JOURNAL OF INFORMATION SYSTEMS & OPERATIONS MANAGEMENT

JISOM
VOL.18 NO.1
MAY 2024

ROMANIAN-AMERICAN UNIVERSITY

ANALYZE, CREATE, SHARE KNOWLEDGE

JOURNAL OF INFORMATION SYSTEMS & OPERATIONS MANAGEMENT

VOL. 18 NO. 1

MAY 2024

ROMANIAN-AMERICAN UNIVERSITY PUBLISHING HOUSE
BUCHAREST

EDITOR-IN-CHIEF NOTE

JISOM is a journal which enables researchers, academia, and professionals to make their voices heard, to share their findings with their peers and the public at large. Being an open-access journal, JISOM aims to increase the all-around level of knowledge in the fields of computer science and economics, to further the society's capabilities to understand new concepts, to see how things are done by cutting-edge technologies implementations, to understand what is in store for us not only at the present time but also in the near, and so near, future.

The published articles focus mainly on IT&C but we also provide a favorable exposure medium for correlated topics, such as economics, management, applied sciences, mathematics, statistics, etc. JISOM encourages cross-disciplinary research of national and international researchers and welcomes the contributions which give new and fresh perspectives to the above-mentioned fields of study.

National and international researchers, professionals, recognized experts, professors who want to share their research results and new ideas, and Ph.D. students who want to improve their knowledge or present their emerging doctoral research are all welcome to join our knowledge-eager community.

We have a history that started back in 2007 and we are permanently striving to bring our community to a higher level of knowledge in the fields we cover. With the constant support of our authors, reviewers, readers, and editorial staff I am sure we are and will live up to the mission we have taken on.

Many thanks to our JISOM community and good luck with your research!

Respectfully,

Alexandru TĂBUȘCĂ, PhD

JISOM Editor-in-Chief

JOURNAL OF INFORMATION SYSTEMS & OPERATIONS MANAGEMENT

JOURNAL OF INFORMATION SYSTEMS & OPERATIONS MANAGEMENT

Journal details

No.	Item	Value
1	Complete title / IDB title	JOURNAL OF INFORMATION SYSTEMS & OPERATIONS MANAGEMENT
2	ISSN (print and/or electronic)	1843-4711
3	Frequency	semestrial
4	Journal website (direct link to journal section)	http://jisom.rau.ro
5	IDB indexation	EBSCO GALE Cengage Learning Index Copernicus ProQuest RePEC/IDEAS

Contact

First name and last name	Alexandru TĂBUŞCĂ, PhD Associate Professor
Phone	+4-0372-120.140
E-mail	alex.tabusca@rau.ro

ISSN: 1843-4711

TABLE OF CONTENTS

ANALYSING THE PERFORMANCE IMPACTS OF LAZY LOADING IN WEB APPLICATIONS

Răzvan-Mihail BÂRA[1]

Costin Anton BOIANGIU[2]

Cătălin TUDOSE[3]

Abstract

This scientific paper explores the effects of lazy loading on the performance of web applications. Lazy loading is a technique employed in web development to defer the loading of non-essential resources until they are effectively needed, thereby optimizing the initial page load times and resource utilization. We'll review a few important sources addressing the problem, we'll examine the most important performance metrics and we'll develop our analysis methodology and experimental setup. This study aims to investigate the impact of lazy loading on various performance metrics, including page load speed, user experience, and overall efficiency of web applications. For comprehensive results, we developed two applications having different architectures (a static website and a website with dynamic content) and we varied the test environment (a high-speed network versus a lower-speed network). The final evaluation would like to conclude about the optimal choice between eager loading and lazy loading in web development and the trade-offs to be made.

Keywords: web development, lazy loading, eager loading, static content, dynamic content, high-speed network, lower-speed network

JEL Classification: O33 Technological Change: Choices and Consequences • Diffusion Processes

1. Introduction

Web applications have become an integral part of our digital lives, serving as dynamic platforms for information dissemination, e-commerce, and interactive user experiences. As user expectations continue to rise, so does the demand for web applications that deliver high

[1] Student, POLITEHNICA National University for Science and Technology of Bucharest, Romania, razvan_mihail.bara@stud.etti.upb.ro

[2] PhD, Professor, POLITEHNICA National University for Science and Technology of Bucharest, Romania, costin.boiangiu@upb.ro

[3] PhD, Lecturer, POLITEHNICA National University for Science and Technology of Bucharest and Luxoft Romania, Romania, catalin.tudose@gmail.com

performance. One critical aspect of optimizing web application performance is the loading time of resources during the initial page load. Users often associate a sluggish or delayed loading experience with poor application performance, which can lead to frustration and decreased user engagement, having as ultimate consequence people no longer accessing the web application. [1]

In response to this challenge, web developers employ various techniques to enhance loading efficiency. Lazy loading has emerged as a prominent strategy, allowing developers to defer the loading of non-essential resources until they are required for user interaction. This deferred loading aims to streamline the initial page load, reduce bandwidth consumption, and improve overall user experience. The concept of lazy loading has gathered significant attention as a means to strike a balance between feature-rich web applications and optimal performance. [2]

The primary objective of this study is to delve into the effects of lazy loading on the performance of web applications. By conducting a systematic analysis, we aim to uncover insights into how lazy loading influences key performance metrics, including page load speed, network usage, and user experience. Through this investigation, we seek to provide web developers with a deeper understanding of the trade-offs associated with lazy loading, enabling them to make informed decisions in the pursuit of creating high-performance web applications.

To achieve this goal, we will employ a rigorous experimental methodology, using real-world web applications and several test scenarios. The study will not only assess the quantitative impact on performance metrics but also explore the qualitative aspects of user interactions and satisfaction. By addressing these aspects, we aim to contribute valuable knowledge that can guide best practices for implementing lazy loading in web development.

In the subsequent sections, we will review existing literature on lazy loading, detail our experimental setup and methodology, present and analyze results, and conclude with recommendations for developers and plans for future research. Through this exploration, we aspire to contribute to the ongoing dialogue on web application optimization and advance the understanding of the practical implications of lazy loading in real-world scenarios.

2. Literature Review

The literature review aims to provide a comprehensive overview of existing research on lazy loading in web applications. It examines key studies and findings related to the impact of lazy loading on web application performance, user experience, and overall efficiency.

2.1 Lazy Loading in Web Development

Lazy loading has gained prominence as a technique for optimizing web application performance by deferring the loading of non-essential resources until they are needed. [3] According to Simon Fray, lazy loading minimizes the initial page load times and enhances user experience by prioritizing critical content. Additionally, lazy loading reduces server loads and conserves bandwidth, contributing to a more sustainable and scalable web infrastructure. [4]

2.2 Performance Metrics in Lazy Loading Studies

Jamie Juviler highlights the importance of considering multiple performance metrics when evaluating the effects of lazy loading. This study analyzes the impact on page load speed, resource utilization, and user engagement. [2]

2.3 User Experience Implications

User experience is a crucial aspect of web application development. The study by Jamie Juviler [2] investigates the correlation between lazy loading and user satisfaction, emphasizing the positive impact on perceived performance. On the opposite side, Felix Arntz and Rick Viscomi [3] suggest potential drawbacks, such as delayed rendering of images during user interaction, raising concerns about the trade-offs associated with lazy loading.

3. Methodology

3.1 Experimental Setup

To assess the effects of lazy loading on web application performance, we conducted experiments on two distinct applications: a static website and a dynamic website representing an online library. Also, we varied the environmental conditions between a high-speed network and a lower-speed network.

3.1.1 Static Website

The first application is a static website designed to showcase content without dynamic interactions. This serves as a baseline scenario to evaluate the impact of lazy loading on straightforward, content-focused web pages. The website consists of one single page, including 5 sections, each section containing a slider of 10 high-quality images.

3.1.2 Full-Stack Application for an Online Library

The second application represents a comprehensive full-stack online library, featuring dynamic content, user authentication, and real-time interactions. This complex application allows us to explore the scalability and adaptability of lazy loading in a more interactive and data-intensive environment.

3.2 Performance Metrics

Our methodology involves measuring various performance metrics to comprehensively evaluate the impact of lazy loading on both applications. The key metrics include:

3.2.1 Initial Load Time

We will measure the time it takes for the initial load page to completely load, starting from the initiation of the request to the rendering of all visible content. This metric serves as a fundamental indicator of the overall responsiveness of the applications.

3.2.2 Bandwidth Usage

To assess the efficiency of lazy loading in conserving bandwidth, we will monitor the amount of data transferred between the server and the client during the page loading process. This measurement is crucial for understanding the potential benefits of lazy loading, particularly in scenarios with limited network resources.

3.2.3 User Experience Metrics

We use the Chrome DevTools and the Lighthouse extension to gather user experience metrics, including First Contentful Paint (FCP), Time to Interactive (TTI), and Cumulative Layout Shift (CLS). These metrics provide insights into the perceived performance and visual stability of web applications.

First Contentful Paint (FCP) is an important metric, centered on the user and measuring the perceived load speed. It indicates the first moment during the loading of the page when the user sees something being displayed. A quick FCP gives the user the impression that things are moving on.

Time to Interactive (TTI) indicates the earliest time after First Contentful Paint (FCP) when the page is ready to interact with the user. This performance metric helps identify situations where a page erroneously looks interactive by measuring a page's load responsiveness.

Cumulative Layout Shift (CLS) examines the unexpected movement of elements in the viewport as the page loads. This is a measure of the visual stability of the content of the page, as unforeseen layout shifts can displease the user.

3.3 Test Cases

For each application, we designed a series of test cases to simulate different user interactions and scenarios. These include:

3.3.1 Navigation Scenarios

- Initial page load
- Navigation between pages
- Loading additional content dynamically

3.3.2 User Interaction Scenarios

- Scrolling through content
- Interacting with dynamic elements (e.g., forms, modals)

3.4 Tools and Technologies

The experiments were conducted using Google Chrome's DevTools and the Lighthouse extension. The DevTools provide in-depth insights into the network activity, rendering performance, and user interactions, while the Lighthouse extension automates the collection of performance metrics and provides a standardized evaluation of web page performance.

3.5 Test Environments

All experiments were conducted in controlled environments to minimize external factors affecting the measurements. The test environments included standardized hardware specifications and network configurations to ensure consistency across experiments.

We will define two environments for testing, to examine the effects of lazy loading in both cases:

- A high-speed network, referenced as HSN
- A lower speed network, referenced as LSN

4. Results and Analysis

4.1 Static Website

This application consists of a single HTML page to showcase different landforms. It has 5 distinct sections, each of them displaying 10 pictures of a landform. We compare the results

of using lazy-loading performance versus eager-loading performance in the following subsections.

4.1.1 Initial Load Time

For all 50 pictures to load, using the eager loading strategy the initial page load time is 5.5 seconds. With the 5G network (a high-speed network), using lazy loading, the initial load time is only 0.5 seconds. The load time decreases drastically as more and more images are lazy-loaded, as described in Figures 1 and 3.

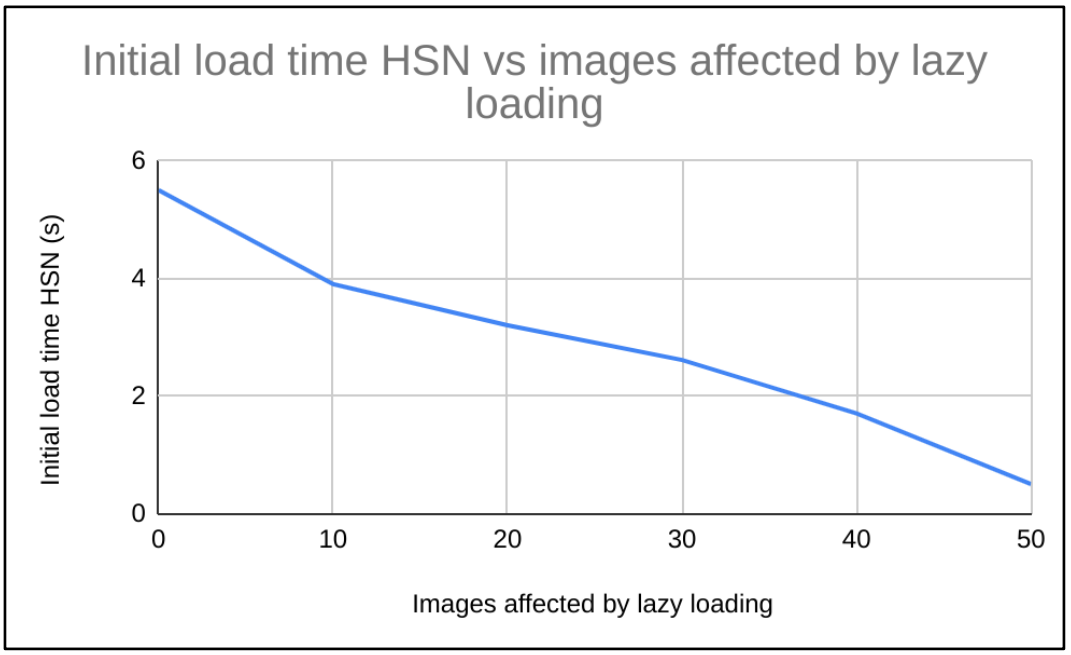


Figure 1. Graph describing the decrease of the initial load time of the HSN

This measurement is heavily impacted by the type of network in use. For example, when the user is connected to a 3G network (a lower-speed network), the total time for eager loading is 1.7 minutes while for lazy loading is only 6.9 seconds.

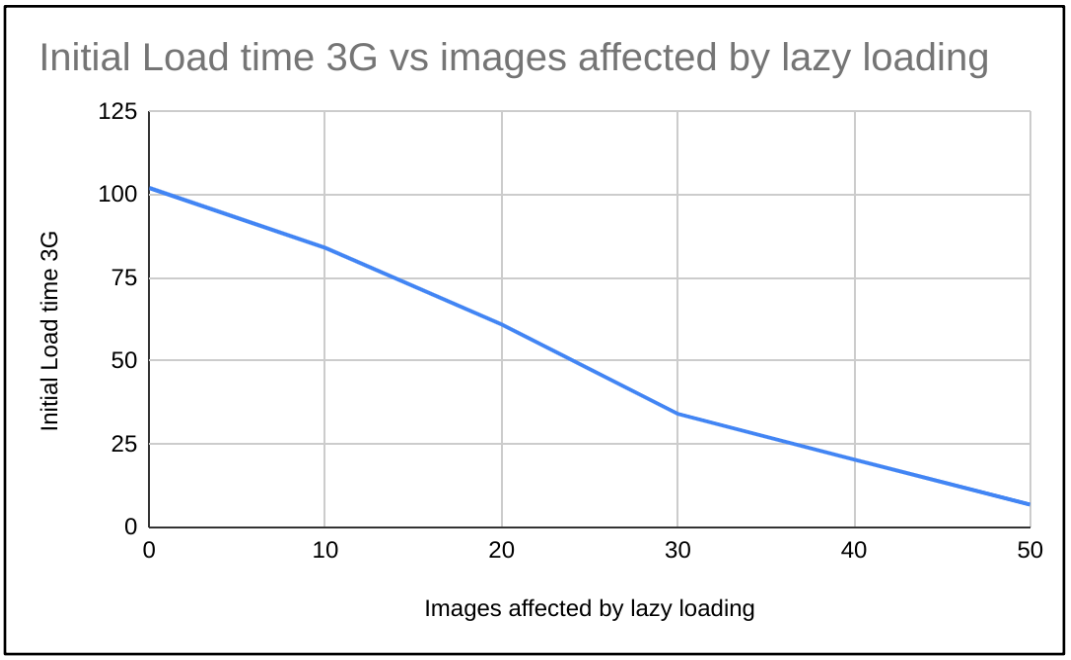


Figure 2. Graph describing the decrease of initial load time for the LSN

In the context of network performance, lazy loading proves to be especially beneficial for a 3G network compared to a 5G network. In a 3G environment, where bandwidth is more limited, lazy loading contributes significantly to a smoother and more responsive user experience. By deferring the loading of images, until they are needed, lazy loading conserves bandwidth and reduces the strain on the network.

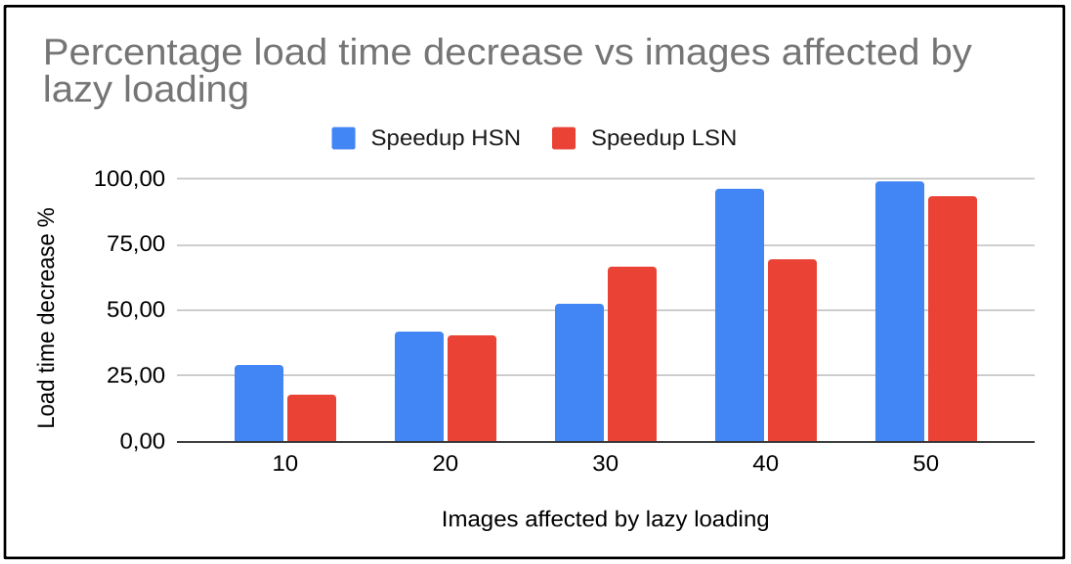


Figure 3. Speedup of HSN and LSN

4.1.2 Network Utilisation

Also, as more and more images are lazy-loaded, the bandwidth usage of the web page decreases. The page transfers about 18 MB of image content through the network. The same amount is transferred when using the lazy loading strategy, but only if the user explores the whole page. Otherwise, the amount of transferred data depends on the percentage of the images requested by the user when navigating to a certain part of the page.

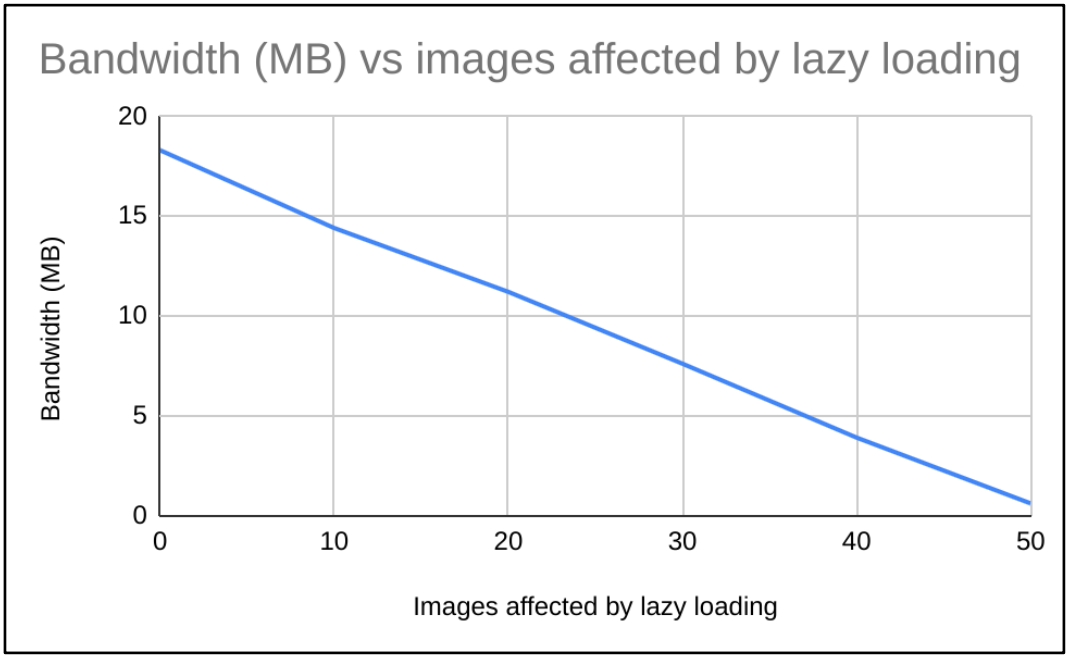


Figure 4. Decrease of initial bandwidth usage

This table displays the data that has been considered for this study:

Lazily loaded images	Bandwidth (MB)	Initial load time 5G (s)	Speedup 5G (%)	Initial load time 3G (s)	Speedup 3G (%)
0	18.3	5.5	0	102	0
10	14.4	3.9	29.09	84	17.65
20	11.2	3.2	41.82	61	40.20

Lazily loaded images	Bandwidth (MB)	Initial load time 5G (s)	Speedup 5G (%)	Initial load time 3G (s)	Speedup 3G (%)
30	7.6	2.61	52.55	34.1	66.57
40	3.9	1.7	95.93	20.37	69.40
50	0.6	0.5	99.05	6.8	93.33

Table 1 – Measurements of the static website

4.1.3 User Experience

While the slow initial page load time of eager loading would make a user leave the website due to the impression that the application isn't working properly, the navigation experience would prove itself to be a pleasant one. All the content is already loaded onto the page, so scrolling and exploring through the different sliders of landforms presents no issue.

On the other hand, for lazy loading, the initial page load time is fast, but the user might encounter some issues while exploring the page. If the network is slow, the user might end up leaving because of the on-demand loading time of a picture. If the network speed is fast enough, the experience of exploring the sliders would be like the one of eager loading. As a result, users can access and interact with the core content of the webpage faster, without being hindered by the need to wait for all images to load simultaneously. This not only enhances the overall performance of the website but also ensures that users can quickly engage with the essential information, creating a seamless and more satisfying browsing experience.

4.1.4 Lighthouse extension reports

For eager loading, both the first contentful paint and the largest contentful paint took about 1.3 seconds to render. For the other strategy, the first contentful paint took 0.5 seconds and the largest contentful paint took 0.6 seconds. The first metric refers to how fast a user can see an element on the page, while the second metric refers to how fast a user can see the largest element.

Because of its static web page nature, the SEO score is the same for both loading strategies.

4.2 Full-Stack Application for an Online Library

This application represents the graphical interface for an online library, which can be used by users to browse the available book collection and place book orders. The books, authors, categories of books, and users can also be managed from the admin panel of the application.

The development of this project was done using Typescript and the Angular framework. It contains two main modules, a public one and a private one. The public one is available to all users, even if they're unregistered, while the private one is dedicated to the managerial needs of the library. There is a total number of 18 Angular modules, containing bundles of HTML, CSS, and Typescript components. As a default setting, the Angular framework uses eager loading.

One of the most obvious uses of lazy loading in the context of this application is to separate the loading of user and admin modules.

4.2.1 Initial Load Time

Breaking down a large codebase into smaller, focused modules, promotes a more manageable and understandable structure, making it easier for developers to collaborate and maintain the code over time. Additionally, the modular approach facilitates code reuse, as individual modules can be utilized across different parts of a project or even in other projects, reducing redundancy and promoting a more efficient development process. Furthermore, by loading only the necessary modules when needed, there is a performance benefit, as it minimizes the initial load time of a web page.

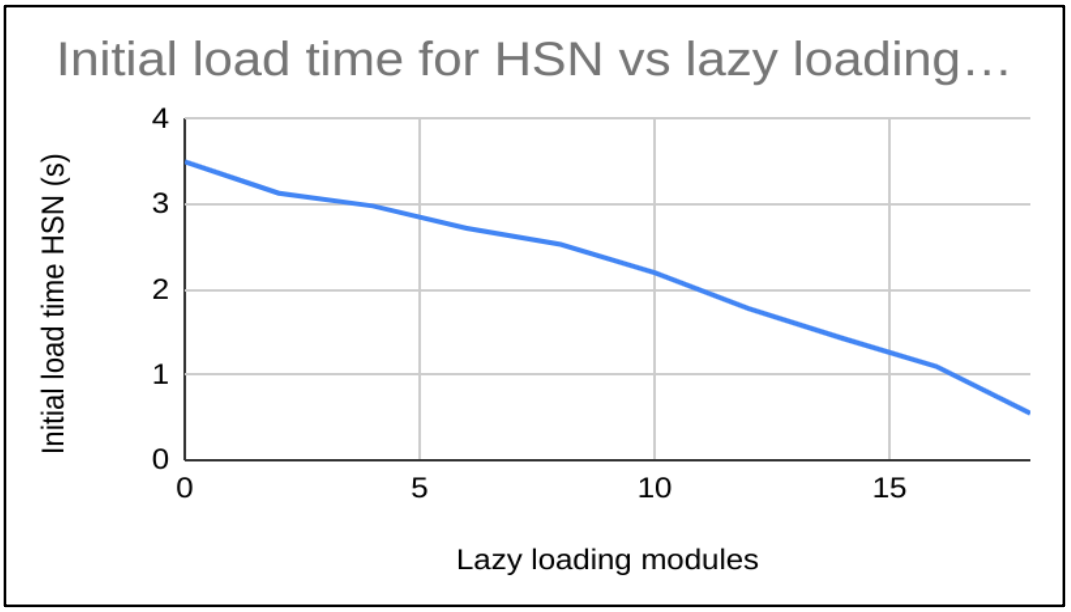


Figure 5. Initial load time HSN vs lazy loading modules

Modules are lazily loaded from the bottom up, starting from the smaller ones. As expected, when lazy loading the admin module, the load time decreases the most, compared to the other measurements, 0.6 seconds.

The slower nature of the 3G network is easily observed here along with the benefits of lazy loading the modules, as the load time decreased from 132 seconds to 40 seconds.

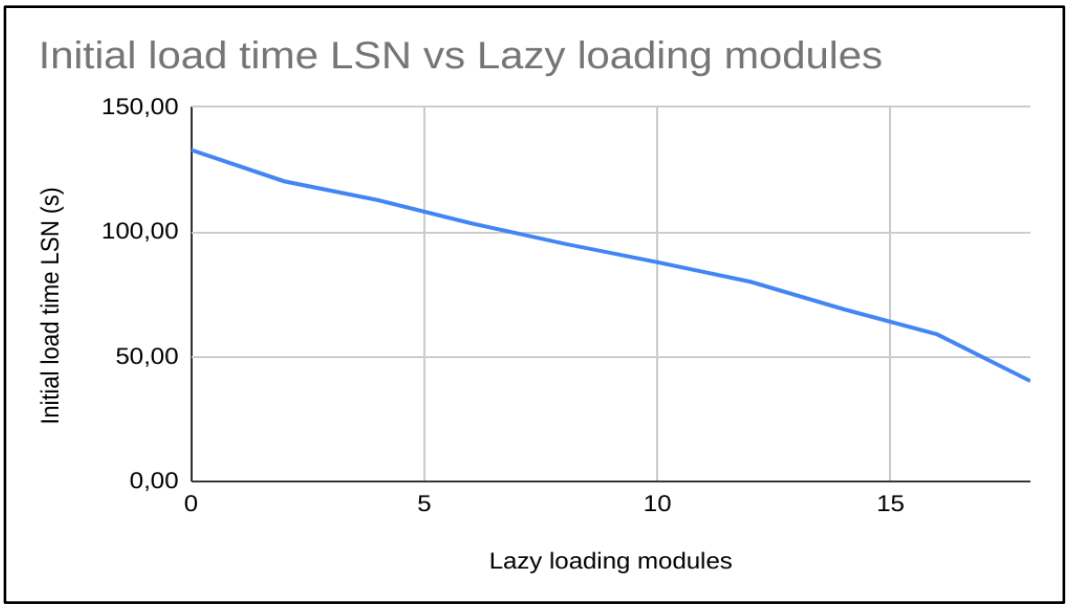


Figure 6. Initial load time for LSN vs lazy loading modules

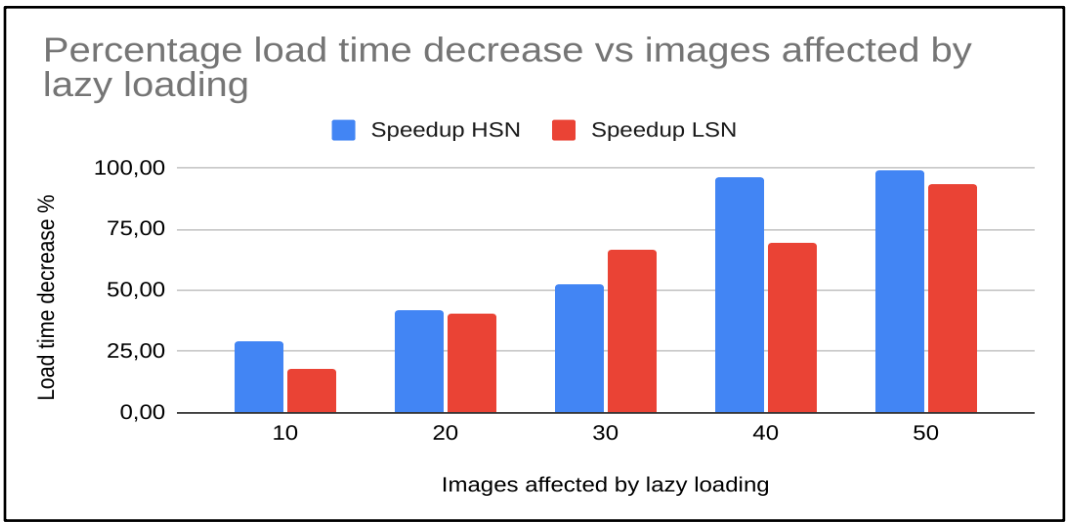


Figure 7. Speedup of HSN and LSN

On the other hand, the constrained bandwidth results in slower data download speeds, leading to delayed loading of web pages, especially those with large or numerous assets such as images, scripts, and stylesheets. So, as a consequence of the slower 3G network, the speedup effect is not as effective as the previous network.

4.2.2 Network Utilization

As anticipated, as the number of lazy-loaded modules rises, so does the initial bandwidth demand. While using an API to retrieve data from a database, this application is not as representative as the previous static website example. Since lazy loading only delays the loading of Angular components that constitute the user interface, the content is therefore dynamic and the amount of data transferred while visiting the same page may vary at different times.

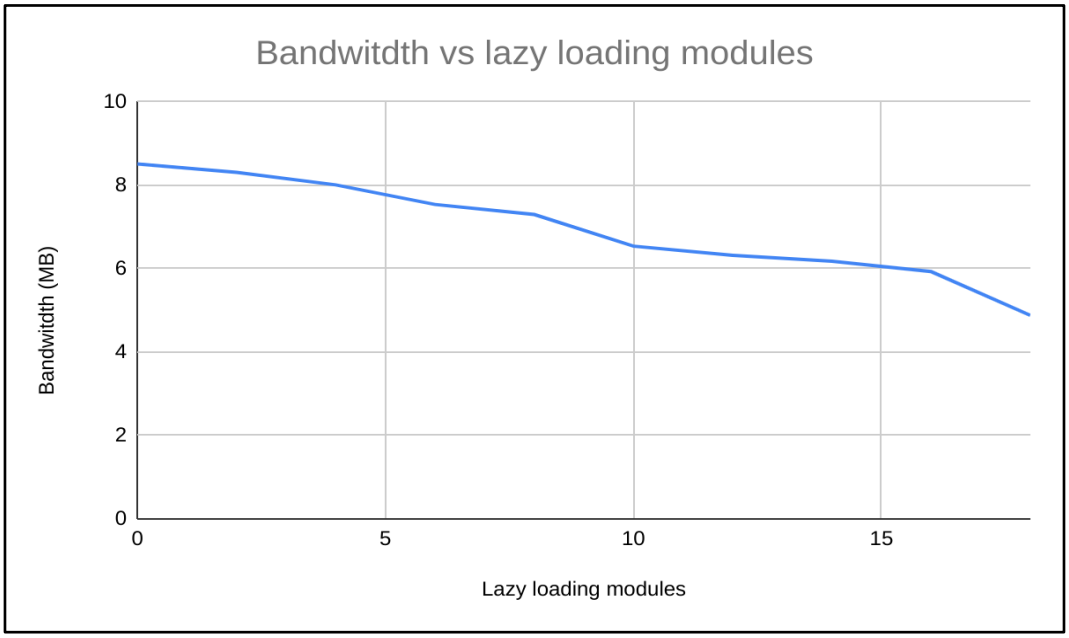


Figure 8. Bandwidth usage

4.2.3 User Experience

The decision to load Angular modules eagerly or slowly has a big impact on how a web application feels to its users. Because only the relevant modules are loaded when needed, lazy loading has the advantage of faster initial page loads, improving user responsiveness. As a result, the connection runs more smoothly overall and bandwidth is used more effectively. However, because the necessary modules are loaded dynamically, users may experience a brief lag when they browse to a new part. Positively, lazy loading reduces the

initial resource footprint and improves scalability, which is especially useful for larger applications.

