Sustainable Procurement Analysis and Visualization of Procurement Orders from Various Provinces

Lili Cao [1, a], Hongjie Tao [2, b], Nana Wang [3, c], Juaner Zheng [3, d] and Jie Geng [2*, e1]

[1] Institute of Standardization Theory and Strategy, China National Institute of Standardization, Beijing, China ;
[2] China Research Institute of Regulation and Public Policy, Zhejiang University of Finance & Economics, Hangzhou, China ;
[2] Institute of Service Standardization, China National Institute of Standardization, Beijing, China .

ABSTRACT.

Procurement is an important link in the supply chain. Sustainable procurement has emerged as a crucial strategy to address environmental and social challenges while promoting responsible sourcing and procurement practices. This study analyzes purchase order data from representative provinces in seven regions of China and evaluates and ranks them using the concept of sustainable procurement. This paper uses natural language processing and visualization techniques to recognize and classify texts, and visually display them on maps. This has reference significance for analyzing supply chain management data in the future.

Keywords: Supply chain; Sustainable procurement; natural language processing; visualization.

1. INTRODUCTION

1.1 Supply Chain and Sustainable Procurement

Procurement is an important link in the supply chain. Procurement is the process of obtaining products and services for a project[1,2].

The sustainability of a project focuses on supply, while sustainable procurement focuses on procurement. Therefore, in order for enterprises or society to achieve sustainable development, they must start from the supply chain and establish various environmental and social standards with suppliers[3,4]. Sustainable procurement has emerged as a crucial strategy to address environmental and social challenges[5]. Sustainable procurement is also a form of green procurement aimed at reducing adverse impacts on the environment, society, and economy[6].

1.2 Visualization

With global development, supply chains are also facing higher challenges[7,8]. Some new technologies will be combined with supply chain technology. For example, Li and Zhao believes that digitization has significant practical significance for supply chain management[9]. Nowadays, many studies use more scientific statistical methods for quantitative analysis.

The use of data visualization as a means of conveying information has been a hot topic in recent years[10-11]. Andrea believes that data visualization can help make decisions. Andrea team has made good progress in visualizing and analyzing the management data process[12].

Overall, visualization technology plays a significant role in processing data, making it intuitive and facilitating further conclusions

1.3 Research objectives

The purpose of this study is to evaluate and analyze sustainable procurement data from various provinces, and to present it using visualization technology.

[1a] caoll@cnis.ac.cn, [b] Leon_722@163.com, [c] wangnn@cnis.ac.cn, [d] zhengje@cnis.ac.cn, *[e] jie.geng@zufe.edu.cn,

The International Conference Optoelectronic Information and Optical Engineering (OIOE2024), edited by Yang Yue, Lu Leng, Proc. of SPIE Vol. 13513, 135131T · © 2025 SPIE · 0277-786X · © The Authors. Published under a Creative Commons Attribution CC-BY 3.0 License · doi: 10.1117/12.3045570

2. DATA SELECTION AND PRE-PROCESSING

2.1 Data sources and data selection

In terms of data sources, this paper aims to cover the entire country as much as possible. China can be divided into 7 regions according to its orientation, namely: North China, East China, Central China, South China, Southwest, Northeast, and Northwest.

Figure. 1 Seven regions of China

This paper selects one province from each region for data collection. The selection list is shown in Table 1.

Table 1. Region and province

region	North China	East China	Central China	South China	Southwest	Northeast	Northwest
province	Beijing	Zhejiang	Hubei	Guangdong	Sichuan	Heilongjiang	Shaanxi

This paper collected 1000 purchase orders from various provinces in 2022, totaling 7000. Each purchase order includes the purchaser, procurement location, procurement content, quantity, unit price, total price, procurement method, procurement time, etc.

2.2 pre-processing

This paper needs to check the completeness of purchase order information and correct some obvious errors for subsequent analysis.

We divide all orders into 4 quarters based on the purchase order date.

