

Find me if you can! Identification of services on websites by human beings and artificial intelligence

Abstract. This paper presents the results of evaluating the number of service offerings marketed on company webpages in the field of advanced manufacturing performed by human annotators, artificial intelligence imitating human behavior, and by advanced in-depth AI analysis. The research focuses on ten different countries and three specific sectors: manufacture of computer, electronic and optical product, manufacture of electrical equipment, and manufacture of machinery and equipment. Even though the artificial intelligence was able to find more services on company webpages than human beings, the average number of services identified on webpages proved to be relatively low. Companies from selected industries focus mostly on product lifecycle services, such as spare parts, repair services or maintenance. On the other hand, the least marketed services are in the field of finances, such as pay per use and instalment payment. The results support the applicability and effectiveness of artificial intelligence tools in the field of procurement. Moreover, the findings are highly relevant for companies in the field of advanced manufacturing as they indicate a great potential for further improvement of service promotion on company webpages.

Keywords: artificial intelligence, servitization, manufacturing, service promotion.

1 Introduction

Companies and scholars have realized the impact of the Internet on the purchasing process. Almost 75% of buyers in the B2B segment prefer searching and selecting goods and services online instead of communicating and engaging with a salesperson [1, 2]. Moreover, as purchasing is an integral part of successful business operations and can provide significant value, researchers and company representatives look for ways to utilize new technology to optimize the process as well as to find suitable suppliers and service providers [3, 4].

As buyers commonly browse the Internet to find suitable offerings and companies move their presence online, it is imperative that suppliers promote their products and service offerings on their webpages and social media.

2 Promoting services on company webpages

As global markets have been changing rapidly and became highly competitive, complex and often unpredictable, companies are looking for new ways to ensure their survival and profitability [5–7]. Current literature points to a servitization trend among manufacturers, where combining product offerings with services is perceived as a stra-

tegic alternative to developing new innovative products [6, 8–13]. Yielding considerable advantages, such as more stable revenue streams, improved profitability, or increased customer satisfaction and loyalty, services help companies to differentiate and create a sustainable source of competitive advantage [14]. With the rapid rate of new technology adoption in organizations, the focus in the context of servitization has been shifting towards potentials and benefits of utilizing the Internet and digital technologies in business processes [15, 16]. Companies have moved their presence and marketing activities online and procurement personnel commonly browses company webpages to obtain information about service characteristics and prices and to search for the most suitable offer [17, 18].

Buyers increasingly rely on information they find on the Internet while identifying new business opportunities, conducting market research or preliminary identification of suppliers [19, 20]. The most significant benefits of online research in procurement are the ability to compare products, services, and prices, the speed and ease of use (finding information they need effortlessly), and convenience [21, 22]. Moreover, reviewing product or service information online increases the accuracy of goods purchased [23].

Manufacturing companies can take advantage of these procurement trends as their websites might reach customer segments, which might not have been possible to target in the past [24]. Proper design of webpages in line with company market positioning enables to promote key quality elements and establish reputation – especially in B2B service segments [24, 25]. Literature suggests that having a company webpage is linked to higher revenues due to increased familiarity with products and services offered [26]. However, in order to exploit the potential presented by company webpages, certain requirements have to be fulfilled. Aspects determining the user intention to use webpages during the procurement process are the accuracy of information provided on the webpage, how easy information can be found, and buyers' perception of the usefulness of the webpage [22, 27]. Trustworthy presentation of information enhances perceived value [18]. The main aim of the webpage design should be customer satisfaction, which can be increased by providing detailed product or service information in a simple yet attractive format. Service providers have to adequately signal on their webpage what services they offer and that they are capable of facilitating high-quality services [25]. Therefore, this paper focusses on how intensively manufacturing companies promote services on their company webpages [28]. To obtain this evaluation of service promotion on webpages, various methods of webpage analysis were employed to fully understand not only if services are offered on company webpages but also if they are offered in a manner which enables human beings to find them.

Hence, this article examines and answers the following research questions:

- How intensively are services promoted on webpages in the field of advanced manufacturing?
- What are the differences in service research on company webpages performed by humans, AI solutions imitating human behavior, and in-depth analysis using improved AI algorithm?