On the other hand, eager loading guarantees that every module loads during the first page load, giving users instant access to all of the application's features. Although this method removes waiting times when navigating between sections, it frequently results in lengthier initial load times, which may affect how quickly the application is perceived, especially for slower network connections. For smaller applications, when the trade-off between initial load time and instant access to functionality is more acceptable, eager loading is beneficial.

To maximize the overall user experience, the two solutions should be carefully assessed depending on the unique features and needs of the Angular application.

4.2.4 Lighthouse extension reports

Single Page Application (SPA) frameworks like the ones developed with Angular can face SEO challenges due to their dynamic content-loading approach. SPAs typically load a minimal HTML shell initially and fetch the content using JavaScript, raising difficulties for search engine crawlers to effectively index content and interpret unique meta information for each page. Traditional websites, which render different HTML files for each page, benefit from clearer link structures and specific SEO elements. While some search engine crawlers have improved their JavaScript execution capabilities, challenges persist, leading to potential delays in indexing and ranking SPAs. Developers often employ strategies like server-side rendering or prerendering to generate static HTML snapshots, making SPAs more SEO-friendly by providing better crawlable and indexable content.

Therefore, after using the Lighthouse extension on different pages of the application, with and without lazy loading the **SEO score** is fairly low compared to a traditional website that renders different HTML files to the user. So, both approaches offer similar results.

5. Conclusions

To sum up, deciding between eager loading and lazy loading in web development requires careful evaluation of trade-offs. Especially in larger applications, lazy loading helps optimize initial page load times, save bandwidth, and improve overall performance with its on-demand module loading approach. Although there may be a small delay when switching between sections, this approach encourages a more responsive user experience.

However, eager loading guarantees that all of an application's features are available right away during the initial load. This might be useful for smaller applications, but it may result in lengthier initial load times.

Deciding about applying lazy loading is a contextual matter. It may be preceded by several experiments, varying the important parameters related to the application (content size,

content type, number of loaded modules) or the environment (network speed and configuration).

References

[1] A. TURCOTTE, S. GOKHALE and F. TIP, *Increasing the Responsiveness of Web Applications by Introducing Lazy Loading*, 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE), Luxembourg, Luxembourg, 2023, pp. 459-470, doi: 10.1109/ASE56229.2023.00192.

[2] J. JUVILER, *Lazy Loading: How It Decreases Load Time and Increases Engagement* https://blog.hubspot.com/website/lazy-loading-eager-loading

[3] X. S. WANG, A. BALASUBRAMANIAN, *Demystifying Page Load Performance with WProf*, Proceedings of the 10th USENIX conference on Networked Systems Design and Implementation, April 2013, Pages 473–486

[4] S. FRAY, *Lazy Loading VS Eager Loading || What Impact Do They Have?*, https://simon-frey.com/blog/lazy-loading-vs-eager-loading

[5] F. ARNTZ, R. VISCOMI, *The performance effects of too much lazy loading*, https://web.dev/articles/lcp-lazy-loading#causal_performance

Bibliography

F. ARNTZ, R. VISCOMI, *The performance effects of too much lazy loading*, https://web.dev/articles/lcp-lazy-loading#causal_performance

A. BLANKSTEIN, S. SEN, M.J. FREEDMAN, *Hyperbolic Caching: Flexible Caching for Web Applications*, 2017 USENIX ANNUAL TECHNICAL CONFERENCE (USENIX ATC '17), Page 499-511

S. FRAY, *Lazy Loading VS Eager Loading || What Impact Do They Have?*, https://simon-frey.com/blog/lazy-loading-vs-eager-loading

F. AL-HAWARI, *Software design patterns for data management features in web-based information systems*, Journal of King Saus University-Computer and Information Sciences, Volume34 Issue10 Page 10028-10043, Part B, doi: 10.1016/j.jksuci.2022.10.003

J. JUVILER, *Lazy Loading: How It Decreases Load Time and Increases Engagement* https://blog.hubspot.com/website/lazy-loading-eager-loading

E. MJELDE, A. OPDAHL, *Load-Time Reduction Techniques for Device-Agnostic Web Sites, Journal of Web Engineering*, Volume 16, Issue3-4, Page 311-346, June 1, 2017

M.D. SALAS-ZARATE, G. ALOR-HERNANDEZ, R. VALENCIA-GARCIA, L. RODRIGUEZ-MAZAHUA, A. RODRIGUEZ-GONZALEZ, J.L.L. CUADRAD, *Analyzing best practices on Web development frameworks: The lift approach*, Science of Computer Programming, Volume 102, Page 1-19, doi: 10.1016/j.scico.2014.12.004

N. TUAYCHAROEN, V. PRODPRAN, B. SRITHONG - *ClassSchedule: a Web-based Application for School Class Scheduling with Real-time Lazy Loading*, Proceedings of 2018 5th International Conference on Business and Industrial Research (ICBIR), Page 210-214

A. TURCOTTE, S. GOKHALE and F. TIP, *Increasing the Responsiveness of Web Applications by Introducing Lazy Loading*, 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE), Luxembourg, Luxembourg, 2023, pp. 459-470, doi: 10.1109/ASE56229.2023.00192.

X. S. WANG, A. BALASUBRAMANIAN, *Demystifying Page Load Performance with WProf*, Proceedings of the 10th USENIX conference on Networked Systems Design and Implementation, April 2013, Pages 473–486

A.K.I. YASARI, A.D. ABBAS, H.A. ATEE, L.A. LATIFF, R.A. DZIYAUDDIN, D.A. HAMMOOD, *Lazy Multi-Level Dynamic Traffic Load Balancing Protocol for Data Center (LMDTLB)*, Journal of Engineering Science and Technology, Volume 16, Issue 3, Page 2439-2453, June 2021

VERSATILE MULTI-AXIS CHASSIS A SOLUTION FOR RESCUE OPERATIONS AND MOBILITY

Andrei BĂZĂVAN[1]

Eusebiu-Rosini IONIȚĂ[2]

Abstract

The development of robots featuring holonomic drivetrains, more exactly swerves, is a comprehensive area of study with numerous potential applications across diverse industries and fields such as metallurgical or car production, where heavy weight mobile transporters are required. Unfortunately, the narrow use of this new technology produced in limited batches and its poor documentation led to its avoidance in some fields, such as hospitality or mining.

These drivetrains can precisely navigate through difficult terrains and environments, making them an optimal fit for search and rescue operations, where the need for quick and efficient mobility in remote, hard-to-reach areas is paramount. Moreover, sensors, LIDARs and cameras can be mounted on these robots to collect parameters providing information about hazardous environments without putting human lives at unnecessary risk.

Due to their high modularity, such chassis can be mounted on patients' beds in hospitals, where there is not enough space for complex maneuvers. Thus, mobility is enhanced in crowded situations where patients must be moved from one room to another, or during an emergency when an efficient evacuation is crucial.

To effectively assess the capabilities of the drivetrain, we designed a scaled-down prototype. We collected data on mobility, off-road capability, modularity, and maintenance requirements for our prototype and three conventional chassis. Upon analyzing the test results, we found that the prototype consistently demonstrated high performance, being 15% more efficient when averaged over all the experiments considered. The findings of this study confirm the effectiveness of this new generation equipment and can support its incorporation in various fields, especially healthcare.

Keywords: holonomic drivetrain, mobility, search and rescue, hazardous environment

JEL Classification: O14

[1] Student, Mihai Viteazul National College, Romania, andreibazavan25@gmail.com
[2] PhD, 3rd Degree Scientific Researcher, National Institute for Laser, Plasma and Radiation Physics, Romania, ionita.rosini@infim.ro

1. Introduction

Swerve-driven robots represent a fundamental discovery in the field of engineering due to the compact and reliable mechanisms that allow independent control of every individual wheel in any direction, having an extra degree of freedom compared to conventional wheels. Their high maneuverability and precision make them optimal for tasks that require precise positioning, such as machine vision systems and robotic arms. The capability to instantly change their direction of movement in response to different types of terrain or obstacles is unmatched by traditional fixed-wheel robots, making them invaluable in unpredictable environments.

Modularity is a particularly important point; swerve type chassis can be easily customized for any type of use ranging from mini scale robots to applications the size of huge industrial machinery. Also, the ease with which components can be replaced or upgraded ensures longevity, adaptability, and simple maintenance [4].

The ability to navigate through difficult terrains and environments where traditional chassis may struggle, makes these chassis a perfect solution for applications such as search and rescue. In military and police operations, the need for quick and efficient mobility in remote or hard-to-reach areas is primary. Conventional vehicles can be difficult to operate in challenging terrain, but a swerve-driven robotic chassis would offer increased accuracy and mobility. These robots can easily cross a range of surfaces including rocks, sand, or snow, which are typically challenging. Moreover, sensors and cameras can be mounted on these robots to collect parameters, making them able to provide information about the situation in hazardous environments while minimizing risking of human lives.

The use of these robots is also helpful in medical emergencies such as natural disasters, where access to locations is critical. In such scenarios a swerve-driven robotic chassis can provide an efficient and safe solution for carrying medical supplies or equipment, transporting patients, and providing first aid, all while being remotely operated from a safe distance, protecting rescuers. In addition, these chassis, being highly modular, can be mounted on patient beds in hospitals, where there is not enough space for complex maneuvers, thus allowing more mobility in crowded situations.

This type of chassis offers a high degree of precision and control, enabling precise operations in various fields such as manufacturing and industrial settings. They can be used for handling and transporting delicate equipment, reducing the risk of damage to the equipment itself but also of injury among workers. Production processes can be optimized by minimizing human error and increasing efficiency.

2. Problem Statement

Although swerve chassis offer numerous benefits that can enhance and optimize various industrial sectors, they are still not widely used, being relatively unexplored and tested. The direct competitor of these chassis, the mecanum ones, have gained popularity in the last

decade, even finding applications in the aeronautical field (Fig. 1). Mecanum chassis can face challenges when moving on non-flat surfaces, a situation that does not affect swerve chassis, providing them with an added versatility. Swerve chassis represent an improvement for any field and can reduce maintenance costs due to their high modularity.[3]

Figure 1. Mecanum transporter for Airbus components[3].

Through this paper, we aim to prove that swerve-driven robots are the future of robotics and have the potential to revolutionize the field of robotics. By demonstrating the superior capabilities of swerve-driven robots, we hope to inspire more research and development in this area and help pave the way for new and innovative applications of this technology.

To prove our hypothesis, we compared swerves' performance to that of traditional omnidirectional chassis to demonstrate how these robots are better equipped to handle tasks with greater speed, accuracy, and efficiency. We designed, built, and accurately documented a small-scale swerve-driven robot, thus being able to test the most precise version of this chassis type.

In our experiments, we evaluated the performance of the swerve-drive by comparing it to three other chassis types of similar size: a six-wheel drive, a mecanum chassis, and a tread drive. We examined mobility by testing robots on uneven and challenging surfaces, such as obstacle terrain and steep inclines. Precision and maneuverability were tested by subjecting the robots to tasks requiring accurate movement and positioning, including navigating through tight spaces, picking up and moving objects, and executing complex maneuvers. Our analysis was conducted using a standardized method, ensuring a comprehensive evaluation of each chassis.

3. Developing a Swerve Chassis

Swerve drivetrains use four or more independently controlled wheels usually mounted on the corners of chassis, each with its own motor and steering mechanism. This allows for

[3] Source: https://roboticsandautomationnews.com/2016/06/30/kukas-monstrous-robotic-vehicle-manoeuvres-gigantic-airbus-components-with-millimetre-precision/6005/

greater flexibility and precision in movement, as each wheel can move and rotate independently of the others.

To have a fair comparison between all four different types of chassis we built a similar sized swerve-driven one as the other three that we already had access to (all of them were First Tech Challenge chassis, dimensions varying from 28cm x28cm to 45cm x45cm).

For this first prototype it was expected that many things would change, thus we opted for a highly modular design, enabling an easy shift from one idea to another without manufacturing new sets of numerous complex components. We tried to use as many parts as possible from the local vendors, thus the iteration time being minimized, without the need to wait for the delivery of new parts.

Designing a robot chassis is a critical step in the development of any robotic system. It is crucial to design the chassis with precision to ensure that the robot operates smoothly and efficiently. Some of the main priorities while designing the first prototype were:

- Compactness: leaving the maximum free space for other systems that could be added to the chassis in the future (ex: a robotic arm).
- Modular design: easy changes between every iteration, being able to take apart every part of the system to study and find its failure points; parts that tend to wear faster can be replaced without disassembling the whole system.
- Easy to manufacture using standard 3D printers and a 3-axis CNC router; need to find solutions to restrict the manufacture process without reducing project's feasibility.
- Lightweight design for improved performance; used topology studies to determine the best weight/strength ratio for selected applications.
- An overall cost under 1500$

3.1 Design

To have a prototype with seamlessly integrated components, the use of 3D design software was mandatory. Having experience working with SolidWorks[4], a computer aided design software who can also provide support for computer simulations, which are useful during optimization process. Using this software, we were able to create and visualize in real-life a virtual model of the robot and test its performance before building the physical prototype.

One of the primary benefits of designing a robot chassis in 3D design software is that it helps identifying any flaws or design errors before building the real prototype. This approach saves time and reduces costs by mitigating the risks of building a faulty robot. In addition, 3D design software allows precise measurement for accurate dimensioning,

[4] https://www.solidworks.com/

having a small margin of error, which was essential for ensuring that all components fit together perfectly.

Moreover, CAD allowed us to test different scenarios and make modifications to the design quickly. This feature is particularly useful when designing complex robotic systems with multiple components and subsystems, like the swerve-driven chassis. We were able to test different configurations and evaluate how changes in one subsystem would have affected the overall performance of the chassis.

For the first iteration we decided to build the smallest possible reliable chassis, so we restricted its dimensions to a 300mm x300mm footprint (Fig. 2). Due to the high modularity, from this point dimensions always can increase just by changing one part, the main 6mm CNC machined aluminum plate that holds the hole assembly together.

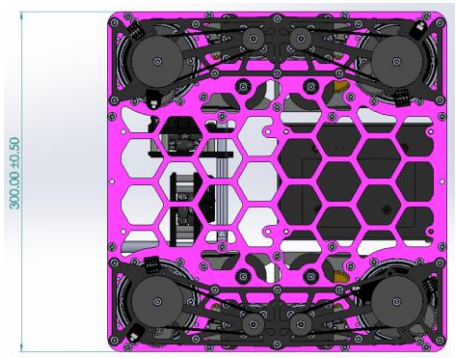


Figure 2. Top view of the swerve chassis showcasing its footprint.

We designed and manufactured our own wheels, thus enabling more freedom during the design process, without being restricted by wheel sizes offered of the vendors. Wheels are composed of two parts: a *3D printed rim*, and a *tire* (like car wheels). Due to the small dimensions, we couldn't fit a pneumatic tire, so we enhanced rims with Andymark's Gray Grippy Tread fixed with 4 bolts into the rim making it a very easy-to-change part. The first prototype wheels are 25mm wide and have a diameter of 72mm, making them able to cross obstacles of moderate size in relation to the size of the entire chassis.

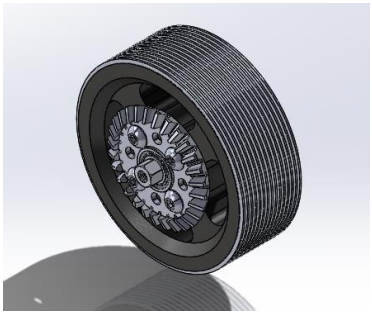


Figure 3. Wheel assembly.

We had two options to mount the main four X-contact bearings that connect swerve modules to the main aluminum plate, one was press-fit into the chassis frame and the other was to put four bolts that intersect the bearing cavity on both parts of the frame, acting like a shim for holding the bearing. The first option required a very thin tolerance on the parts and didn't seem as reliable and modular as the second one where any bearing can be taken apart if it is damaged only by removing four M4 bolts, so we stuck to the second option.

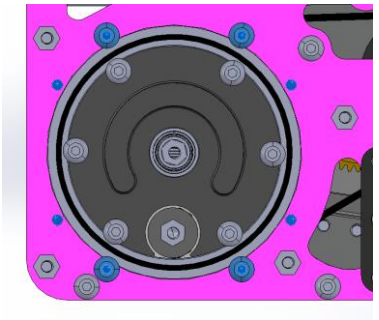


Figure 4. Top view of X-contact bearing fixed with 4 bolts highlighted in blue.

We had a variety of options for the motors that run every individual wheel; the main vendor sold bare motors with different gearboxes mounted to them. The stock gearbox offered by the vendor could wear out over time due to the inconsistencies during production phase, being very difficult to replace. Moreover, to accomplish our priority of having a modular design we chose to buy just the 6000RPM bare motors, without gearboxes attached to them.

Figure 5. Left[5] – Motor with Gearbox; Right[6] – Motor without Gearbox.

[5] Source: https://www.gobilda.com/copy-of-5202-series-yellow-jacket-planetary-gear-motor-3-7-1-ratio-1620-rpm-3-3-5v-encoder/
[6] Source: https://www.gobilda.com/modern-robotics-matrix-12vdc-motor-with-8mm-rex-pinion-shaft/

We manufactured a custom reduction gearbox so we could change its ratio accordingly to the mission. For the first prototype we chose a 1:10 reduction composed of three stages (Fig 6):

- The first stage: 1:2.5 by using a 3M HTD belt (was the best option because spur gear reductions would put very much resistance at the speed of 6000RPM that the motor shaft has)
- The second stage: 1:2 using spur gears.
- The second stage: 1:2 using 90-degree bevel gears.

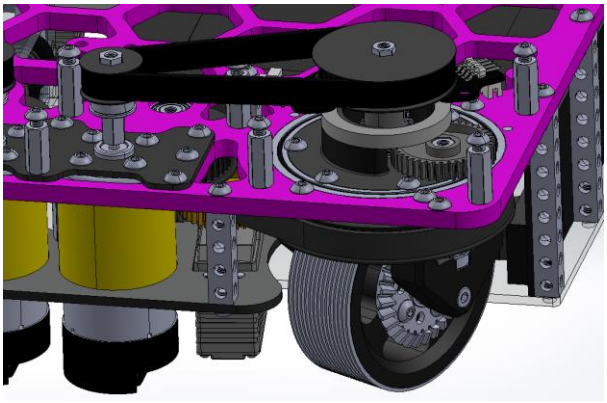


Figure 6. Motor and the 3 transmissions through which it powers to the wheel.

The module rotation must be fast and precise. The biggest problem was caused by the friction created between the wheels and running surface during the scrub generated while modules are rotating. Besides fast servomotors to rotate every module we needed some that can provide enough torque, thus having nearly infinite module rotation acceleration. To connect the servomotor with the module we chose belt drive (3M HTD) because a spur gear transmission would have passed all wheel shocks caused by the rough terrain directly to the servo shaft, causing its failure rapidly, thus we increased reliability. Servomotors spin modules to a 1:1 ratio composed by a first 2:1 spur gear stage and then a 1:2 belt stage, we chose these ratios to save up more space.

Figure 7. Servomotor connected to the module through two transmissions.

To ensure that the module rotation is as precise as possible we chose not to use servos with internal absolute encoders because, due to the two-stage transmission, inaccuracies could occur. Instead, we mounted an absolute magnetic encoder right above the module, thus monitoring its motion with outstanding precision.

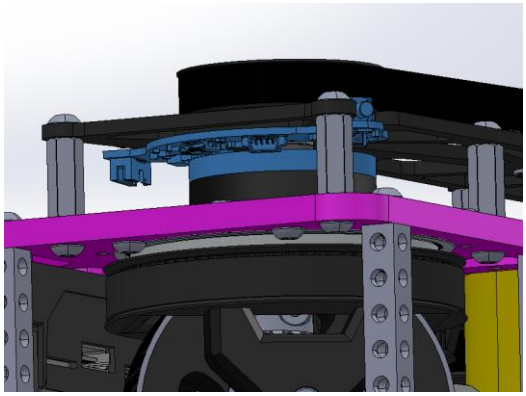


Figure 8. The external absolute encoder is mounted above wheel module.

Localization is a means for being able to locate the position of the chassis at any point in time, for the precision tests we opted to add two OpenOdometry[7] modules, that use REV relative encoders, which, together with the IMU sensor inside the Control Hub, can precisely track robot's position. Odometry is a form of localization that uses data from encoders to derive an estimated position relative to a starting point.

[7] https://openodometry.weebly.com/

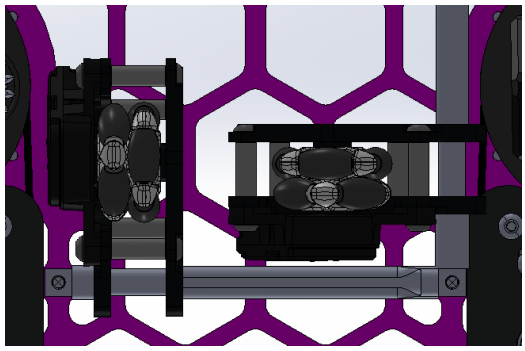


Figure 9. Position of the two odometry pods.

3.2 Mechanical Simulations

To facilitate the understanding of the functioning of the mechanical simulations and how they are performed, we have chosen to detail the process through which a standard part goes through its evaluation**.**

As shown in Fig. 10, the first step in performing a mechanical simulation in SolidWorks is to select which parts will remain fixed during the action of some forces. Basically, these parts will remain rigid.

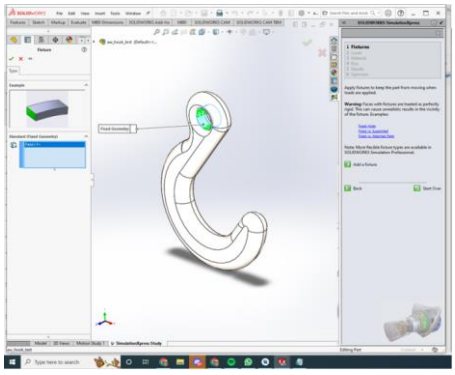


Figure 10. Selecting fixed points.

The next step is to apply the forces we want. In Fig. 11 we have to select the faces where the force will be applied, the plane on which the force acts and also its magnitude.

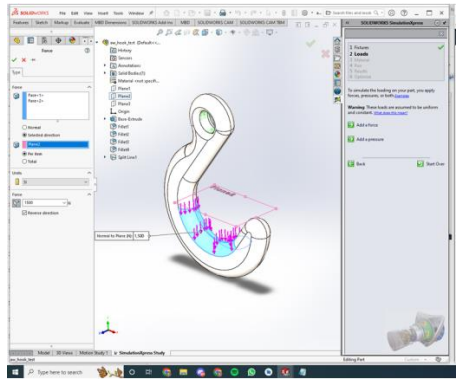


Figure 11. Applying forces.

The third step is selecting the material. SolidWorks gives us a multitude of materials and we can create our own materials, for example if we use a special plastic filament for 3D printers. The material selection interface is shown in Fig. 12.

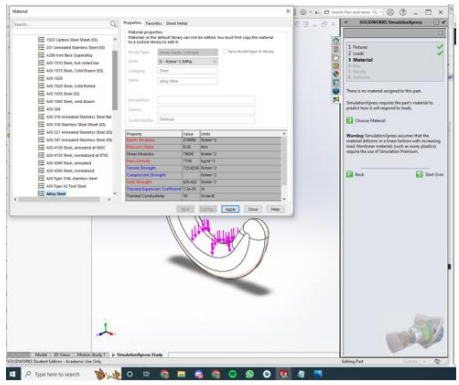


Figure 12. Selecting Material.

Once we have prepared all the necessary data for the simulation, we run it and get 4 things as a result:

- An animation showing the deformation of the part.
- A Von Mises stress graph.
- A graph showing how much each point of the part has moved.
- A graph that shows us where the safety factor is below the value we select.

The Von Mises Stress Graph For this part is the Fig. 13. Stress is represented as the multitude of internal forces generated by a flexible body when it interacts with an external force. The von Mises method of stress measurement is one of the most widely used in industry, measuring equivalent stress.

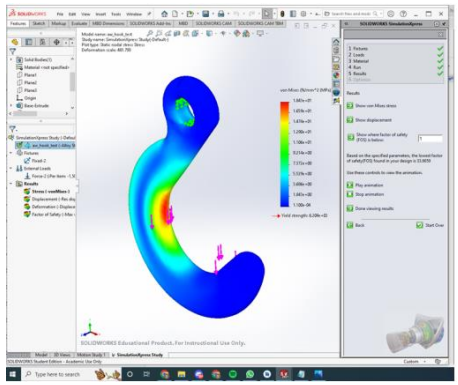


Figure 13. Von Mises Stress Graph.

On the right of the part, it can be seen the von Mises stress scale, red indicates the maximum calculated stress and blue the minimum stress. This way we can see which segments of the part are most stressed and use this information to stiffen them.

Factor of Safety for the part analyzed above in Fig 14, it shows which are the points where the safety factor has a value below 1.

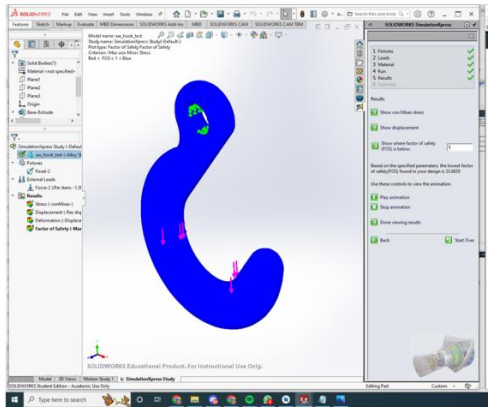


Figure 14. Factor of Safety.

The Factor of Safety is calculated by dividing the yield strength of the material by the equivalent stress at that point. The Von Mises criterion says that a point in a part fails when the equivalent stress at a point is greater than the material strength factor. If a point has a yield strength factor value below 1 it means that the material has started to break.

With this measurement we can tell if we have vulnerable points much more easily. We can also tell the maximum pressure a material can withstand.

For example, if we apply a force of N newtons at a point and the factor of safety at that point, estimated by SolidWorks, is K, the maximum force that can be applied is N*K. So,

we can figure out what forces we can put on that system or part to get the maximum performance out of it.

For most of the manufacturing process we were able to use 6082 series aluminum plates and PLA filament for the 3D printed parts. To determine what material to use for every part we ran computer simulations to make sure that the desired material provides enough strength for its application.

Bearing in mind that for CNC milling we have access only to a 3-axis machine, we could only manufacture plates made of aluminum, no more complex geometry parts because they would require a 5-axis machine. So, deciding which parts to be made of aluminum was straightforward to us.

For the main plate we needed a 300mm x300mm x6mm aluminum plate which would be very heavy, nearly 1kg, so we were required to effectuate material removal to make it lightweight, the disadvantage of this solution is that it can affect its strength. To find the perfect relation between weight and strength we used a Factor of Safety bigger than 1.3, this way we were sure that it won't fail besides its low weight. The final item reduced its weight by more than 50%.

For this simulation we applied a force of 10000N on the plate, equivalent to a weight of 1000kg. The fixed points of this simulation were the holes where the modules will be mounted.

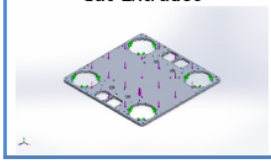
Solid Bodies			
Document Name and Reference	Treated As	Volumetric Properties	Document Path/Date Modified
Cut-Extrude6 	Solid Body	Mass:1.03874 kg Volume:0.000387589 m^3 Density:2,680 kg/m^3 Weight:10.1796 N	D:\Robotica\SWERVE GAME\Placa swerve.SLDPRT Jul 12 01:19:27 2023

Figure 15. Main plate without weight reduction properties.

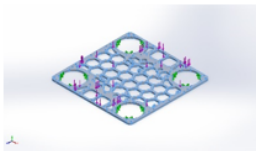
Solid Bodies			
Document Name and Reference	Treated As	Volumetric Properties	Document Path/Date Modified
Fillet7 	Solid Body	Mass:0.565642 kg Volume:0.00021106 m^3 Density:2,680 kg/m^3 Weight:5.54329 N	D:\Robotica\SWERVE GAME\Placa swerve.SLDPRT Jul 12 01:19:27 2023

Figure 16. Main plate with weight reduction properties.

Although after the main plate's weight has been reduced to almost half due to the pocketing process, its structural integrity can still withstand our application (Fig. 17).

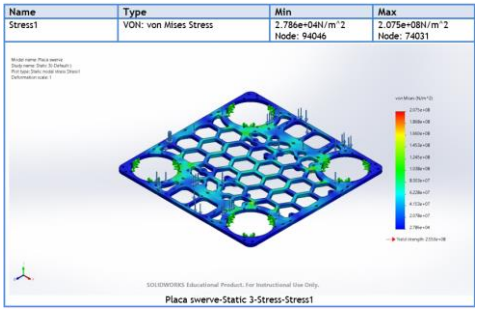


Figure 17. Von Misses stress graph for the main plate.

To determine the best shape for rims we implemented topology studies in the design process to find the part with maximum efficiency. Weight played a crucial role in this design because a light wheel has reduced momentum while spinning and improves performance. While running this study we applied a torque of 10Nm to the holes where the power shaft is attached and fixed the outer part of the rim, thus we ensured that the wheel won't break under big loads caused by the fast changing of spinning direction. The resulting item was a 3-spoke rim with 60% reduced weight.

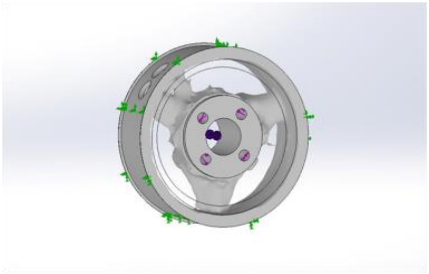


Figure 18. Wheel rim after running the topology study.

3.3 Manufacturing

A key aspect setting this drivetrain apart is that 90% of its structure is composed of custom parts, unfortunately, having a limited manufacturing capacity, this implied a relatively long fabrication time of one week.

For the custom parts manufacturing process, we had access to standard 3D printers (Prusa MK3S+) and a 3-axis CNC (Shapeoko Pro XL). 3D printers were very important during this project, due to their speed and ability to generate parts with complex geometries.

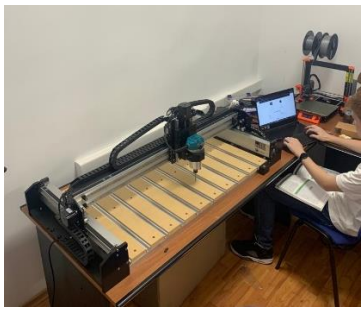


Figure 19. The CNC machine and one 3D printer used in the manufacturing process.

To ensure that all components are seamlessly integrated, we have 3D printed all custom components first. This not only ensured a high level of convenience during the prototyping process, but also guaranteed that all components would fit correctly before manufacturing them out of aluminum.

However, 3D printing is not without its limitations. For instance, the strength of PLA plastic that we used is insignificant in comparison to metals such as aluminum, thus making 3D printed parts less suitable for the high-stress applications during this project.

The CNC machine, while requiring more setup time, has the edge in precision and versatility in material choices. For instance, aluminum, due to its high strength, corrosion resistance, and which can be easily found at local suppliers, was the ideal material for CNC machining the prototype plates.

Some 3D printed parts, for example the wheels, required heat-insert nuts mounted to have a safe mode of fixing bolts on them. Fig. 20 shows how the heat-inserts were mount on the wheels to help fix the grey grippy tread that is secured with screws.

Figure 20. Threaded heat inserts mounted in the 3D printed wheel rim.

