engineering proceedings

MDPI

Proceeding Paper

Indirect Assessment of Implementation of Industry 4.0 Technologies in Regional Companies †

Rubén Nicolás Ibáñez *, Antonio Guerrero González and Juan Carlos Molina Molina

Universidad Politécnica de Cartagena, poctcode city, country; antonio.guerrero@upct.es (A.G.G.); jcarlos.molina@upct.es (J.C.M.M.)

* Correspondence: ruben.nicolas@edu.upct.es

† Presented at the 10th International Electronic Conference on Sensors and Applications (ECSA-10), 15–30 November 2023; Available online: https://ecsa-10.sciforum.net/.

Abstract: This article evaluates the implementation of Industry 4.0 technology in companies in a region using indirect methods such as web scraping and examination of publicly available information. By analyzing online data and reports, the level of adoption and integration of 4.0 technology is determined. This provides valuable information on the technological progress of companies to help policy makers promote widespread adoption, businesses to benchmark and make informed technology investments, and researchers to analyze the impact on regional economies. The use of online data sources to assess Industry 4.0 implementation is essential to understand the technological progress and growth potential of these technologies in various industries, contributing to the formulation of policies that encourage innovation.

Keywords: Industry 4.0; Web Scrapping; Implementation of Industry 4.0; digital transformation; business competitiveness

1. Introduction

In a world in constant evolution, driven by the digital revolution, Industry 4.0 stands as a beacon guiding entrepreneurs and companies through uncharted waters towards an increasingly digitized and automated future. The work by Michela Piccarozzi, Barbara Aquilani and Corrado Gatti [1] seeks to shed light on this concept, offering a precise definition and pointing out gaps in existing research.

Industry 4.0, more than a simple evolution of manufacturing, represents a revolution that connects all aspects of the production chain, infusing intelligence into every phase. V. Alcácer and V. Cruz-Machado [2] explore how enabling technologies are transforming management and decision making in industry, moving away from traditional centralized applications to a more distributed and agile approach.

Roland Ortt, Stolwijk, and Matthijs Punter [3] embark on a mission to summarize and combine numerous articles to address this crucial question. Their work provides a comprehensive overview of the implementation, evaluation methods, and current status of Industry 4.0.

However, implementation is not uniform everywhere. In developing countries, competitive pressure and government support are key drivers of technology adoption, as a study of manufacturing companies in developing economies reveals.

Industry 4.0 not only transforms manufacturing, but also has a significant impact on business sustainability, as the study by Julian Marius Müller, Daniel Kiel and Kai-Ingo Voigt [4] shows. Opportunities drive adoption, but challenges vary by company size and sector.

However, a persistent issue in this field is the lack of a clear definition of Industry 4.0. Mario Hermann, Tobias Pentek and Boris Otto [5] set out to solve this problem by establishing a precise definition and presenting key implementation principles.

Citation: Ibáñez, R.N.; González, A.G.; Molina, J.C.M. Indirect Assessment of Implementation of Industry 4.0 Technologies in Regional Companies. *Eng. Proc.* **2023**, *56*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s): Name

Published: 15 November 2023

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

As countries and companies struggle to adapt to Industry 4.0, the adoption of national strategies becomes essential. The article by Batchkova I.A, Popov G.T, Ivanova Ts. A. and Belev Y.A. [6] discusses how different approaches assess countries' readiness for this industrial revolution.

In the European Union, the adoption of digital technologies becomes a key competence in the single market. Isabel Castelo-Branco, Frederico Cruz-Jesus and Tiago Oliveira [7] explore key indicators of readiness and the importance of detailed data.

Meanwhile, in Turkey [8], the specific opportunities and challenges facing the country in adopting Industry 4.0 are assessed through the perceptions of managers in the home appliance industry.

But not all regions of the world are advancing at the same pace. In Kathmandu, Nepal [9], the limited availability of highly skilled labor poses significant barriers to Industry 4.0 adoption, while in Ethiopia [10], information technology implementation is a critical challenge.

The transformation of manufacturing through cyber-physical systems, as highlighted by the work of Joseph Flynn, Steven Dance and Dirk Schaefer [11], underscores the importance of preparing the workforce to deal with the changes that this new industrial era brings.

To access the valuable information that fuels this revolution, we turn to web scraping, a crucial technique for extracting data from online sources. Chaimaa Lotfi, Swetha Srinivasan, Myriam Ertz, and Imen Latrous [12] delve into this essential technique, covering its design and the tools and technologies involved.