3 Methodology

3.1 Empirical context – Service typology in the field of advanced manufacturing

The European Union defines advanced manufacturing as “the use of knowledge and innovative technology to produce complex products [...] and improve processes to lower waste, pollution, material consumption and energy use.” [29] Important elements are artificial intelligence, robotics, 3D and 4D printing, and high-performance computing for modelling. The industry is an important segment of the economy and labor market in the European Union, with 14.5 million jobs and a strategic focus in many regional and national strategies [29–31].

The research presented in this article focuses on the webpages of companies in the field of advanced manufacturing in three specific sectors: manufacture of computer, electronic and optical product (NACE 26), manufacture of electrical equipment (NACE 27) and manufacture of machinery and equipment (NACE 28). To identify and structure services, which would be analyzed on company webpages, the research consortium adjusted the service typology presented by Partanen [32]. The main service categories utilized in this research are pre-sales services, product support services, product lifecycle services, as well as R&D, operational, and financial services. A detailed overview of the categories assessed can be found in Table 1.

The research focused on and was conducted for the following countries: Austria, Germany, Czech Republic, Hungary, Italy, Slovakia, Slovenia, France, United Kingdom, and the United States of America. The lists of companies from each country were obtained from national and international databases such as Europages, Aurelia and other. Only companies operating in the specified sectors (NACE 26,27,28) were analyzed. To ensure accurate representation of companies within each industry, companies with more than 20 employees were selected regardless of whether they operate nationally or internationally.

3.2 Applied methods for analyzing webpages

In order to be able to assess the degree of service promotion on company webpages in the field of advanced manufacturing as well as to identify differences in various methods of company webpage service search, the consortium developed a research process comprising of three steps. Firstly, the traditional market research was performed using human annotators to investigate selected company webpages (see part 3.3). The data set gathered in this step also formed the basis for the training of the developed artificial intelligence algorithms. This was accomplished by manually extracting specific information from company webpages in selected industries. To standardize the process, a scheme to perform the annotations as well as a structured template were designed to extract and classify the data. In the second research phase, the artificial intelligence algorithm was developed in order to imitate human behavior and a large-scale automated webpage research was executed. In the third step, an advanced artificial intelligence algorithm was created and implemented to in-depth analyze service offerings on

company webpages. The final artificial intelligence algorithm should be able to answer the general question: “Which companies promote which kind of services on their webpages?”. In the final stage of the research, the results from all three methods were compared.

3.3 Human behavior: Manual annotations

In total 1.487 company webpages were manually analyzed through annotations between March and September 2020 with the objectives to understand the webpage promotion of services and to generate training data for AI algorithms. The analysis was conducted in English, German, French, Italian, Slovak, Slovenian, Hungarian, and Czech languages by twenty-two native-speakers.

In the first method, data from company webpages was manually extracted according to a comprehensive structure, categorized, and subsequently analyzed. In order to collect the data in a structured manner, a detailed manual was designed. The researchers from seven different countries were carefully selected and trained. They were graduate or postgraduate students from the field of business and/or technology, had domain knowledge, and were native speakers of the languages specified in the text above (with the only exception being the English language). The scheme used for the analysis consisted of specific information, which was to be retrieved from company webpages - company name, URL, country, language, location of headquarters, number of employees, NACE code, identified services from the list of services (Table 1).

In order to standardize the process and minimize the discrepancies between different researchers, a trial process of annotation was applied. After all researchers collected company information according to the predefined structure, the data was cross-checked, differences among the results were identified, and a comprehensive feedback was presented to the researchers in order to improve the manual research process. This process was repeated in three loops to ensure high quality of the data collection.

For the quantitative measurement of the inter annotator agreement, we had 15 cases when the same company website was annotated by a pair of independent annotators. The annotation process was open ended, i.e., annotators were not provided with a fixed set of websites from a given URL to check but a base URL of a company was given to them and then they were allowed to check any of the in-domain webpages on the given URL. As such, a pair of annotators could find evidence for a particular service being offered by some company on two distinct pages of the same company website. In order to account for that, we aggregated the services found by the individual annotators at the company level, i.e., the result of each annotation with respect a company website consisted of binary values, each one indicating whether the annotator deemed a service to be offered based on the contents of a company website.