3. DESCRIPTIVE ANALYSIS

3.1 Classification by Province

Table 2. Basic information of purchase orders from 7 provinces

	parameter				
Province	Sum(mil.)	Amount	Mean(k)	Min(k)	Max(mil.)
Guangdong	182.8594	1000	182.9	0.049	9.685
Heilongjiang	163.4346	1000	163.4	0.001	11.914
Zhejiang	1051.2914	1000	1051.3	0.418	97.0
Beijing	4175.8964	1000	4175.9	0	558.00
Hubei	1998.4724	1000	1998.5	2.76	143.20
Shaanxi	381.2977	1000	381.3	0.027	14.54.6
Sichuan	1810.6360	1000	1810.6	0.65	72.98

As shown in Table 2, each province has 1000 purchase orders, totaling 7000. Among the purchase orders from 7 provinces, Beijing has the highest total value at 4.1758964 billion yuan, followed by Hubei and Sichuan provinces at 1.9984724 billion yuan and 1.8106360 billion yuan, respectively. Heilongjiang province has the lowest total value at 163.4346 million yuan.

3.2 Classify by procurement method

Table 3. Classification based on procurement methods

	parameter				
procurement method	Sum(mil.)	Amount	Mean(k)	Min(k)	Max(mil.)
Single source	500.0216	149	3355.8	25	72.98
Electronic shopping malls	308.3106	2615	117.9	0.0092	97.00
Designated procurement	4.3035	39	110.3	0.55	2.5653
Public bidding	7076.8202	3092	2288.8	0	153.98
Competitive consultation	1697.3316	867	1957.7	0.001	558.00
Competitive negotiation	149.9031	157	954.8	4.098	4.795
Agreement supply	3.2569	27	120.6	1.94582	0.49966
Inquiry	20.4989	47	436.1	2.55	1.7996
Invited Bidding	3.4413	7	491.6	51.476	1.91

According to Table 3, the total value of “Public bidding” is the highest, reaching 7.0768202 billion yuan. The second is significantly different from the first, and it is a “competitive consultation” with a total value of 1.6973316 billion yuan. “Agreement supply”, “Designated procurement” and “Invited Bidding” are the least, with 3.2569 million, 4.3035 million, and 3.4413 million, respectively.

In terms of quantity, the number of "Public bidding" is still the highest, which is 3092. The minimum number of “Invited Bidding” is 7.

The average price of "Single source" orders is the highest, at 3.3558 million, followed by "Public bidding", with an average price of 2.2888 million.

Overall, it can be seen from the table that the main methods of procurement are “Public bidding” and “Electronic shopping malls”. Moreover, the expensive purchases often use “Public bidding” methods, while the cheaper purchases use Electronic shopping malls. There is a small amount of expensive procurement through “single source” and “competitive consultation”.

4. HETEROGENEITY ANALYSIS OF DIFFERENT PROVINCES

4.1 T-test

T-test is used to test two types of categorical variables. For example, T-test can be used to evaluate whether two provinces have differences in purchase orders.

Taking Guangdong Province and Heilongjiang Province as examples, we obtain the following Table 4&5.

Table 4. Group statistics

	Provence	amount	mean	Std.	Mean standard error
Contract amount (10k)	Guangdong	1000	18.28593840200	56.262349295819	1.779171702868
	Heilongjiang	1000	16.34345614475	77.034275783647	2.436037693779

Table 5. Independent sample testing

		Levin's test for homogeneity of variance		Mean isotropy t-test						
									95% confidence interval for difference	
		F	Significance	t	The degree of freedom	Significance (Double tailed)	Mean value difference	Standard error difference	lower limit	upper limit
Contract amount (10k)	Assuming equal variance	.382	.536	.644	1998	.520	1.942482257248	3.016576137576	-3.973482112868	7.858446627364
	Not assuming equal variance			.644	1828.693	.520	1.942482257248	3.016576137576	-3.973814128912	7.858778643408

Based on the above table, Table 5 provides a simple data description, while Table 6 focuses on analyzing whether the differences are significant. The first step is to test whether the variance is equal, with a significance level of 0.536, indicating that the variance condition is met. Therefore, only the first row needs to be examined. From the above table, it can be seen that the significance level is 0.520, which is much greater than 0.05, indicating that there is no significant difference in purchase orders between Guangdong Province and Heilongjiang Province.