After all CNC parts were produced some final adjustments were needed due to the machine's incapacity of making threads for the M4 bolts that hold the X-contact bearings. Unfortunately, one tap had an internal crack and during tapping it failed, making one hole unusable, luckily it didn't affect prototype's performance.

Figure 21. All CNC parts produced and a failed tap.

To prevent aluminum from oxidizing we chose to paint it. Having in mind that any paint chip that would enter in the bearing cavity will affect the tolerances, we decided to cover the important areas with tape.

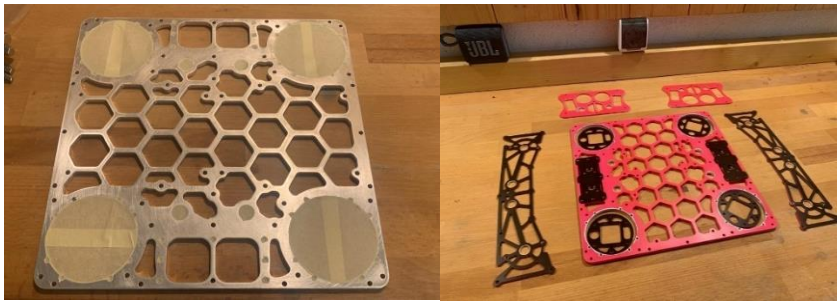


Figure 22. All parts before and after spray painting

3.4 Electrical Components

One of the key components in the first prototype are the four GoBilda 6000RPM 12V bare motors, which power the wheels. These motors are known for their impressive performance in any scenario, furthermore, the output shaft is an 8mm HEX to what transmission can be easily attached without being aware that it will have inaccuracies. These motors also provide an internal encoder that can be used for tracking the wheel's motion, making them the perfect choice for this drivetrain.

We have incorporated Axon Robotics MAX servos for individual module heading control, allowing for fast, powerful, and accurate movement. These standard sized servos provide 34kgcm torque and 0.115 sec/60° speed at 6.0V with titanium gears inside, making them a long-lasting solution.

Additionally, to track module movement we used Lamprey2 Absolute Encoders. Before selecting these encoders for our project, we tested them in comparison to other two magnetic and respectively optic encoders to different speeds of the rotating shaft varying from 5RPM to 6000RPM; the Lamprey2 proved to have the best accuracy.

To power these components, we have used the REV Robotics Control Hub control system thanks to its high-power capabilities in powering robot components and its big number of analog and digital inputs that can be used for sensors and encoders. A REV Robotics Servo Power Module that can provide constant 6.0V output for every servo was added. Teamed up with a 3000 mAh 12V NiMh battery, to ensure a fair amount of power supply for all 4 motors and 4 servos, for any situation.

3.5 Building

The building process required a high level of attention due to the slight tolerances that are featured in this demo version. The modular design enabled a straightforward way to build it by assembling all modules separately and then mounting them on the main assembly.

No additional interventions to the custom parts were necessary during the building process, showcasing the rigor with which the chassis was designed.

The first step was to mount the X-contact bearings that support the modules. Securing the bearings with four bolts on each side proved to be highly efficient, facilitating their process of mounting and removing without any additional speial tools.

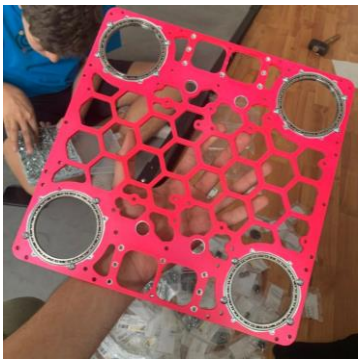


Figure 23. X-contact Bearings are mounted to the main plate.

The next step was to mount the motors and servomotors. At this point we observed that some servos, while mounted on their gear transmission, spined slightly restrained than others. After a fast debug we found that the servomotors that we used had a bigger mounting hole diameter (>4M) which let a slight play for adjusting servos position, thus we were able to adjust the pressure between gears in the transmission without having excessive friction, solving the problem.

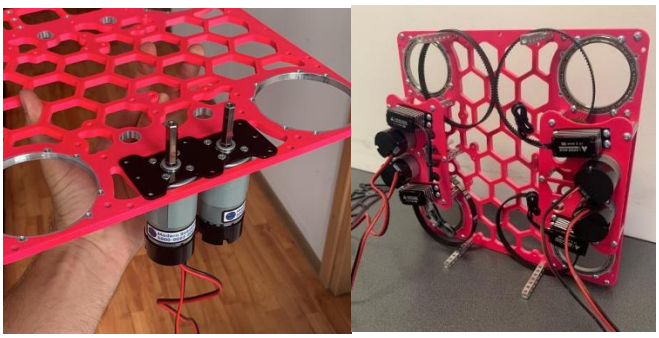


Figure 24. Left – Motors are mounted; Right – Servomotors are mounted.

Mounting wheel modules on the chassis. Here additional +0.25mm shims were necessary to maintain a constant spacing between the 90-degree bevel gears for all four modules, thus minimizing the performance differences between them.

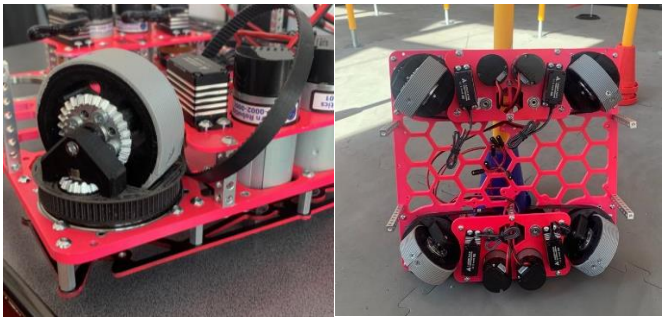


Figure 25. Wheel modules are mounted to the chassis' main structure plate.

After having the wheels mounted and secured it was time to mount the last mechanical components: the upper main transmission and the lamprey encoders for module's rotation tracking.

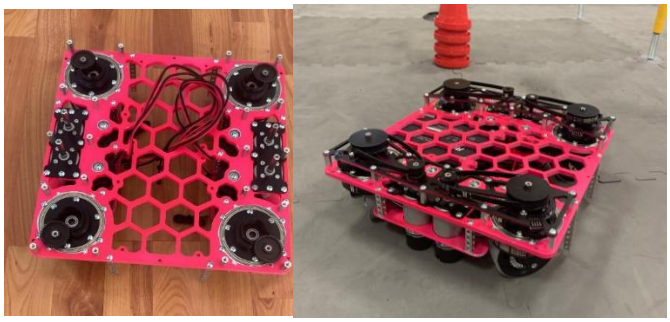


Figure 25. The upper motor transmission and absolute encoders are mounted.

Cable management: having a large number of cables, eight for each module, the best way to retain them and to avoid any hazardous situations was to add a polycarbonate plate under the system to protect them from external interactions.

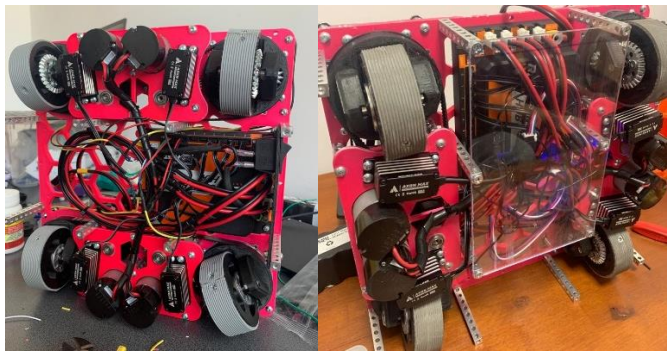


Figure 26. Covering all cables with a guarding polycarbonate plate.

The final step was adding the odometry modules, that, for customization reasons, can be detached using only three bolts and without making changes to the whole assembly.

Figure 27. The odometry subassembly

3.6 Costs

A large part of the cost of this project was allocated to the electronic components as follows:

- REV Robotics Control HUB – 350$ x1
- REV Robotics Servo Power Module – 57.5$ x1
- REV Robotics Through Bore Encoder – 48$ x2
- Modern Robotics/MATRIX 12VDC Motors – 25$ x4
- Axon Max Servomotors – 75$ x4
- Lamprey2 Absolute Encoder – 50$ x4

- Matrix 12V 3000mAh NiMH Battery – 50$ x1

Costs of belts, bearings and other mechanical components bought from vendors – 150$

Aluminum sheets and 3D printers' filament – 150$

Total: 1453.5$

Approximately 2/3 of this amount is allocated towards the acquisition of essential electrical components. The remaining budget is dedicated to the procurement of components that cannot be manufactured in-house and materials for manufacturing. These include items such as belts, bearings, aluminum plates, and Polylactic acid (PLA), utilized in the manufacturing of custom parts through CNC machines and 3D printers.

Owning such machinery for production of the custom parts may not be a feasible option for all entities looking to implement a swerve drivetrain. An alternative route for such cases is to engage the services of local businesses that specialize in such manufacturing processes. The cost of these services will vary depending on the locality and market conditions, and it is advisable to research and negotiate with multiple providers to secure the most cost-effective solution.

4. Studying all chassis

In this section, we conduct a comprehensive analysis of the different chassis types presented in our survey, including the swerve-driven prototype and three conventional chassis: a six-wheel drive, a mecanum chassis, and a tread drive.

Our evaluation focuses on key criteria that are crucial for assessing the effectiveness of each chassis type. The criteria include:

- Mobility: Ability to traverse challenging terrains, including uneven surfaces, obstacles, and steep inclines.
- Precision and Maneuverability: Capability to perform tasks requiring accurate movement and positioning, such as navigating through tight spaces, picking up and moving objects, and executing complex maneuvers.
- Modularity: The extent to which the chassis can be customized for different applications, considering ease of modification, replacement, and upgrade of components.
- Maintenance Requirements: The ease with which maintenance can be performed, including the replacement of worn-out or damaged parts.

Each criterion has been measured using a standardized scale ranging from 0 to 5 (0 representing the lowest performance and 5 the highest).

	Swerve Drive	Mecanum	Six-Wheel Drive	Tread Drive
Mobility	5	3	1	2
Precision and Maneuverability	5	4	2	2
Modularity	4	5	5	4
Maintenance Requirements	4	4	5	4

Table 1. Results obtained by every chassis type for different tests.

Following a closer analysis of all the chassis, we were able to determine certain weak points and strong points for each one.

The mecanum suffered loss of grip due to the specific construction of the wheels with rollers placed at 45 degrees. There is a risk that residues get stuck between the rollers, drastically reducing performance. The maintenance of the rollers was difficult, but at the same time the possibility of quick change of the wheels allowed a high modularity. The precision and maneuverability were similar to those of the swerve type chassis, with small redundancies on challenging surfaces.

Six-wheel and tread drives have shown impressive capabilities to transport heavy objects due to the increased traction they offer. Maintenance for both was very easy, having a small number of moving parts that can wear out during use. Modularity was a strong point, allowing the chassis to be adapted to many conditions. However, maneuverability and mobility suffered, these chassis not having the ability to move holonomically, making them unusable in areas with reduced freedom.

Swerve drives managed to combine all qualities of the chassis presented above. With traction equivalent to six-wheel or tread drives and with the ability to move holonomically like mecanum chassis, swerves become the perfect choice for applications that require a high degree of versatility. Although the system requires a more complex construction to benefit of all these attributes, when implemented correctly, it acquires modularity and easy maintenance, similar to common chassis. The subsystems allow for isolated replacement of faulty parts without requiring the change of the entire assembly as in the case of mecanum wheels, becoming much more efficient in terms of costs over an extended period of time. These chassis offer a variety of options through which can adapt to any environment, for example incorporation of a suspension in module's design [5].

5. Conclusions

The exploration of versatile multi-axis chassis, particularly the swerve-driven robots, represents a groundbreaking solution with far-reaching implications for various industries. The study conducted here affirms the exceptional capabilities of swerve chassis in terms of mobility, precision, modularity, and maintenance efficiency.

The swerve-driven robots showcase unparalleled maneuverability, allowing independent control of each wheel in any direction. An attribute invaluable in tasks that demand precise positioning, making them optimal for applications such as search and rescue operations. The ability to navigate challenging terrains and hazardous environments, coupled with the incorporation of sensors and cameras, enhances their utility in scenarios where human safety is a top priority.

The research undertaken involved the development of a scaled-down prototype, which underwent rigorous testing alongside conventional chassis types. The results clearly demonstrate the superior performance of the swerve-driven prototype, outperforming other chassis in terms of efficiency. This affirms the potential of this new generation equipment, emphasizing its efficacy in diverse fields, particularly in healthcare settings where space constraints and the need for quick, efficient mobility are crucial.

Despite the evident advantages, it is noted that swerve chassis technology is still relatively unexplored and underutilized. The study highlights the need for increased research and development in this area, encouraging broader adoption of swerve-driven robots across various industries. The comparison with other chassis types, including mecanum, six-wheel drive, and tread drive, emphasizes the unique strengths of the swerve-driven design, positioning it as a versatile and efficient choice.

In conclusion, swerve-driven robots present a paradigm shift in the field of robotics, offering a solution that combines the traction of traditional drives with the holonomic capabilities of mecanum wheels. The study aims to inspire further exploration and application of swerve chassis technology, paving the way for innovative and impactful advancements in the realm of robotics and automation.

References

[3] https://www.freshconsulting.com/insights/blog/how-to-build-a-swerve-drive-robot/ - Details about differences between mecanum and swerves, an overview of swerve drives.

[4] https://technology.nasa.gov/patent/MSC-TOPS-74 - Modular Robotic Vehicle (MRV)

[5] https://vikramramanathan.com/files/Senior_Design_Final_Report.pdf - Improved Swerve Drive Design with Suspension.

Bibliography

https://www.freshconsulting.com/insights/blog/how-to-build-a-swerve-drive-robot/ - Details about differences between mecanum and swerves, an overview of swerve drives.

https://technology.nasa.gov/patent/MSC-TOPS-74 - Modular Robotic Vehicle (MRV)

https://vikramramanathan.com/files/Senior_Design_Final_Report.pdf - Improved Swerve Drive Design with Suspension.

https://www.chiefdelphi.com/t/swerve-drive-vs-mecanum-wheel-drive/136385/21?page=2 – Swerve and Mecanum Comparison

https://www.first1684.com/uploads/2/0/1/6/20161347/chimiswerve_whitepaper__2_.pdf

https://team1640.com/wiki/index.php/Swerve_Central

VUCA AND BANI WORLDS - CHALLENGES FOR NOWADAYS BUSINESS MODELS. COULD CIRCULAR ECONOMY AND VARIOUS DIGITAL INSTRUMENTS LIKE INTERNET OF THINGS HELP A BUSINESS MODEL TO RESPOND TO SUCH CHALLENGES? AN EXAMINATION OF SOLUTIONS AND FUTURE STRATEGIES THROUGH MARKETING LENS

Denisa-Roxana BOTEA-MUNTEAN[1]

Raluca CONSTANTINESCU[2]

Abstract

The present paper constitutes a research endeavour grounded in a fundamental question posed by its title: 'Could circular economy and various digital instruments like Internet of Things help a business model to respond to such challenges?' This inquiry is intrinsically linked to the broader context of the last decades, which have witnessed a rapid technological revolution, reshaped global industries, and nurtured a digital ecosystem. While this transformation has yielded considerable advantages, it has concurrently introduced formidable challenges, encompassing climate change, resource depletion, large-scale migrations, pandemic, and conflicts, all of them provoking profound contemplation on the way forward for both businesses and humanity.

In the context of this research, we will delve into the crucial topic of how VUCA, a framework for risk assessment and response, has played a central role in adapting and developing new business marketing strategies to effectively navigate complexity. Additionally, we will explore the significance of BANI, emphasizing the vital need for flexible marketing strategies when confronting unforeseen global crises. The paper focuses on the role of VUCA and BANI frameworks in guiding businesses' marketing strategies. Therefore, the case study will explore how these frameworks have influenced marketing strategies, and what they mean in response to unforeseen global crises.

In our interconnected world, IoT is vital for understanding consumer behaviour and creating data-driven marketing strategies. The transition to circular business models, influenced by environmental concerns, fosters sustainable marketing through eco-friendly products. To navigate the intricate VUCA and BANI landscape, marketers need innovative strategies blending digital technologies, big data, and AI with IoT insights and circular economy principles. This paper will provide essential insights for understanding whether circular economy and various digital instruments like the Internet of Things can empower a business model to effectively respond to the multifaceted challenges of the modern era.

[1] PhD Candidate, Bucharest University of Economic Studies, Romania, munteandenisa21@stud.ase.ro

[2] PhD Candidate, Bucharest University of Economic Studies, Romania, constantinescu1raluca21@stud.ase.ro

Keywords: Sustainable Development, Technological Change, International Business, Marketing, Environmental Economics

JEL Classification: Q01, O33, F23, M31, Q55

1. Introduction

21st century came with many changes either good or bad. Technology evolved exponentially, basically in all areas, whether we talk about healthcare, automotive, retail or even tourism. People get more and more information each and every day, being more connected to each other through digital means, living in a world evolving quickly. This evolution touching all layers came with its downsize effects. Climate change, depletion of resources, migration experienced in certain areas of the globe started to pose questions about the direction of humanity and ways forward.

During the 20th century and over the last decades, humanity faced many challenges. Two World Wars, a Cold War, many and various changes in the geo-politics that had a huge impact on economies and many initiatives in response to these challenges like the United Nations, the World Bank and International Monetary Fund. In this tumultuous world, the worldwide business environment was forced to find solutions, not only to survive, but to thrive, being in a continuous search to find the right business model as a response to what was going on outside their big or small doors.

In the late 80s, two American academics put together a model that could have helped companies for a better and faster response to challenges and assessment of risks. VUCA (Volatility, Uncertainty, Complexity and Ambiguity) concept has been firstly adopted by the U.S. Army as a way to help them to assess the situation and developments of the Cold War. The concept became quite common and used in companies being helpful in quantifying various risks and put together mitigation solutions and strategies.

And when everybody thought that humanity has started to understand its challenges and launched a comprehensive process of finding solutions, March 2020 came and with it a totally new series of challenges: a pandemic which affected the entire world in its deepest layers and brought the world economy on the verge of a crisis, followed by a war which jolted an energy crisis and serious movements in the East. VUCA already felt like an obsolete notion and the term BANI *(Brittle, Anxious, Nonlinear, and Incomprehensible*) has been brought to the spotlight. BANI could be seen as a "the end is near" scenario, but it sheds enough light to put some order into a chaotic world.

In addition to the challenges presented by the VUCA and BANI frameworks, the 21st century has also witnessed a critical shift toward circular business models, sustainability, and eco-conscious marketing practices. With concerns about climate change and resource depletion becoming even more pressing, businesses explore circular economy principles. This paradigm shift allows marketers to create eco-friendly products and campaigns that

resonate with environmentally conscious consumers. Moreover, in our rapidly evolving world, IoT serves as a pivotal tool in understanding consumer behavior and formulating data-driven marketing strategies. This paper will explore how these elements - circular economy, sustainable marketing, and IoT - intertwine with the VUCA and BANI frameworks, offer new possibilities and strategies for businesses to thrive amidst evolving challenges.

2. VUCA and ways it helped companies to build long lasting and sustainable business models

Volatility, Uncertainty, Complexity and Ambiguity or in short VUCA is a notion firstly used by Warren Bennis and Burt Naus, American academics who dedicated their work to management and business administration (Bennis and Nanus, 1985). At that time, it was just a theory without any concrete applications. In 1987, 2 years later, the U.S. Army adopted and started to apply the model, more concretely by describing the multilateral military world after the end of the Cold War in 1991 and the war in Afghanistan (U.S. Army Heritage and Education Centre, 2022).

But what does VUCA mean in practice?

Volatility (V) means repeated, fast and, usually, big changes. For example, in a volatile market, prices may change considerably overnight and nobody could set a trend (Wright and Wigmore, 2023).

Uncertainty (U) comes when events become unpredictable. Situations are not entirely understood, and previous experiences' lessons learnt do not apply. In such an environment investments, growth and development could not take place as nobody would be able to predict the way forward (*VUCA - Leaders with Vision, Understanding, Clarity, Agility!*, 2015)

Complexity (C) is made up of multiple layers of issues, challenges, factors and causes which are complexly connected, their understanding being quite difficult and implicitly the decision-making process being almost impossible. As an example, the gas price could be considered an important distorting factor, given that its change may affect other products and services that are not directly linked or related (Wright and Wigmore, 2023).

Ambiguity (A) is in fact the lack of clarity, with information being misinterpreted, where "one size fits all" and "best practice" could not be applied and the decision-making process is an act of courage, open to risks and even failures (Wright and Wigmore, 2023; *VUCA - Leaders with Vision, Understanding, Clarity, Agility!*, 2015).

It is good that challenges have been identified and reduced to short and approachable notions, but is it possible to be prepared for such challenges and even build a business model that could compensate for their effects?

For *Volatility*, the challenge is utterly sudden without a defined timeline, but information should be available and the challenge not hard to translate. In the above example about prices, one could adopt the strategy of stockpiling, but usually the investment should match the risk (Bennett and Lemoine, 2014). Questions should be put and asked in order to find the right model of business and balance between investments and risks. They should address basic information that could be collected, but crucial for the future, like the high and low limit of the change, the timeline and speed of the change and the resilience of the company (Wright and Wigmore, 2023)

For *Uncertainty*, lack of information is the main cause, but usually there is information about the cause and effect. That is why companies should invest more in collecting, interpreting, and sharing intelligence, also operating structural changes by improving through intelligence networks (Bennett and Lemoine, 2014). For this it is also important for companies to establish the base lines: What might change, and which are the signs of change? Would the company be aware of the change and how fast could it react? (Wright and Wigmore, 2023)

For Complexity, layers and implications are multiple. Some information is unknown, but not all of it. Some info is accessible and other could be predicted and the real obstacle could be the amount of data and too much effort to put in processing with the existing resources. Companies should invest more in developing either human and/or technology resources in order to address more efficient the complexity (Bennett and Lemoine, 2014). It is also important for the companies to start understanding the layers involved and their interconnection, but also its capability to stop a domino effect (Wright and Wigmore, 2023).

For *Ambiguity*, it is a contest of try and learn, of conducting experiments, generating hypotheses, and testing them. It could be costly on a shorter term, but quite useful and add profit on a longer term (Bennett and Lemoine, 2014). At the same time an assessment on knowledge and visibility of various internal and external factors should be performed in order to understand to what degree events could be misunderstood or misinterpreted. It should also be a process of finding clearer direction and the real amount of information needed to take a decision (Wright and Wigmore, 2023)

In a VUCA world, other parameters have been identified which should be monitored in order for companies to have a faster reaction with better results. Choyon in the Financial Express considers the impact of technology on businesses one of these parameters. Technology trends and tracking various shifts like big data collected through IoT, automations built with the help of big data and even AI and chatbots should be assessed and used to predict and respond to sudden changes (Choyon, 2021). Consumer behaviour is another such parameter extremely important especially for businesses driven by customers. Updated SWOT analysis, mapping of resources' skills, assessment of markets, both consumers and supplies, should be considered by all companies when facing the VUCA world (Choyon, 2021).

Leadership and leaders of companies play the crucial part in navigating a business through such distortions. Vision, focused more on short-term initiatives and mid-term goals and data

should be at the forefront of their strategies. Leaders should rely more on data, and this means bringing on more digitalisation into a company to produce more precise trend analysis. Using IoT would ease their jobs in determining consumer behaviour and react quickly to any change that could undermine their companies (Choyon, 2021).

3. BANI or a way to shed some light into a chaotic world

We may say that nowadays we live in a perfect VUCA world. Who would have thought of a pandemic that lasted over 2 years? Or a war potentiating an economic crisis started because of the pandemic and on all of it adding an energy crisis? Many challenges organised in multiple layers which companies had and still have to face without any warning.

In 2020, Jamais Cascio thought that VUCA was already obsolete, and he started to speak about BANI (*Brittle, Anxious, Nonlinear, and Incomprehensible*) trying to find a way to understand the chaos and find ways to respond to it.

The world that we lived in seemed advanced, in a continuous development, ready to face any tasks and to have answers to any challenges. The Pandemic proved that actually our world is *brittle*, rigid and non-resilient, victim of a potential catastrophe at any time (Prasad, 2022). Many of the systems on which human life depends on, like energy, food or global trade already experienced brittleness and only thinking of this induces a lot of *anxiety* (Cascio, 2022). In a world full of choices, making the right one became the hardest thing to do and from here comes the *anxiety* and it seems that VUCA started this feeling of uneasiness of taking decisions (Prasad, 2022). An *anxious* state is detrimental to anyone, as it could lead to an over-reaction, making the things worse than they were, under-reaction, mostly when the bigger picture is missed, or even no-reaction, waiting for everything to collapse (Cascio, 2022). On top of these two notions comes *non-linear*, where the decision or the cause and the outcome are usually unbalanced and out of the known grid, and *incomprehensible*, where answers that everybody tries to find have no meaning and sense (Prasad, 2022).

Even if BANI world sounds gloomy and the novelty and implications of the term are quite deep, academia, think-tanks and various professionals started to look into this and tried to find solutions. The subject is yet to be further studied, but better business planning, resilience, empathy, transparency, and data have already emerged as possible solutions (Grabmeier, 2020)

The industrial revolution and all its discoveries brought us also the depletion of resources and the environmental challenges that we are facing nowadays. Grabmeier in his study BANI versus VUCA: a new acronym to describe the world considers this as one of the starters of our BANI world. But recent years and studies showed that companies found a way to respond to this challenge by applying new business models and transforming their linear models into circular ones.

The qualitative context is based on Grabmeier highlights, where the interconnectivity of global systems as a precursor of a ripple effect concludes that a failure in one country could have catastrophic consequences worldwide. He illustrated this concept by referring to the resource curse, wherein large regions overly rely on their natural resources, only to be superseded by technological advancements. Moreover, Grabmeier emphasizes that in today's interconnected world, a disastrous event in one country can trigger a chain reaction that impacts the entire planet (Grabmeier, 2020). Contrasting with the complexity of the VUCA framework, BANI offers a concise and productive framework for understanding the world. It aims to provide a fresh perspective, enhance comprehension of cause-and-effect relationships, and establish a stable structure for interpreting global dynamics. Additionally, Grabmeier suggests that technologies such as Artificial Intelligence, Big Data, and Data Science offer effective solutions to address misunderstandings within this framework.

4. Navigating the VUCA and BANI worlds: challenges for contemporary business models

In this part of the paper, we embark on a journey through the intricate landscapes of the VUCA (Volatility, Uncertainty, Complexity, Ambiguity) and BANI (Brittle, Anxious, Nonlinear, Incomprehensible) worlds, exploring their implications for modern business models (Forbes, 2023). Against the backdrop of an ever-evolving global landscape characterized by rapid technological advancements and unpredictable socio-economic shifts, we delve into the pressing question: Could the principles of circular economy and the integration of digital tools like the Internet of Things (IoT) empower businesses to effectively navigate the challenges posed by the VUCA and BANI environments? Through a detailed case study and analysis of innovative solutions, this chapter aims to shed light on this critical inquiry and provide insights for shaping the future of business strategies.

Exploring the VUCA and BANI context: the premises were set by dissecting the VUCA and BANI frameworks, tracing their origins and evolution in response to the dynamic forces shaping our world. By examining the concepts of volatility, uncertainty, complexity, and ambiguity, as well as their brittle, anxious, nonlinear, and incomprehensible counterparts, we paint a comprehensive picture of the challenges confronting contemporary businesses (Bushuyev, S., Piliuhina, K., & Chetin, E., 2023).

Challenges faced by modern business models: an analysis of the specific challenges encountered by today's business models within the VUCA and BANI contexts was highlighted. From the disruptive effects of rapid market changes to the pervasive sense of uncertainty permeating global markets, we elucidate the obstacles hindering sustainable growth and adaptation. Additionally, we explore how the inherent complexities of these environments can impede the implementation and scalability of innovative solutions, like digital transformation, e-commerce, data-driven or tech savy businesses (Fonseca, M. P., et al., 2024).

The role of circular economy and digital instruments: central to our inquiry is the exploration of whether circular economy principles and digital instruments such as IoT could offer viable pathways for businesses to overcome VUCA and BANI challenges. Drawing on real-world examples and empirical evidence, we assess the effectiveness of integrating sustainability practices and leveraging digital technologies to enhance resilience and responsiveness (Nijkamp, P., Ţigănaşu, R., Bănică, A., & Pascariu, G. C., 2024).

5. Case study – solutions and future prospects

Through an analytical examination of some explanatory case studies, we present businesses within education, retail, and healthcare that have adeptly navigated through the challenges of VUCA and BANI environments. These businesses have achieved success by adopting circular economy principles and leveraging digital innovations. Our analysis delves into the strategies, best practices, and lessons acquired, offering actionable insights for businesses aiming to fortify their operations and excel amidst uncertainty.

Therefore, we reflect on the overarching themes and findings presented in this paper, emphasizing the critical importance of adopting adaptive, forward-thinking approaches in the face of VUCA and BANI challenges. By embracing the principles of circular economy and leveraging digital tools effectively, businesses can not only survive but thrive in an increasingly complex and uncertain world.

5.1. Higher education in the VUCA and BANI World: IoT and Circular Economy Perspectives

The higher education industry is undergoing significant transformations in response to the challenges posed by the VUCA and BANI environments, with the integration of IoT and circular economy principles offering promising solutions.

By considering some various global and technological trends, including Artificial Intelligence (AI) and Machine Learning, Internet of Things (IoT), Blockchain, Augmented Reality (AR), and Robotics, we all witness the reshaping of the education landscape for organizations. IoT, for instance, facilitates the creation of intelligent systems that enhance efficiency and safety, while Blockchain ensures transparency and security in data management. AR and VR technologies provide immersive learning experiences, and robotics offer opportunities for automation and collaboration. Additionally, biotechnologies and genetic engineering contribute to advancements in medical education and personalized learning.

Digitalization serves as a key driver in implementing these trends, enabling educational organizations to digitize processes, utilize digital platforms for enhanced management, leverage cloud computing for data storage and scalability, and harness AI and data analytics for informed decision-making (Bushuyev, S., Bushuyeva, N., Murzabekova, S., &

Khusainova, M., 2023). Embracing digitalization not only transforms customer interactions and business processes, but also fosters innovation and the development of new business models. Educational organizations that successfully adapt to these trends stand to gain a competitive edge in the evolving market landscape.

The research delves into the intricacies of the BANI world in comparison with the VUCA realm, characterized by increasing process chaotization and heightened vulnerability to unexpected adverse events, often resulting in severe consequences (Adzhienko, V. L., Soboleva, S. Y., Knyazev, S. A., Shestakova, I. V., & Zavolochkina, K. A., 2023). In such circumstances, the traditional approach of extrapolating past events to predict the future, commonly used in strategic planning, proves insufficient. An essential aspect of organizational strategic management in the era of Industry 4.0 is to strengthen anti-fragility. The desk research study outlines key priorities for universities in this context, including anticipating future labour market demands, promoting seamless integration of science and education, and embracing interdisciplinary research approaches. These objectives could be achieved through strategies such as the barbell or lateral strategy, the latter challenging traditional patterns of strategic thinking. The investigation exemplifies how universities align their strategies with global and local agendas through lateral development strategies. However, to enhance their antifragility and effectively address the challenges of the BANI world, universities must intensify their efforts to implement lateral strategies, focusing on strengthening horizontal connections in social space, delivering personalized education, and adopting an edutainment approach.

As a result, the integration of IoT and circular economy principles emerges as a pivotal opportunity for higher education institutions. By embracing these principles, universities can effectively address the complexities of the VUCA and BANI environments while advancing sustainability goals (Garcia-Alvarez, M., Rekalde-Rodríguez, I., & Gil-Molina, P., 2023). As educational organizations adapt to global and technological trends such as AI, IoT, Blockchain, AR, and Robotics, they gain the capability to reshape the educational landscape and enhance efficiency and safety. However, in the face of escalating process chaotization and vulnerability to unforeseen events, traditional strategic planning methods prove inadequate. To thrive in the era of Industry 4.0, universities must prioritize antifragility by anticipating future labour market demands, seamlessly integrating science and education, and embracing interdisciplinary research. Implementing lateral strategies, including strengthening horizontal connections, delivering personalized education, and adopting an edutainment approach, will be imperative for universities to effectively navigate the challenges of the BANI world. Through these concerted efforts, educational institutions can not only gain a competitive edge but also contribute significantly to building a sustainable and resilient future.