We calculated the Jaccard coefficient over the 15 pairs of annotations, more specifically, we took the fraction of the aggregated number of such services that were identified in consensus by a pair of annotators and that of such services that were found to be relevant by at least one of the annotators [33]. Besides this empirical quotient, we also quantified the expected value of the Jaccard coefficient, more specifically, the quantity that we would obtain if the annotators were choosing the same number of services as

they did during the annotation, except for the fact that they chose the selected number of services randomly. Assuming that a pair of annotators identified m and n services for a given company website out of S potential services being distinguished, the expected value of their Jaccard coefficient can be given as $\frac{nm}{Sn+Sm-nm}$.

The observed Jaccard coefficient ended up 0.377, whereas the expected value calculated as described above was 0.143. A typical way to contrast observed (o) and expected (e) values of random variables is to calculate the fraction $\frac{o-e}{1-e}$, resulting in 0.272 in our particular case. This value is not particularly high, which illustrates the complexity of the annotation process. This value, however, also indicates that the annotators agreed to a non-trivial extent.

Twenty-two researchers manually extracted data from 1487 company webpages; 613 annotations were performed by the Austrian researchers, 193 by the Hungarian, 198 by the German, 198 by the Italian, 97 by the Slovenian, 94 by the Slovak and 94 by Czech investigators. Depending on the amount of information presented, each company webpage analysis took on average 20 minutes to complete.

Discrepancies among the investigators were overcome and limited due to the quality loop controls, which standardized the research process and the interpretation methods. During the investigation process, differences among the webpages regarding their professionalism were identified. Moreover, certain service categories (such as “customer seminars” or “technical user training”) were not sufficiently distinguished on the webpages by companies. These hurdles were diminished through the domain knowledge of the researchers and peer-reviews, discussions and interpretations of these challenging webpages or service presentations. Furthermore, a critical mass of company webpages with higher content quality was guaranteed in this process, in order to generate the basis for the second method and to prevent and limit vagueness of information provided.

3.4 Analysis using artificial intelligence

In the second step, the data obtained from the manual webpage analysis was used as training data to perform identification of offered services on webpages. Based on the annotations, we trained a logistic regression classifier, which assigns a weight for every combination of words from the training data and service categories investigated. A higher weight indicated that some word was more likely to serve as evidence for a particular kind of service being offered.

The input features of our logistic regression model were derived from the individual words included on the analyzed websites. In order to overcome the limited size of the training data and to be able to process websites in multiple languages, we replaced every content word found on the websites by their 5 most similar English synonyms/translations.

In order to do so, we performed an automatic language detection on the contents of the websites and relied on cross-lingual word embeddings (CLWEs) to find the most likely synonym/translation of a word [34]. Word embeddings assign a vector to every word form in a way that words with similar meaning get assigned to vectors with similar

orientations. CLWEs also have the desirable property, that they are insensitive to the particular language a word is written in, e.g., it assigns a similar vector to the German word ‘*Hund*’ and the English word ‘*dog*’, as they have the same meaning. We used our logistic regression classifier for the individual webpages of a company website traversing the websites in a priority queue. The priority queue contained at most 200 websites that could be reached via at most clicking 3 hyperlinks from the start page. The priority queue contained the URLs ranked by an additional model which tries to predict the likelihood that the given URL contained some evidence being offered by the company based on the URL itself, e.g., an URL containing */our-services/* is more likely to contain mentions of services *a priori* compared to another website which contains */contacts/*. The limit for checking up at most 200 websites per company websites at most 3 hops away from the starting page was introduced for performance reasons and also as human annotations generally behaved like that as well. As these choices resembled that of the annotators, we refer to this variant of our approach as our model imitating human behavior.

We also made a more permissive variant of our approach which also considered the analysis of those webpages that were up to 10 clicks away from the initial website to start the analysis from. We refer to this kind of analysis as the in-depth analysis hereinafter.

The logistic regression classifier can assign a probability value of every webpage promoting a certain service. For a company website we calculate all these probability values for all the services towards all the (at most 200) webpages our ranking module selects for analysis, then for each service category choose the highest probability value obtained and take it as the probability of the given company offering that given service. We conclude that a given service is offered by a company whenever the corresponding probability exceed 0.5, i.e., it is more likely to be offered than not.