This paper conducted pairwise analysis on 7 provinces using the same independent sample T-test, and obtained the following Table 6:

Table 6. Significance of 7 provinces

	Guangdong	Heilongjiang	Zhejiang	Beijing	Hubei	Shaanxi	Sichuan
Guangdong							
Heilongjiang	0.52						
Zhejiang	0.001	0.001					
Beijing	0.001	0.001	0.001				
Hubei	0.001	0.001	0.001	0.001			
Shaanxi	0.001	0.001	0.001	0.001	0.001		
Sichuan	0.001	0.001	0.001	0.001	0.396	0.001	

It can be seen that except for "Guangdong Province" and "Heilongjiang Province", "Hubei Province" and "Sichuan Province" do not have significant differences, while the purchase orders of other provinces have significant differences.

4.2 F-test (analysis of variance)

This article conducts ANOVA analysis on purchase orders from 7 provinces simultaneously, and obtains the following Tables 7&8.

Table 7. ANOVA

Contract amount (10k)					
	Sum of squares	degree of freedom	mean square	F	Significance
Inter group	124020829.362	6	20670138.227	32.375	.000
Within the group	4464736775.846	6993	638457.997		
total	4588757605.208	6999			

Table 8. Multiple comparisons

Dependent variable: Contract amount (10k)						
LSD						
(I) Purchasing Province	(J) Purchasing Province	Mean value difference(I-J)	Standard error	Significance	95% confidence interval	
					lower limit	upper limit
Guangdong	Heilongjiang	1.942482257248	35.733961364618	.957	-68.10691931991	71.99188383440
	Zhejiang	-86.843200391000*	35.733961364618	.015	-156.89260196815	-16.79379881385
	Beijing	-399.303698963400*	35.733961364618	.000	-469.35310054055	-329.25429738625
	Hubei	-181.561298596000*	35.733961364618	.000	-251.61070017315	-111.51189701885
	Shaanxi	-19.843831116000	35.733961364618	.579	-89.893232693 15	50.205570461 15
	Sichuan	-162.777662917000*	35.733961364618	.000	-232.82706449415	-92.728261339 85
Heilongjiang	Guangdong	-1.942482257248	35.733961364618	.957	-71.991883834 40	68.106919319 91
	Zhejiang	-88.785682648248*	35.733961364618	.013	-158.83508422540	-18.736281071 09
	Beijing	-401.246181220648*	35.733961364618	.000	-471.29558279780	-331.19677964349
	Hubei	-183.503780853248*	35.733961364618	.000	-253.55318243040	-113.45437927609
	Shaanxi	-21.78631337324 8	35.733961364618	.542	-91.8357149504 0	48.2630882039 1
	Sichuan	-164.720145174248*	35.733961364618	.000	-234.76954675140	-94.670743597 09

Zhejiang	Guangdong	86.843200391000*	35.733961364618	.015	16.79379881385	156.89260196815
	Heilongjiang	88.785682648248*	35.733961364618	.013	18.73628107109	158.83508422540
	Beijing	-312.460498572400*	35.733961364618	.000	-382.50990014955	-242.41109699525
	Hubei	-94.718098205000*	35.733961364618	.008	-164.76749978215	-24.668696627785
	Shaanxi	66.999369275000	35.733961364618	.061	-3.05003230215	137.04877085215
	Sichuan	-75.934462526000*	35.733961364618	.034	-145.98386410315	-5.88506094885
Beijing	Guangdong	399.303698963400*	35.733961364618	.000	329.25429738625	469.35310054055
	Heilongjiang	401.246181220648*	35.733961364618	.000	331.19677964349	471.29558279780
	Zhejiang	312.460498572400*	35.733961364618	.000	242.41109699525	382.50990014955
	Hubei	217.742400367400*	35.733961364618	.000	147.69299879025	287.79180194455
	Shaanxi	379.459867847400*	35.733961364618	.000	309.41046627025	449.50926942455
	Sichuan	236.526036046400*	35.733961364618	.000	166.47663446925	306.57543762355
Hubei	Guangdong	181.561298596000*	35.733961364618	.000	111.51189701885	251.61070017315
	Heilongjiang	183.503780853248*	35.733961364618	.000	113.45437927609	253.55318243040
	Zhejiang	94.718098205000*	35.733961364618	.008	24.66869662785	164.76749978215
	Beijing	-217.742400367400*	35.733961364618	.000	-287.79180194455	-147.69299879025
	Shaanxi	161.717467480000*	35.733961364618	.000	91.66806590285	231.76686905715
	Sichuan	18.783635679000	35.733961364618	.599	-51.26576589815	88.833037256155
Shaanxi	Guangdong	19.843831116000	35.733961364618	.579	-50.20557046115	89.893232693155