For those immersed in academic research, the CRISP-DM framework, proposed by Hossam El-Din Hassanien [13], offers a valuable solution for analyzing literature in scientific web repositories, addressing challenges related to unstructured data.

Ultimately, web scraping proves to be a vital bridge between the vast ocean of online information and its application in various fields, facilitating informed decision making in an increasingly digital world. Industry 4.0 continues to advance, and those who understand its concepts and challenges are better prepared to navigate this Fourth Industrial Revolution.

2. Materials and Methods

Within the framework of our project, we have used the SABI (Sistema de Análisis de Balances Ibéricos) database. SABI, provided by the Spanish firm Informa D&B, represents a fundamental tool used predominantly in Spain and Portugal to carry out financial analysis of companies. It stands out as one of the most prominent sources of business information on the Iberian Peninsula, as it compiles and stores detailed financial and commercial data on both Spanish and Portuguese companies. This information covers aspects such as balance sheets, income statements, shareholders, managers, economic activity, and other key indicators.

The percentage of companies with websites has been determined using the SABI database through an exhaustive analysis of the information collected in the database. The process involved the identification of companies with website records in their profiles, followed by a calculation that related the number of companies with websites to the total number of companies included in the SABI database. This assessment provided an accurate picture of the extent of the online presence of the companies in the dataset and allowed obtaining a percentage that reflects the level of adoption of web presence in the business context analyzed.

The process of determining the level of Industry 4.0 implementation in a company through a visit and survey was carried out as follows: first, a thorough preparation was carried out and clear objectives were defined. During the visit, an interview was conducted with management, a tour of the facility was conducted, employees were spoken to, and systems and data were evaluated. A targeted survey was then used to gather additional information. The data collected was then analyzed to rank the level of

implementation in key areas. Finally, recommendations for improving Industry 4.0 adoption were provided and tailored to the specific needs of the assessed company.

In our research process to assess the level of Industry 4.0 implementation in companies, we carefully identified and defined a number of key terms related to this emerging industry. These key terms were used as a basis for formulating questions and conducting targeted searches in the analyzed companies. By focusing on essential concepts such as "Internet of Things (IoT)", "Artificial Intelligence (AI)", "Automation", "Big Data" and "Cybersecurity", we were able to address key aspects of Industry 4.0 and more accurately assess the degree of technological adoption and sophistication in the organizations studied.

Web scraping is used in a variety of applications, such as data collection for market research, online price monitoring, news aggregation, information extraction for data analysis, and many other areas. However, it is important to note that web scraping must be carried out in an ethical manner and in compliance with the websites' terms of service, as unauthorized access or excessive requests can have legal and ethical implications.

Web scraping is a technique that automates the extraction of data from websites in order to collect information efficiently and use it in various applications.

In the project we developed, we applied a number of fundamental Python techniques and libraries to carry out web scraping effectively. To extract data from web pages, we employed HTML Document Object Model (DOM) navigation techniques using BeautifulSoup and Scrapy. These tools were essential for analyzing the structure of the web pages and extracting relevant information in a systematic way. In addition, we made use of other key libraries such as Requests to handle HTTP requests, Selenium to automate interaction with web browsers, PyQuery to perform jQuery-like parsing of HTML documents and Lxml for efficient processing of XML and HTML data. Thanks to these techniques and libraries, we were able to perform accurate and efficient data extraction for our project successfully.

Methodology

In the development of our project, we have followed a specific methodology to analyze the level of implementation of 4.0 technologies in companies. The general steps we have applied are as follows:

- Identification of information sources: we started the process by identifying relevant online information sources that host data on the adoption of 4.0 technologies in companies. This includes news websites, research reports, specialized blogs and social media platforms where companies share their technological advances.
- Selection of target pages: We identified specific web pages or sections of websites that contained relevant information on the implementation of 4.0 technologies in companies. This included press releases, annual reports, case studies, company blogs and other resources that provided meaningful data.
- Use of web scraping libraries: To automate data extraction from the selected web pages, we opted for Python web scraping libraries, such as BeautifulSoup and Scrapy. These tools facilitated the navigation through the HTML code of the pages and the extraction of relevant data.
- Development of a web scraping script: We created a Python script that used the selected library to collect information from the target web pages. This involved sending HTTP requests to download the content of the pages, parsing the HTML to extract relevant data, and then storing it in a data structure, such as a CSV file or a database.
- Definition of analysis criteria: We established clear criteria to assess the level of implementation of 4.0 technologies in companies. This included identifying keywords and phrases related to Industry 4.0, such as "Internet of Things (IoT)", "Artificial Intelligence (AI)", "Automation", and more.
- Data processing and analysis: Once we collected data from various sources, we proceeded to process and analyze the information collected to assess the level of adoption

of 4.0 technologies in companies. We employed word processing, text mining and data analysis techniques to quantify and classify relevant information.