4 Findings

Table 1. Results of service research on company webpages based on the service typology developed by Partanen et al. 2017

No.	Service category	Human Annotations		Prediction (AI imitating human behavior)		Prediction (AI imitating human behavior)		Prediction (in-depth AI analysis)		Prediction (in-depth AI analysis)	
		n		n		n		n		n	
	Sample size	n = 1487		n = 1487		n = 7740		n = 1487		n = 7740	
	Number (%) of identified services	#	%	#	%	#	%	#	%	#	%
1	Pre-sales services										
11	Product	41	2.76	141	9.48	694	8.97	168	11.30	821	10.61

	Demonstrations										
12	Customer Seminars	57	3.83	199	13.38	1125	14.53	244	16.41	1340	17.31
2	Product support Services										
21	Warranty	157	10.56	343	23.07	1133	14.64	376	25.29	1295	16.73
22	Technical user training	288	19.37	415	27.91	1493	19.29	459	30.87	1670	21.58
23	Customer consulting and support by phone	251	16.88	460	30.93	1427	18.44	489	32.89	1579	20.40
24	Testing, test rigs, quality assurance	135	9.08	322	21.65	1320	17.05	351	23.60	1437	18.57
3	Product lifecycle Services										
31	Installation Services	294	19.77	512	34.43	2172	28.06	546	36.72	2372	30.65
32	Repair service	335	22.53	527	35.44	2255	29.13	562	37.79	2460	31.78
33	Spare parts	387	26.03	513	34.50	1998	25.81	554	37.26	2213	28.59
34	Maintenance	380	25.55	548	36.85	2088	26.98	578	38.87	2257	29.16
35	Retrofit, modernization, upgrades	200	13.45	353	23.74	1204	15.56	379	25.49	1356	17.52
4	R&D services										
41	Research service	86	5.78	319	21.45	1433	18.51	354	23.81	1575	20.35
42	Prototype design and development	227	15.27	470	31.61	2365	30.56	510	34.30	2499	32.29
43	Feasibility studies	63	4.24	237	15.94	1309	16.91	280	18.83	1471	19.01
5	Operational services										
51	Project management	118	7.94	290	19.50	1664	21.50	327	21.99	1829	23.63
52	Service for operating the product for the customer	7	0.47	23	1.55	85	1.10	29	1.95	104	1.34
53	Service for operating customer's processes	28	1.88	138	9.28	684	8.84	169	11.37	809	10.45
6	Financial services										
61	Pay per use	3	0.20	9	0.61	34	0.44	14	0.94	52	0.67
62	Installment Payment	16	1.08	25	1.68	77	0.99	31	2.08	93	1.20
63	Leasing	23	1.55	43	2.89	108	1.40	50	3.36	137	1.77
64	Rental system	56	3.77	74	4.98	299	3.86	88	5.92	349	4.51

Table 2. The average number of services identified on company webpages based on different research methods applied

Service category	Human Annotations	Prediction (AI imitating human behavior)	Prediction (AI imitating human behavior)	Prediction (in-depth AI analysis)	Prediction (in-depth AI analysis)
Sample size	n = 1487	n = 1487	n = 7740	n = 1487	n = 7740
Average number of services identified on company webpages	2.12	4.01	3.23	4.41	3.58

The results in Table 1 show considerable differences between the website research performed by human annotators and by the AI algorithm imitating human behavior. Even greater disparity can be observed when comparing human annotations with the results provided by the in-depth AI analysis. In all the service categories analyzed, the AI was more likely to find service offerings on company webpages compared to the annotations performed by humans. These discrepancies might be the result of ineffective and imprecise promotion of services on the webpages resulting in lower visibility of offerings for human researchers.

The service category, which is promoted by companies on the webpages in the highest rate is product lifecycle services. Within this service group, humans found the most service offerings in spare parts category (26.03%), whereas the artificial intelligence detected the highest percentage of services in maintenance (38.87%). On the contrary, the category with the lowest rate of service offerings as identified both by humans and the AI is the financial services, more specifically pay per use services, where human annotations identified that only 0.2% of companies offer these services, compared to 0.94% identified by the in-depth AI analysis.

