MARGARETHE ÜBERWIMMER
ROBERT FÜREDER
MICHAEL SCHMIDTHALER

PROCEEDINGS
CCBC 2021

CROSS-CULTURAL BUSINESS CONFERENCE 2021

May 6th, 2021

School of Business and Management, Steyr Campus

Intercultural or International Perspectives in

- Global Business and Export Management
- Marketing, Sales and Service Management
- Higher Education Research, Teaching and Learning
- Innovation and Entrepreneurship
- Human Resource Management

www.fh-ooe.at/gsm
www.global-business.at

IAUP
International Association of University Presidents

UNIVERSITY OF APPLIED SCIENCES UPPER AUSTRIA

PROCEEDINGS 2021 CROSS-CULTURAL BUSINESS CONFERENCE 2021 Überwimmer et al. (Eds.)

Automatic information retrievement for exporting services: First project findings from the development of an AI based export decision supporting instrument

David Aufreiter[1], Doris Ehrlinger[1], Christian Stadlmann[1], Margarethe Überwimmer[1], Anna Biedersberger[2], Christina Korter[2], Stefan Mang[2]

(1) University of Applied Sciences Upper Austria, Austria;
(2) University Passau, Germany

ABSTRACT

On the servitization journey, manufacturing companies complement their offerings with new industrial and knowledge-based services, which causes challenges of uncertainty and risk. In addition to the required adjustment of internal factors, the international selling of services is a major challenge. This paper presents the initial results of an international research project aimed at assisting advanced manufacturers in making decisions about exporting their service offerings to foreign markets. In the frame of this project, a tool is developed to support managers in their service export decisions through the automated generation of market information based on Natural Language Processing and Machine Learning. The paper presents a roadmap for progressing towards an Artificial Intelligence-based market information solution. It describes the research process steps of analyzing problem statements of relevant industry partners, selecting target countries and markets, defining parameters for the scope of the tool, classifying different service offerings and their components into categories and developing annotation scheme for generating reliable and focused training data for the Artificial Intelligence solution. This paper demonstrates good practices in essential steps and highlights common pitfalls to avoid for researcher and managers working on future research projects supported by Artificial Intelligence. In the end, the paper aims at contributing to support and motivate researcher and manager to discover AI application and research opportunities within the servitization field.

1 INTRODUCTION

Emerging technologies, such as Artificial Intelligence, prove to have the potential to give companies a competitive advantage in an increasingly competitive global market situation. Artificial Intelligence applied in marketing enables companies to increase revenues and reduce costs. Businesses are able to make faster business decisions, which are based on outputs of cognitive technologies. Improved marketing decision may further lead to increased revenue streams by identifying and maximizing sales opportunities. Whereas due to the automation and optimization potential of simple marketing tasks, customer services and structured market transactions costs may decrease (Davenport et al. 2020). Moreover, Artificial Intelligence may facilitate the customers' product and service offerings. Servitization is the shift of manufacturers from a product focused business model towards more customer-centric service offerings (Oliva and Kallenberg 2003). This transformation can benefit of the deployment of Artificial Intelligence. AI tools help decision makers to navigate the future path in business. Through AI and Machine learning algorithms, enterprises are equipped to process huge amounts of data

in an efficient way, which allow them to formulate the better business decision (Verma et al. 2020).
This paper is based upon the first results of an international research project, which aims at using AI possibilities in the field of service export. One of the main project tasks is to develop a tool, which finally supports Advanced Manufacturers in their service export decision by automatically providing critical service market information with the use of Artificial Intelligence. Academic research on machine learning modeling limitedly addresses the transfer into business. Difficulties are among others how to define business use cases for an AI application or how to convert business requirements into the world of data scientists (Akkiraju et al. 2020). Based on this research project, the authors aim at contributing to a better transfer of AI possibilities into the field of service export by presenting a procedure how to develop an AI solution step-by-step. Therefore, this paper targets to answer the following research questions:

- What market information is needed to support the decision making on the service export of Advanced Manufacturers?
- How can this respective information be collected with the help of Artificial Intelligence?