- Data visualization: In order to effectively present our findings, we created graphs and visualizations that showed the level of implementation of 4.0 technologies in companies in a clear and concise manner. For this, we used tools such as Matplotlib, Seaborn and D3.js.
- Data validation: To ensure the accuracy and consistency of the data collected, we validated and verified the information. This involved manual review of some results and comparison of data from multiple sources.
- Results report: We documented our findings in a report detailing the level of implementation of 4.0 technologies in the companies analyzed. The report included information on the methodology used, the analysis criteria and any significant observations or trends identified.
- Continuous update: Given the constantly evolving nature of 4.0 technology implementation, we considered automating the web scraping and analysis process to continuously track trends and changes in the level of technology adoption by companies.

3. Results

Prior to applying the web scraping techniques, the number of Murcian companies with a website was analyzed from the SABI database. A total of 5913 Murcian companies were counted out of the 19,921 companies with a website in 2021. This corresponds to 29.68% of the companies. In this study the level of implementation was analyzed by economic activities, locations, income, age, employees and export level.

Thirteen essential KETs have been identified and evaluated, including Internet of Things (IoT), Cloud/Edge Compu-ting, Artificial Intelligence (AI), Systems Integration, Visualization and Data Management Systems, Augmented, Virtual and Mixed Reality (AR/VR/MR), Drones, Cybersecurity, Additive Manufacturing, Collaborative Robotics, Smart Factory, Building Information Modeling (BIM), Big Data and Digital Twin. These KETs have been for web scraping search on their website.

This paper shows the results obtained for a sample consisting of two 4.0 enabler companies, which are referents in the drone sector, and another in the collaborative robotics sector. In the first company, 254 websites were analyzed. From the second company, 178 websites were analyzed using the web scraping technique.

Table 1. Web-scrapping results versus interviews and company visits.

KETs	Enabler 1	Enabler 2
IoT	0 (0%) [0]	48 (11.40%) [2]
Cloud/Edge Computing	380 (23.7%) [1]	52 (12.3%) [2]
IA	237 (14.7%) [1]	36 (8.5%) [1]
vertical and horizontal integration	14 (8.7%) [1]	159 (37.7%) [3]
Data Visualization	465 (29.8%) [3]	206 (48.9%) [4]
RA, RV, RM	45 (2.8%) [1]	14 (3.3%) [1]
Drones	1606 (100%) [5]	35 (8.31%) [1]
Cibersecurity	54 (3.36%) [1]	43 (10.1%) [1]
Additive Fabrication	23 (1.4%) [1]	13 (3.01%) [1]
Collaborative Robotics	0 [0]	421 (100%) [5]
Smart Factory	0 [0]	68 (16.15%) [1]
BIM	154 (9.5%) [2]	0
Big Data	[0]	15 (3.5%) [1]
Digital Twin	[0]	12 (2.8%) [1]

The tables show the number of times the keyword or terms related to KET were found on their website, and in parentheses the percentage measured with respect to the number of times the main topic was found. Personal interviews were conducted as indicated in the methodology. The interview findings are presented within parentheses on a scale from 1 to 5, (with 0 denoting no implementation and 5 representing high implementation). The percentages of keyword encounters on the web are related to the encounters on the company's main topic. Correlation is found between the results obtained from the interviews and those obtained from the web scraping techniques.

4. Discussion

This study demonstrates the usefulness of using indirect methods, such as web scraping and public information analysis, to assess the adoption of Industry 4.0 technologies in companies in a region.

The results obtained through web scraping correlate well with the direct assessments made through company visits, interviews and surveys. This validates the use of web mining techniques to extract relevant data from corporate websites and other online sources.

The analysis of the SABI database was also very useful to obtain an overview of the level of online presence of companies in the region. This information allows focusing web scraping efforts on the companies most likely to disclose data on their digital transformation.

An important finding is the considerable variability in the adoption of 4.0 technologies depending on the economic activity, location, revenue and other factors of the companies. This underscores the need for customized digital transformation strategies according to the characteristics of each organization.