	Heilongjiang	21.786313373248	35.733961364618	.542	-48.2630882039 1	91.8357149504 0
	Zhejiang	-66.99936927500 0	35.733961364618	.061	-137.048770852 15	3.05003230215
	Beijing	-379.4598678474 00*	35.733961364618	.000	-449.509269424 55	-309.410466270 25
	Hubei	-161.7174674800 00*	35.733961364618	.000	-231.766869057 15	-91.6680659028 5
	Sichuan	-142.9338318010 00*	35.733961364618	.000	-212.983233378 15	-72.8844302238 5
Sichuan	Guangdong	162.77766291700 0*	35.733961364618	.000	92.72826133985	232.827064494 15
	Heilongjiang	164.72014517424 8*	35.733961364618	.000	94.67074359709	234.769546751 40
	Zhejiang	75.934462526000 *	35.733961364618	.034	5.88506094885	145.983864103 15
	Beijing	-236.5260360464 00*	35.733961364618	.000	-306.575437623 55	-166.476634469 25
	Hubei	-18.78363567900 0	35.733961364618	.599	-88.8330372561 5	51.2657658981 5
	Shaanxi	142.93383180100 0*	35.733961364618	.000	72.88443022385	212.983233378 15

*. The significance level of the difference in mean values is 0.05.

Table 8 shows the results of one-way ANOVA, from which it can be seen that the F-value is 32.375, with a significance of 0.0001, indicating significant differences in purchase orders among different provinces.

Table 9 is for Post comparison, which first examines whether the significance level is significant. If significant, the magnitude is explained by looking at the difference in mean values. Except for "Guangdong Province" and "Heilongjiang Province", "Hubei Pro vince" and "Sichuan Province" do not have significant differences, purchase orders from other provinces have significant differences. This result is consistent with the t-test, and the data in this paper is reasonable and can be further analyzed.

5. ANALYSIS OF PROCUREMENT CATEGORIES BY PROVINCE

5.1 Classification of procurement categories based on natural language processing

Natural language processing is mainly applied in machine translation, automatic summarization, public opinion monitoring, question answering, viewpoint extraction, text classification, speech recognition, text semantic comparison, and other aspects[13,14,15,16]. Natural language processing technology can process a large amount of text information and improve data parsing efficiency[17].

This paper uses natural language processing to classify the procurement content of 7000 purchase orders.In this paper, natural language processing can extract keywords from longer text content in purchase orders and classify them. For example, this paper categorizes "property services" and others as "management services"; Classify "consultation",

"discussion", "consultation with experts", etc. as "consulting services"; Classify "law" and "lawyers" as "legal services"; Classify "information platforms", "technical training", etc. as "technical services".

In addition, natural language processing algorithms can also analyze whether the procurement content is "leased" or "purchased".

5.2 Sustainable procurement evaluation based on procurement categories

Among multiple **procurement** categories, some are relatively “green”, such as "management services", "consulting services", and "technical services". In addition, leasing is more environmentally friendly compared to purchasing, so this paper specifically calculates the order value and proportion of “rental services”.

We found that the proportion of printing services cannot be ignored, and the paper separately counted the order values and proportion of "Printing services" and "Printing equipment".