To answer the questions, the paper is structured as follows: Aligned with the project process, first the target group of Advance Manufacturers is defined. Second the field of service export decisions is separated into operational aspects and strategical questions. Further, the geographic markets and related languages, which shall be supported by the AI tool are selected. Then the methods of Artificial Intelligence being used are specified and the publicly available sources for continuous data collection are determined. Ultimately, the necessary steps to develop an Artificial Intelligent solution to support managers in the export service decision making are described and the project learnings summarized.
The European Union determines advanced manufacturing as the use of knowledge and innovative technologies to produce complex products and improve processes to lower waste, pollution, material consumption and energy use. Important elements in advanced manufacturing are robotics, 3D and 4D printing, artificial intelligence and high performance computing for modelling (Izsak, Perez, and Kroll 2020). Therefore, Advanced Manufacturer are defined as the enablers of advanced manufacturing solutions. To integrate an operational perspective, the classification of companies according to the NACE code can be taken, which comes with a variety of benefits. Using an international recognized classification facilitates comparing the individual companies in the different countries. The enablers of advanced manufacturing solutions are companies that operate mainly in three business sectors according to the Statistical Classification of Economic Activities in the European Community (or NACE): in Code 26, Manufacture of computer, electronic and optical products; Code 27, Manufacture of electrical equipment and Code 28, Manufacture of machinery and equipment (European Communities 2008). The paper follows the project team from the initial idea to the collection of trainings data for the Artificial Intelligence.

2 RESEARCH METHODOLGY

According to Block and Block (2005) every research project follows a series of decisions and steps and therefore follows some sort of process. A research process in a very general form is outlined in the figure 1. Usually, the process progresses from deciding what is to be researched, to select a research design, defining the sampling parameters and data collecting instruments, to conduct the study, as well as preparing and analyzing the data and ultimately reporting the

results to answer the research objectives are formulated in the beginning. The research project to develop an AI based solution to support managers in the decision making process regarding the export of services followed a similar process.