5.2. Integrating IoT and circular economy principles in healthcare industry

In the VUCA and BANI world, the healthcare industry faces unprecedented challenges, necessitating innovative approaches to ensure effective patient care and operational

resilience. This case study delves into how healthcare organizations are embracing digital transformation through the integration of IoT and circular economy principles. By leveraging IoT technologies, healthcare providers can enhance patient monitoring, streamline data collection, and optimize resource utilization. Circular economy principles further promote sustainability by minimizing waste and maximizing resource efficiency in healthcare operations. Through the adoption of these strategies, healthcare organizations can navigate the complexities of the VUCA and BANI environments, ensuring quality care delivery while fostering sustainability and resilience in the healthcare sector.

Starting from the idea of some specialized authors (Steffen, B., Braun von Reinersdorff, A., & & Rasche, C., 2023), the healthcare landscape grapples with two significant disruptors: hyper-dynamic competition and evolving customer demands. Digitalization, while causing industry turbulence, also offers opportunities to enhance services, making them smarter and more user-friendly. However, closing the digitalization gap is crucial for healthcare providers to transition into VUCA service organizations. Digitalization serves as a catalyst for overcoming challenges arising from traditional hierarchies and evidence-based healthcare approaches. Moreover, the shift towards value-based healthcare emphasizes the importance of patient engagement and proactive health management. Yet, many hospitals face obstacles due to outdated practices and governance structures, hindering their ability to adapt to changing dynamics. This vulnerability makes them targets for digital giants like SAP, Google, and Amazon, who challenge the status quo with innovative solutions tailored to value-based care. Digitalization, encompassing AI, machine learning, and the Internet of Medical Things, holds immense potential to revolutionize healthcare delivery and empower individuals to make informed health decisions. In this context, ICT competencies emerge as vital assets for navigating the uncertainties of the VUCA landscape, enabling strategic and operational resilience. However, true success lies in aligning technology with people and processes to foster adaptability and responsiveness in the face of disruption.

As a result, it is evident that the healthcare industry is undergoing profound changes driven by digitalization and the challenges of VUCA conditions. These disruptions necessitate a paradigm shift in established healthcare organizations towards dynamic and agile service delivery models. Also, many healthcare organizations are struggling to bridge the digitalization gap and adapt to the evolving landscape (Stein Kleppestø). To address these challenges, we propose leveraging IT-based multi-perspective analysis processes, such as enabling holistic understanding and decision-making. This approach facilitates the development of customized digitalization strategies tailored to each organization's unique needs and context. By integrating IoT and circular economy principles into marketing strategies, healthcare organizations can harness the power of data analytics to derive valuable insights, improve patient outcomes, and lower healthcare costs. Embracing these innovative approaches not only enhances organizational resilience but also fosters sustainable growth in an ever-changing healthcare landscape.

5.3. Retail industry

In 2017, Prachi Gupta in their *How to Survive in the VUCA World of Retail* (Prachi Gupta, How to Survive in the VUCA World of Retail, 2017) exposed the challenges faced by the retail industry in taking the decision to move from physical retail to online. Such decisions were not influenced only by consumers who were more inclined towards shopping from their own couches, with various platforms becoming more attractive and appealing and offering more tailored services, but also by the costs involved in managing physical stores and the race to become sustainable.

Some companies and brands took this challenge seriously and started to plan accordingly, moving part of their operations in an online environment, starting to experience the benefits of their decisions: fewer costs, less consumption, more attention towards the online consumer who was more present and more vocal, ready to provide feedback and help companies to improve. Logistics had to be changed, personnel to be re-trained, and all these to re-adapt their businesses and answer to the challenges posed by the digital environment and the necessity to adopt modern circular economy principles.

In 2020, when pandemic hit, those companies who took seriously their pledge on digital transformation, had only benefits, because their business model was already adapted to an online environment. Their response to a VUCA world has actually been a wise preparation for a BANI world, which did not take them by surprise. The same happened also with environmental challenges, as they were already prepared for a stricter stance to take.

6. Methodology

In this exploratory research regarding two major subjects VUCA and BANI worlds and their implications for contemporary business models, this research adopts a multi-faceted approach encompassing literature review, critical analysis, and examination of case studies. (Tavanti, M., 2023).

The methodology begins with an extensive review of academic articles, papers, and internet sources pertaining to VUCA and BANI frameworks. This review aims to identify key concepts, challenges, and strategies associated with navigating volatile, uncertain, complex, and ambiguous business environments.

By synthesizing insights from diverse sources, the research seeks to establish a comprehensive understanding of the theoretical foundations underlying these frameworks.

Drawing upon the insights garnered from the literature review, the research employs critical thinking to interpret and contextualize the concepts of VUCA and BANI within the contemporary business landscape. Through critical analysis, the study aims to elucidate the practical implications of these frameworks for businesses, including their relevance, limitations, and potential applications in addressing modern challenges.

To provide empirical evidence and illustrate theoretical concepts in real-world contexts, the methodology incorporates the analysis of relevant case studies. These case studies offer insights into how businesses have responded to VUCA and BANI challenges, leveraging

strategies such as circular economy principles and digital technologies like Internet of Things (IoT). By examining concrete examples of successful adaptation and innovation, the research aims to identify actionable insights and best practices for contemporary business models.

In addition to synthesizing existing knowledge, the research adopts qualitative research aspects to explore emerging themes related to circular economy and digital instruments like IoT. Qualitative approaches, including interviews, surveys, and thematic analysis, may be employed to gather perspectives from academic experts, industry practitioners, and stakeholders. These qualitative insights contribute to a nuanced understanding of the complexities and opportunities inherent in addressing contemporary business challenges.

Central to the methodology is the exploration of interconnections between VUCA and BANI frameworks, circular economy principles, and digital technologies like IoT. By synthesizing these disparate elements, the research aims to uncover synergies and potential pathways for businesses to effectively respond to multifaceted challenges. This synthesis underscores the holistic and integrated nature of contemporary business strategies, highlighting the importance of interdisciplinary approaches in navigating complex environments.

Overall, the methodology adopts a comprehensive and integrative approach to investigate the research questions posed, leveraging diverse sources, analytical frameworks, and empirical evidence to generate meaningful insights and contribute to the scholarly discourse on modern business models and strategies.

7. Results

The results effectively communicate the idea that alongside the challenges posed by the VUCA and BANI frameworks, there has been a notable transition toward circular business models, sustainability, and eco-conscious marketing practices in the 21st century. The premises of the work are conveying the message that businesses are increasingly embracing circular economy principles and leveraging IoT for understanding consumer behaviour and developing data-driven marketing strategies. Overall, the paper effectively sets the stage for exploring how these elements intersect with the VUCA and BANI frameworks to provide new opportunities for businesses facing evolving challenges.

8. Conclusions

The findings of the research underscore the central theme of the study, which is dedicated to delving into the vast potential offered by the integration of circular economy principles and digital technologies such as the Internet of Things (IoT) in assisting business models to effectively address the multifaceted challenges of the contemporary landscape. This research seeks to meticulously examine the intricate interplay between these innovative

approaches and their impact on reshaping business strategies in response to the complex demands of the modern era.

One notable aspect of the research is the well-structured presentation of its premises, which facilitates a coherent and logical progression of ideas throughout the study. From the outset, the context is established, effectively introducing the research topic and highlighting its significance within the broader context of today's rapidly evolving business environment. This serves as a foundational framework upon which subsequent explorations of the VUCA and BANI frameworks, IoT applications, and circular economy principles are built.

As the research unfolds, it systematically navigates through the complexities of each component, offering in-depth analyses and insights into their respective roles and contributions to business strategies. The exploration of VUCA and BANI frameworks provides a nuanced understanding of the challenges and uncertainties businesses face, while also emphasizing the need for adaptive and resilient marketing approaches in navigating turbulent times.

Furthermore, the integration of IoT technologies emerges as a pivotal aspect of the research, showcasing its transformative potential in deciphering consumer behavior patterns and enabling data-driven marketing strategies. This technological advancement not only enhances businesses' ability to understand market dynamics but also empowers them to tailor their offerings to meet evolving consumer preferences effectively.

In parallel, the study sheds light on the growing significance of circular economy principles in fostering sustainable business practices. By embracing circularity and minimizing waste through innovative product design and resource management, businesses can not only reduce their environmental footprint but also appeal to an increasingly eco-conscious consumer base. This shift towards sustainability aligns with broader societal trends and presents opportunities for businesses to differentiate themselves in the market while contributing positively to environmental conservation efforts.

Overall, the research offers a comprehensive exploration of the intricate relationship between circular economy principles, IoT technologies, and business strategies in the face of contemporary challenges. Through its rigorous analysis and insightful observations, the study provides valuable insights and actionable recommendations for

References

[1] Adzhienko, V. L., Soboleva, S. Y., Knyazev, S. A., Shestakova, I. V., & Zavolochkina, K. A. (2023). The Lateral Strategy of University Development in the BANI World. In Anti-Crisis Approach to the Provision of the Environmental Sustainability of Economy, 53-62.

[2] Bennett, N. and Lemoine, G.J. (2014) 'What VUCA Really Means for You', Harvard Business Review, 1 January. Available at: https://hbr.org/2014/01/what-vuca-really-means-for-you (Accessed: 29 July 2023).

[3] Bennis, W.G. and Nanus, B. (1985) Leaders: the strategies for taking charge. New York: Harper & Row. Available at: http://archive.org/details/leadersstrategi00benn (Accessed: 29 July 2023).

[4] Bushuyev, S., Piliuhina, K., & Chetin, E. (2023). Transformation of values of the high technology projects from a VUCA to a BANI environment model. Innovative Technologies and Scientific Solutions for Industries, (2 (24)), 191-199.

[5] Bushuyev, S., Bushuyeva, N., Murzabekova, S., & Khusainova, M. (2023). INNOVATIVE DEVELOPMENT OF EDUCATIONAL SYSTEMS IN THE BANI ENVIRONMENT. Scientific Journal of Astana IT University, 14.

[6] Cascio, J. (2022) 'BANI and Chaos – BANI', 26 April. Available at: https://ageofbani.com/2022/04/bani-and-chaos/ (Accessed: 29 July 2023).

[7] Choyon, S.A. (2021) Businesses in the VUCA world, The Financial Express. Available at: https://thefinancialexpress.com.bd/views/analysis/businesses-in-the-vuca-world-1619200860 (Accessed: 29 July 2023).

[8] De la Cruz, C. D., & Valdivieso, R. E. P. (2023). Competencies for sustainability: Insights from the encyclical letter Laudato Si.

[9] Dziadkiewicz, A., & Jakubowski, M. (2022). Managerial determinants in design management as a means of developing organisational resilience in an uncertain economic environment. Annales Universitatis Mariae Curie-Skłodowska, Sectio H Oeconomia, 56(5)

[10] Fonseca, M. P., de Lima, F. L. P., Craveiro, E. G., da Silva Pessoa, S., Colares, C. B., Fonseca, M. P., & Correia Filho, W. L. (2024). Analyzing economic conditions and operating dynamics in global markets: a comprehensive study. Caderno Pedagógico, 21(2), e2636-e2636.

[11] Garcia-Alvarez, M., Rekalde-Rodríguez, I., & Gil-Molina, P. (2023). No Transition Without Transformation: Educating Sustainability. SDGs in the European Region, 1-35.

[12] Godoy, M.F. de, & Ribas Filho, D. (2021). Facing the BANI World. International Journal of Nutrology, 14(2), 33.

[13] Grabmeier, S. (2020) 'BANI versus VUCA: a new acronym to describe the world', Stephan Grabmeier, 28 July. Available at: https://stephangrabmeier.de/bani-versus-vuca/ (Accessed: 15 August 2023).

[14] Grucza, B., & Tomszys, P. (2021). The concept of organizational resilience model and adequate measurement mechanism. Challenges in Economic Policy, Business and Management in the COVID-19 Era, 221.

[15] Gläser, W. (2023). VUCA as a practical model: A commentary. Journal of Applied Journalism & Media Studies, 12(2), 267-275.

[16] Jeyakodeeswari, R., & Nithila, R. C. J. DIGITAL TRANSFORMATION AS A KEY TO UNCERTAINTY WITH SPECIAL REFERENCE TO THE SHIFT FROM VUCA TO BANI WORLD. Convergence of Business and Technology, 318.

[17] Nandi, A., Agarwala, N., & Sahu, T. N. (2023). Towards a Sustainable Future: The Interaction of Corporate Governance and Sustainable Policies with Corporate Social Responsibility. Indian Journal of Corporate Governance, 16(2), 149-176.

[18] Nijkamp, P., Ţigănaşu, R., Bănică, A., & Pascariu, G. C. (2024). Institutional adaptability: Its relevance for enhancing resilience and smart specialization capacity of the European Union's lagging regions. Eurasian Geography and Economics, 65(1), 1-33.

[19] Obradović, V., Todorović, M., & Bushuyev, S. (2019). Sustainability and Agility in Project Management: Contradictory or Complementary? Advances in Intelligent Systems and Computing.

[20] Prasad, S.S. (2022) BEAT the VUCA or BANI world: A responsive framework, MarketExpress. Available at: http://www.marketexpress.in/2022/04/beat-the-vuca-or-bani-world-a-responsive-framework.html (Accessed: 29 July 2023).

[21] Prachi Gupta (2017). How to Survive in the VUCA World of Retail. Retrieved from https://www.linkedin.com/pulse/how-survive-vuca-world-retail-prachi-gupta/

[22] Tavanti, M. (2023). The impact challenge for global sustainability. In Developing sustainability in organizations: A values-based approach (pp. 427-437). Cham: Springer International Publishing.

[23] Wright, G. and Wigmore, I. (2023) What is VUCA (Volatility, Uncertainty, Complexity and Ambiguity)? | Definition from TechTarget, WhatIs.com. Available at: https://www.techtarget.com/whatis/definition/VUCA-volatility-uncertainty-complexity-and-ambiguity (Accessed: 29 July 2023).

[24] Zachosova, N. (2023). PATTERNS OF INDIVIDUALS'SECURITY-ORIENTED ECONOMIC BEHAVIOR UNDER THE RISK OF DISPLACEMENT IN BANI WORLD AND WAR CONDITIONS. Economics, Finance and Management Review, (3), 4-16.

Bibliography

Adzhienko, V. L., Soboleva, S. Y., Knyazev, S. A., Shestakova, I. V., & Zavolochkina, K. A. (2023). The Lateral Strategy of University Development in the BANI World. In Anti-Crisis Approach to the Provision of the Environmental Sustainability of Economy, 53-62.

Bennett, N. and Lemoine, G.J. (2014) 'What VUCA Really Means for You', Harvard Business Review, 1 January. Available at: https://hbr.org/2014/01/what-vuca-really-means-for-you (Accessed: 29 July 2023).

Bennis, W.G. and Nanus, B. (1985) Leaders: the strategies for taking charge. New York: Harper & Row. Available at: http://archive.org/details/leadersstrategi00benn (Accessed: 29 July 2023).

Bushuyev, S., Piliuhina, K., & Chetin, E. (2023). Transformation of values of the high technology projects from a VUCA to a BANI environment model. Innovative Technologies and Scientific Solutions for Industries, (2 (24)), 191-199.

Bushuyev, S., Bushuyeva, N., Murzabekova, S., & Khusainova, M. (2023). INNOVATIVE DEVELOPMENT OF EDUCATIONAL SYSTEMS IN THE BANI ENVIRONMENT. Scientific Journal of Astana IT University, 14.

Cascio, J. (2022) 'BANI and Chaos – BANI', 26 April. Available at: https://ageofbani.com/2022/04/bani-and-chaos/ (Accessed: 29 July 2023).

Choyon, S.A. (2021) Businesses in the VUCA world, The Financial Express. Available at: https://thefinancialexpress.com.bd/views/analysis/businesses-in-the-vuca-world-1619200860 (Accessed: 29 July 2023).

De la Cruz, C. D., & Valdivieso, R. E. P. (2023). Competencies for sustainability: Insights from the encyclical letter Laudato Si.

Dziadkiewicz, A., & Jakubowski, M. (2022). Managerial determinants in design management as a means of developing organisational resilience in an uncertain economic environment. Annales Universitatis Mariae Curie-Skłodowska, Sectio H Oeconomia, 56(5)

Forbes. (2023). The 7 Most Successful Business Models Of The Digital Era. Retrieved from Forbes.com: https://www.forbes.com/sites/bernardmarr/2023/03/14/the-7-most-successful-business-models-of-the-digital-era/?sh=2979988e5617

Fonseca, M. P., de Lima, F. L. P., Craveiro, E. G., da Silva Pessoa, S., Colares, C. B., Fonseca, M. P., & Correia Filho, W. L. (2024). Analyzing economic conditions and operating dynamics in global markets: a comprehensive study. Caderno Pedagógico, 21(2), e2636-e2636.

[Garcia-Alvarez, M., Rekalde-Rodríguez, I., & Gil-Molina, P. (2023). No Transition Without Transformation: Educating Sustainability. SDGs in the European Region, 1-35.

Godoy, M.F. de, & Ribas Filho, D. (2021). Facing the BANI World. International Journal of Nutrology, 14(2), 33.

Grabmeier, S. (2020) 'BANI versus VUCA: a new acronym to describe the world', Stephan Grabmeier, 28 July. Available at: https://stephangrabmeier.de/bani-versus-vuca/ (Accessed: 15 August 2023).

Grucza, B., & Tomszys, P. (2021). The concept of organizational resilience model and adequate measurement mechanism. Challenges in Economic Policy, Business and Management in the COVID-19 Era, 221.

Gläser, W. (2023). VUCA as a practical model: A commentary. Journal of Applied Journalism & Media Studies, 12(2), 267-275.

Jeyakodeeswari, R., & Nithila, R. C. J. DIGITAL TRANSFORMATION AS A KEY TO UNCERTAINTY WITH SPECIAL REFERENCE TO THE SHIFT FROM VUCA TO BANI WORLD. Convergence of Business and Technology, 318.

Nandi, A., Agarwala, N., & Sahu, T. N. (2023). Towards a Sustainable Future: The Interaction of Corporate Governance and Sustainable Policies with Corporate Social Responsibility. Indian Journal of Corporate Governance, 16(2), 149-176.

Nijkamp, P., Ţigănaşu, R., Bănică, A., & Pascariu, G. C. (2024). Institutional adaptability: Its relevance for enhancing resilience and smart specialization capacity of the European Union's lagging regions. Eurasian Geography and Economics, 65(1), 1-33.

Obradović, V., Todorović, M., & Bushuyev, S. (2019). Sustainability and Agility in Project Management: Contradictory or Complementary? Advances in Intelligent Systems and Computing.

Prasad, S.S. (2022) BEAT the VUCA or BANI world: A responsive framework, MarketExpress. Available at: http://www.marketexpress.in/2022/04/beat-the-vuca-or-bani-world-a-responsive-framework.html (Accessed: 29 July 2023).

Prachi Gupta (2017). How to Survive in the VUCA World of Retail. Retrieved from https://www.linkedin.com/pulse/how-survive-vuca-world-retail-prachi-gupta/

Tavanti, M. (2023). The impact challenge for global sustainability. In Developing sustainability in organizations: A values-based approach (pp. 427-437). Cham: Springer International Publishing.

Team, M.J.V. (2021) From a VUCA world to a BANI one: how uncertainty has changed and how your company can prepare - MJV, MJV Technology & Innovation. Available at: https://www.mjvinnovation.com/blog/from-a-vuca-world-to-a-bani-one/ (Accessed: 29 July 2023).

U.S. Army Heritage and Education Centre (2022) Who first originated the term VUCA (Volatility, Uncertainty, Complexity and Ambiguity)? - USAHEC Find Your Answer. Available at: https://usawc.libanswers.com/faq/84869 (Accessed: 29 July 2023).

VUCA - Leaders with Vision, Understanding, Clarity, Agility! (no date) VUCA-WORLD. Available at: https://www.vuca-world.org/ (Accessed: 1 December 2022).

Wright, G. and Wigmore, I. (2023) What is VUCA (Volatility, Uncertainty, Complexity and Ambiguity)? | Definition from TechTarget, WhatIs.com. Available at:

https://www.techtarget.com/whatis/definition/VUCA-volatility-uncertainty-complexity-and-ambiguity (Accessed: 29 July 2023).

Zachosova, N. (2023). PATTERNS OF INDIVIDUALS'SECURITY-ORIENTED ECONOMIC BEHAVIOR UNDER THE RISK OF DISPLACEMENT IN BANI WORLD AND WAR CONDITIONS. Economics, Finance and Management Review, (3), 4-16.

QUALITATIVE STUDY REGARDING THE PERCEPTION OF STUDENTS FROM BUCHAREST REGARDING THE ROLE AND IMPORTANCE OF TELEVISION

Elisabeta Andreea BUDACIA[1]

Lucian Constantin Gabriel BUDACIA[2]

Marian Florin BUSUIOC[3]

Abstract

Marketing research has a very important role within any organization and with regard to any of the investigated markets. The present research belongs to the category of qualitative research, being exploratory research, carried out with the aim of allowing a deeper understanding of the specific problems of the television market. The obtained results are qualitative and cannot be extrapolated to the level of the researched population. The problem addressed is a special one of current interest, because the audiovisual in general, and television in particular, determine contradictory and vehement opinions. The current trends of the television market in Romania are quite uncertain in the context of increased digitization. The young respondents appreciate that the public's thirst for information and entertainment is satisfied, rather, by modern communication channels, in particular social networks and the Internet, in general.

Keywords: marketing research, qualitative research, semi-structured interview, audiovisual.

JEL Classification: M 21, M 31

1. Literature review

Marketing research represents the activity through which, with the help of specific investigation methods and techniques, the specification, measurement, collection, analysis and interpretation of marketing information is carried out, necessary for the management of the units to know the environment in which they operate, identify opportunities, evaluate alternatives, marketing actions and their effects. Marketing research has an extremely wide scope [1].

[1] PhD, Senior Lectuer, Romanian – American University, andreea.budacia@rau.ro
[2] PhD, Junior Lectuer, Romanian – American University, lucian.constantin.budacia@rau.ro
[3] PhD, Senior Lectuer, Romanian – American University, marian.florin.busuioc@rau.ro

From the multitude of variants of marketing research, television stations turn, in particular, to quantitative research - through the people meter system. Through qualitative research, perceptions, opinions, attitudes, feelings are recorded, which lead to certain behaviors [2]. For this purpose, I considered it extremely useful to carry out a discussion through semi-structured interviews among young students from Bucharest regarding the role and importance of television within media channels.

Specialists appreciate that the role of marketing research in the decision-making process is very important, because they allow: the identification of favorable opportunities and potential dangers in the environment in which the company operates, the substantiation of the choice of decision-making alternatives, the evaluation of the effectiveness of marketing decisions [3].

Fulfilling this role, marketing research creates the necessary conditions for the practice of proactive [4] management, of continuous adaptation to changes in the economic, social and competitive environment.

2. Organizing and conducting a research through a semi-structured interview among students from Bucharest

The problem addressed is a special one and of current interest, because the audiovisual in general, and television in particular, determine contradictory and vehement opinions. The main aspects that cause controversies are those regarding the role of television and the audience, the quality of broadcasts, the assumption of a cultural - educational role, interference with the political sphere, etc., which we have proposed to address through this research.

2.1 Methodological aspects regarding the semi-structured interview

The present research falls into the category of qualitative research, being an exploratory research, carried out with the aim of allowing a deeper understanding of the specific problems of the television market. The obtained results are qualitative and cannot be extrapolated to the level of the researched population. For a better understanding of the importance and relevance of this method, we present a series of methodological aspects.

The semi-structured / semi-directed interview is based on the principles of non-directivity, the fact that the investigator's attitude is neutral, and the interview is subject to a content analysis; basically this is a variant of the undirected in-depth interview. The difference between the in-depth interview and the semi-directed interview is that the latter is carried out by means of an interview guide (conversation guide), designed in advance. This guide covers the main topics that should be addressed during the discussion (the interviewer does not have a set of predetermined questions, the interview is not conducted on the basis of a standardized questionnaire or an open-ended questionnaire). No particular order is required; the themes will be discussed depending on the degree of proximity, in relation to the themes

addressed by the subject. The interview has a total duration of 30-60 minutes. The technique involves the use of small samples, whose competence must reflect the structure of the reference community as well as possible [5].

The organization of the interview is very important, because the quality of the information obtained largely depends on it. As far as we are concerned, during the research we staged our activity, taking into account the specific requirements, requirements that take into account: defining the problem and establishing the objectives, creating the sample, recruiting the participants, ensuring the material conditions necessary to conduct the interview, certain procedural manners allowed and the analysis of information is prohibited [6].

- Establishing the objectives has as its starting point the investigated phenomenon.
- The sample, which is small in size, is not representative from a statistical point of view, but it requires a rigorous selection of the people participating in the study.
- The recruitment of participants is usually carried out by using a recruitment questionnaire, which must be designed in such a way that the persons contacted in order to participate in the interview do not know precisely the topic of the study.
- Ensuring the necessary material conditions for conducting the interview refers in particular to the fact that it is necessary for the interviewer to have a notebook available to write down certain words or topics addressed by the subject, with the aim of bringing them back into the discussion later for deepening them.
- The conception of how to conduct the interview starts from the premise that the interviewer has an active role during its conduct and must respect the principles of non-directivity; the interviewee having to be helped to express himself so that the investigation carried out is as deep as possible; achieving the proposed objectives largely depends on the investigator's attitude. In this sense, we emphasize that procedural manners, such as: interrupting the interviewed person, are prohibited; making judgments about the answers; approval or disapproval; interpretation of answers; eliciting responses; interventions in a directive manner, in the sense of orienting the discussion, imposing a certain structure, which is not related to the discourse of the investigated subject. We also emphasize that it is allowed: encouraging the subject to deepen certain aspects; intervention in the discussion with the aim of facilitating the discourse and deepening the topics; relaunching the discussion by repeating a word or a phrase; to show the investigated person that he is being listened to carefully, using expressions such as: "yes, I understand!", "aha!"; to show the subject that he is understood without interrupting him or making interpretations of what he said; to make a synthesis of what the subject said, taking over his words and respecting his logic.
- The analysis of the information refers to the aspects regarding the content of the interview, based on a grid developed starting from the problem of the study and taking into account the objectives of the analysis. The text was "cut" into units of analysis, depending on the main themes addressed, the units thus identified being grouped into homogeneous,

exhaustive and exclusive categories; then it was evaluated according to certain rules of the frequency of occurrence of each unit.

2.2 The stages followed in carrying out the research through a semi-structured interview

The stages followed in carrying out the research are: establishing the objectives and hypotheses, designing the research, including here the establishment of the sample, the place and period of data collection, methodological aspects regarding the actual interview, the elaboration of the interview guide, the presentation of the research results.
Regarding the objectives, we record the fact that considering the purpose of the research stated previously, we proposed a series of objectives:
● identifying young people's perception of the main problems in the Romanian television market;
● knowing the opinions regarding the current trends specific to the television market in Romania;
● outlining the role and importance of television in the current context of Romanian society and the television market;
● identifying the perception regarding the role and importance of television within media channels;
● identifying opinions regarding the typology of broadcasts;
● values promoted by television versus values that should be promoted;

I established the hypotheses of the research starting from certain public information and, in this context, the hypotheses from which I proceeded in carrying out the research were:
● the main problems at the level of the television market concern the questionable quality of broadcasts, the information provided, which is often distorted, the promotion of common characters, without special merits, of questionable quality, etc.;
● the place, role and importance of television are very obvious through the lens of the values that this institution must promote in the current stage in which Romanian society is, but the general public is not aware of these elements;
● young people's perception of the importance of television within media channels is negative;

The design of the research is a very important stage, because it is more extensive and includes a series of sub-stages such as: establishing the sample, the place and period of conducting the research, methodological aspects regarding the actual interview. That's why we paid special attention to the punctual solution of each problem.

● Regarding the sample, we specify the fact that our intention was to investigate a number that varies between 30 and 50 people who meet the established criteria, but in the end we managed to discuss with 41 subjects.
● Regarding the place and period of data collection, we note that the interviews took place between March 4 and 26, 2024, in Bucharest.
● The methodological aspects of the actual interview received special attention from us, both in terms of the preparation and conduct of the interview, as well as in terms of its phasing.
In preparing and conducting the interview, I took into account the following rules:

- careful preparation of the interview guide;
- the subjects were asked additional questions if it was necessary to deepen the ideas;
- ensuring a pleasant and comfortable interview framework;
- when it was the case, specific questions were also asked.

The interview was structured in four phases, with a total duration that varied from 30 to 60 minutes, as shown in the table below.

Research phase	**Approximate duration**	**Key elements of the phase**
The introduction phase	5-10 minutes	It had as its objective the creation of a state of confidence of the subject in the topic addressed, but also the gathering of information about the position of the subject and his activity;
The phase of focusing on the researched problem	10-20 minutes	had as its objective the evaluation of the television market in Romania, emphasizing the main problems that can be identified at its level and theextent of opinions regarding the current perceptions and trends specific to the television market in Romania;
Deepening phase	15 – 20 minutes	with the main objective establishing the role of television in the current context, outlining opinions regarding the typology of broadcasts; identifying the values promoted by television versus values that should be promoted by it; outlining some ideas about the place and importance of television within the set of communication channels; canalelor de comunicare;
Conclusion phase	10 minutes	with the objective of drawing some conclusions regarding the perceptions of young people in relation to

		television, as well as other comments that the respondents would like to make about the discussed topic and which were not previously captured.

Table no: 1 Structure of the interview

● Regarding the rules that we had in mind as operators, we should mention: observing a waiting interval of 10-15 seconds when switching from one theme to another; leaving it up to the subject to approach the provided topics; avoiding interrupting the subject; the resumption of unaddressed or partially addressed topics; avoiding comments on the answers and expressing surprise at the subject's remarks; relaunching some discussions in case the subject finished one idea and did not restart another, when previous aspects remained unclear or to reconfirm some important clarifications made by the subject.

The interview guide that was developed by us in order to carry out the respective research included the main questions that were addressed to the interviewed subjects, following the phases considered and the corresponding theme, previously established.

Research phase	Main questions formulated and addressed
The introduction phase	~ I would like you to talk to me about your concerns. ~ Helping questions: ♦ Have you previously participated in other studies/market research? ♦ How do you spend your free time? ♦ Do you find out about aspects of everyday life?
The phase of focusing on the researched problem	~ How do you rate Romanian television? ~ Helping questions: ♦ What are the main problems in the Romanian television market? ♦ What is your perception of television in Romania, in general?
Deepening phase	~ In your opinion, what is the place of television within the media channels in Romania? ~ Helping questions: ♦ What is your perception of media channels in general? ♦ What are the values that television should promote? ~ What should be the models promoted by the televisions? ~ Helping questions: ♦ Who are the people you followed/follow on TV channels? ♦ How would you sketch the portrait of a TV star? ~ What are your favorite shows? What are your reasons for watching them?

	~ What would ideal television look like in Romania?
Conclusion phase	~ What advice would you have for televisions in general? ~ In conclusion, how do you briefly characterize television? ~ Do you have any other comments to make?

Table no: 2 The main questions formulated and addressed during the interview

Research results

Regarding certain aspects regarding the activity carried out by the interviewed persons, we must specify the fact that we are talking about students (35 persons) and master's students (6 persons) of the Romanian-American University.

Regarding the opinions resulting from the questions formulated and the topics addressed in the introduction phase, they are presented according to the table below.

Opinions/Perceptions of respondents collected in the introduction phase	**Frequency of occurrence**
"your concerns": - Going out with friends - Watching movies with friends (Netflix, especially) - The music - The read - The faculty - Finding (changing) a job	 41 32 27 16 28 18
"previous participation in other studies/market research" - Yes (especially those regarding the evaluation of university teachers) - Not	 31 10
"spending free time" - the sport - social networks - outings with friends (at the Mall, especially) - on TV	 27 36 38 18

"information on aspects of everyday life"	
- Yes, from the net	11
- Yes, (also) from TV	8
- Not really	26
- Not	5

Table no: 3 Respondents' perceptions collected in the research introduction phase

Regarding the opinions resulting from the questions formulated and the topics addressed in the phase of focusing on the researched problem, they are presented according to the table below.

