besides, a more comprehensive analysis can be conducted using other data mining techniques to investigate the relationship between OM concepts even further. Moreover, keyword data might be used for other tasks, such as to define the focal topics of journals. In fact, the methodology of our study provides an example of analytics on article meta-data with data mining.

REFERENCES

Agrawal, R., Imieliński, T. and Swami, A. (1993). Mining Association Rules Between Sets of Items in Large Databases. *In ACM SIGMOD Record*, 22 (2), pp. 207-216.

Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. *In Proceedings of 20th International Conference on Very Large Data Bases*, 1215, pp. 487-499.

Amoako-Gyampah, K. and Meredith, J. R. (1989). The Operations Management Agenda: An Update. *Journal of Operations Management, 8* (3), pp. 250-262.

Bayardo Jr., R.J. and Agrawal, R. (1999). Mining the Most Interesting Rules. *In Proceedings of the 5th International Conference on Knowledge Discovery and Data Mining*, pp. 145-154.

Filippini, R. (1997). Operations management research: some reflections on evolution, models and empirical studies in OM. *International Journal of Operations & Production Management, 17* (7), pp. 655-670.

Fry, T.D. and Donohue, J.M. (2013). Outlets for operations management research: a DEA assessment of journal quality and rankings. *International Journal of Production Research, 51* (23-24), pp. 7501-7526.

HaCohen-Kerner, Y. (2003). Automatic extraction of keywords from abstracts. *In International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, (pp. 843-849). Springer Berlin Heidelberg.

Maimon O. and Rokach L. (2005): *Data Mining and the Knowledge Discovery Handbook*. New York, USA: Springer.

Pannirselvam, G. P., Ferguson, L. A., R. Ash, C. and Siferd, S. P. (1999). Operations Management Research: An Update for the 1990s. *Journal of Operations Management, 18* (1), pp. 95-112.

Pilkington, A. and Fitzgerald, R. (2006). Operations Management Themes, Concepts and Relationships: A Forward Retrospective of IJOPM. *International Journal of Operations and Production Management, 26* (11), pp. 1255-1275.

Schniederjans, M. J., Schniederjans, A. M. and Schniederjans, D. G. (2009). Operations Research Methodology Life Cycle Trend Phases as Recorded in Journal Articles. *Journal of the Operational Research Society, 60* (7), pp. 881-894.

Shang, G., Saladin, B., Fry, T. and Donohue, J. (2015). Twenty-six years of operations management research (1985–2010): authorship patterns and research constituents in eleven top rated journals. *International Journal of Production Research, 53* (20), pp. 6161-6197.

Silverstein, C., Brin, S. and Motwani, R. (1998). Beyond Market Baskets: Generalizing Association Rules to Dependence Rules. *Data Mining and Knowledge Discovery, 2*, pp. 39-68.

Sridhara G., Hill, E., Pollock, L. and Vijay-Shanker, K. (2008). Identifying Word Relations in Software: A Comparative Study of Semantic Similarity Tools. *The 16th IEEE International Conference on Program Comprehension*. pp. 123-132.

Tan P.N., Steinbach M. and Kumar, V. (2005). *Introduction to Data Mining.* Boston: Addison-Wesley.

Taylor A. and Taylor, M. (2009). Operations Management Research: Contemporary Themes, Trends and Potential Future Directions. *International Journal of Operations and Production Management, 29* (12), pp. 1316-1340.

The United States Census Bureau. (2017). North American Industry Classification System, Retrieved February 03, 2017, from http://www.census.gov/eos/www/naics/.

ThoughtWorks. (2017). Selenium – Web Browser Automation, Retrieved February 03, 2017, from http://www.seleniumhq.org/.

Tsiptsis, K. and Chorianopoulos, A. (2009). *Data Mining Techniques in CRM: Inside Customer Segmentation.* Chichester, West Sussex: John Wiley & Sons.