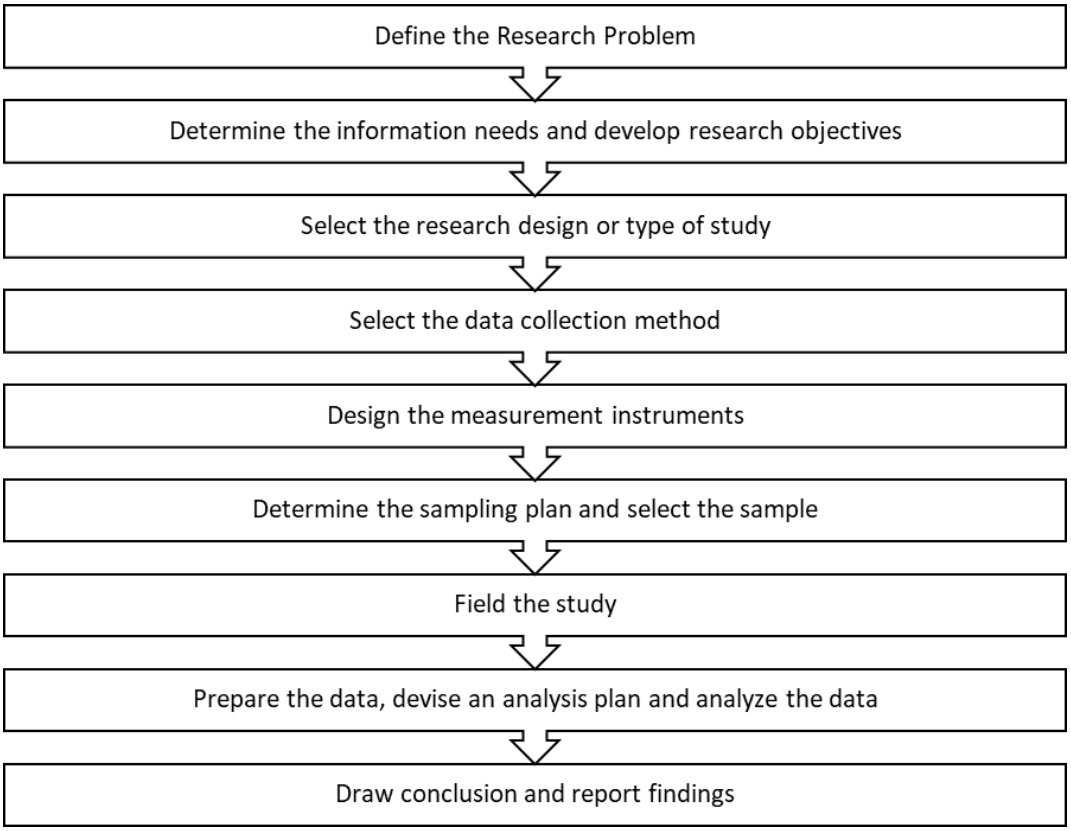

Figure 1. The research process (Block and Block 2005)

After establishing the research questions, the project team had to decide on various aspects to further determine the information needs. First, the focus was specified on those companies, which already sold their products in specific countries and want to expand the export by their service offerings. Secondly, the automatic retrieving of information should support the solution process of strategic questions, and not of operational ones, although both fields are highly relevant for companies in their service export. However, professional and public institutions already support enterprises with operational information as latest relevant legislative requirements. After intensive discussions with ten export managers as well as representatives of business support institutions (as the Chamber of Commerce), a list of parameters relevant for exporting services/service export was created. Due to the focus on strategic aspects for exporting services and the lack of an existing solution, the focus was determined on information about the service competition and the Servitization level of companies in the target markets. Therefore, the AI based solution should support answering the following research questions:

- Which services should the manufacturer offer in a specific country?
- With which partners could the manufacturer collaborate?
- Which of the manufacturers' competitors are active in the specific target market and which services do they offer?

After having established the target companies and defining the questions which should be answered by the tool, the process how to gather and display the relevant information have to be defined as well as the targeted countries have to be selected. As the project team consists of seven countries in Central Europe, the target countries represent these seven countries and the most important export markets of Advanced Manufacturers located in these CE countries. These additional export markets have been identified through desk research, studying economic reports, statistics and outlooks from the chambers of commerce. Additionally, expert reviews from the different markets have been conducted in close cooperation with the business supporting organizations to determine the most promising markets for service export in the next five to ten years. The final focus is hence, put on the following countries: Austria, Czech Republic, France, Hungary, Germany, Italy, Slovakia, Slovenia, United Kingdom and United Stated of America.

In general, Artificial Intelligence enables many applications in business and market research. Based on the project needs, as well as the expertise of project members, it was decided to use Machine Learning techniques and Natural Language Processing (NLP). Natural Language Processing enables machines to understand and interpret written or spoken human languages. Most commonly, NLP is used for speech recognition, language translation, summarization, question responses, speech generation and search applications. Machine Learning algorithms enable insights into structures and patterns within large datasets. If provided with datasets, the algorithms can be used to create prediction models, which predict or forecast outcomes or behavior. (Rebala, Ravi, and Churiwala 2019) Using Natural Language Processing as well as Machine learning presupposes that the project team has to provide the system with many examples of texts from which the system can learn (training data). This enables the system to analyze other texts on its own in the future.

Similar to the process in Wilson et al. (2016) an annotation scheme was developed on a literature analysis. The gathered annotations will serve as training and validation datasets for the Artificial Intelligence solution. Besides the general information of a company, the annotation scheme consists of a service specific part. Therefore, we adapted a service categorization by Partanen et al. (2017) with inputs by Olivia and Kallenberg (2003) and Homburg, Fassnacht and Günter (2003). The annotation scheme and the scale of industrial service offerings is shown in figure 2.