Opinions/Perceptions of the respondents collected in the phase of focusing on the researched problem	**Frequency of occurrence**
"appreciations/perceptions regarding Romanian television"	
- "too many TV channels"	28
- "uninteresting"	34
- "few shows for young people"	31
- "presents only negative aspects: scandals, crimes, etc."	29
"problems at the level of the Romanian television market"	
- "too much politics on TV"	26
- "posts copy each other"	33
- "lots of weird people on TV"	38
- "bad series"	35
- "unprepared people/moderators"	32

Table no: 4 Perceptions of the respondents collected in the phase of focusing on the researched problem

Regarding the opinions resulting from the questions formulated and the topics addressed in the in-depth phase, they are presented according to the table below.

Opinions/Perceptions of the respondents collected in the deepening phase	**Frequency of occurrence**

"the place of television within the media channels in Romania"	
- relatively important	21
- TV makes the biggest audience	26
- "will be smaller and smaller"	37
"the people you followed/are following on TV channels"	
- Selly	
- Smily + Pavel Bartos	8
	10
"favorite TV shows"	
- Do I know you	20
- Romanians have talent	29
- Survivor	32
- Movies on Pro TV	19
"reasons that make you watch them"	
It's fun	24
It is interesting	21
She is funny	27
"Ideal TV":	
It has nice/interesting/funny shows	27
Streams new movies	22
Debate topics for young people	16
It has interactive shows	33

Table no: 5 Perceptions of the respondents collected in the deepening phase of the research

Regarding the opinions resulting from the questions formulated and the topics addressed in the conclusion phase, they are presented according to the table below.

Opinions/Perceptions of the respondents collected in the conclusion phase on the researched problem	**Frequency of occurrence**
"television tips"	
- To be more dynamic	12
- To attract young people through interesting topics	7

"characterization of Romanian television in one word"	
- Old	26
- Uninteresting	13
- Fake	8

Table no: 6 Respondents' perceptions collected in the conclusion phase on the researched problem

3. Conclusions

Analyzing the problems identified at the level of the Romanian television market, following the discussions, a series of more important aspects were outlined.

The unanimous opinion is that at this moment, the television market is suffocating for the public because of the too large number of televisions, but even so, in the conditions where there are many investors, the big TV channels are concentrated in the sphere of a few "moguls" of the media. The highest frequency was the problem regarding the quality of the broadcasts, which is obviously in a direct dependence relationship with the personnel policy of the TV stations and with their possibility and availability to invest in serious projects of interest to the young audience.

The current trends of the television market in Romania are quite uncertain, even gloomy in the context of increased digitization. The young respondents appreciate that the public's thirst for information and entertainment is satisfied, rather, by modern communication channels, in particular social networks and the Internet, in general.

The place, role and importance of television have changed in recent years, including the fact that television is the cheapest means of entertainment for the poor and middle class. Today, the Internet is the easiest means of diversion.

The audience of TV stations is an important issue in the context in which this means of communication is no longer so relevant for the young population. Regarding the possibilities of improving the television audience, the respondents' opinions focused on the improvement of the shows, their attractiveness for the young audience, the elimination of fake stars and non-values and on the observance of journalistic principles aimed at equidistance, editorial independence, the elimination of political interference, etc.

Also, the idea emerged that, although an ideal public television is impossible to achieve, even more so in the context of diversity and digitization, there are still some elements that could direct television towards this desired, among which we specify: the broadcasting of interesting shows , interactive, new formats, the presentation of authentic national, European and international models, etc.

The deepening, but especially the detailing and understanding of the investigated phenomenon, can be achieved through further research, especially through those of a quantitative nature in order to be able to have the representativeness of the ideas. Also, in

the future, research can be carried out among specialists to capture other nuances of the researched phenomenon.

References

[1] I. Smedescu (coordonator), *Marketing*, Ed. Universitară, Bucureşti, 2004, p. 563.

[2] M. Papuc, *Cercetări de marketing*, Ed. Universitară, Bucureşti, 2007, ediţia a doua revăzută, p.29.

[3] C. Bălan, *Cercetări de marketing*, *Conţinutul, rolul, tipologia şi procesul cercetării ce marketing intern şi internaţional*, Ed. ASE, Bucureşti, 2001, p. 24.

[4] C. McDaniel, Jr., R. Gates, *Contemporary Marketing Research*, second edition, West Publishing Company, St. Paul, MN, 1993, p. 12 – 14.

[5] I. Cătoiu (coordonator), *Cercetări de marketing*, Ed. Uranus, Bucureşti, 2002, p. 207

[6] V. Balaure (coordonator), Marketing ediţia a - II – a revăzută şi adăugită, Ed. Uranus, Bucuresti, 2002, p. 260 – 263

Bibliography

[1] C. Bălan, Cercetări de marketing, Conţinutul, rolul, tipologia şi procesul cercetării ce marketing intern şi internaţional, Ed. ASE, Bucureşti, 2001

[2] Cătoiu I.(coord.), Cercetări de marketing, Ed. Uranus, Bucureşti, 2009.

[3] Datculescu P., Cercetarea practică de marketing, Bucuresti, Brandbuilders, 2013

[4] Malhotra, N.; Nunan, D.; Birks, D. – Marketing Research. An Applied Approach (5th ed.), Pearson, 2017

[5] C. McDaniel, Jr., R. Gates, Contemporary Marketing Research, second edition, West Publishing Company, St. Paul, MN, 1993

[6] McDaniel, C.; Gates, R. – Marketing Research (11th ed.), Wiley, 2018

[7] M. Papuc, Cercetări de marketing, Ed. Universitară, Bucureşti, 2007

[8] Smedescu (coordonator), Marketing, Ed. Universitară, Bucureşti, 2004

ESTABLISHING COLLABORATIVE RESEARCH NETWORKS: A LITERATURE REVIEW

Oana Daniela BUGAN[1]

Sorin IONESCU[2]

Abstract

This study delves into the complex dynamic within collaborative research networks (CRNs), with the aim of examining theoretical perspectives and common issues within the existing literature. CRNs have emerged as crucial mechanisms for addressing complex scientific challenges, fostering interdisciplinary collaboration, and enhancing knowledge dissemination. Through a comprehensive analysis, this review provides valuable recommendations for policymakers, academic institutions, and researchers seeking to establish and nurture successful collaborative research networks. Through synthesizing insights from various scholarly articles, we explore the components, mechanisms, practices, challenges, and opportunities encountered within CRNs bolstered by digital technologies and innovative practices, and their impact on sculpting the future landscape of manufacturing, technology transfer, and enhancing organizational performance. The article concludes with recommendations for future research directions, emphasizing the need for further exploration of CRNs' long-term impacts on the scientific community and society at large, addressing a notable gap for robust frameworks and practical guidelines essential for the effective development and implementation of CRNs.

Keywords: Literature review, Collaborative research, Networks, Innovation, Framework

JEL Classification: O32 Management of Technological Innovation and R&D

1. Introduction

In the current era, marked by rapid technological advancements and a pressing need for innovative solutions to complex global challenges, CRNs have emerged as vital platforms for facilitating interdisciplinary and cross-sectoral cooperation. However, the absence of a standardized or universally adopted framework for managing these collaborations often leads to inefficiencies, misunderstandings, and missed opportunities.

[1] PhD Student, Corresponding author, National University of Science and Technology Politehnica Bucharest, Romania, oana.bugan@upb.ro

[2] Prof. Dr. Ing, Co-author, National University of Science and Technology Politehnica Bucharest, Bucharest, Romania, sc.ionescu@gmail.com

In recent years, research has begun to delve deeply into how networks behave, evolve, and influence various phenomena, from biological systems to social networks and urban infrastructures. Theories presuppose an interaction between individuals and their networks, focusing on their relationships and structures, with many different theoretical concepts defining networks [1].

Numerous types of networks are identified: internet, industrial, social, scientific, organizational, transportation, electrical energy, distribution, telecommunication, trophic, neural, aerial, linguistic, etc., they are represented by graphs. The advent of the Internet has led to an exponential increase in the ways people can connect and has brought about essential changes in the way information is circulated and relationships are formed.

Most businesses that cooperate with higher education institutions that are also research and development (R&D) institutions also collaborate with other businesses or have their own R&D capabilities [2]. Collaboration in research plays a fundamental role in encouraging valorization activities and increasing the innovation capacity of organizations. Through partnerships between universities, research institutes, companies, and other entities, the exchange of knowledge and technologies is facilitated, accelerating the transfer of innovations from the research stage to the market.

As noted in an OECD study [3], collaborative scientific research platforms face a series of issues: the need for a robust data infrastructure, global material data management, the existence of scalable data repositories, and the implementation of efficient data cleaning strategies; lack of coordination, redundancy, or dispersion of equipment and technical expertise; the need for interdisciplinary research, development, and training; the absence of ecosystems that facilitate the building of new supply chains. Therefore, valorization activity, defined as the process of transforming research results into tangible benefits, is closely linked to innovation capacity, and both are enhanced through strategic and efficient collaborations in research.

The industry is increasingly moving towards an open innovation process, which includes collaboration with other businesses and higher education institutions, recognizing the value brought by this collaboration. Collaboration in the field of research plays a fundamental role in encouraging valorization activity and in increasing the innovation capacity of organizations. Through partnerships between universities, research institutes, companies, and other entities, the exchange of knowledge and technologies is facilitated, accelerating the transfer of innovations from the research stage to the market.

Sharing research resources through platforms optimizes the use of expensive equipment and specialists' time, encourages intercultural collaboration in research, promotes the adoption of open and universal formats for data storage and transmission, and facilitates access to cutting-edge technologies and equipment, involving all participants in collaborative studies [4].

The aim of this study is to enhance the methodological framework of collaborative research network systems and to outline directions and strategies for boosting the synergistic development of industrial innovation and collaborative ecosystems. We set to explore the components, mechanisms, and practices within CRNs, aiming to identify and analyze CRNs and their characteristics, examine the challenges and opportunities encountered by scientific literature tackling collaborative research networks and evaluate the benefits and outcomes of CRNs.

By addressing our objectives, we intend to add to the body of work by proposing guidelines and recommendations for the development and implementation of CRNs for further research. The research questions below have emerged:

1. What are the main barriers and opportunities to effective collaboration in the context of CRNs?

2. What models or frameworks for CRNs have demonstrated success in overcoming these barriers, and what can be learned from them?

2. Research Methodology

2.1. Research design

The research method deployed of content analysis was useful for detecting theoretical perspectives and common issues within the existing literature regarding collaborative research networks. This research used a semi-systematic literature review methodology, as it is best suited to our research objectives. The method used to analyze and synthesize the findings from the review was content analysis as it can be broadly used for identifying, analyzing, and reporting patterns related to our research questions.

2.2. Data collection

To ensure the validity of the study, the Web of Science database was selected as it is one of the largest databases of the relevant scientific and research literature. Because of its diverse and inclusive nature, Web of Science can ensure that a large number and a broadened perspective on collaborative research networks is captured, and that state-of-the-art findings and emerging topics are reviewed.

Inclusion criteria for the review was guided by the selected research questions. The search within the database was performed by a Boolean expression applied in terms of the title, abstract and keywords of papers: TITLE-ABS-KEY (collaborative AND research AND networks) AND (LIMIT-TO (SUBJAREA, "ENGI")) AND (LIMIT-TO (LANGUAGE , "English")) for the time span of 2020-2024.

The initial search on Web of Science, performed on the 15th of March 2024, was thus limited to Engineering Multidisciplinary, Engineering Industrial, Engineering Manufacturing and English, for the period 2020-2024, and it produced 881 results. As the initial literature searches yielded many articles, a strategy was needed to identify which are relevant. In terms of research quality, we decided on exclusion criteria based on (engineering industrial) – this produced 330 results.

We decided to further refine our search on exclusion criteria based on (industry innovation and infrastructure) which produced 75 results, and finally (open source) which produced 19 results, presented in the table below.

Article Title & Authors	Publication Year
"Micro dynamics and macro stability in inventor networks" Fritsch M. et al. [5]	2022
"A complexity assessment framework with structure entropy for a cloud-edge collaborative manufacturing system" Li, JJ. et al. [6]	2023
"Exploring self-organization and self-adaption for smart manufacturing complex networks" Guo, ZG. et al. [7]	2023
"Commercialization networks in emerging technologies: the case of UK nanotechnology small and midsize enterprises" Salehi, F. et al. [8]	2022
"Collaborative innovation in emerging innovation systems: Evidence from Central and Eastern Europe" Stojcic, N. et al. [9]	2021
"Collaborative modes with Cultural and Creative Industries and innovation performance: The moderating role of heterogeneous sources of knowledge and absorptive capacity" Santoro, G. et al. [10]	2020
"Innovation and innovator assessment in R&I ecosystems: the case of the EU Framework Programme" Nepelski, D. et al. [11]	2021
"Design Decisions and Interactions: A Sociotechnical Network Perspective" Pirzadeh, P. et al. [12]	2021
"Collaboration in BIM-based construction networks: a qualitative model of influential factors" Oraee, M. et al. [13]	2022
"Collaborations for Digital Transformation: Case Studies of Industry 4.0 in Brazil" Rocha, C. et al. [14]	2023
"Promoting academic engagement: university context and individual characteristics" Zhao, ZY. et al. [15]	2020
"Blockchain-based Shared Additive Manufacturing" Lupi, F. et al. [16]	2023
"Exploring dyadic relationships between Science Parks and universities: bridging theory and practice" Löfsten, H. et al. [17]	2024
"The effectiveness of interactive virtual reality for furniture, fixture and equipment design communication: an empirical study" Prabhakaran, A. et al. [18]	2021
"FLEAM: A Federated Learning Empowered Architecture to Mitigate DDoS in Industrial IoT" Li, JH. et al. [19]	2020
"Open data for open science in Industry 4.0: In-situ monitoring of quality in additive manufacturing" Gronle, M. et al. [20]	2023
"Developing distributed manufacturing strategies from the perspective of a product-process matrix" Kumar, M. et al. [21]	2020

"Impact of the changing business environment on performance measurement and management practices" Nudurupati, SS. et al. [22]	2021
"Buyer-supplier collaboration during emerging technology development" Moradlou, H. et al. [23]	2022

Table 1. Web of Science research articles results based on our inclusion and exclusion criteria.

3. Research Results

The researched articles provide in-depth information about collaborative research on a diverse range of contexts, including inventor networks, cloud-edge collaborative manufacturing systems, smart manufacturing, commercialization networks, collaborative innovation, and digital transformation in various sectors. The highest occurrence keywords resulting in our selected articles are below.

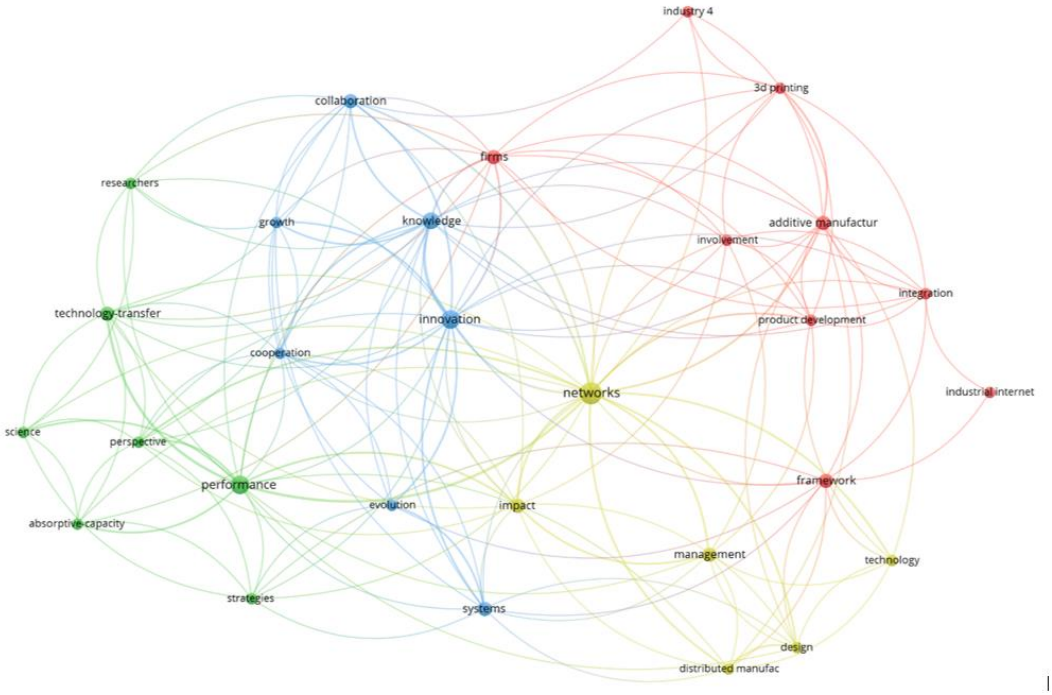


Fig. 1. Co-occurring key words in the 19 research articles.[3]

Based on the co-occurrence of keywords, the themes we researched for the purpose of the study are:

3.1. Emerging Technologies / Innovation:

Lupi et al. and Li et al. investigates blockchain-based shared additive manufacturing, presenting a new network paradigm that ensures transparency, security, and efficiency in

[3] Realized using VOSviewer, www.vosviewer.com

collaborative manufacturing and federated learning to mitigate DDoS in Industrial IoT. These articles point towards the role of emerging technologies in securing and optimizing manufacturing and IoT systems [6], [16].

Rocha et al. and Lupi et al. discuss partnering for digital transformation, including case studies from Brazil occurring in the context of Industry 4.0 and blockchain-based shared additive manufacturing. These contributions reflect on the transformative potential of digital technologies and collaborative approaches in industry practices [14], [16].

Salehi et al. and Moradlou et al. explore commercialization networks and buyer-supplier collaboration in the context of emerging technologies. They emphasize the significance of networks in facilitating the commercialization of new technologies, with a specific focus on nanotechnology SMEs in the UK context [8], [23].

Prabhakaran et al., Kumar et al., and Nudurupati et al. examine the effectiveness of interactive virtual reality in developing distributed manufacturing strategies [18], [21], [22].

3.2. Impact of Collaboration Networks on Performance and Innovation:

Zhao et al. discuss strategies to promote academic engagement, drawing attention to the influence of the university context and individual characteristics within academic networks on engagement levels [15].

Kumar et al.'s analysis on distributed manufacturing strategies from the perspective of a product-process matrix emphasizes the strategic importance of network configurations for achieving manufacturing efficiency and flexibility [21], while Moradlou et al. study buyer-supplier collaboration in emerging technology development, indicating how such networks are critical for managing risks and fostering innovation [23].

Similarly, Guo et al.'s exploration of smart manufacturing networks emphasizes the critical self-organization and self-adaptation mechanisms that allow for enhanced efficiency and dynamic adaptability [7].

Prabhakaran et al., Li et al., and Nudurupati et al. highlight how collaborative technologies and practices can significantly affect organizational performance and innovation capability [18], [19], [22].

Stojcic and Santoro et al. provide evidence of collaborative innovation in new innovation ecosystems in Central and Eastern Europe and the impact of collaborative models with Cultural and Creative Industries on innovation outcomes pointing out how these networks are vital for innovation in emerging systems by bridging gaps between diverse stakeholders. These studies underscore the value of heterogeneous knowledge sources and absorptive capacity in enhancing innovation performance [9], [10].

Fritsch and Kudic delve into the micro dynamics and macro stability in inventor networks, providing insights into how individual relationships contribute to broader innovation

ecosystems' stability and productivity [5]. The research highlights inventor networks as examples of complex adaptive systems, despite considerable micro-level changes. An interesting insight from the study suggests that the structural stability of inventor networks can coexist with high levels of individual and tie fluidity due to mechanisms that operate at an intermediary level, such as the transfer of key player roles within the network.

3.3. Framework Programmes:

Oraee et al. focuses on collaboration within construction networks utilizing Building Information Modeling (BIM) technology, identifying influential factors, and proposing a qualitative model for effective collaboration [13]. The article offers a model highlighting how BIM technology facilitates or hinders collaboration in construction projects, addressing technical, organizational, and interpersonal aspects of collaboration, and emphasizes the Innovation Radar's role in fostering innovation within the EU's research and innovation ecosystem.

Pirzadeh et al. offer a sociotechnical network perspective, highlighting the crucial interplay between social and technical elements in shaping design decisions and interactions [12]. The document also emphasizes the development of the Innovation Radar (IR), a platform that offers insights into the innovation processes within extensive collaborative research and innovation initiatives. It aids in monitoring these projects more effectively and offering tailored support to help in the commercialization of results. Moreover, it helps external actors find collaborative partners or investment opportunities by utilizing the public IR data platform.

Li et al. proposes a complexity assessment framework for cloud-edge collaborative manufacturing systems, addressing the challenges of managing complexity and ensuring efficient operation within these collaborative environments and also highlights the importance of network structure in optimizing collaboration and efficiency [6].

Moradlou et al. propose a framework that emphasizes the importance of early-stage incubation in research universities, followed by collaborative efforts either through traditional buyer-supplier relationships or within catapult centres, to successfully develop and adopt emerging technologies [23].

Santoro et al. extend this discussion into the realms of Cultural and Creative Industries (CCIs), pointing out the significant impact of collaborative modes on innovation performance, particularly when networks are enriched with diverse knowledge sources and an absorptive capacity [10].

This is complemented by Nepelski and Van Roy's focus on the importance of networks in facilitating collaboration and knowledge exchange among innovators within Research and Innovation (R&I) ecosystems, highlighting the evolution of the EU's Framework Programme into a comprehensive support system for innovation, emphasizing the role of

the Innovation Radar in identifying, managing, and supporting innovations and innovators within this ecosystem [11].

Löfsten & Klofsten and Gronle et al. provide a comprehensive analysis of the symbiotic relationships between science parks that offer infrastructural, technological, and organizational support, while universities provide academic expertise and access to research and development (R&D) capabilities. The studies highlight the importance of strategic management, alignment of goals, and effective communication in maximizing the benefits of these collaborations [17], [20].

4. Discussion and managerial implications

The research findings provide valuable insights on the importance of collaborative networks, digital technologies, and innovative practices in shaping the future of manufacturing, technology transfer, and organizational performance in driving efficiency, innovation, and transformation in contemporary industries that need to be taken into account at managerial level when establishing CRNs.

The articles demonstrate the varied nature of CRNs, from tightly knit groups focusing on specific challenges [5] to expansive networks seeking to leverage broad expertise across fields [3]. The technological backbone, often highlighted in the discussions on smart manufacturing [7] and digital transformation, plays a crucial role in enabling these networks by facilitating communication, data sharing [20], and collaborative problem-solving [17]. Furthermore, the adoption of emerging technologies such as blockchain for shared manufacturing [16] and federated learning for IoT security [19], points to new avenues for addressing prevalent challenges, improving data interoperability, and safeguarding against threats.

Despite the potential of CRNs, several challenges persist. Communication and coordination issues are prevalent, particularly in geographically dispersed networks. Intellectual property rights and trust emerge as significant concerns, especially in networks involving multiple organizations with varying interests. Moreover, securing funding and aligning the diverse cultural and organizational practices of participants pose additional hurdles.

Yet, CRNs present immense opportunities. They allow access to a wide array of expertise and resources, driving innovation through diverse perspectives and knowledge. The potential for increased research impact and the development of new research avenues also stand out, as collaborations can lead to groundbreaking discoveries and the exploration of uncharted scientific territories.

According to the articles researched, there are various barriers and opportunities to effective collaboration in the context of CRNs that we present in the figure below.

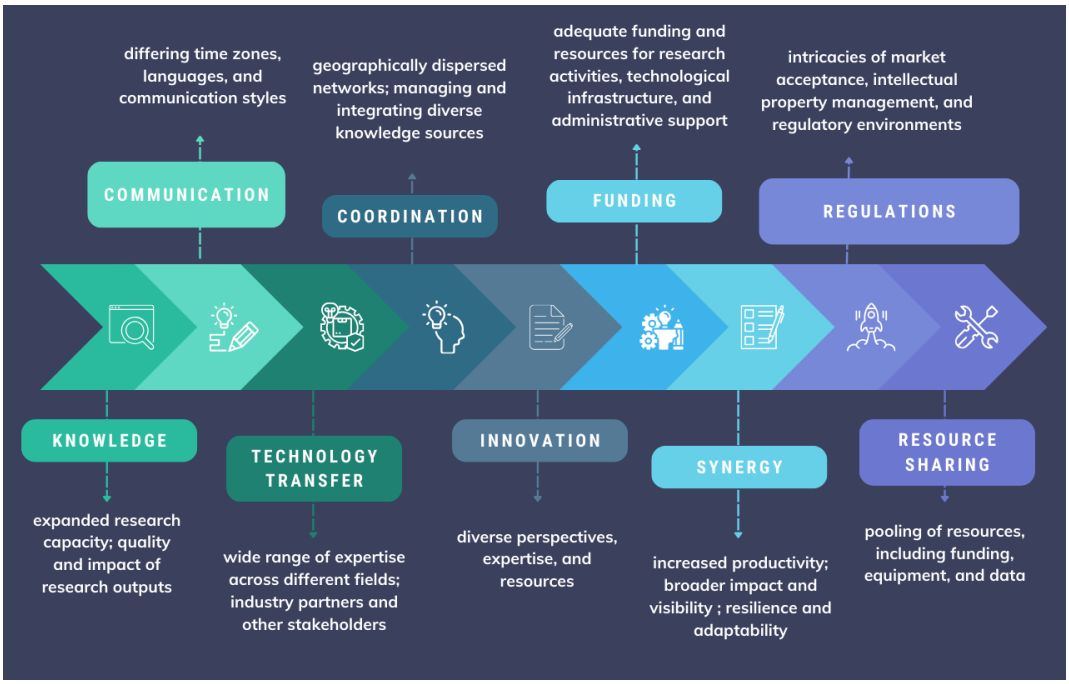


Fig. 2. Barriers and opportunities for CRNs[4]

There are various models or frameworks for CRNs in our research that have demonstrated success, like Cloud-Edge Collaborative Systems, which is a model that successfully addresses efficiency and security concerns in CRNs [6]. By leveraging the strengths of both cloud and edge computing, this model offers a solution to the challenge of processing large volumes of data securely and efficiently, demonstrating the importance of strategic technology integration. Federated Learning for Security Enhancements represents a novel approach to enhancing security in CRNs, especially in the context of Industrial IoT [19]. This model demonstrates the effectiveness of federated learning in addressing cybersecurity threats, a significant barrier for collaborative networks.

Self-Organization and Adaptation in Smart Manufacturing is a framework that allows for flexibility and adaptability, critical for the survival and competitiveness of CRNs in rapidly changing environments [7]. This approach underscores the value of promoting self-organization and adaptive leadership within CRNs.

Blockchain for IP Protection and Resource Coordination is used to overcome barriers related to intellectual property protection and resource coordination [16]. This suggests that secure and transparent technologies are crucial for fostering trust and collaboration in CRNs.

Collaborative Innovation in Emerging Systems show that structured collaboration modes, along with the integration of heterogeneous sources of knowledge and absorptive capacity,

[4] Realized using Canva, www.canva.com

are effective in enhancing innovation performance [8], [9]. This highlights the importance of fostering dynamic interactions and leveraging diverse contributions within CRNs.

5. Conclusions

Our research, besides the synthesis of the state of knowledge regarding collaborative research networks, is proposing recommendations for the development and implementation of CRNs. This contribution not only highlights the foundational importance of collaborative networks in driving industrial advancement but also underscores the potential of CRNs in fostering a more innovative, efficient, and transformative industrial ecosystem.

The synergistic collaboration among various actors in the innovation ecosystem stimulates the development of new products, services, and processes, thus contributing to the enhancement of economic competitiveness [22]. Furthermore, interdisciplinary and intersectoral collaboration paves the way for addressing complex challenges, facilitating the creation of sustainable solutions, and exploiting innovative potential. Therefore, valorization activity, defined as the process of transforming research outcomes into tangible benefits, is closely linked to innovation capacity, and both are enhanced through strategic and effective collaborations in the field of research [8], [9].

A strategy integrating diverse contributions, cutting-edge technologies, equitable benefits, self-organization, flexibility, and stringent security protocols is key to enhancing Collaborative Research Networks (CRNs). Essential to this strategy is the adept incorporation of technologies such as cloud and edge computing, and blockchain, which are instrumental in elevating efficiency, security, and transparency [6], [16].

Promoting self-organization and flexibility [7] allows CRNs to swiftly adapt to new challenges and opportunities, maintaining their relevance and impact. This adaptability is supported by adaptive leadership and governance structures, which enable the network to evolve operationally and managerially.

Key to the success of CRNs is the establishment of clear governance structures, defining roles and decision-making processes to unify participants towards shared objectives. Cultivating a culture of trust and mutual respect is vital for encouraging open sharing and collective problem-solving [12]. Strategic financial and resource management, stakeholder engagement from industry to government, and aligning research with societal needs are all pivotal in amplifying the network's relevance and impact. Mechanisms for ongoing monitoring and improvement ensure the network's adaptability and long-term success [20], while navigating legal and ethical issues upholds its integrity and trustworthiness. These principles are fundamental to developing robust, effective, and impactful CRNs, setting a solid foundation for collective research endeavors.

Expanding on the intention to contribute to the body of work with guidelines and recommendations for the development and implementation of Collaborative Research

Networks (CRNs), the proposed approach involves several key strategies aimed at maximizing the effectiveness and impact of these networks. The recommendations are derived from synthesized research findings, aiming to address common challenges and leverage best practices identified in the literature.

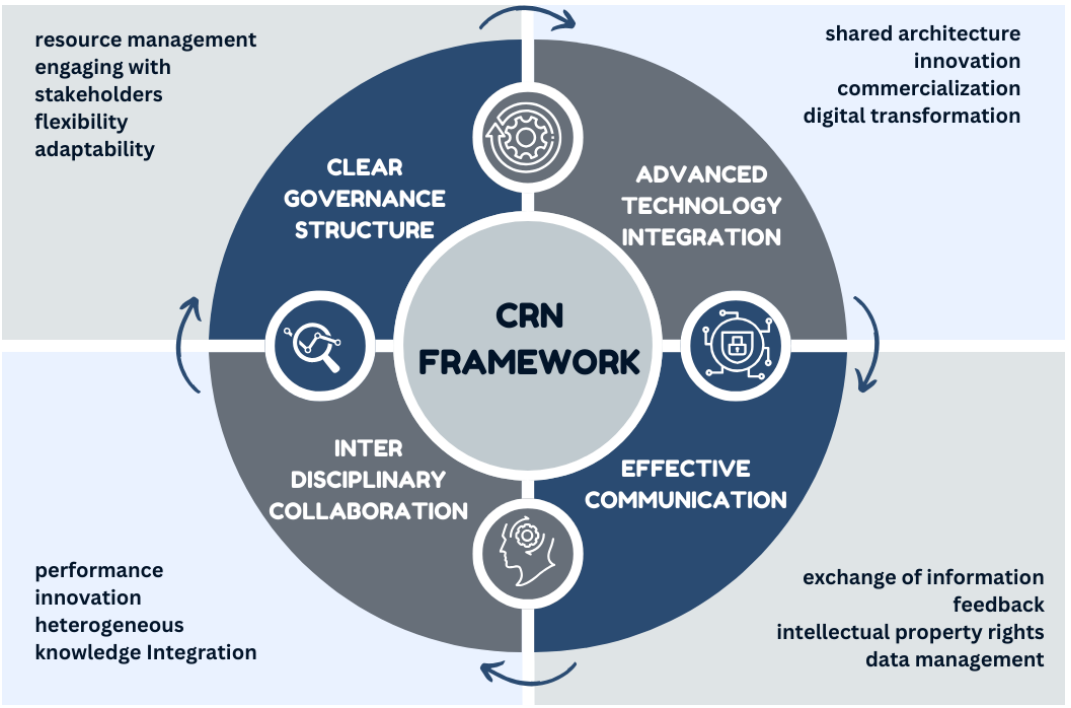


Fig. 3. A framework for CRNs[5]

Considering the scope of this research, a few limitations are identified. Firstly, we must acknowledge limitations in our data collection methods due to the extensive amount of information in the existing literature, and also mention a key limitation related to the highly specific field of study – industrial engineering, which reduces the relevance to broader research or practice. Other manufacturing industries and companies should be explored to obtain generalized results.

Additionally, future research on Collaborative Research Networks (CRN) could delve into several areas to further elaborate on their dynamics, impact, and optimization, focusing on the need for solid frameworks that support the development and functioning of CRNs, particularly in cross-border collaboration.