Table 9. Classification by procurement category

	parameter				
Procurement category	Sum(mil.)	Amount	Mean(k)	Min(k)	Max(mil.)
Rental services	65.4589	35	1870.3	15.	11.205
Management services	466.6856	375	1244.5	0.8	34.896
Technical service	99.3385	66	1505.1	1.25	14.249675
Consulting service	27.5530	23	1198.0	11.88	2.965
Legal service	3.4895	45	77.5	3	0.50
Printing services	8.3990	142	59.1	0.049	3.705066
Printing equipment	32.8463	317	103.6	0.13	4.848
Procurement	9084.4496	6015	1510.3	0	558.00

From the perspective of procurement category, the quantity and total value of "purchases" are the highest. However, the average value of “Rental services” is the highest.

5.3 procurement categories by province

As shown in Table 10, This paper presents procurement type data and proportion of purchase orders from 7 provinces.

Table 10. Analysis of procurement categories by province

region	South China		Northeast		East China		North China		Central China		Northwest		Southwest	
province	Guangdong		Heilongjiang		Zhejiang		Beijing		Hubei		Shaanxi		Sichuan	
Sum(10k)	18285.94		16343.46		105129.14		417589.6374		199847.24		38129.77		181063.60	
Procurement category	value	%	value	%	value	%	value	%	value	%	value	%	value	%
Rental services	295.20	1.61%	60.70	0.37%	420.82	0.40%	2679.16	0.64%	1787.79	0.89%	57.10	0.15%	1245.12	0.69%
Management services	7774.61	42.52%	81.69	0.50%	2284.60	2.17%	19823.94	4.75%	12445.35	6.23%	409.41	1.07%	3848.97	2.13%
Technical service	8.09	0.04%	86.68	0.53%	952.92	0.91%	5054.72	1.21%	1331.48	0.67%	316.49	0.83%	2183.47	1.21%

Consulting service	8.09	0.04%	0.00	0.00%	19.50	0.02%	2341.88	0.56%	211.50	0.11%	124.34	0.33%	50.00	0.03%
Legal service	319.00	1.74%	0.00	0.00%	29.95	0.03%	0.00	0.00%	0.00	0.00%	0.00	0.00%	0.00	0.00%
Printing services	229.03	1.25%	23.39	0.14%	0.00	0.00%	456.41	0.11%	108.96	0.05%	22.12	0.06%	0.00	0.00%
Printing equipment	214.46	1.17%	28.30	0.17%	15.92	0.02%	1987.99	0.48%	364.58	0.18%	164.35	0.43%	509.02	0.28%
Procurement	9445.56	51.65%	16062.70	98.28%	101405.42	96.46%	387410.42	92.77%	183783.08	91.96%	37110.76	97.33%	173227.02	95.67%

6. VISUALIZATION

6.1 Comparison and ranking of data

For the sake of comparison, this paper selected multiple parameters, including the proportion of rental services, the proportion of purchases, the ratio of leases to purchases, and the proportion of printing services.

This study uses visualization technology, where "green" represents sustainable procurement and "red" represents non sustainable procurement. The depth of the color represents the degree, which makes the data very intuitive.

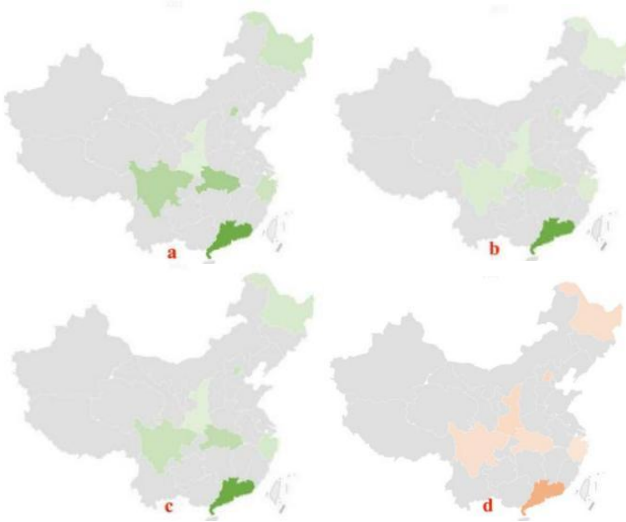

Figure. 2 Score chart of each province a) The proportion of rental services; b) The proportion of Procurement; c) The ratio of leasing to purchasing; d) The proportion of Printing

As shown in Figure. 2 a), this chart shows the proportion of rental services, with Guangdong being the greenest among all, followed by Hubei and Sichuan.