In order to conclude the sampling method, one mayor question is still unanswered. Which sources are publicly available and contain the relevant information to answer the given questions? The information has to be publicly available in order to guarantee a continuous future automatic collection. Another important criterion was that the information is readable and interpretable from a Natural Language Processing point of view. Nevertheless, the most important criterion is that the sources contain the relevant information for the Advanced Manufacturers and ideally, there are only a few sources containing most of the relevant information. Otherwise, if the tool has to combine the relevant information from many different sources, this would translate into significantly more processing effort and decrease accuracy of the results. In the end, after having weighed options such as business news articles,

professional networking platforms for businesses, etc. the most fitting sources are company webpages, as they check most of the boxes. Such as almost all companies, have publicly accessible webpages containing company specific information. Webpages usually offer manifold details relevant for the projects purpose such as products and services offered, the company size, geographic export activities, customer references, distribution or service partners, etc.

3 DATA COLLECTION AND VALIDATION METHOD

The following chapter describes the collection of the required training and validation data set. Within 2 phases, the project team collected more than 1800 annotations of company webpages in eight months. Depending on the goal, it is recommended to use a variety of different companies, which fit to the investigated industry. The training data must have enough representation of the world that the model wants to approximate (Akkiraju et al. 2020). In this respect, the training data should include small and big sized companies, as well as national and international companies to provide an as comprehensive as possible overview of the targeted market to the Artificial Intelligence and differences between the webpage professionalism of the different sized companies. As the AI tool will screen multilingual webpages, annotations were collected in Czech, English, French, German, Hungarian Italian, Slovakian and Slovenian language by twenty-two annotators from seven countries. With the collected annotations, the Artificial Intelligence should learn to analyze webpages on its own and provide answers to the research questions.

3.1 Annotation scheme

Before analyzing the webpages, an annotation scheme was developed. The scheme is based on a literature research and divided into two major sections and extracts thirty-one pieces of information. On the one hand, rather general information such as the company name, URLs, country, the NACE-code the company is operating in, revenue, number of employees, if the company has a product and/or service webpage, if the webpage indicates different customer segments, as well as customer references were collected. On the other hand, service specific information were collected, too. Therefore, a scale for industrial service offerings was developed based upon service taxonomies (Baines et al. 2013; Gebauer et al. 2010; Homburg, Fassnacht, and Günther 2003; Mathieu 2001a; Mathieu 2001b; Partanen et al. 2017; Oliva and Kallenberg 2003). Figure 2 demonstrates the finally selected six main service categories with twenty-one sub categories.

1: pre-sales services
11: product demonstrations 12: customer seminars
2: product support services
21: warranty 22: technical user training 23: customer consulting and support by phone 24: testing, test rigs, quality assurance
3: product lifecycle service
31: installation services 32: repair service 33: spare parts 34: maintenance 35: retrofit, modernization, upgrades
4: R&D services
41: research service 42: prototype design and development 43: feasibility studies
5: operational services
51: project management 52: service for operating the product for the customer 53: service for operating customer's processes
6: financial services
61: pay-per-use 62: instalment payment 63: leasing 64: rental system

Figure 2. Taxonomy of Industrial Services Offerings

3.2 Annotation process

In order to enable the AI solution to learn, a huge amount of annotations is necessary. The annotation process has been split up into two phases. At the beginning of the first phase of annotation, each of the annotators received a training, in which the goal of the annotations was communicated, the domain knowledge on services was transferred and the annotators did examples together with the trainer. In order to ensure a high quality of the annotations, every annotator had to finish the same set of five test annotations. After finishing the sample companies, the results were discussed together. At this feedback meetings, issue, which came up during the annotations, were discussed and misconception regarding different service categories cleared. Additionally, during the feedback meetings good practice experience from