Among the limitations of the study is the impossibility of covering all the companies in the region through manual web scraping. However, the automation of this process using scripts would make it possible to significantly expand the sample analyzed.

Another limitation is the reliance on the information that companies publish on their websites and other online media. Not all companies disclose their techno-logical advances, which may underestimate their level of Industry 4.0 adoption.

In conclusion, this work lays the groundwork for a continuous monitoring of the progress of digital transformation in the business fabric using web mining techniques. This information is essential for the formulation of policies that encourage in-novation and for companies to assess their capabilities in relation to those of their competitors.

Author Contributions: Conceptualization, R.N.I.; methodology, A.G.G. and J.C.M.M.; software, R.N.I., A.G.G., and J.C.M.M.; validation, A.G.G. and R.N.I.; formal analysis, R.N.I. investigation, R.N.I., A.G.G., A.C.L., and J.C.M.M.; resources, R.N.I.; data curation, J.C.M.M.; writing—original draft preparation, R.N.I.; writing—review and editing, A.G.G.; visualization, R.N.I. and A.G.G.; supervision, R.N.I. and A.G.G.; project administration, A.G.G.; funding acquisition, A.G.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Comunidad Autónoma de la Región de Murcia, Project: Cátedra Industria 4.0.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the Instituto de fomento de la Región de Murcia for the financial support for this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Piccarozzi, M.; Aquilani, B.; Gatti, C. Industry 4.0 in management studies: A systematic literature review. *Sustainability* **2018**, *10*, 3821.
2. Alcácer, V.; Cruz-Machado, V. Scanning the industry 4.0: A literature review on technologies for manufacturing systems. *Eng. Sci. Technol. Int. J.* **2019**, *22*, 899–919.
3. Ortt, R.; Stolwijk, C.; Punter, M. Implementing industry 4.0: Assessing the current state. *J. Manuf. Technol. Manag.* **2020**, *31*, 825–836.
4. Müller, J.M.; Kiel, D.; Voigt, K.I. What Drives the Implementation of Industry 4.0? The Role of Opportunities and Challenges in the Context of Sustainability. *Sustainability* **2018**, *10*, 247.
5. Herman, M.; Pentek, T.; Otto, B. *Design Principles for Industry 4.0 Scenarios: A Literature Review*; Business Engineering Institute St. Gallen: St. Gallen, Switzerland, 2015; Volume 10, p. 16.
6. Batchkova, I.A.; Popov, G.T.; Ivanova, T.A.; Belev, Y.A. Assessment of readiness for "Industry 4.0". *Int. Sci. J.* **2021**, *6*, 288–291.
7. Castelo-Branco, I.; Cruz-Jesus, F.; Oliveira, T. Assessing Industry 4.0 readiness in manufacturing: Evidence for the European Union. *Computers in Industry* **2019**, *107*, 22–32.
8. Demirbağ, K.Ş.; Yıldırım, N. Getting the measure of the fourth industrial revolution: Advantages and challenges of Industry 4.0 in the Turkish white goods industry. *Cent. Eur. Manag. J.* **2022**, *31*, 82–101.
9. Rajbhandari, S.; Devkota, N.; Khanal, G.; Mahato, S.; Paudel, U.R. Assessing the industrial readiness for adoption of industry 4.0 in Nepal: A structural equation model analysis. *Heliyon J.* **2022**, *10*, 1–11.
10. Beza, M. *Readiness Assessment of Ethiopian Industrial Park for Industry 4.0*; College of Social Science Addis Ababa Science and Technology University: 2020; pp. 1–118.
11. Flynn, J.; Dance, S.; Schaefer, D. Industry 4.0 and its potential impact on employment demographics in the UK. *Adv. Transdiscipl. Eng.* **2017**, *6*, 239–244.
12. Lotfi, C.; Srinivasan, S.; Ertz, M.; Latrous, I. Web Scraping Techniques and Applications: A Literature Review. ResearchGate. 2021. Available online: https://www.researchgate.net/profile/Myriam-Ertz/publication/367719780_Web_Scraping_Techniques_and_Applications_A_Literature_Review/links/64c3b8600fae1319bfbed20e/Web-Scraping-Techniques-and-Applications-A-Literature-Review.pdf (accessed on).
13. Hassanien, H.E.D. Web Scraping Scientific Repositories for Augmented Relevant Literature Search Using CRISP-DM. *Appl. Syst. Innov.* **2019**, *2*, 37.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.