As shown in Figure. 2 b), this graph shows the proportion of purchases, and Guangdong is far "green" compared to other provinces.

As shown in Figure. 2 c), this chart shows the proportion of leasing and purchasing, with Guangdong being the greenest among all, followed by Hubei and Sichuan.

As shown in Figure. 2 d), this Figureure shows the proportion of printing, and Guangdong is far more popular than other provinces.

6.2 Score and sum based on ranking

This paper is evaluated on a scale of 1-7 based on the ranking of 7 provinces. And ultimately sum up to obtain the overall ranking.

Table 11. Province Score and Ranking Table

	Guangdong	Heilongjiang	Zhejiang	Beijing	Hubei	Shaanxi	Sichuan
The proportion of rental services	1.61%	0.37%	0.40%	0.64%	0.89%	0.15%	0.69%
score	1	6	5	4	2	7	3
The proportion of Procurement	51.65%	98.28%	96.46%	92.77%	91.96%	97.33%	95.67%
score	1	7	5	3	2	6	4
The ratio of leasing to purchasing	3.12%	0.38%	0.41%	0.69%	0.97%	0.15%	0.72%
score	1	6	5	4	2	7	3
The proportion of Printing	2.42%	0.31%	0.02%	0.59%	0.23%	0.49%	0.28%
score	7	4	1	6	2	5	3
the total score	10	23	16	17	8	25	13
Final ranking	2	6	4	5	1	7	3

As shown in Table 11, The higher the proportion of rental services, the higher the score for sustainable procurement. Guangdong ranks first, followed by Hubei, Shaanxi and Heilongjiang, which have the lowest ranking.

The evaluation of procurement is opposite, and the lower the proportion of procurement, the higher the evaluation of sustainable procurement. Guangdong ranks first, followed by Hubei, Heilongjiang and Shaanxi, which have the lowest ranking.

The ratio of leasing to procurement represents a visual representation of the proportion of the two procurement categories within the province. Guangdong ranks first, followed by Hubei, Shaanxi and Heilongjiang, which have the lowest ranking.

The proportion of printing has a negative impact on sustainable procurement, so this paper believes that those with a lower proportion of printing have a higher ranking. Zhejiang ranks highest, followed by Hubei and Sichuan. Guangdong has the lowest ranking.

The province that ultimately ranks first is Hubei, followed by Guangdong, by Sichuan, Zhejiang, Beijing, Heilongjiang, and Shaanxi.

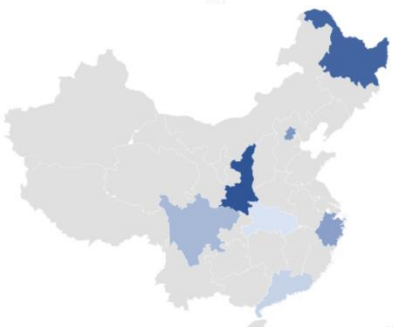

Figure. 3 Score chart of each province

As shown in Figure 3, the darker the color, the higher the score, and the poorer the effectiveness of sustainable procurement.

7. CONCLUSION

This paper analyzes purchase orders from multiple provinces and introduces the evaluation of sustainable procurement, analyzing data from each province to obtain a satisfactory result.

This paper uses SPSS software for data analysis, NLP algorithm for text recognition and classification, and visualization technology to display the results. This paper simplifies and organizes a pile of irregular and disorderly data to make it easy and intuitive results. This has reference significance for analyzing supply chain management data in the future.

From the results, Hubei(Central China) and Guangdong(South China) are the most in line with sustainable procurement, while Heilongjiang(Northeast) and Shaanxi(Northwest) perform the worst.

In future research, we will increase the number of provinces and expand the data volume for each province.

8. ACKNOWLEDGMENT

This paper is fund by The Dean's Fund Project of China National Institute of Standardization "Sustainable procurement model, instrucment, and standard de-velopment for central enterprises" (Proect Number:602023Y-10399).