To corroborate the findings and conclusions of this research, subsequent studies may consider adopting methodologies like surveys and the validation of frameworks.

[5] Realized using Canva, www.canva.com

References

[1] M. Gamper, "Social Network Theories: An Overview," in Social Networks and Health Inequalities, Springer International Publishing, 2022, pp. 35–48. doi: 10.1007/978-3-030-97722-1_3.

[2] T. Davey et al., "The State of University-Business Cooperation in Europe," Luxembourg, 2018. doi: 10.2766/676478.

[3] D. E. Winickoff, L. Kreiling, M. Borowiecki, H. Garden, and J. Philp, "Collaborative Platforms for Emerging Technology: Creating Convergence Spaces," 2021. [Online]. Available: http://www.oecd.org/termsandconditions.

[4] T. Schultz et al., "Linking Labs: Interconnecting Experimental Environments," pp. 1–13, Feb. 2021, doi: 10.1101/2021.02.09.430407.

[5] M. Fritsch and M. Kudic, "Micro dynamics and macro stability in inventor networks," Journal of Technology Transfer, vol. 47, no. 2, pp. 353–382, Apr. 2022, doi: 10.1007/s10961-021-09851-8.

[6] J. Li, Y. Shi, X. Sun, and D. Liu, "A complexity assessment framework with structure entropy for a cloud-edge collaborative manufacturing system," IET Collaborative Intelligent Manufacturing, vol. 5, no. 2, Jun. 2023, doi: 10.1049/cim2.12077.

[7] Z. Guo, Y. Zhang, S. Liu, X. V. Wang, and L. Wang, "Exploring self-organization and self-adaption for smart manufacturing complex networks," Frontiers of Engineering Management, vol. 10, no. 2, pp. 206–222, Jun. 2023, doi: 10.1007/s42524-022-0225-1.

[8] F. Salehi, P. Shapira, and J. Zolkiewski, "Commercialization networks in emerging technologies: the case of UK nanotechnology small and midsize enterprises," Journal of Technology Transfer, 2022, doi: 10.1007/s10961-022-09923-3.

[9] N. Stojčić, "Collaborative innovation in emerging innovation systems: Evidence from Central and Eastern Europe," Journal of Technology Transfer, vol. 46, no. 2, pp. 531–562, Apr. 2021, doi: 10.1007/s10961-020-09792-8.

[10] G. Santoro, S. Bresciani, and A. Papa, "Collaborative modes with Cultural and Creative Industries and innovation performance: The moderating role of heterogeneous sources of knowledge and absorptive capacity," Technovation, vol. 92–93, Apr. 2020, doi: 10.1016/j.technovation.2018.06.003.

[11] D. Nepelski and V. Van Roy, "Innovation and innovator assessment in R&I ecosystems: the case of the EU Framework Programme," Journal of Technology Transfer, vol. 46, no. 3, pp. 792–827, Jun. 2021, doi: 10.1007/s10961-020-09814-5.

[12] P. Pirzadeh, H. Lingard, and N. Blismas, "Design Decisions and Interactions: A Sociotechnical Network Perspective," J Constr Eng Manag, vol. 147, no. 10, Oct. 2021, doi: 10.1061/(asce)co.1943-7862.0002136.

[13] M. Oraee, M. R. Hosseini, D. Edwards, and E. Papadonikolaki, "Collaboration in BIM-based construction networks: a qualitative model of influential factors," Engineering, Construction and Architectural Management, vol. 29, no. 3, pp. 1194–1217, Mar. 2022, doi: 10.1108/ECAM-10-2020-0865.

[14] C. F. Rocha, C. O. Quandt, F. Deschamps, and S. Philbin, "R&D collaboration strategies for industry 4.0 implementation: A case study in Brazil," Journal of Engineering and Technology Management - JET-M, vol. 63, Jan. 2022, doi: 10.1016/j.jengtecman.2022.101675.

[15] Z. Zhao, A. Broström, and J. Cai, "Correction to: Promoting academic engagement: university context and individual characteristics," Journal of Technology Transfer, vol. 45, no. 1. Springer, p. 338, Feb. 01, 2020. doi: 10.1007/s10961-018-9687-z.

[16] F. Lupi et al., "Blockchain-based Shared Additive Manufacturing," Comput Ind Eng, vol. 183, Sep. 2023, doi: 10.1016/j.cie.2023.109497.

[17] H. Löfsten and M. Klofsten, "Exploring dyadic relationships between Science Parks and universities: bridging theory and practice," Journal of Technology Transfer, 2024, doi: 10.1007/s10961-024-10064-y.

[18] A. Prabhakaran, A. M. Mahamadu, L. Mahdjoubi, P. Manu, C. K. I. Che Ibrahim, and C. O. Aigbavboa, "The effectiveness of interactive virtual reality for furniture, fixture and equipment design communication: an empirical study," Engineering, Construction and Architectural Management, vol. 28, no. 5, pp. 1440–1467, 2020, doi: 10.1108/ECAM-04-2020-0235.

[19] J. Li, L. Lyu, X. Liu, X. Zhang, and X. Lyu, "FLEAM: A Federated Learning Empowered Architecture to Mitigate DDoS in Industrial IoT," Dec. 2020, [Online]. Available: http://arxiv.org/abs/2012.06150

[20] M. Gronle, M. Grasso, E. Granito, F. Schaal, and B. M. Colosimo, "Open data for open science in Industry 4.0: In-situ monitoring of quality in additive manufacturing," Journal of Quality Technology, vol. 55, no. 2, pp. 253–265, 2023, doi: 10.1080/00224065.2022.2106910.

[21] M. Kumar, N. Tsolakis, A. Agarwal, and J. S. Srai, "Developing distributed manufacturing strategies from the perspective of a product-process matrix," Int J Prod Econ, vol. 219, pp. 1–17, Jan. 2020, doi: 10.1016/j.ijpe.2019.05.005.

[22] S. S. Nudurupati, P. Garengo, and U. S. Bititci, "Impact of the changing business environment on performance measurement and management practices," Int J Prod Econ, vol. 232, Feb. 2021, doi: 10.1016/j.ijpe.2020.107942.

[23] H. Moradlou, S. Roscoe, and A. Ghadge, "Buyer–supplier collaboration during emerging technology development," Production Planning and Control, vol. 33, no. 2–3, pp. 159–174, 2022, doi: 10.1080/09537287.2020.1810759.

Bibliography

Davey, T., Meerman, A., Galan Muros, V., Orazbayeva, B., Baaken, T., Allinson, R., Plewa, C., Melonari, M., Tercanli, H., Paula Troutt, M., Japarashvili, N., Dreier, A., Serbin, D., Rosalinde Hofer, A., Perkmann, M., & Potter, J. (2018). The State of University-Business Cooperation in Europe. https://doi.org/10.2766/676478

Fritsch, M., & Kudic, M. (2022). Micro dynamics and macro stability in inventor networks. Journal of Technology Transfer, 47(2), 353–382. https://doi.org/10.1007/s10961-021-09851-8

Gamper, M. (2022). Social Network Theories: An Overview. In Social Networks and Health Inequalities (pp. 35–48). Springer International Publishing. https://doi.org/10.1007/978-3-030-97722-1_3

Gronle, M., Grasso, M., Granito, E., Schaal, F., & Colosimo, B. M. (2023). Open data for open science in Industry 4.0: In-situ monitoring of quality in additive manufacturing. Journal of Quality Technology, 55(2), 253–265. https://doi.org/10.1080/00224065.2022.2106910

Guo, Z., Zhang, Y., Liu, S., Wang, X. V., & Wang, L. (2023). Exploring self-organization and self-adaption for smart manufacturing complex networks. Frontiers of Engineering Management, 10(2), 206–222. https://doi.org/10.1007/s42524-022-0225-1

Kumar, M., Tsolakis, N., Agarwal, A., & Srai, J. S. (2020). Developing distributed manufacturing strategies from the perspective of a product-process matrix. International Journal of Production Economics, 219, 1–17. https://doi.org/10.1016/j.ijpe.2019.05.005

Li, J., Lyu, L., Liu, X., Zhang, X., & Lyu, X. (2020). FLEAM: A Federated Learning Empowered Architecture to Mitigate DDoS in Industrial IoT. http://arxiv.org/abs/2012.06150

Li, J., Shi, Y., Sun, X., & Liu, D. (2023). A complexity assessment framework with structure entropy for a cloud-edge collaborative manufacturing system. IET Collaborative Intelligent Manufacturing, 5(2). https://doi.org/10.1049/cim2.12077

Löfsten, H., & Klofsten, M. (2024). Exploring dyadic relationships between Science Parks and universities: bridging theory and practice. Journal of Technology Transfer. https://doi.org/10.1007/s10961-024-10064-y

Lupi, F., Cimino, M. G. C. A., Berlec, T., Galatolo, F. A., Corn, M., Rožman, N., Rossi, A., & Lanzetta, M. (2023). Blockchain-based Shared Additive Manufacturing. Computers and Industrial Engineering, 183. https://doi.org/10.1016/j.cie.2023.109497

Moradlou, H., Roscoe, S., & Ghadge, A. (2022). Buyer–supplier collaboration during emerging technology development. Production Planning and Control, 33(2–3), 159–174. https://doi.org/10.1080/09537287.2020.1810759

Nepelski, D., & Van Roy, V. (2021). Innovation and innovator assessment in R&I ecosystems: the case of the EU Framework Programme. Journal of Technology Transfer, 46(3), 792–827. https://doi.org/10.1007/s10961-020-09814-5

Nudurupati, S. S., Garengo, P., & Bititci, U. S. (2021). Impact of the changing business environment on performance measurement and management practices. International Journal of Production Economics, 232. https://doi.org/10.1016/j.ijpe.2020.107942

Oraee, M., Hosseini, M. R., Edwards, D., & Papadonikolaki, E. (2022). Collaboration in BIM-based construction networks: a qualitative model of influential factors. Engineering, Construction and Architectural Management, 29(3), 1194–1217. https://doi.org/10.1108/ECAM-10-2020-0865

Pirzadeh, P., Lingard, H., & Blismas, N. (2021). Design Decisions and Interactions: A Sociotechnical Network Perspective. Journal of Construction Engineering and Management, 147(10). https://doi.org/10.1061/(asce)co.1943-7862.0002136

Prabhakaran, A., Mahamadu, A. M., Mahdjoubi, L., Manu, P., Che Ibrahim, C. K. I., & Aigbavboa, C. O. (2020). The effectiveness of interactive virtual reality for furniture, fixture and equipment design communication: an empirical study. Engineering, Construction and Architectural Management, 28(5), 1440–1467. https://doi.org/10.1108/ECAM-04-2020-0235

Rocha, C. F., Quandt, C. O., Deschamps, F., & Philbin, S. (2022). R&D collaboration strategies for industry 4.0 implementation: A case study in Brazil. Journal of Engineering and Technology Management - JET-M, 63. https://doi.org/10.1016/j.jengtecman.2022.101675

Salehi, F., Shapira, P., & Zolkiewski, J. (2022). Commercialization networks in emerging technologies: the case of UK nanotechnology small and midsize enterprises. Journal of Technology Transfer. https://doi.org/10.1007/s10961-022-09923-3

Santoro, G., Bresciani, S., & Papa, A. (2020). Collaborative modes with Cultural and Creative Industries and innovation performance: The moderating role of heterogeneous sources of knowledge and absorptive capacity. Technovation, 92–93. https://doi.org/10.1016/j.technovation.2018.06.003

Schultz, T., Putze, F., Fehr, T., Meier, M., Mason, C., Ahrens, F., & Herrmann, M. (2021). Linking Labs: Interconnecting Experimental Environments. 1–13. https://doi.org/10.1101/2021.02.09.430407

Stojčić, N. (2021). Collaborative innovation in emerging innovation systems: Evidence from Central and Eastern Europe. Journal of Technology Transfer, 46(2), 531–562. https://doi.org/10.1007/s10961-020-09792-8

Winickoff, D. E., Kreiling, L., Borowiecki, M., Garden, H., & Philp, J. (2021). Collaborative Platforms for Emerging Technology: Creating Convergence Spaces. http://www.oecd.org/termsandconditions

Zhao, Z., Broström, A., & Cai, J. (2020). Correction to: Promoting academic engagement: university context and individual characteristics. In Journal of Technology Transfer (Vol. 45, Issue 1, p. 338). Springer. https://doi.org/10.1007/s10961-018-9687-z

https://www.canva.com/ - Canva official website. 22-03.2024

https://www.vosviewer.com/ - VOSviewer official website. 20-03.2024

https://clarivate.com/products/scientific-and-academic-research/research-discovery-and-workflow-solutions/webofscience-platform/ - Web of Science official website. 15-03.2024

THE USE OF NANOTECHNOLOGY IN MEDICINE

Florentina BURLACU[1]

Grațiela Doina GHIOALDĂ LUȚĂ[2]

Abstract

Nanotechnology is a new field of science that produces extremely small structures from organic or inorganic materials. These nanoparticles can be used in many fields: medicine, foods, cosmetics, clothing, constructions, furniture, not only in the making of electronics or in the automotive industry. Because of their special properties nanoparticles, also known as nanomaterials, can revolutionize all fields in which they are used. But the domain in which they can bring the most important benefits for human life is medicine. Nanoparticles can be used in the medicine field in the production of drugs and vaccines, and also in diagnostic and investigative tools, etc. Nanomaterials are used with great success in imaging and in the construction of medical devices. Many European citizens have legitimate concerns about the safe use of nanoparticles in medicines. All these concerns can be put aside only through thoroughly testing all the medicines that contain nanoparticles before they will be put on the market for general use.

Keywords: Nanotechnology, State of the Art, New Technology, Medicine, Health

JEL Classification: O32

1. Introduction

Nanotechnology is not good or bad, it only is. The way it is used for the benefit of humanity or against humanity can determine if this discovery is The Holy Grail or the Nemesis for humankind. Nanotechnology can change the world bringing benefits in many important fields, but in medicine it can do miracles, and it can really make the difference between life and death.

Before we can discuss the use of nanotechnology in the medicine field, we must first define the term of nanotechnology.

[1] PhD candidate at The Bucharest University of Economic Studies, Bucharest, Romania, e-mail address: florentina.burlacu@yahoo.com

[2] PhD candidate at The Bucharest University of Economic Studies, Bucharest, Romania, e-mail address: gratiela101@yahoo.com

Artem Oppermann in his article „What Is Nanotechnology?" explains that nanotechnology is the „field of science and engineering that focuses on the design and manufacture of extremely small devices and structures" [2].

These extremely small devices that are smaller than 100 nanometers (nm) can be used in many fields from electronics to medicine, from energy to textiles, their range of applications is very wide [2].

Oppermann says that „nanotechnology refers to any device or product that is created or modified by manipulating matter at the nanoscale. By controlling atoms and molecules, companies can develop nanomaterials no more than 100 nanometers thick and apply these materials to industries like healthcare, sports and electronics" [2].

Nanotechnology is so important because it can be utilized to create materials and to develop devices and systems that have unique properties and functions. The extremely small size of the materials that are used in nanotechnology allows them to have different physical and chemical proprieties than the same materials have when they are at a larger scale [2].

The author of this article explains that „Due to their small size, nanomaterials have a large surface area-to-volume ratio, which can lead to increased reactivity, strength and conductivity" [2].

Because nanomaterials are exceptionally small in size, they can be incorporated very easy in a vast variety of products and processes such as medical treatments, electronic devices, energy production, environmental remediation, etc. Nanomaterials are used not only for improving the already existing products, but also for creating new ones that have an increased durability [2].

We are using nanotechnology in our daily life and some of us don't even know it. Nanotechnology is used in the making of electronics - smartphones, laptops, televisions, etc., cosmetics - foundations, moisturizers, etc., sporting goods - equipment, etc., clothing - outdoor gear, athletic wear, etc., sunscreen, furniture, adhesives, automotive, etc. [2].

Many are wondering how nanotechnology is made and how exactly nanomaterials are created. Nanomaterials are developed using the following methods:

- *Top-down approaches* – are using procedures like lithography, laser, chemical etching or mechanical milling to carve into a larger piece of material and to reduce it to nanoscale. They are used for making integrated circuits in electronics.
- *Bottom-up approaches* – are utilizing techniques such as chemical synthesis or self-assembly to build up materials from smaller components such as atoms or molecules. This is the way the synthesis of nanoparticles is created.
- *Self-assembly* – uses techniques such as template-assisted self-assembly or directed self-assembly to create materials or structures that are able to organize themselves into the desired nanoscale structure. For example, these techniques are used in the production of block copolymers.

- *Physical vapor deposition* – vaporizing a material to form a thin film used in the coating of various objects such as cutting tools.
- *Chemical vapor deposition* – uses a technique that makes a gas to react to a surface in order to create a thin film of the desired material. This procedure is used in the production of thin films for solar cells [2].

There are four main types of nanomaterials that are used in nanotechnology:

1. ***Carbon-based nanomaterials*** – are being made up of carbon nanotubes produced through carbon-based vapor deposition. In order to obtain such materials carbon is heated and added up after a reaction between a surface and a catalyst takes place.
2. ***Metal-based nanomaterials*** – they consist of quantum dots that are created by growing in a solution, under very specific conditions, nanoscale crystals of two different elements.
3. ***Dendrimers*** – are nanoparticles that are composed of a core, inner shell and outer shell. They can be constructed in two ways: starting from the core or starting from the outer shell.
4. ***Nanocomposites*** – are formed of either multiple nanomaterials or a blend of nanomaterials and much larger materials. The result is stronger metals, plastics and other substances [2].

Using nanotechnology brings a lot of benefits: enhanced medical treatments - creating more targeted and effective drugs, diagnostic tools and medical devices, improved materials, increased energy efficiency, improved water filtration and purification, improved food safety and agriculture [2].

But we must never forget the potential risks of nanotechnology: health and environmental risks - we can't predict with accuracy the long-term health effects of exposure to nanomaterials, nanotechnology in food could be harmful for humans and also for the environment, economic risks - nanotechnology can produce economic and financial inequality and job displacement and ethical risks - using nanotechnology as a weapon in the military field or in surveillance [2].

Nanotechnology can change the world as we know it. Because it can be applied in many fields it can really improve the quality of the human life. But in our humble opinion in the medicine field nanotechnology can bring the most important and valuable thing in the world: health, healing and life for all human beings.

2. Problem Statement

In his book „Current Advances in the Medical Application of Nanotechnology" Mark Slevin explains that in the medical field nanotechnology will help implement „new therapies, more rapid and sensitive diagnostic and investigative tools for normal and diseased tissues, and new materials for tissue engineering" [5].

Nanotechnology plays an important role in biomedical research and clinical medicine. Due to their small size nanoparticles can deliver drugs exactly into the targeted diseased cells or into exact locations within the tissue, which increases drug concentrations locally and reduces systemic toxicities. Nanoparticles can also be used in achieving sustained drug release after only one injection. Nanotechnology can help improving medical imaging and diagnosis techniques [5].

Mark Slevin explains that „nanomedicine is the medical use of molecular-sized particles to deliver drugs, heat, light or other substances to specific cells in the human body. Engineering particles to be used in this way allows detection and/or treatment of diseases or injuries within the targeted cells, thereby minimizing the damage to healthy cells in the body." [5].

In order to increase the efficiency and to reduce the side effects it is important to direct the drugs exactly where the disease is localized. This is possible with the help of some nanoparticles that have only 150 nm in diameter, they are made from biodegradable polymer poly (d,1-lactic-co-glycolic acid) and poly (ethylene glycol) and have been developed especially for encapsulating drugs, specific proteins or antibodies which can be attached to artificial RNA strands – aptamers. These nanoparticles are picked up by specific cells in which they dissolve in order to discharge the drug or the protein [5].

Slevin says that the treatment of a disease depends on the identification of a target and the delivery of a therapeutic agent which either causes a function to be restored, switches off inappropriate activity, or destroys the diseased cell in the case of cancer [5].

Nanotechnology is used not only in the production of drugs, but also in the development of vaccines for humans.

In the book „Micro- and Nanotechnology in Vaccine Development" edited by Mariusz Skwarczynski and István Tóth, the authors of chapter 5 Neeland, de Veer and Scheerlinck explain that nanoparticles can function as both a delivery system and an adjuvant to increase immunity [4].

The more frequently used nanoparticles in the development of vaccines are the following:

- *Virus-like particles (VLPs)* – 20-800 nm – are self-assembling nanoparticles that morphologically resemble infectious virions but lack infectious viral nucleic acid.
- *Aluminum salts and related inorganic particles (eg. $CaPO_4$, silicates, or gold)* – 25-1000 nm – many conventional vaccines use as adjuvant aluminum salts because they bind antigens through charge and hydrostatic interactions in order to enhance absorption by dendritic cells.
- *Immune stimulating complexes (ISCOMs)* – 40 nm – are nanoparticles that resemble with a cage. They are made up of Quil A (saponin adjuvant), cholesterol and phospholipids.

- *Viral sized inert polystyrene rigid nanoparticles* – 40-50 nm – they enhance the absorption by the dendritic cells at 40-50 nm viral-like size and can conjugate antigen.
- *Micelles* – 40-200 nm – the hydrophobic core is able to hold hydrophobic antigens or peptides coupled to the outer soluble layer.
- *Emulsions* – 50-600 nm – are oil-in-water formulas that transport antigens inside their core.
- *Liposomes* – 100-400 nm – are phospholipid bilayer vesicles with a water core in which the antigen is encapsulated.
- *Polymer nanoparticles [chitosan, poly (lactic-co-glycolic acid) - (PLGA), polyglycolic acid (PGA)]* – 40 nm-large gels – are being evaluated for development, they form gels and stable spheres to absorb or conjugate [4].

The authors of this study are convinced that the manipulation of the innate response through nanoparticles has a deep and important effect on the induced adaptive immune reaction. The addition of pathogen-associated molecular patterns (PAMPs) as adjuvants frequently results in serious inflammation at the injection site and the local lymph node [4].

It was discovered that associating these pathogen-associated molecular patterns to nanoparticles can result in a lower reactogenicity and at the same time maintaining or enhancing immunogenicity. Consequently, nanoparticles could really improve the clinical result of the next generation of effective adjuvants [4].

The researchers give a fair warning that if the pharmaceutical companies want to use these new technologies in the development of commercial products, they all need to improve their understanding of these new adjuvants. And the most important thing is that they have to address issues relating to safety including biodistribution and degradation of the nanoadjuvants [4].

In the last 4 years nanotechnology was used on a large scale in the development of vaccines. Nanoparticles were used in both anti-Covid-19 vaccines developed by Pfizer-BioNTech and Moderna [1].

The goal of the powerful pharmaceutical companies is to develop a multivalent vaccine for all the 20 known influenza A and B virus sub-types. This new generation of influenza vaccine is developed with the help of modified messenger RNA (mRNA) and lipid nanoparticles [3].

But the vaccine industry is not the only one that utilizes with great success nanotechnology. Another medical field in which nanoparticles can bring great change and start a revolution is imaging, diagnostics and medical devices.

The use of nanotechnology in medicine led to advances in early detection, imaging, and treatment of disease. The possibility to utilize nanoparticles in order to diagnose, treat and monitor disease progression without surgery, biopsy, or other invasive methodologies may increase the rate of survival, improve the quality of life and also reduce the costs [6].

The advances in bioconjugation and self-assembly made possible the development of „smart" targeted nanoparticles, which combine targeting molecules with therapeutic and/or imaging agents all in one extremely small particle. These nanoparticles make possible the delivery of imaging agents and drugs to the disease sites, resulting in the production of high contrast images with low levels of noise and also in an increased efficiency of the therapy with reduced adverse side effects [6].

Some nanoparticles have been developed from various naturally occurring organic building blocks such as: lipids, proteins, polymers. And others were produced from inorganic materials: gold, silica, iron [6].

More than 25 nanoparticles or platforms how they are also referred as in this study have been approved for clinical use. We mention here only one of them that has been used for imaging: the superparamagnetic iron oxide nanoparticles that were developed as contrast agents for magnetic resonance imaging (MRI) [6].

Although inorganic nanoparticles have some advantages like wide availability, controlled shape and size, and easy surface manipulation, but at the same time they have poor stability under aqueous conditions and low cellular transfer efficiency. Instead, lipid-based nanoparticles like liposomes and micelles can carry significant payloads of drugs, but they experience poor stability in vivo [6].

Apparently, the solution for all these limitations are the nanomaterials that are made up of plant viruses. Viral nanoparticles (VNPs) have many qualities and come in many shapes and sizes. In nature the viral capsid has as a main function the protection of its genome under various environmental conditions, therefore the VNPs are naturally stable and monodisperse. The viral nanoparticles are able to combine the strengths of inorganic and lipid-based materials [6].

The viral nanoparticles are derived from plants or bacteria and are able to deliver an ideal basis for the development of targeted imaging agents and „drug delivery vehicles". Virus-like particles (VLPs) are a subset of viral nanoparticles (VNPs) that can be produced at a large scale. VLPs lack any genomic replicative information and because of that they are safe, being non-infectious and non-hazardous for humans and animals [6].

The authors of this study explain that „VNPs and VLPs are self-assembling systems that are highly symmetrical, dynamic, polyvalent, and monodisperse, rendering them one of the most advanced nanomaterials produced in nature. They offer the advantages of biocompatibility and biodegradability over synthetic inorganic nanoparticles. In addition, they are extremely robust and well characterized and can be produced in large quantities within a short period of time." [6].

The structure of the viral nanoparticles makes it possible for them to be modified in several ways in order to allow the loading of drugs, imaging agents and other nanoparticles in their

internal cavity, as well as chemical conjugation of targeting ligands on the external surface for the delivery in a specific tissue [6].

The application of nanotechnology in medicine has the potential to bring health and to cure all the incurable diseases. Also, this new technology can increase the life expectancy of all people on the planet.

But for this massive and revolutionary change in medicine to happen it is necessary for two things:

1. to test the long-term effects that nanoparticles have on human health and if they can be utilized without generating serious side effects and health problems,
2. the small group of rich people that have the patent for nanotechnology to wish to use the nanoparticles for the right causes and in the benefit of all humans.

We continue this article with the research questions that this paper is trying to answer.

3. Research Questions/Aims of the research

The questions we would like to find answers for are the following:

Have Europeans heard about nanomaterials?

Do Europeans have concerns about the negative impact of nanomaterials?

Do Europeans want to be informed about the products that contain nanomaterials especially the medicines?

We believe the answer to all the above questions is affirmative.

The objective of this article is to demonstrate that the citizens from the member states of the European Union are aware of the existence of nanotechnology and nanomaterials (nanoparticles). Europeans are to some extent informed about this subject and have legitimate concerns and questions when it comes to the use of nanoparticles especially in food, medicines and cosmetics.

4. Research Methods

We used in this article mainly the qualitative method in order to obtain the extensive data about the use of nanotechnology in the medicine field. Quantitative analysis is also used, especially when we present and interpret the statistical data. The following techniques are used in this paper: a case study on the use of nanoparticles in medicine, and the analysis of the theoretical works in the nanotechnology field.

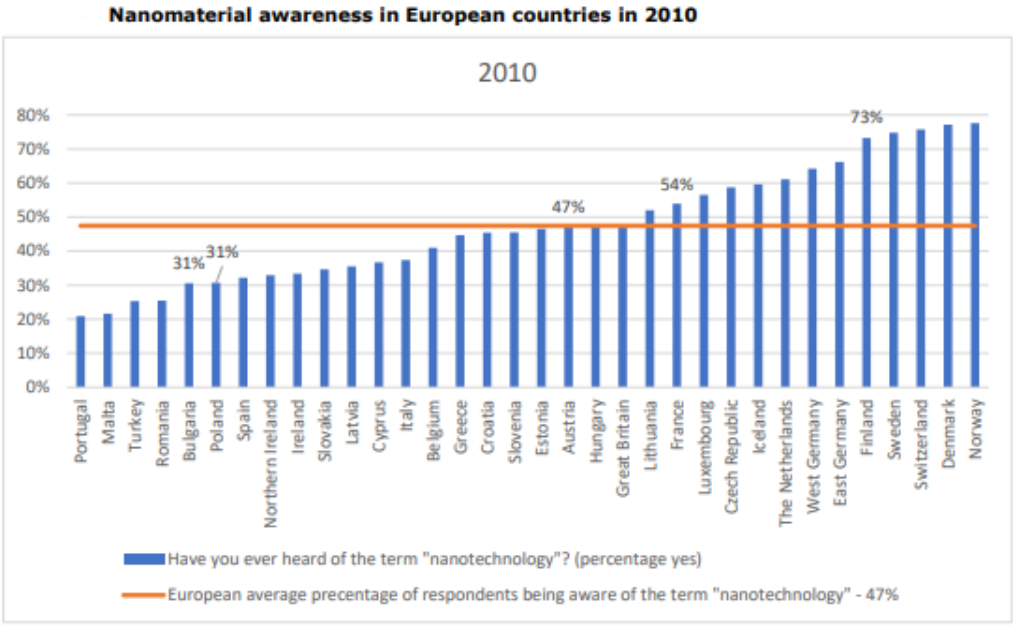
5. Findings

When it comes to nanotechnology statistical data is a little hard to find especially on open access journals. We don't understand why but ordinary people are also reserved in answering some questions about this subject.

The statistical data we are using come from a study that was commissioned by the European Chemicals Agency (ECHA) and was published in November 2020. The previous mentioned study „Understanding Public Perception Of Nanomaterials And Their Safety In The EU. Final report" is published on The European Union Observatory for Nanomaterials (EUON).

EUON is an entity funded by the European Commission and is being hosted and kept in existence by the European Chemicals Agency (ECHA). The European Union Observatory for Nanomaterials provides information about the existing nanomaterials on the EU market [7].

In the study „Understanding Public Perception Of Nanomaterials And Their Safety In The EU. Final report" it is presented a graphic that shows the European citizens awareness regarding the nanomaterials in the year 2010.



Graphic 1. Nanomaterial awareness in European countries in 2010[3]

[3] Source: *Understanding Public Perception Of Nanomaterials And Their Safety In The EU Final report,* November 2020, p. 42.

At the question „Have you ever heard of the term „nanotechnology"?" we can see illustrated with the blue columns the percentages of the „yes" answers in every European state. The red horizontal line fixed at 47% shows the average percentage of the European respondents that are aware of the term „nanotechnology".

From Graphic 1 we can easily see that the five countries that have the highest percent of knowledge about nanotechnology (over 70%) are with only one exception from the Northern Europe (Norway, Denmark, Switzerland, Sweden, Finland).

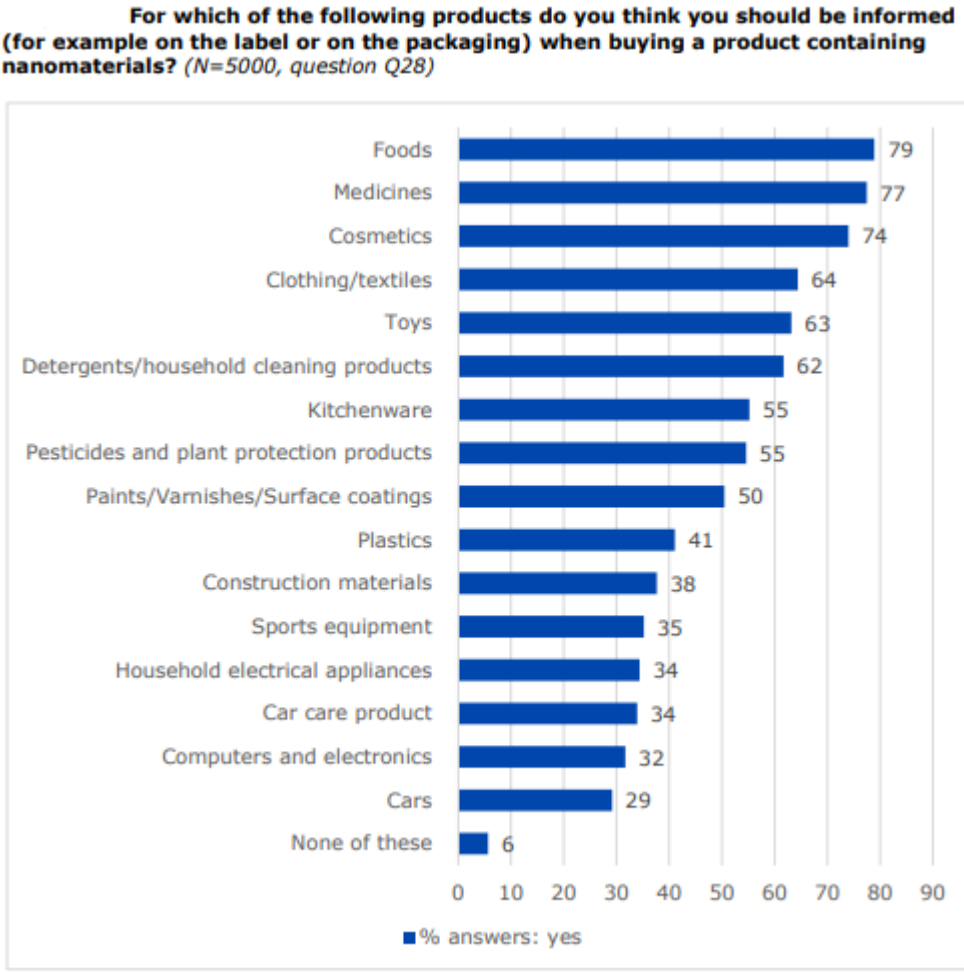
On the other hand, the European states that have the lowest level of awareness when it comes to nanotechnology (under 30%) are Portugal, Malta, Turkey and our country Romania.