Table 2 shows that the average number of service offerings identified on a single webpage analyzed by humans is 2.12, compared to the results of 4.01 and 4.41 found by human-imitating AI and in-depth AI analysis respectively. Even though manufacturing is considered a sector with a high rate of servitization compared to other industries, all three figures show a relatively low number of services offered by companies, indicating a need for improvement and increase in service promotion in the industry.

5 Outlook and limitations

The article contributed to literature in a relatively new and rapidly developing field. However, further research and verification is advised, such as a qualitative investigation analyzing the reasons of underrepresentation of services. Moreover, further quantitative research could provide insight and empirical understanding of companies' point of view and verify the results. This research and article also highlight the importance and necessity of action in the field of management and marketing of services, as the lack of service offerings on company webpages or their poor visibility pose a significant challenge for procurement performed by human employees. The results presented support

the applicability and effectiveness of artificial intelligence tools in the field of procurement due to the low quality of information provided on webpages of suppliers. Further research should also establish if the lack of service offerings on company webpages stems from insufficient marketing or if companies are indeed not offering many services and are, therefore, not presented on webpages.

Certain biases might have affected the results in the article. Insufficient or noisy annotations by humans might have led to skewed results, which is the particularly important for services with large gaps between the different stages of the research process. Annotators might have had different labelling preferences and styles, which could have resulted in discrepancies of information collected. Moreover, a sampling bias might have occurred due to the selection of particular types of companies.

References

1. Koponen, J.P., Rytty, S.: Social presence and e-commerce B2B chat functions. *EJM* (2020).
2. Arli, D., Bauer, C., Palmatier, R.W.: Relational selling. Past, present and future. *Industrial Marketing Management* (2018).
3. Wynstra, J.Y.F.: Past, present and future trends of purchasing and supply management: An extensive literature review. A journey through manufacturing and supply chain(199–228) (2016)
4. Pulles, N.J., J. Veldman, H. Schiele: Identifying innovative suppliers in business networks: An empirical study. *Industrial Marketing Management* **43**(3), 409–418 (2014)
5. Anderson, J.C., Narus, J.A.: Capturing the Value of Supplementary Services. *Harvard Business Review* **73**, 75–83 (1995)
6. Vandermerwe, S., Rada, J.: Servitization of business. Adding value by adding services. *European Management Journal* (1988).
7. Dimache, A., Roche, T.: A decision methodology to support servitisation of manufacturing. *Int Jnl of Op & Prod Mngemnt* (2013).
8. Carlborg, P., Kindström, D., Kowalkowski, C.: The evolution of service innovation research. A critical review and synthesis. *The Service Industries Journal* (2013).
9. Baines, T.S., Lightfoot, H.W., Evans, S., Neely, A., Greenough, R., Peppard, J., Roy, R., Shehab, E., Braganza, A., Tiwari, A., Alcock, J.R., Angus, J.P., Bastl, M., Cousens, A., Irving, P., Johnson, M., Kingston, J., Lockett, H., Martinez, V., Michele, P., Tranfield, D., Walton, I.M., Wilson, H.: State-of-the-art in product-service systems. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* (2007).
10. Bustinza, O.F., Bigdeli, A.Z., Baines, T., Elliot, C.: Servitization and Competitive Advantage. The Importance of Organizational Structure and Value Chain Position. *Research-Technology Management* (2015).
11. Parida, V., Rönnerberg Sjödin, D., Wincent, J., Kohtamäki, M.: Mastering the transition to product-service provision: Insights into business models, learning