the different annotators were shared and improved the overall understanding of the task. Then the annotators received companies to annotate in tranches of 25-50 companies. A trainer regularly checked the annotations and feedback was provided to the annotators.
In the first phase of annotations, fourteen annotators conducted approximately 1.300 annotations of company webpages operating in NACE codes 26, 27 and 28. To gather the necessary training data, a critical aspect of annotations is the factor time. As annotations demand a precise way of working, it may appear as a tedious task and cannot be done for more than a couple of hours. Otherwise, the quality of annotations suffers significantly by dropping concentration levels of annotators. The annotation process of one webpage took 25 minutes on average, as the annotators had to work precise to extract the relevant information from the webpages. In order to have a sufficient amount of examples per information category a threshold of at least 100 annotations per category of service types was set.
After the first phase of annotation, the intermediate results showed that services especially in the categories product support services and product lifecycle services are well established in the service market and in contrast to the other categories sufficient examples for the training data were found. Therefore, the second phase of annotations focused on the missing training data for eleven sub-categories. With focus on these categories to annotate, the project team also switched to a different process of collecting the annotations. Whereas before the annotators extracted the trainings data by gathering the different web addresses from a company webpage together with the phrases and text passages, which gave the indication of the search service category, now the project team used the browser extension tool “SingleFile”. With the tool, annotators were able to save the company webpage in an html-format and highlight the respective text passages in the file. This way of annotation and the focus on the missing service categories, the annotations process was significantly accelerated, averaging with fifteen minutes per annotation.
The process in the second phase of the annotations was similar to the one in the first phase. At the beginning of the second phase, annotators received a training, in which the goal and focus of the annotations was communicated, the domain knowledge on services was transferred and the annotators did examples together with the trainer. Afterwards the annotators received a list of five identical company webpages to annotate, which was different from in the first phase. This served two main purposes. On the one hand, it made the work comparable and enabled the control of the understanding of the task and the quality of the annotators work. The annotations were then compared with a heat map, which showed if the annotators highlighted the same passages and phrases on the webpage. It showed that it is key to highlight just the essential parts on the webpage to reduce the “noise” in the algorithm.
In total, twenty-two different annotators from Austria, Germany, Hungary, Italy, Slovakia and Slovenia have been extracting information from the company websites. Approximately 1.800 annotations have been collected in Czech, English, French, German, Hungarian Italian, Slovakian and Slovenian language. Figure 3 displays the adapted version of the research process up to the point, where the collected trainings data got prepared for further use.