In Table 1 we can see illustrated in percentages the concerns about the potential negative impact of nanomaterials on respondents' lives. These numbers must be interpreted having in mind the fact that this view is influenced by the level of people's knowledge about nanomaterials.

VIEW BY LEVEL OF WHAT PEOPLE HEARD ABOUT NANOMATERIALS				
		What people heard about nanomaterials		
Concern level related to nanomaterials	Total	Nothing at all	A little	A lot
Number of respondents	5000	1752	2703	545
I am concerned about possible negative impacts on my life	25%	21%	27%	30%
I am not worried about possible negative impact on my life	38%	23%	45%	57%
I do not know + I do not care	37%	57%	28%	14%

Table 1. Concerns about the potential negative impact of nanomaterials on respondents´ lives[4]

We can see from Table 1 that people with a higher knowledge about nanomaterials have a lower percentage of concern about the negative impact of nanotechnology in their life and at the same time they have a higher degree of acceptance of the new technology. The response „ I don't know" or "I don't care" is used primarily by the people that have no awareness concerning nanomaterials [8].



[4] Source: *Understanding Public Perception Of Nanomaterials And Their Safety In The EU Final report,* November 2020, p. 64.

Graphic 2. For which of the following products do you think you should be informed (for example on the label or on the packaging) when buying a product containing nanomaterials?[5]

In Graphic 2 is shown the list of products or category of products that European citizens want to know if they contain nanomaterials. We can observe that medicines are on the second place with 77% right after foods that are on the first place with 79%. On the third place are situated with 74% cosmetics. On the last two positions in this graphic, we can find computers and electronics with only 32% and cars with 29%.

From Graphic 2 we can deduce the importance that is given by Europeans to the use of nanoparticles in medicines and in the medicine field in general. We are convinced that the majority of the Europeans have heard by now of nanotechnology and nanoparticles and are open to the advances of the medical science. Their only request is to be informed when medicines contain nanoparticles and of course we can presume that they want for this new technology to be thoroughly tested before its usage on a mass scale in medicines and in other consumption products.

6. Conclusion

Nanotechnology is a new field of science that can bring many benefits in the life of humans. Nanoparticles can be used in many domains with stupendous results, but their usage in the medical field has the potential to make miracles happen and to bring health, healing and longevity to all the human beings on this planet. But for this to happen it is very important to be meticulously tested before it can be used on humans.

In our opinion nanotechnology has the power to change the world. The only question is if this change will be in good or in bad. That depends on the small group of people that have the power and control over nanotechnology and the research in this field. Their way of thinking and their interests will shape our future.

Acknowledgment

First, we would like to thank to our supervisor Professor Toma Roman for his guidance, great support and kind advice throughout our PhD research studies. Second, we would like to thank to our families for the continuous support they have given us throughout our PhD studies.

[5] Source: *Understanding Public Perception Of Nanomaterials And Their Safety In The EU Final report,* November 2020, p. 118.

References

[1] Amit KHURANA, Prince ALLAWADHI, Isha KHURANA, Sachin ALLWADHI, Ralf WEISKIRCHEN, Anil Kumar BANOTHU, Deepak CHHABRA, Kamaldeep JOSHI, Kala Kumar BHARANI - *Role of nanotechnology behind the success of mRNA vaccines for COVID-19* - Nano Today, Volume 38, June 2021, https://www.sciencedirect.com/science/article/pii/S1748013221000670?via%3Dihub – 24.05.2024

[2] Artem OPPERMANN - *What Is Nanotechnology*? - Built In, Jan 16, 2024, https://builtin.com/hardware/nanotechnology - 18.05.2024

[3] Claudia P. AREVALO, Marcus J. BOLTON, Valerie LE SAGE, Naiqing YE, Colleen FUREY, Hiromi MURAMATSU, Mohamad-Gabriel ALAMEH, Norbert PARDI, Elizabeth M. DRAPEAU, Kaela PARKHOUSE, Tyler GARRETSON, Jeffrey S. MORRIS, Louise H. MONCLA, Ying K. TAM, Steven H. Y. FAN, Seema S. LAKDAWALA, Drew WEISSMAN, Scott E. HENSLEY - *A multivalent nucleoside-modified mRNA vaccine against all known influenza virus subtypes* – Science, Volume 378, Issue 6622, November 23, 2022, https://www.science.org/doi/10.1126/science.abm0271 - 24.05.2024

[4] Mariusz SKWARCZYNSKI, István TÓTH, (Ed.) - *Micro- and Nanotechnology in Vaccine Development* – [page 87 - 93]. ISBN: 9780323400299, ISBN: 0323400299. Elsevier Science, Oxford, UK, 2017.

[5] Mark SLEVIN, (Ed.) – *Current Advances in the Medical Application of Nanotechnology* – [page i - 5]. ISBN: 9781608051311, ISBN: 1608051315. Bentham Science Publishers, Sharjah, U.A.E, 2012.

[6] Mikhail Y. BEREZIN (Ed.) – *Nanotechnology for Biomedical Imaging and Diagnostics From Nanoparticle Design to Clinical Applications* – [page 401 - 402]. ISBN: 9781118121184, ISBN: 111812118X. Wiley, Hoboken, New Jersey, USA.

[7] The European Union Observatory for Nanomaterials (EUON) https://euon.echa.europa.eu/about-us - 29.05.2024

[8] *Understanding Public Perception Of Nanomaterials And Their Safety In The EU Final report, November 2020,* ISBN: 978-92-9481-730-3. This study was commissioned by the European Chemicals Agency (ECHA) and was carried out by SC&C s.r.o., EcoMole s.r.o. and ReachSpektrum s.r.o, Helsinki, Finland, https://euon.echa.europa.eu/documents/2435000/3268573/nano_perception_study_en.pdf/f7d43c54-8a78-061b-49b0-de619454e137?t=1605523237212 – 29.05.2024

Bibliography

Amit KHURANA, Prince ALLAWADHI, Isha KHURANA, Sachin ALLWADHI, Ralf WEISKIRCHEN, Anil Kumar BANOTHU, Deepak CHHABRA, Kamaldeep JOSHI, Kala Kumar BHARANI - *Role of nanotechnology behind the success of mRNA vaccines for COVID-19* - Nano Today, Volume 38, June 2021, https://www.sciencedirect.com/science/article/pii/S1748013221000670?via%3Dihub – 24.05.2024

Artem OPPERMANN - *What Is Nanotechnology*? - Built In, Jan 16, 2024, https://builtin.com/hardware/nanotechnology - 18.05.2024

Claudia P. AREVALO, Marcus J. BOLTON, Valerie LE SAGE, Naiqing YE, Colleen FUREY, Hiromi MURAMATSU, Mohamad-Gabriel ALAMEH, Norbert PARDI, Elizabeth M. DRAPEAU, Kaela PARKHOUSE, Tyler GARRETSON, Jeffrey S. MORRIS, Louise H. MONCLA, Ying K. TAM, Steven H. Y. FAN, Seema S. LAKDAWALA, Drew WEISSMAN, Scott E. HENSLEY - *A multivalent nucleoside-modified mRNA vaccine against all known influenza virus subtypes* – Science, Volume 378, Issue 6622, November 23, 2022, https://www.science.org/doi/10.1126/science.abm0271 - 24.05.2024

Federica Ragucci, Francesca Sireci, Francesco Cavallieri, Jessica Rossi, Giuseppe Biagini, Giovanni Tosi, Chiara Lucchi, Rodolfo Molina-Pena, Natalia Helen Ferreira, Mariana Zarur, Alba Ferreiros, William Bourgeois, François Berger, Miguel Abal, Audrey Rousseau, Frank Boury, Carmen Alvarez-Lorenzo, Emmanuel Garcion, Anna Pisanello, Giacomo Pavesi, Corrado Iaccarino, Luca Ghirotto, Maria Chiara Bassi, Franco Valzania – *Insights into Healthcare Professionals' Perceptions and Attitudes toward Nanotechnological Device Application: What Is the Current Situation in Glioblastoma Research?* – Biomedicines, Volume 11, Issue 7, 2023, Published: 28 June 2023, https://www.mdpi.com/2227-9059/11/7/1854 - 31.05.2024

Helena Silva Costa, Sebastian Sethe, Ana P Pêgo, I Anna S Olsson – *Scientists' perception of ethical issues in nanomedicine: a case study* – Nanomedicine, Volume 6, 2011, Issue 4, Published online: 30 Jun 2011, https://www.tandfonline.com/doi/full/10.2217/nnm.11.9 - 31.05.2024

Mariusz SKWARCZYNSKI, István TÓTH, (Ed.) - *Micro- and Nanotechnology in Vaccine Development* – [page 87 - 93]. ISBN: 9780323400299, ISBN: 0323400299. Elsevier Science, Oxford, UK, 2017.

Mark SLEVIN, (Ed.) - *Current Advances in the Medical Application of Nanotechnology* - [page i - 5]. ISBN: 9781608051311, ISBN: 1608051315. Bentham Science Publishers, Sharjah, U.A.E, 2012.

Massimo Bottini, Nicola Rosato, Fulvia Gloria, Sara Adanti, Nunziella Corradino, Antonio Bergamaschi, Andrea Magrini – *Public optimism towards nanomedicine* –

International Journal of Nanomedicine, Volume 6, 2011, Published online: 20 December 2011, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3260040/ - 31.05.2024

Mikhail Y. BEREZIN (Ed.) – *Nanotechnology for Biomedical Imaging and Diagnostics From Nanoparticle Design to Clinical Applications* – [page 401 - 402]. ISBN: 9781118121184, ISBN: 111812118X. Wiley, Hoboken, New Jersey, USA.

Najib Nassani, Youssef El-Douaihy, Yana Khotsyna, Thinzar Shwe, Suzanne El-Sayegh – *Knowledge, Perceptions, and Attitudes of Medical Residents Towards Nanomedicine: Defining the Gap* – Medical Science Educator, Volume 30, 2020, Published online: 20 October 2019, https://link.springer.com/article/10.1007/s40670-019-00837-8 - 31.05.2024

Ranjita SHEGOKAR (Ed.) – *Nanopharmaceuticals Volume 1: Expectations and Realities of Multifunctional Drug Delivery Systems* – [page 135-150]. ISBN: 978-0-12-817778-5. Elsevier, Oxford, UK, 2020.

Shuying CHEN, Xiangang HUANG, Yonger XUE, Ester ÁLVAREZ-BENEDICTO, Yesi SHI, Wei CHEN, Seyoung KOO, Daniel J. SIEGWART, Yizhou DONG, Wei TAO - *Nanotechnology-based mRNA vaccines* - Nature Reviews Methods Primers, Volume 3, Article number: 63, August 17, 2023, https://www.nature.com/articles/s43586-023-00246-7 - 24.05.2024

The European Union Observatory for Nanomaterials (EUON) https://euon.echa.europa.eu/about-us - 29.05.2024

Understanding Public Perception Of Nanomaterials And Their Safety In The EU Final report, November 2020, ISBN: 978-92-9481-730-3. This study was commissioned by the European Chemicals Agency (ECHA) and was carried out by SC&C s.r.o., EcoMole s.r.o. and ReachSpektrum s.r.o, Helsinki, Finland, https://euon.echa.europa.eu/documents/2435000/3268573/nano_perception_study_en.pdf/f7d43c54-8a78-061b-49b0-de619454e137?t=1605523237212 – 29.05.2024

Victoria Xenaki, Daniela Elena Costea, Mihaela Cuida Marthinussen, Mihaela Roxana Cimpan, Anne Nordrehaug Åstrøm – *Use of nanomaterials in dentistry: covariates of risk and benefit perceptions among dentists and dental hygienists in Norway* – Acta Odontologica Scandinavica, Volume 78, 2020, Issue 2, Published online: 27 September 2019, https://www.tandfonline.com/doi/full/10.1080/00016357.2019.1668055 - 31.05.2024

PROSPECTS FOR MIGRATION IN THE DEMOGRAPHIC CONTEXT OF THE EUROPEAN UNION

Elena Raluca CRISTIAN[1]

Carmen Dalia ȚÎRDĂ[2]

Laura Georgeta BĂRĂGAN[3]

Anda Veronica DAN[4]

Abstract

Studies in specialized literature on labor migration have started from cases of voluntary immigration, where the motivation of immigrants is economic/social or related to family reunification. This is often generated by other causes, such as more favorable environmental conditions (environmental migration) or temporary jobs (seasonal immigration). The migration phenomenon, especially immigrants, also generates a series of negative effects having an impact on the social plane, on the labor market and on the living standards of the native population.

The aim of this study is to observe how the number of immigrants influences the population of the twenty-seven European Union member states, as, it is assumed, it would be much easier for these states or specialized institutions to manage the negative effects of labor migration.

Immigration is part of the phenomenon of international migration, a phenomenon that is also extremely relevant in relation to the native population of the countries involved, as well as having economic and social effects in those countries. In the host country, immigration increases overall demand for goods and services, creates new opportunities for consumption, intensifies the supply of goods and services and makes the labor market more flexible, while also raising the number of unemployed people in various fields of activity.

Keywords: immigrants, the migration phenomenon, refugees, labor market, political asylum

JEL Classification: I24, J15, J61, O15, R23

[1] PhD, Teaching Assistant at Romanian American University, Bucharest, Romania. elena.raluca.cristian@rau.ro
[2] PhD, Lecturer at Romanian American University, Bucharest, Romania. carmen.dalia.tirda@rau.ro
[3] PhD, Lecturer at Romanian American University, Bucharest, Romania. laura.georgeta.baragan@rau.ro
[4] PhD, Teaching Assistant at Romanian American University, Bucharest, Romania. anda.veronica.dan@rau.ro

1. Introduction

The economic and social effects of migration have been identified over the year by studying immigration flows in the immigrants' favorite destination countries and in the European Union – Spain, Italy, Germany, the UK, and France. These studies and analyses over several years have shown that migration has both positive and negative implications at the level of the destination countries. Thus, in the host country, immigration increases overall demand, generates new opportunities for consumption, improves the variety of goods and services and makes the labor market more flexible in the medium and long terms. On the other hand, immigrants generate negative effects socially, on the labor market, on the living standards of the native population and on decisions to assimilate a new culture, traditions, and customs.

The aim of this study is to observe how the number of immigrants influences the population of the twenty-seven member countries of the European Union, as well as the management of the negative effects of the migration phenomenon by the institutions of the countries involved in this broad phenomenon. This topic has been chosen because immigration is a current phenomenon, which is extremely sensitive for some EU Member States as well as for immigrants.

According to the latest demographic statistics published officially by Eurostat, the population of the European Union has decreased by more than half a million inhabitants compared to 2020. Only seventeen EU countries have seen an increase in population after 2020, while seven countries have registered a significant decrease, among which we may mention Italy, Poland, Greece, and Croatia.

On the 1st of January 2022, the population of the European Union was 446.7 million, in a slight decrease from 2021. Although the population increased by 4% compared to 2001, the population decreased for the second consecutive year.

In January 2022, there were 585,000 fewer people in the European Union compared to the first day of 2020. The population of the European Union continued to increase annually until January 2020 and then began to resume the negative trend, which is attributed to casualties caused by the COVID-19 pandemic.

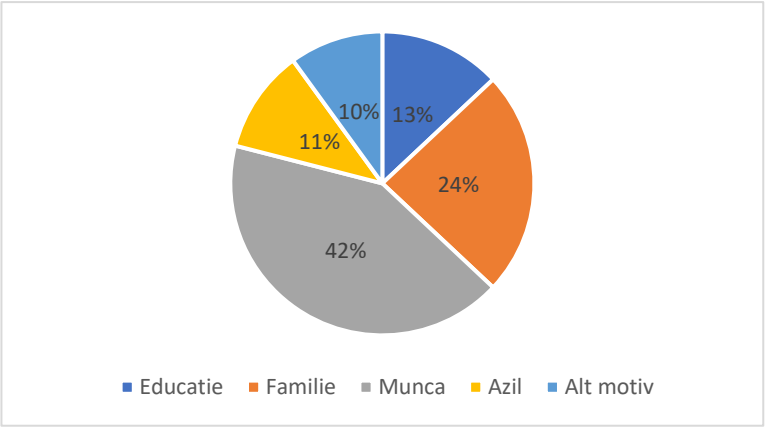
In 2022, 9.93 million non-EU citizens were working in the EU, out of a total of 193.5 million people aged between twenty and sixty-four, representing 5.1% of the total. The total number of employees in the age group twenty-sixty-four is 197.458 million, of which 9.927

million are non-EU nationals. In 2022, the employment rate among the EU working population was higher for European citizens (77.1%) than for non-EU citizens (61.9%).[5]

According to Eurostat figures from 2022, 2.25 million people immigrated to the European Union and 1.12 million people emigrated from the European Union, thus registering 1.14 million people as total net immigration into the European Union.

In 2022, almost 3.4 million first residence permits were issued in the EU, compared to 2.9 million in 2021, hence exceeding the figures recorded before the COVID-19 pandemic started (3 million in 2019). Compared to 2021, the number of first permits issued by Germany (+ 190%), Malta (+ 164%) and Ireland (+ 146%) more than doubled. 2022 saw a particularly large increase in asylum-related permits (+ 60%), but the number of permits issued for education (+ 29%), family (+ 28%) and other reasons (+ 47%) also increased compared to the previous year. On the other hand, the number of permits issued for work decreased compared to 2021 (-6%). In 2022, the first residence permits were issued in the European Union for foreign immigrants, as follows[6]:

Chart 1. Residence permits for immigrants in the European Union for 2022



Source: the authors' processing based on 2022 Eurostat data

The total population of the European Union is 446.7 million, out of which 23.8 million are citizens of the European Union and 38 million are citizens born outside the European Union, i.e. foreign immigrants from other continents: Asia, Africa, North or South America or other European countries (Ukraine, Russia, Republic of Moldova etc.).

Population density, i.e. the number of persons per square kilometer (km²), averaged 109 people per square kilometer in the EU in 2019, a value which varied significantly between Member States. The highest population density was observed in Malta (1 595 persons per

[5] https://ec.europa.eu/eurostat/databrowser/bookmark/a2e95b51-c08a-462c-b4b7-c0a11bab32a7?lang=en
[6] https://ec.europa.eu/eurostat/databrowser/bookmark/9d961e67-a618-4485-8353-c0ec8b12808a?lang=en

square kilometer), followed at a distance by the Low Countries (507) and Belgium (377). At the other end there were Finland (18 persons per square kilometer), and Sweden (25), with the lowest population density. Comparing 2019 with 2001, there was an increase in population density in around two-thirds of the Member States, with the largest increases in Malta (from 1 245 in 2001 to 1 595 in 2019), Luxembourg (from 171 to 240) and Belgium (339 to 377). Romania (96 to 83) and Lithuania (55 to 45) saw the largest decreases. In the EU, population density increased on average from 104 people per square kilometer to 109 over this period.[7]

2. Immigrant labor force assimilation in the European Union labor market

On May 13, 2015, the European Commission adopted a new European Agenda for Migration, a document proposing a series of measures needed to meet the current challenges in the area of migration of persons, as well as a series of medium and long-term initiatives to provide solutions for better management of the migration phenomenon, in all its political, social, economic, and religious aspects at state level.

The implementation of the EU Agenda for Migration started with the adoption, on May 27, 2015, by the European Commission, of a first package of measures focusing on three strands of action: legal migration, temporary measures regarding the resettlement of persons in need of international protection, and the consolidation of operational capacity (search and rescue at sea). A second package of measures was adopted on September 9, 2015, and includes measures on return policy, the list of safe countries of origin, the EU emergency trust fund for Africa, the refugee resettlement scheme, funding for migration and security, balancing responsibility, and solidarity in migration issues.[8]

Migration policy should become a common policy rather than a communitarian policy, to ensure that all Member States have the same approaches, rules and regulations towards both internal migrants (those within the European Union) and external migrants (those from third countries). Such a comprehensive and unified approach can only be to the benefit of the European Union, especially in the social context strongly shaken by Brexit. [9]

The EU and its Member States are stepping up their efforts to establish an effective, humanitarian, and safe European migration policy. The European Council plays an important role in this respect by setting strategic priorities. Based on these priorities, the EU Council establishes lines of action and gives negotiating mandates to countries outside the EU. Moreover, it adopts legislation and defines specific programs. Over the past few

[7] https://insse.ro/cms/demography-in-europe/bloc-1a.html?lang=ro
[8] http://www.mae.ro/termeni-conditii
[9] Buță, Viorel; Radu, Andreea Mihaela (2019), Migration Of Labor Force Within The European Union, Revista Academiei de Științe ale Securității Naționale, Volume 1, p73.

years, the Council and the European Council have built a strong response to the pressures determined by migration[10].

On December 20, 2023, the Council and the European Parliament reached an agreement on the five key regulations [11]:

- ✓ a new set of regulations on the management of asylum and migration situations
- ✓ new rules governing migration and *force majeure* situations.
- ✓ updating the fingerprint database
- ✓ a new set of regulations on the screening procedure
- ✓ a common asylum procedure

On February 8, 2024, the representatives of the EU Member States approved the three laws that had already been agreed upon between the Council and the Parliament in 2022[12]:

- ✓ uniform rules on asylum applications
- ✓ better reception conditions
- ✓ a new EU resettlement framework

In 2022, 9.93 million non-EU citizens were working in the EU, out of a total of 193.5 million people aged between twenty and sixty-four, representing 5.1% of the total. The total number of employees in the age group twenty-sixty-four is 197.458 million, of which 9.927 million are non-EU citizens. In 2022, the employment rate among the EU active population was higher for EU nationals (77.1%) than for non-EU nationals (61.9%).[13]

Since the beginning of Russia's military aggression in Ukraine in February 2022, Europe has received the greatest number of refugees since the Second World War.

On the European Union labor market, in 2022, non-EU citizens or immigrants were predominantly represented in the following economic sectors: accommodation and food services (HORECA) – 11.3%, administrative and support services – 7.6%, domestic activities – 5.9%, and constructions – 9.1%.[14]

As far as the labor market occupations are concerned, immigrants from the European Union are represented as follows: cleaning and housekeeping 11.4%, personal service workers 7.3%, care workers 5.5%, construction workers 6.1%, mining/construction/manufacturing and transport workers 6.0%, kitchen helpers 2.6%, agricultural and fishing workers 2.4%.[15]

[10] https://www.consilium.europa.eu/ro/policies/ipcr-response-to-crises/
[11] https://www.consilium.europa.eu/ro/policies/ipcr-response-to-crises/
[12] https://ec.europa.eu/eurostat/databrowser/bookmark/a2e95b51-c08a-462c-b4b7-c0a11bab32a7?lang=enidem
[13] https://ec.europa.eu/eurostat/databrowser/bookmark/a2e95b51-c08a-462c-b4b7-c0a11bab32a7?lang=en
[14] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/promoting-our-european-way-life/statistics-migration-europe
[15] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/promoting-our-european-way-life/statistics-migration-europe

Immigrants have a much higher share in some sectors than in the total labor force. In Germany and Italy, more than a quarter of the foreign workforce is employed in mining and industry. In Austria, Belgium, France, Italy and Spain, foreigners work predominantly in constructions.

Western European countries have seen a positive natural increase over the past two years, as well as a large arrival of immigrants from outside the space of the European Union. This last factor decisively explains the population growth in Spain (20%), Austria (15%), Belgium (14%), Italy (7%) and Greece (6%).

However, one country in the European Union is an exception. France, where the natural balance, favored by high fertility (18%), has contributed more to population growth than net migration[16].

Migration data can be collected in destination countries, countries of origin or upon crossing the borders. The methods used to quantify migration rely on both official (administrative) and unofficial data (interviews with different people). The definition of migration varies greatly depending on the context in which it is treated: political, economic, social, and cultural. However, regardless of the contextual approaches to migration, there are two defining variables of international migration – time and space – which all definitions of migration include.[17]

The methodology employed in our research analysis.

To better highlight the relationship between the migration phenomenon and the population of the European Union, econometric research was carried out using a simple regression model.

In this econometric research, several aspects were investigated:

To determine the relationships between the variable "total population" as the dependent variable and the variable "number of immigrants" as the independent/explanatory variable in this research.

To build a linear econometric model in order to analyze the extent to which it can answer the formulated research question.

The validation of the results through specific tests to see how well the "Output" type reports answer the formulated research question.

The prediction of the dependent variable "total population" according to the variable "number of immigrants" as an independent/explanatory variable.

The method proposed in this paper to answer the question under investigation is the analysis of a unifactorial econometric model for cross-sectional data.[18]

[16] https://adevarul.ro/stiri-externe/europa/prapastia-demografica-s-a-adancit-intre-europa.

[17] Pescaru, C. M. - *International migration and european integration – factor of globalization [Migratia internationala si integrarea europeana-factor al globalizarii].* In Rogojanu, D. C. (coord.) (2014). Regionalizare si politici regionale, Iasi, Romania: Editura Lumen, pp.189-190

[18] https://ec.europa.eu/eurostat/web/lfs/data/database

From an economic point of view, both variables, respectively the number of immigrants and the total population, can be considered as variables of interest in their quality of dependent variables. Thus, we select in our research the following variables:

Y – the dependent variable: the total population, expressed as the total population of the country concerned, measured in thousands/millions of inhabitants, i.e. the population of each Member State of the European Union.

X – the independent/explanatory variable: the number of immigrants, expressed as the sum of all immigrants in the territory of the country concerned, measured in thousands/millions of persons.

In order to build a simple linear regression model of the correlation between the number of immigrants and the total population, we will analyze the evolution of these two indicators for the twenty-seven Member States of the European Union in the year 2023, using data sets published online by Eurostat.

Table 1 - Number of immigrants and population in the twenty-seven European Union countries

Country	X - No. of immigrants	Y- Total population
Austria	105.633	9.104.772
Belgium	137.860	11.754.004
Bulgaria	29.559	6.447.710
Cyprus	23.442	920.701
Czech Republic	65.910	10.827.529
Germany	893.886	84.358.845
Denmark	64.669	5.932.654
Estonia	17.547	1.365.884
Greece	119.489	10.394.055
Spain	643.684	48.059.777
Finland	31.106	5.563.970
France	386.911	68.070.697
Croatia	26.029	3.850.894
Hungary	82.937	9.597.085
Ireland	97.712	5.194.336
Italy	332.324	58.850.717
Lithuania	28.914	2.857.279
Luxembourg	24.644	660.809
Latvia	10.909	1.883.008
Malta	26.444	542.051

Low Countries	194.306	17.811.291
Poland	214.083	36.753.736
Portugal	43.170	10.467.366
Romania	172.578	19.051.562
Sweden	132.602	10.521.556
Slovenia	28.455	2.116.792
Slovakia	7.253	5.428.792

Source: https://ec.europa.eu/eurostat/web/lfs/data/database[19]

Table no. 2 – Statistical description for the number of immigrants from the 27-EU

X - Immigration	
Mean	146002,0741
Standard Error	39704,43945
Median	65910
Mode	#N/A
Standard Deviation	206310,3192
Sample Variance	42563947817
Kurtosis	6,827248747
Skewness	2,550785862
Range	886633
Minimum	7253
Maximum	893886
Sum	3942056
Count	27
Confidence Level (95.0%)	81613,64413

Source: the authors' processing in Eviews

Table no. 3 – Statistical description for the 27-EU population in Eviews

Y- Populație	
Mean	16549057,93
Standard Error	4286514,982
Median	8858775

[19] https://ec.europa.eu/eurostat/web/lfs/data/databas

Mode	#N/A
Standard Deviation	22273385,21
Sample Variance	4,96104E+14
Kurtosis	2,769311052
Skewness	1,916182829
Range	82525654
Minimum	493559
Maximum	83019213
Sum	446824564
Count	27
Confidence Level (95.0%)	8811057,736

Source: the authors' processing in Eviews

In the tables generated above, we have the presentation of the mean of the variables, median, sum, sample number, minimum, maximum, and other statistical indicators.

Table no. 4. Anova description of the simple regression model in Eviews

	X - No. of immigrants	*Y- Total population*
X - No. of immigrants	1	
Y- Total population	0,910063977	1

Source: the authors' processing in Eviews

Source: the authors' processing in Eviews

We may notice from the chart that the distribution of the points (x_i, y_i) can be approximated very well with a straight line (trendline), so it can be assumed that the econometric model describing the relationship between the two variables is a linear model: $y = E(Y|X(= \alpha + \beta x + \varepsilon), where\ \alpha\ şi\ \beta - parameters\ of\ the\ model.$

ANOVA

	df	*SS*	*MS*	*F*
Regression	1	1,06829E+16	1,06829E+16	120,5319727
Residual	25	2,21578E+15	8,86314E+13	
Total	26	1,28987E+16		

It may be noticed that $\beta > 0$ (slope/slope of the line), which confirms the hypothesis that the two variables.

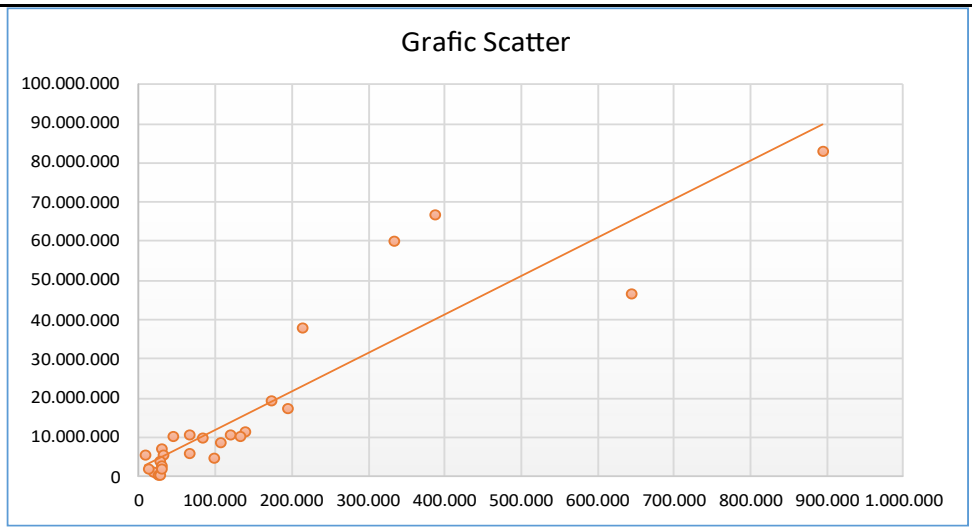
are in a direct positive relationship: increasing x leads to increasing y.

Thus, the correlation function allows us to add an extension to the Scatter plot, demonstrating how related the variables chosen in our research are to each other. The value below is a positive value, which therefore shows the close relationship between the two variables.

Chart No. 1 – Scatter between the two variables analyzed: the number of immigrants and the population in the 27-EU

ANOVA

	df	*SS*	*MS*	*F*
Regression	1	1,06829E+16	1,06829E+16	120,5319727
Residual	25	2,21578E+15	8,86314E+13	
Total	26	1,28987E+16		



Source: the authors' processing in Eviews

Table No. 5. Regression Statistics

Multiple R	0.910063977
R Square	**0.828216442**
Adjusted R Square	0.8213451
Standard Error	9414422.702
Observations	27

Source: the authors' processing in Eviews

Table No. 6. – The t-Statistic test

	Coefficients	*Standard Error*	*t Stat*
Intercept	2204200,817	2233800,689	0,986749099
X - Immigration	98,2510502	8,949238224	10,97870542

Source: the authors' processing in Eviews

For the regression to be valid, the value of R Square must be above 50%, or close to 1. In the case of this analysis, R Square has a value of 82%, significantly higher than the validity criterion, respectively very close to 1, which means that the relationship between variable X and variable Y is strongly linear.