9. REFERENCES

[1] R. Ruparathna, K. Hewage, Sustainable procurement in the Canadian construction industry: current practices, drivers and opportunities, J. Clean. Prod. 109(2015) 305 – 314, https://doi.org/10.1016/j.jclepro.2015.07.007.

[2] R. Ruparathna, K. Hewage, Review of contemporary construction procurement practices, J. Manag. Eng. 31 (2013), 04014038, https://doi.org/10.1061/(ASCE) ME.1943-5479.0000279.

[3] S. Sharma, I. Henriques, Stakeholder influences on sustainability practices in the Canadian forest products industry, Strat. Manag. J. 26 (2) (2005) 159 – 180, https://doi.org/10.1002/smj.439..

[4] D. Simpson, D. Power, D. Samson, Greening the automotive supply chain: a relationship perspective, Int. J. Oper. Prod. Manag. 27 (1) (2007) 28 – 48, https://doi. org/10.1108/01443570710714529.

[5] Donatus Ebere Okonta, 2023. The scientometric analysis and visualization of sustainable procurement. Heliyon.

[6] M. Overvest, Sustainable Procurement - the Ultimate Guide of 2023. Procurement Tactics, Retrieved, 2023. from, https://procurementtactics.com/sustainableprocurement/. (Accessed 22 February 2023).

[7] Tay, H.L., Loh, H.S., 2021. Digital transformations and supply chain management: a Lean Six Sigma perspective. J. Asia Bus. Studies 16 (2), 340 – 353.

[8] Vu, K., Asongu, S., 2023. Patterns and drivers of financial sector growth in the digital age: insights from a study of industrialized economies. Res. Int. Bus. Finance 66, 102075.

[9] Pingrui Li, Xu Zhao, 2024. The impact of digital transformation on corporate supply chain management: Evidence from listed companies. Finance Research Letters.

[10] Finch, W. H., & Hern´and ez Finch, M. E. (2020). Poverty and covid-19: Rates of incidence and deaths in the United States during the first 10 weeks of the pandemic. Frontiers in Sociology, 5, 47.

[11] Andrea V azquez-Ingelmo, Alicia García-Holgado, Sonia Verdugo-Castro, Roberto Ther´on, Francisco Jos´e García-Pe˜nalvo, Data visualization and domain knowledge: Insights through focus groups of researchers in Spain. Computers in Human Behavior.

[12] Daniel Schuster, Francesca Zerbato, Sebastiaan J. van Zelst, Wil M.P. van der Aalst. Defining and visualizing process execution variants from partially ordered event data. Information Sciences.

[13] Barber, E. L., Garg, R., Persenaire, C., & Simon, M. (2021). Natural language processing with machine learning to predict outcomes after ovarian cancer surgery. Gynecologic Oncology, 160(1), 182-186. https://doi.org/10.1016/j.ygyno.2020.10.004

[14] Falissard, B., Simpson, E. L., Guttman-Yassky, E., Papp, K. A., Barbarot, S., Gadkari, A., Saba, G., Gautier, L., Abbe, A., & Eckert, L. (2020). Qualitative Assessment of Adult Patients' Perception of Atopic Dermatitis Using Natural Language Processing Analysis in a CrossSectional Study. Dermatology and Therapy, 10(2), 297-305. https://doi.org/10.1007/s13555-020-00356-0

[15] Karmen, C., Hsiung, R. C., & Wetter, T. (2015). Screening Internet forum participants for depression symptoms by assembling and enhancing multiple NLP methods. Computer Methods and Programs in Biomedicine, 120(1), 27-36. https://doi.org/10.1016/j.cmpb.2015.03.008

[16] Řehůřek R, Sojka P. Software Framework for Topic Modelling with Large Corpora. In: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks. Valletta, Malta: ELRA; 2010, p. 45 – 50, http://is.muni.cz/publication/884893/en.
[17] Esther Lázaro, Jean-Carel Yepez, Vanessa Moscardó, Patricia Marín, Patricia LópezMasés, Tamara Gimeno, Samuel de Paúl. Efficiency of natural language processing as a tool for analysing quality of life in patients with chronic diseases. A systematic review. Computers in Human Behavior Reports. 2023.