- activities, and capabilities. *Research Technology Management* **57**(3), 44–52 (2014)
12. Kamal, M.M., Sivarajah, U., Bigdeli, A.Z., Missi, F., Koliouisis, Y.: Servitization implementation in the manufacturing organisations. Classification of strategies, definitions, benefits and challenges. *International Journal of Information Management* (2020).
 13. Rabetino, R., Kohtamäki, M., Brax, S.A., Sihvonen, J.: The tribes in the field of servitization. Discovering latent streams across 30 years of research. *Industrial Marketing Management* (2021).
 14. Guo, A., Li, Y., Zuo, Z., Chen, G.: Influence of organizational elements on manufacturing firms' service-enhancement. An empirical study based on Chinese ICT industry. *Technology in Society* (2015).
 15. Coreynen, W., Matthyssens, P., van Bockhaven, W.: Boosting servitization through digitization. Pathways and dynamic resource configurations for manufacturers. *Industrial Marketing Management* (2017).
 16. Gago, D., Rubalcaba, L.: Innovation and ICT in service firms. Towards a multi-dimensional approach for impact assessment. *J Evol Econ* (2007).
 17. Bienhaus, F., Haddud, A.: Procurement 4.0. Factors influencing the digitisation of procurement and supply chains. *BPMJ* (2018).
 18. Chang, Y.-Y., Lin, S.-C., Yen, D.C., Hung, J.-W.: The trust model of enterprise purchasing for B2B e-marketplaces. *Computer Standards & Interfaces* (2020).
 19. Konuk, F.A.: The role of store image, perceived quality, trust and perceived value in predicting consumers' purchase intentions towards organic private label food. *Journal of Retailing and Consumer Services* (2018).
 20. Davila, T., Gupta, M.R., Palmer, R.J.: Moving Procurement Systems to the Internet. The Adoption and Use of E-Procurement Technology Models. *SSRN Journal* (2002).
 21. Bridges, E., Goldsmith, R.E., Hofacker, C.F.: Attracting and Retaining Online Buyers. In: Clarke, I., Flaherty, T.B. (eds.) *Advances in Electronic Marketing*, pp. 1–27. IGI Global (2005)
 22. Merrilees, B., Fenech, T.: From catalog to Web. B2B multi-channel marketing strategy. *Industrial Marketing Management* (2007).
 23. Vaidyanathan, G., Devaraj, S.: The role of quality in e-procurement performance. An empirical analysis. *Journal of Operations Management* (2008).
 24. Williams, Z., Lueg, J.E., Hancock, T., Goffnett, S.P.: Positioning through B2B carrier signals. Understanding how service quality is communicated via websites. *Industrial Marketing Management* (2019).
 25. Connelly, B.L., Certo, S.T., Ireland, R.D., Reutzel, C.R.: Signaling Theory. A Review and Assessment. *Journal of Management* (2011).
 26. Liu, C., Arnett, K.P., Capella, L.M., Beatty, R.C.: Web sites of the Fortune 500 companies. Facing customers through home pages. *Information & Management* (1997).
 27. Zhang, Y., Fang, Y., Wei, K.-K., Ramsey, E., McCole, P., Chen, H.: Repurchase intention in B2C e-commerce—A relationship quality perspective. *Information & Management* (2011).

28. Liu-Thompkins, Y.: A Decade of Online Advertising Research. What We Learned and What We Need to Know. *Journal of Advertising* (2019).
29. European Commission: Why the EU supports advanced manufacturing. https://ec.europa.eu/info/research-and-innovation/research-area/industrial-research-and-innovation/key-enabling-technologies/advanced-manufacturing_en (2021). Accessed 21 July 2021
30. European Commission: Horizon 2020. <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/nanotechnologies-advanced-materials-advanced-manufacturing-and-processing-and> (2020). Accessed 21 July 2021
31. European Commission: Regional Innovation Monitor Plus. <https://ec.europa.eu/growth/tools-databases/regional-innovation-monitor/news/challenges-ris3-implementation-and-advanced-manufacturing-region-bratislava> (2021). Accessed 21 July 2021
32. Partanen, J., Kohtamäki, M., Parida, V., Wincent, J.: Developing and validating a multi-dimensional scale for operationalizing industrial service offering. *JBIM* (2017).
33. Deza, M.M., Deza, E.: *Encyclopedia of Distances*. Springer, London (2016)
34. Joulin, A., Bojanowski, P., Mikolov, T., Jégou, H., Grave, E.: Loss in Translation. Learning Bilingual Word Mapping with a Retrieval Criterion. In: Riloff, E., Chiang, D., Hockenmaier, J., Tsujii, J.'i. (eds.) *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, 01-Oct-18 - 01-Nov-18*, pp. 2979–2984. Association for Computational Linguistics, Stroudsburg, PA, USA (2018).