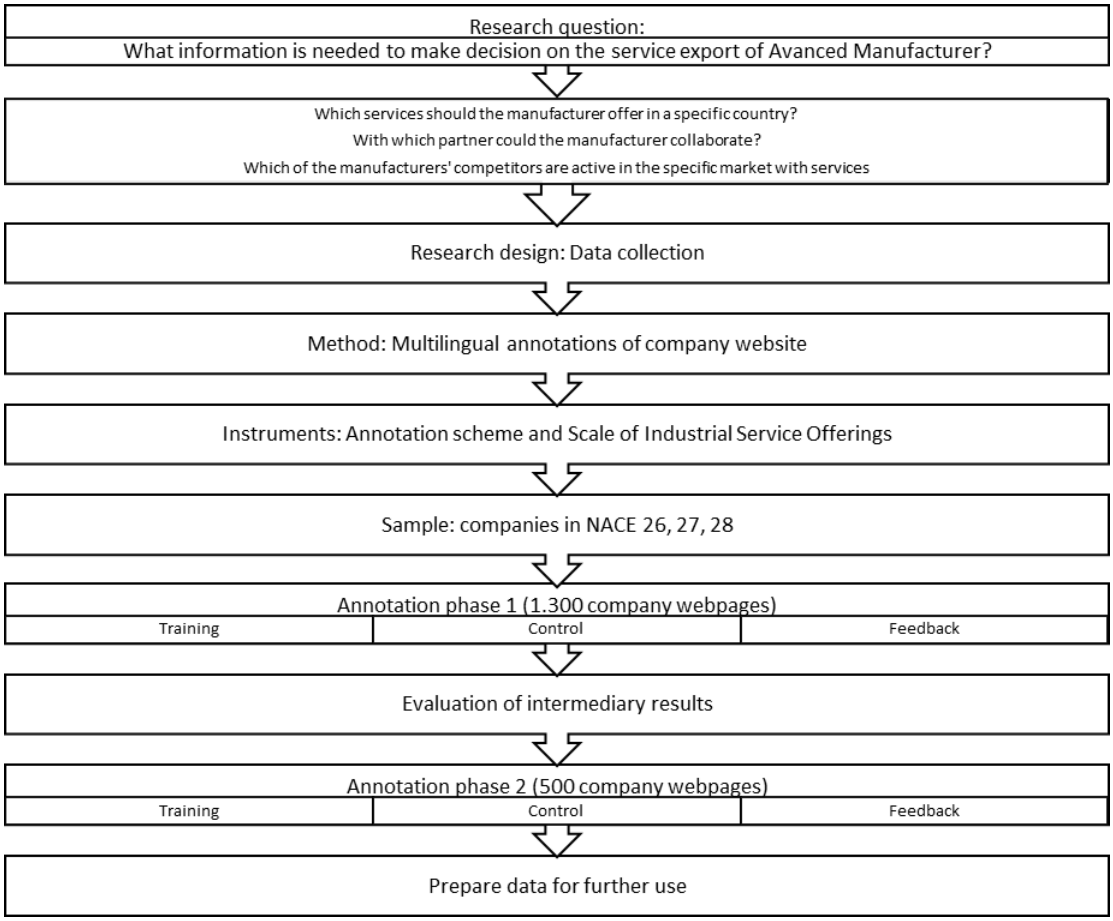

Figure 3. Adapted research project process

4 PROJECT LEARNINGS

This chapter describes the major learnings, which were gained during the first phase of the project and especially during the annotation process. One of the most important aspects of annotation is to be as precise as possible. This starts at the definition of the scope of the service categories, and the investigated topic. There has to be a clear understanding of what type of action is included in which service category. If there is a misconception among the service categories, this will result in “noise” in the machine learning algorithms and distort the results. Similar, the annotators have to have a clear understanding of the domain knowledge of the investigated topic, so they do not create “noise” with different interpretations of the text. In general, it is preferable to have only a few well-trained annotators with comprehensive understanding of the domain knowledge to gather the training data for the AI solution. On the one hand, the time it takes for one annotation decreased by the amount of repetitions, meaning there is a learning curve for annotators. On the other hand, it is easier to convey the domain knowledge to fewer people, so they have the same understanding of what action by the companies is included in what service category. Furthermore, individual annotator-specific

errors are more easily uncovered, discussed and fixed with a limited number of annotators. If the annotators have difficulties to distinguish between categories, the AI solution will have the same problem. Additionally there should be one unified way of working among the annotators, as this reduces the effort to prepare or clean the data as well as to control the annotations. Regular communication with the annotators is necessary to clarify misunderstandings, clear misconceptions and disputes, as annotations strongly depend on the interpretation of the text and language on the webpages. In general, text is always open for interpretation in the eye of the reader. This means that the language proficiency of the annotator also plays an important role and may has an influence on the overall annotation success.
Additionally, the source of information is critical to the annotations process. Regarding the company webpages, the effort and success of the annotations strongly depends on the professionalism of the webpages. The annotations can only capture the content, which is publicly available at the company webpages. Therefore, the extracted information is limited by the effort companies put into their websites.

5 SUMMARY AND OUTLOOK

The paper follows an international research team, which develops a tool to support Advanced Manufacturer in their decision making process regarding the export of their services giving them relevant information with the use of Artificial Intelligence. Therefore, the team adapted the research process of Block and Block (2005). The paper describes the project process from the initial idea until the finished collection of 1.800 annotated company webpages, which will serve as trainings and validation data for the Machine Learning algorithm. Further, project learnings are summarized to help future Artificial Intelligence project in their development.
The next steps in the project are the detailed specification of use cases and testing of various machine learning algorithms to find the best suited one for these use cases. Additionally, a front-end design for the users of the tool needs to be developed and ultimately the deployment of the collected trainings data to train the Artificial Intelligence. Another aspect in the project will be the validation regarding the quality assurance of the collected annotations.

6 ACKNOWLEDGEMENTS

The work described in this document has been conducted as part of a project “prosperAMnet”, which is supported by the Interreg CENTRAL EUROPE Programme funded under the European Regional Development Fund.

REFERENCES

Akkiraju, Rama, Vibha Sinha, Anbang Xu, Jalal Mahmud, Pritam Gundecha, Zhe Liu, Xiaotong Liu, and John Schumacher (2020), “Characterizing Machine Learning Processes: A Maturity Framework,” *Business Process Management*. *Lecture Notes in Computer Science*, Dirk Fahland, Chiara Ghidini, Jörg Becker and Marlon Dumas, eds. Cham: Springer International Publishing, 17–31.

Baines, Tim, Howard Lightfoot, Palie Smart, and Sarah Fletcher (2013), “Servitization of manufacture,” *Journal of Manufacturing Technology Management*, 24 (4), 637–46, DOI: 10.1108/17410381311327431.

Block, Martin P. and Tamara S. Block (2005), *Business-to-business market research*: *[a value-based approach],* 2. ed., 1. [print.]. Mason OH u.a.: Thomson.

Davenport, Thomas, Abhijit Guha, Dhruv Grewal, and Timna Bressgott (2020), “How artificial intelligence will change the future of marketing,” *Journal of the Academy of Marketing Science*, 48 (1), 24–42, DOI: 10.1007/s11747-019-00696-0.

European Communities (2008), *NACE rev. 2,* Revision 2, English edition. Luxembourg: Office for Official Publications of the European Communities.

Fahland, Dirk, Chiara Ghidini, Jörg Becker, and Marlon Dumas (2020), *Business Process Management. Lecture Notes in Computer Science.* Cham: Springer International Publishing.

Gebauer, Heiko, Bo Edvardsson, Anders Gustafsson, and Lars Witell (2010), “Match or Mismatch: Strategy-Structure Configurations in the Service Business of Manufacturing Companies,” *Journal of Service Research*, 13 (2), 198–215, DOI: 10.1177/1094670509353933.

Homburg, Christian, Martin Fassnacht, and Christof Guenther (2003), “The Role of Soft Factors in Implementing a Service-Oriented Strategy in Industrial Marketing Companies,” *Journal of Business-to-Business Marketing*, 10 (2), 23–51, DOI: 10.1300/J033v10n02_03.

Izsak, Kincsö, Maialen Perez, and Henning Kroll (2020), “Advanced Technologies for Industry - EU Report. Technological trends and policies,” European Commission.

Mathieu, Valérie (2001a), “Product services: from a service supporting the product to a service supporting the client,” *Journal of Business & Industrial Marketing*, 16 (1), 39–61, DOI: 10.1108/08858620110364873.

Mathieu, Valérie (2001b), “Service strategies within the manufacturing sector: benefits, costs and partnership,” *International Journal of Service Industry Management*, 12 (5), 451–75, DOI: 10.1108/EUM0000000006093.

Oliva, Rogelio and Robert Kallenberg (2003), “Managing the transition from products to services,” *International Journal of Service Industry Management*, 14 (2), 160–72, DOI: 10.1108/09564230310474138.

Partanen, Jukka, Marko Kohtamäki, Vinit Parida, and Joakim Wincent (2017), “Developing and validating a multi-dimensional scale for operationalizing industrial service offering,” *Journal of Business & Industrial Marketing*, 32 (2), 295–309, DOI: 10.1108/JBIM-08-2016-0178.

Rebala, Gopinath, Ajay Ravi, and Sanjay Churiwala (2019), *An Introduction to Machine Learning.* Cham: Springer International Publishing.

Verma, Sanjeev, Rohit Sharma, Subhamay Deb, and Debojit Maitra (2021), “Artificial intelligence in marketing: Systematic review and future research direction,” *International Journal of Information Management Data Insights*, 100002, DOI: 10.1016/j.jjimei.2020.100002.

Wilson, Shomir, Florian Schaub, Aswarth A. Dara, Frederick Liu, Sushain Cherivirala, Pedro G. Leon, Mads S. Andersen, Sebastian Zimmeck, Kanthashree M. Sathyendra, N. C. Russell, Thomas B. Norton, Eduard Hovy, Joel Reidenberg, and Norman Sadeh, “The Creation and Analysis of a Website Privacy Policy Corpus”.