Significance F
4.71571E-11

Significance F is below 5%, which shows us that it is valid, and that the analysis is performed correctly in our research.

From the graphical representation of the points Xi and Yi, we notice that their distribution can be approximated with a straight line. Thus, the econometric model describing the relationship between the two variables is the linear model.

Yi = α + βxi + ε_i , where α and β are the parameters of the model.

Parameters of the model:

$\sum_{i=1}^{n} yi$ = a · n + b$\sum_{i=1}^{n} xi$

$\sum_{i=1}^{n} xi \cdot yi = a\sum_{i=1}^{n} xi + b\sum_{i=1}^{n} xi^2$, where a and b are determined from the system of equatons

446.284.564 = a · 27 + b · 3.942.056

173968E+14 = a · 3.942.056 + b · 168221E+12

Based on the data generated in Excel, in order to calculate the regression, it shows us the coefficients, respectively and b.

Table No. 6 – Coefficients of a and b values for the simple regression model

	Coefficients
Intercept	2204200.817
X - Immigration	98.2510502

Source: the authors' processing in Excel

Thus, we find that the value of a is 2204200.817, and b is 98.2510502.

$\widehat{Y}_i$ = a + b_{xi} ; n= 27

The econometric analysis of these two variables used confirms the hypothesis that the total population level is explained by a direct relationship between it and the number of immigrants variable. These two variables are directly proportional, so that the increase in the number of immigrants in the European Union automatically leads to an increase in the total population of the countries concerned, which are also more involved in the migration process.

Conclusions

The phenomenon of migration involves movements of popular masses of various nationalities and ethnicities and has been known since ancient times; it has grown in importance as society has experienced economic and social development.

The concept of migration is based on the free movement of people, which is synonymous with the freedom of individuals to move from one territory to another or from one country to another. In the European Union, the free movement of workers was one of the first rights acknowledged by Community law. The European Union was founded on a philosophy of free circulation of citizens from all over the world. The Member States of the European Union have faced specific situations regarding migration, adopting over the years a few different sectoral and social policies and positions on the migration phenomenon.[20]

[20] Tudorache, Carmen (2019), Evoluția fenomenului migrației în Europa, Revista Economie Teoretică și Aplicată, București, p.99.

The 1957 Treaty of Rome, which laid the foundations of the European Economic Community, was based on certain principles, among which ensuring the free movement of people between the then six founding countries.

The Treaty on the European Economic Community stipulates the right of every citizen of the European Union to move and reside freely within the territory of the Member States, subject to the conditions laid down in the Treaty of Rome.

The underlying reasons for the migration phenomenon are economic (the improvement of living standards through increased income), social grounds (the desire to provide a better life for children, family reunification), professional goals (the desire for personal and professional career development) and political or religious rationales (the need to seek political asylum).

At the same time, labor migration provides a framework for an intercultural community, a permanent exchange of values, customs, habits, and traditions taking place between the host population and immigrants.

Also, because of the migration phenomenon, a transfer of technology and technical/scientific information between immigrants (the foreign labor force) and the host country occurs.

For the country of origin involved, labor immigration has a few major positive and negative consequences, as it brings about a series of changes in the numbers and structure of the labor force, influences consumption and investment in the economy, as well as the social behavior of household members on the labor market for the country involved.

According to the New Migration Pact that will be formally adopted in 2024 by the European Parliament, countries that are not on the external border of the European Union will have the choice between accepting refugees or paying into an EU fund for accepting immigrants to transit their country.

The main aim of the border control system will be to distinguish between those who need international protection and those who do not, i.e. social refugees who are forced to seek political asylum for political, social, or religious reasons. People whose applications for political asylum are unlikely to be granted are from countries such as India, Tunisia, and Turkey.

Immigrants from these countries can be prevented from entering the European Union and can also be detained at the border, as can people considered to be a security threat who enter the European Union with terrorist intentions.

The 2024 Pact states that migrants who stand the lowest chances of receiving asylum will be detained in centers so that they can be sent back to their country of origin or transit more quickly. This procedure will apply to people from countries where the average rate of acknowledgement of their refugee status in the EU is less than 20%. The EU is currently

facing an increase in unauthorized arrivals and asylum applications, to more than one million people by the 31st of December 2023, according to the EU Asylum Agency (EUAA).[21]

The current phenomenon of migration, whether it takes a legal or illegal form, is an extremely important problem nowadays, because of the spread of globalization of consumer markets and the dysfunctional imbalances that have arisen in the European Union labor market.Strictly economic migration has led to more medium and long-term investments, increased competition, advanced entrepreneurship, and the fulfilment of EU employment goals. However, there are, of course, as always, apart from advantages, a series of negative aspects that become more numerous with the spread of the migration phenomenon, namely a higher crime rate and more terrorism acts in various countries in the European Union.

Romania has been a country of emigration for the last thirty years, providing workforce for the countries of the European Union. A large part of Romanian emigrants has preferred to leave for work in countries from the European Union (Italy, Spain, France, Germany, UK), where there are no labor restrictions. The main reasons why Romanians emigrate are of an economic nature: an improvement in the standard of living for those who remain at home, and the possibility of finding a workplace that is in line with their professional competences and which is paid according to the amount of work that they have put in.

References

[1] *Comisia Europeană.- Uniunea Europeană - despre țări.* Retrieved from Europa., 2023. eu: https://europa.eu/european-union/about-eu/countries_ro

[2] EUROSTAT - *Eurostat - your key to European statistics. Retrieved from Eurostat - your key to European statistics. 2023:* https://ec.europa.eu/eurostat/web/main/home.

[3] BARBU, T.C, Vuță, M., Străchinaru, A.I. and Cioacă, S.I. - *An Assessment of the Immigration Impact on the International Housing Price.* Amfiteatru Economic, 19(46) - pp. 682-695, 2017

[4] ARTNER Annamaria, SOREG Krisztina - *Working Labour Market Tendencies in the European Union, in Romanian* Journal of European Affairs, Vol. 18. No.1 - June 2018.

[5] BUȚĂ, Viorel; RADU, Andreea Mihaela - *Migration of Labor Force within the European Union,* Revista Academiei de Științe ale Securității Naționale - Volume 1, p.73, 2019.

[21] https://www.bursa.ro/aderarea-la-schengen-mai-aproape-dupa-acordul-european-privind-noul-pact-pentru-migratie-si-azil

[6] POPA, Mădălina, UNGUREANU, Dragoș, ONEAȘCĂ, Iulian (coord.). - *Politica de migrație a Uniunii Europene: implicații pentru piața muncii - Buzău*, Alpha Mdn, 2013.

[7] TUDORACHE, Carmen (2019) - *Evoluția fenomenului migrației în Europa* - Revista Economie Teoretică și Aplicată, București - paper accessed on 19.03.2024.

[8] https://ec.europa.eu/eurostat/databrowser/bookmark/a2e95b51-c08a-462c-b4b7-c0a11bab32a7?lang=en -link accessed on 19.03.2024

[9] https://ec.europa.eu/eurostat/databrowser/bookmark/9d961e67-a618-4485-8353-0ec8b12808a?lang=en – link accessed on 19.03.2024

[10] http://www.mae.ro/termeni-conditii - link accessed on 19.03.2024.

[11] https://insse.ro/cms/demography-in-europe/bloc-1a.html?lang=ro – link accessed on 20.03.2024.

[12] https://www.bursa.ro/aderarea-la-schengen-mai-aproape-dupa-acordul-european-privind-noul-pact-pentru-migratie-si-azil- link accessed on 20.03.2024.

[13] http://www.euractiv.ro/uniunea-europeana - link accessed on 20.03.2024.

[14] https://www.consilium.europa.eu/ro/policies/ipcr-response-to-crises/ - link accessed on 19.03.2024.

[15] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/promoting-our-european-way-life/statistics-migration-europe - link accessed on 20.03.2024

[16] https://adevarul.ro/stiri-externe/europa/prapastia-demografica-s-a-adancit-intre-europa. – link accessed on 20.03.2024

Bibliography

ARSENE, Andreea - *Travellers into the Giant Space of Globalization – Migration Phenomenon and its Implications at the Family Level* - Revista de Administraţie Publică şi Politici Sociale Anul I, Nr. 2 / Martie 2010- Paper accessed on 25.03.2024.

BOSWELL Christina - *Addressing the causes of migratory and refugee movements: the role of the European Union* – December 2002

CHRISTL,M.; A. BÉLANGER, A. CONTE, J. MAZZA, E. NARAZANI - *Projecting the fiscal impact of immigration in the European Union, Fiscal Studies* The Journal of Applied Public Economics, Volume 44, Issue 3 (2021), https://wileyonlinelibrary.com/journal/fisc, paper accessed on 25.03.2024

CRISTIAN Elena Raluca, DĂNCIULESCU Alexandra Irina, DAN Anda Veronica, ȚÎRDĂ Carmen Dalia - *The Relationship Between the Human Development Index (HDI)*

and International Migration, Proceedings of the Journal of Information Systems & Operations Management (JISOM), Vol.15, No.2, p.46-p.60, ISSN: 1843-4711, BDI journal indexed in EBSCO, GALE Cengage Learning, Index Copernicus, ProQuest, RePec / Ideas, www.jisom.rau.ro, December (2022)

JUHYUN Lee - *A comparative study of eight European countries: how life course events affect female migrant labour market integration under the perspective of welfare and production regimes,* - Journal of International and Comparative Social Policy (2022), 38: 3, 254–274 https://doi.org/10.1017/ics.2022.17

PESCARU, C. M. - *International migration and European integration – factor of globalization [Migratia internationala si integrarea europeana-factor al globalizarii].* In Rogojanu, D. C. (coord.) (2014). Regionalizare si politici regionale, pp.187-196. Iasi, Romania: Editura Lumen. Paper accessed on 24.03.2024.

RYAKITIMBO, Crispin; HOSSAIN, Babul. - *Factors of International Migration of Haitians to the Dominican Republic in 2010-2015*. International Journal of Global Sustainability. 3.88. 10.5296 /ijgs. v3i1.15841, 2019.

SUSSMUTH, Rita. WEIDENFELD. Werner - *Challenge integration: Living together in a Europe of diversity* - 2004.

NOVARA. C., M.J. MARTOS-M´ENDEZ , L. GOMEZ-JACINTO, I. HOMBRADOS-MENDIETA , L. VARVERI, C. Polizzi - *The influence of social support on the wellbeing of immigrants residing in Italy: Sources and functions as predictive factors for life satisfaction levels, sense of community and resilience*, International Journal of Intercultural Relations 92 (2023) 101743, http://www.elsevier.com/locate/ijintrel

https://doi.org/10.1093/oxrep/grn027 - link accessed on 20.03.2024.

https://ec.europa.eu/eurostat/ark/ -link accessed on 19.03.2024.

https://www.consilium.europa.eu/ro/policies/ipcr-response-to-crises/ - link accessed on 19.03.2024.

https://ec.europa.eu/eurostat/ - link accessed on 20.03.2024.

https://www.unhcr.org/refugee-statistics/ - link accessed on 20.03.2024.

https://home-affairs.ec.europa.eu/policies/migration-and-asylum/migration-management/migration- management-welcoming-refugees-ukraine_ro - link accessed on 20.03.2024.

https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/promoting-our-european-way- life/statistics-migration-europe - link accessed on 20.03.2024.

DATA REPORTING IN THE ENERGY SYSTEM OF ROMANIA. BUILDING COMPETITIVE ADVANTAGES IN A GLOBAL ENVIRONMENT

Ana Maria IANOSI[1]

Abstract

In an ever-expanding world, the development of an organization depends on its ability to innovate and modernize. Companies gain competitive advantages over the world's best competitors thanks to pressures and challenges from the external environment. For example, the unpredictability and vagueness of the legislation that forces them to identify the necessary measures to overcome critical moments in order to reduce as much as possible the material damage, represents the ability of each organization to develop continuously, even if forced by circumstances.

In a labor market dominated by globalization and dynamism, competition has become fierce. Thus, organizations must identify the necessary tools to be competitive. Data analysis has become an indispensable tool for the organization that aims to achieve a competitive advantage in today's business environment.

So, effective data reporting is essential to maintain and improve a company's competitive position in the market. It provides the information needed to make strategic decisions, optimize operations, understand customers, and drive innovation.

Keywords: globalization, organization, analysis, digitization, reporting, energy, strategy

JEL Classification: K32, L10, M38

Introduction

Michael Porter who proposed the theory of competitive advantage in 1985 suggests that states and businesses should pursue policies that create high quality goods to sell at high prices in the market. Porter emphasizes productivity growth as the focus of national strategies.

He states that the industry is like a competition in arena for which the result is the advantage of won or lost. Firms, through competitive strategy, define and establish an approach to competition in their industry that is both profitable and sustainable. There is no one-size-

[1] PhD Student, Economics Engineer, University Politehnica of Bucharest, Faculty of Entrepreneurship, Business Engineering and Management, Romania, anamaria161996@yahoo.ro

fits-all competitive strategy, only strategies are tailored to a particular industry and to a particular firm's skills and assets.

Porter claims that there are two central concerns underlying the choice of a competitive strategy. The first is the structure of the industry in which the firm competes. Industries differ greatly in the nature of competition, thus not all industries can offer equal opportunities for sustained profitability.

The second central concern in strategy is the position in an industry. Some positions are more sustainable than others, regardless of the industry average return.

The attractiveness of the industry and the competitive position represent elements that can be shaped by a company. Successful businesses not only respond to their environment, but also try to influence it in their favor. Indeed, changes in the structure of the industry, or the emergence of new elements for competitive advantage, this is the basis for substantial changes in the competitive position [1].

Competitive analysis involves identifying direct and indirect competitors using research to reveal their strengths and weaknesses relative to those of the organization under analysis.

Competitive analysis has multiple advantages because it helps the organization to develop its long and medium-term strategies, to determine its position in the sector in which it operates, but also to identify weak points to be able to determine the necessary measures that must be taken and implemented to evolve as business.

In order to carry out a correct market research, it is necessary to collect the information that is being pursued. Thus, the data collected must be real, subjective, but also objective.

All this collected information will be found in the activity report, which is a structured and analytical document that respects precise rules.

It is a reference document that provides a broad and complete view of all the actions carried out within an institution or company during a certain period (often a quarter, a semester, a year) and presents the results obtained in relation to the established objectives in the activity program for the past year. It is the key document for making mature decisions oriented towards the success of the organization. Going beyond simple information, the activity report is, in fact, the main document that supports the quality of the decision-making process.

The activity report is a document that includes the figures and strategies developed during a certain period regarding the expected objectives.

In this sense, the activity report of an organization should be structured in such a way as to present the information as concisely as possible, but without losing its essence [2].

Data reporting in the energy system of Romania

The energy sector has a central role in Romania's economic and security policy. Through its climate and natural resources, Romania benefits from an enormous potential in terms of the development of the energy system. It is one of the most prolific sectors for investigations and the most promising for the Romanian economy [3].

Because they are an important system for the functioning of a country, but also for the good functioning of the economy, it is necessary for this industry to be monitored, to avoid market manipulation, but also unfair competition. This monitoring can be done by drawing up a report containing the information in a concise form, but without losing the essence of the subject.

Writing reports is part of every professional's job, as this process also has a strategic, not just an organizational purpose. For many professions, such as accounting, business or engineering, reporting is a core activity as it drives decision making based on documented information.

There are many different formats and styles of writing reports. Reports can vary from informative or evaluation reports to official reports addressed to the head of public institutions of a state. Each report has a specific goal, to communicate with the target audience. A good report that can be defined as an effective report, is a report that produces the expected results. If the author directly presents the objectives pursued and manages to write a well-structured document, there is a better chance that the documents will catch the attention of the reader, be understood and be successful. A report written afterwards gives the publisher and the organization a good professional image and makes them credible in front of the target audience [4].

Each industry must report on the activities carried out. Teams of professionals working together throughout the life cycle of a project must write reports on decisions, activities or results of studies, visits, discussions, and practical work, thus encouraging the communication and preservation of important information over a long period of time, according to current legislation of each state. A well-written and easy-to-read report provides an overview of the project and credibility for the reviewer.

Data reporting in the energy system of Romania aims to monitor the electricity and natural gas market, in order to evaluate the level of performance, efficiency, competition and transparency on the market, also, in order to prevent anti-competitive practices and those that may harm national security in the supply of electricity.

The objectives are:

- determining the degree of efficiency and competition on the electricity and natural gas market and identifying situations that can lead to a decrease in performance in terms of the activity of supplying electricity and natural gas to final customers;
- increasing the level of transparency, and ensuring a fair competitive environment, in order to protect end customers through correct information;

- identification of anti-competitive behaviors.

The process of evaluating the efficiency of the energy market, its competition and performance and determining the behavior of electricity suppliers on this market, is carried out by the specialized department within Romanian Energy Regulatory Authority – ANRE, through a set of well-defined indicators through a set of laws, which is based, mainly, on the collection of data, specific information that characterizes the activity of supplying electricity on the territory of Romania under a competitive regime to final customers [5].

Good practices for a competitive framework in the energy sector in Romania

Competitiveness, although it is frequently used, is still an ambiguous notion, in the sense that there is no unanimously accepted definition. At the microeconomic level, it refers to a firm's ability to compete, grow and be profitable.

In a broad sense, it is the ability of an organization/ enterprise to develop a competitive advantage over the competition. Thus, this concept remains a relative one, according to the definitions of the specialized literature [6].

Competitive analysis involves identifying direct and indirect competitors using research to reveal their strengths and weaknesses relative to those of the organization under analysis.

Competitive analysis has multiple advantages because it helps the organization to develop its long-term and medium-term strategies, to determine its position in the sector in which it operates, but also to identify weak points in order to be able to determine the necessary measures that must be taken and implemented to evolve as business.

In order to carry out a correct market research, it is necessary to collect the information that is being pursued. Thus, the data collected must be real, subjective, but also objective.

In the energy sector, data collection is done under the careful supervision of Romanian Energy Regulatory Authority – ANRE, respecting a well-defined legislative framework, in order to maintain a transparent and fair environment for all market participants.

The activity of monitoring the energy and natural gas market is carried out based on the provisions of the Electricity and Natural Gas Law no. 123/2012 with subsequent amendments and additions, as well as by the Commission Implementing Regulation (EU) no. 1348/2014 of 17 December 2014 on data reporting in order to maintain the integrity and transparency of the wholesale energy market.

All the regulations in force both at the national and international level are aimed at discouraging market abuse, encouraging a transparent framework that seeks to determine a set of indicators that are clear and well defined for all participants in the energy market.

Thus, the entire activity carried out on the territory of a state must be reported, and this process is carried out through the platforms provided by the European stock exchanges that hold an „*Organised Market Places*" license [7].

The transactions are reported at the level of Romania both to the Agency for the Cooperation of Energy Regulators – ACER, but also to the Romanian Energy Regulatory Authority – ANRE and are based on a set of international and national regulations. The obligation to submit these reports belongs with the REMIT department (the acronym comes from the "Regulation on Wholesale Energy Market Integrity and Transparency) within each organization that provides energy services, which must ensure that the reports have been submitted correctly and without errors [8].

On the territory of Romania, each stock market product traded on the basis of the standard contract established by the Romanian Commodities Exchange – BRM and by Romanian gas and electricity market operator – OPCOM, the European stock exchanges that hold an "Organized Market Places" license, according to the regulations in force, has a name and a specific coding that will be mentioned in the trading report. All these codifications have the role of preventing unfair competition and the development of the energy market at the national and international level, so that there are as many participants as possible in the energy market, and that demand and supply are as diversified as possible.

Example of transaction code for natural gas and electricity:

Transaction code for natural gas	**Description**
BRMM_month-aaaa_xxxx (the name of that month)	BRM – Romanian Commodities Exchange MM – month Month – the name of the month (January, February, and so on) aaaa – the year (example: 2025, 2025, and so on) xxxx – the identification number assigned to company "Z", being the verification key

Table 1 Transaction code for natural gas[2]

Transaction code for electricity	**Description**

[2] Apendix 1, Procedure for the organization and functioning of the market for medium and long-term products administered by the company Romanian Commodities Exchange, pp. 21

Monthly Delivery: PCSU_B_L PCSU_V1_L PCSU_V2_L	PCSU – the centralized market for universal service B – energy consumption in the band V – peak energy consumption L – month

Table 2. Transaction code for electricity[3]

These codifications are based on clear procedures, prepared under the careful supervision of ANRE - the National Energy Regulatory Authority, as follows:

Procedure for the organization and functioning of the market for medium - term and long - term products administered by the company Romanian Commodities (for Natural Gas) [9].

The operational procedure for conducting auctions on the centralized market for the universal service administered by Romanian gas and electricity market operator (for electricity [10].

These operational procedures, aim is to ensure a high degree of competitiveness, greater customer orientation, flexibility and non-discrimination on the European electricity market, with supply prices based on demand and supply.

In this sense, they strengthen and expand the rights of customers and energy communities, contribute to the fight against energy poverty, clarify the roles and responsibilities of market participants and regulatory authorities, and address the issue of security of supply of electricity, gas and oil, as well as the development of trans -European networks for the transport of electricity and gas.

Also, the Operators on the centralized markets are obliged at the end of the trading day to send a report in electronic format to Romanian Energy Regulatory Authority – ANRE. This report will be according to the table below, in editable format (excel), and will be sent no later than the 20th of the current month "L" for the previous month "L-1". As can be seen below, the form of an electronic report contains the following information:

Auction session date
The name of the standard product offered
The name of the holder of the tender obligation
The quantity offered (MWh)
Offered price (MWh)
The quantity traded (MWh)
Traded price (MWh)

[3] The operational procedure for conducting auctions on the centralized market for the universal service administered by Romanian gas and electricity market operator, pp. 8

Although it could be stated that this information is minimal, for the organizations responsible for monitoring the activity, the information is sufficient to identify if there are attempts to manipulate the energy market.

Auction session date	The name of the standard product offered	The name of the holder of the tender obligation	The quantity offered (MWh)	Offered price (MWh)	The quantity traded (MWh)	Traded price (MWh)
Total/Product						

Table 3. Offer obligation report

It can be stated that in order to be able to monitor the reporting process of the activities carried out in each field of activity, the structure of the reports must be similar for each industry, so that the centralization and analysis of the data to be an easier process. The creation of layouts/ models, which could be made available to each institution that is obliged to report information about the activity it has, would contribute to a well-defined and organized framework. Therefore, both parties involved in the reporting process would save time, a valuable asset these days [11].

Improving the reporting process to the competent authorities

The purpose of the research was to identify the opinion of adults from various sectors of activity regarding the efficiency of the reporting process, but also to identify solutions for a functional system at a high level, according to European requirements. In order for the activities carried out on the territory of a state to be successful in such a changing environment, it is necessary to implement some quality services and, at the same time, to hire a competent staff, who can cope with all the changes and meet the demands of the customers.

Also, in order to be able to determine the necessary measures to improve the reporting process, we must first draw the objectives and establish the assumptions of the problem.

The objectives of the research are:

Objective 1: Determining the problems encountered by staff from various fields of activity in the relationship with the state when certain reports are requested regarding the activity carried out within the company in which they work;

Objective 2: Determining the methods/ measures needed to improve the reporting process;

Objective 3: Determining the impact of periodic reporting to the competent institutions, for the implementation of regulations to protect small entrepreneurs.

The research hypotheses are:

Hypothesis 1: We assume that the level of staff satisfaction increases when the statute offers quality services and well-trained staff to be in the reporting process, thus reducing the problems they still face;

Hypothesis 2: We assume that the implementation of a professional training program could lead to the improvement of the reporting process. Also, simplifying the reporting process and eliminating red tape can be big steps towards evolution;

Hypothesis 3: We assume that the digitization of the reporting process could lead to periodic regulations, so that there is a legal framework that offers equal opportunities to all entrepreneurs.

The research was carried out to see what the problems people are face most often when they have to make periodic reports to the competent authorities, in order to find the best measures to improve the reporting process. Thus, for this research, the analysis method was "Survey based on questionnaire", through which the hypotheses established at the beginning of the study could be confirmed or denied.

Analysis method: Survey based on questionnaire.

Target group: 80 people from all fields of activity.

Statistical survey results:

Gender: 66.25% women, 33.75% men
Age: 60% of people who took part in this survey are over 30 years old, 15% are between 31 and 40 years old, 14% between 31 and 40 years old and only 8.8% over 51 years old.
Professional category: medical (15%), economic (27.5%), technical (20%), IT (10%), administrative (13.8%), and other professional categories (13.7%).
Education: high school (3.75%), university (61.25%), postgraduate (35%).
Starting from hypothesis 1, in which we assume that the level of staff satisfaction increases when the state offers them quality services and well-trained staff to support them in the reporting process, thus reducing the problems they face, we can state that this hypothesis it is confirmed given the answers received from the interviewed persons. They stated that the lack of communication is the biggest problem in the relationship with state organizations. In addition to the previously mentioned major problem, others such as:

- The bureaucratic system;
- Lack of digitization;
- Lack of transparency and cooperation.

To eliminate these problems that we all find ourselves in every day, there would be solutions, but the costs are high and the implementation time is long.

The next hypothesis assumed the implementation of a professional training program that could lead to the improvement of the reporting process. Just like the previous hypothesis, this one was confirmed by the answers sent by the people participating in the investigation.

The solutions identified to improve the reporting process were the following:

Digitization and regulation of some situations by means of clear laws that leave no room for interpretation;
Hiring young and well-trained staff;
Trainings for employees;
The existence of online platforms for data processing;
Clear presentation of the requirements for the content of the report;
Drafting of clear procedures;
Using applications to fill in the information required in reports and create a general report template, respectively
Elimination of bureaucracy.

As we can see, in order to develop a system of maximum efficiency and sustainability, innovation must be supported by a well-trained staff, the digitization of reporting activities, but also the existence of well-defined procedures.

In addition to the two hypotheses, followed the hypothesis that concerned the digitization process through which we assumed that the digitization of the reporting process could lead to periodic regulations, so that there is a legal framework that offers equal opportunities to all entrepreneurs.

So, statistically, according to the results, 83.75% agree that the periodic reporting of data to state institutions helps to implement some regulations, so that the state also protects small entrepreneurs, and 16.25% disagree, being divided into categories: 5% consider that reporting does not help to implement some regulations, and 8.75% do not think that it helps and only 2% consider this aspect irrelevant.

This statistical survey sought to determine the best measures to improve the reporting process.

Through this research, it was possible to identify the situations faced by the employees in the relationship with the state, but also the necessary measures that could lead to better collaboration.

The first objective was to determine the problems faced by staff from various fields of activity in the relationship with the state when certain reports are requested regarding the activity carried out within the company where they work.

It could be observed from the statistical survey that the lack of communication is the biggest problem faced by the staff involved in the reporting activities. This lack of communication

also occurs because of poor professional training, an outdated system, but also because of bureaucracy.

Thus, we can say that the level of satisfaction of the staff responsible for reporting increases when the state offers them quality services and well-trained staff to support them in the reporting process, thus reducing the problems they still face.

The second objective aimed to determine the methods/measures needed to improve the reporting process. The answers received from the respondents were mostly related to the implementation of training programs, but also to the simplification of the reporting process and elimination of bureaucracy.

The last objective was related to determining the impact of periodic reporting to the competent institutions, for the implementation of regulations to protect small entrepreneurs.

Thus, it can be concluded from the answers collected that the digitization of the reporting process could lead to periodic regulations, so that there is a legal framework that offers equal opportunities to all entrepreneurs.

Conclusions

Competitiveness is a broad concept, which cannot be concretely defined, but which requires companies to continuously innovate, adapt and differentiate themselves in order to be successful.

In a constantly expanding world, the development of an organization depends on its ability to innovate and modernize, to surpass itself. Companies gain competitive advantages over the world's best competitors due to pressures and challenges from the external environment.

Thus, we can conclude the following: the concept of competitiveness associated with an organization must suggest efficiency, adaptability, productivity, trust and safety, but also transparency.

Transparency strengthens ethical conduct and integrity by promoting openness, honesty and accountability in organizational practices and decision-making. By establishing a culture of monitoring and transparency, organizations can increase their resilience, sustainability and long-term success.

By reporting the data, the efficiency of the activities is achieved. Thus, data collection, grouping, monitoring enable analytical processes of abstraction, calculation, modeling and classification of information and knowledge. At the same time, the collected data also ensures a fair competition of activities not only in the energy system, but in all fields of activity, thanks to the monitoring process that ensures transparency and integrity at the level of each state.

It is important to know what the problems are in the process of reporting information, because reporting is the activity by which both participants and consumers of products and services are assured that they can trust the integrity of the markets, thus there is a balanced and competitive interaction between demand and supply, not being able to make profits through market abuse.

Considering the results from the statistical survey, it can be stated that a real problem is the lack of communication which significantly affects the reporting process, as much as bureaucracy.

This lack of communication comes from a lack of professional training, so the first step to improve the reporting process must be the implementation of training, because people are the most important asset in a company.

Another problem identified in this research was related to the lack of digitization of the system.

Access to and use of cutting-edge digital technologies can drive efficiency in any field of activity, but progress in integrating technology into daily activities is slow, partly because of very high costs.

In today's society, information is an important asset that gives "power" to its owner, whether we are talking about written or oral information. Therefore, more and more organizations have been looking for alternatives to protect their information. Although the information itself may not be very informative, it forms the basis of all reporting and is a key element of the business environment.

References:

[1] Porter M. - *The competitive advantage of nation* - Free Press. ISBN 0-684-84147-9. pp 1 – 56. 1990

[2] https://www.enssib.fr/bibliotheque-numerique/documents/66455-produire-un-rapport-d-activite-pourquoi-comment.pdf - Thomas V. - Produire un rapport d'activité; pp 11-13. 15.05.2024

[3] https://www.ccifer.ro/actualites/n/news/le-marche-de-lenergie-en-roumanie.html - The Energy Market in Romania, CCI Française en Roumanie. 15.05.2024

[4] https://wedc-knowledge.lboro.ac.uk/resources/booklets/G009FR-La-redaction-d-un-rapport-le-guide-online.pdf - Reed B. - La redaction d'un rapport le guide online, Loughborough University. ISBN 978 1 911252 00 9. 2016. 15.05.2024

[5] https://anre.ro/despre/statutul-si-rolul/ - ANRE: Statutul si rolul. 15.05.2024

[6] https://ibn.idsi.md/sites/default/files/imag_file/148_152_Competitive%20potential%20of%20the%20enterprises.pdf – Isachi E. - Competitive potential of the enterprises.15.05.2024

[7] https://legislatie.just.ro/Public/DetaliiDocument/139677 - Legea energiei electrice si a gazelor naturale nr. 123/2012. 15.05.2024

[8] https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A32014R1348 - Regulamentul de punere in aplicare (UE) nr. 1348/2014 al Comisiei din 17 decembrie 2014 privind raportarea de date, pentru punerea in aplicare a articolului 8 alineatele (2) si (6) din Regulamentul (UE) nr. 1227/2011 al Parlamentului European si al Consiliului privind integritatea si transparenta pietei angro de energie Text cu relevanta pentru SEE. 15.05.2024

[9] https://www.brm.ro/wp-content/uploads/2020/09/Procedura-de-tranzactionare-a-produselor-pe-termen-mediu-si-lung-11.09.2020-1.pdf - Procedura de organizare si functionare a pietei produselor standardizate pe termen mediu și lung administrata de societate. 15.05.2024

[10] https://www.opcom.ro/uploads/doc/PCCB/GP/Procedura%20PCSU_02.2018_obs_clean.pdf. – OPCOM - Procedura operationala de desfasurare a licitatiilor, 15.05.2024

[11] https://lege5.ro/Gratuit/gm3tqojwgeya/ordinul-nr-143-2020-privind-obligatia-de-a-oferta-gaze-natural - ANRE - Ordinul nr. 143/2020 privind obligatia de a oferta gaze naturale pe pietele centralizate a producatorilor de gaze naturale a caror productie anuala realizata in anul anterior depaseste 3.000.000 MWh. 15.05.2024

Bibliography

Porter M.*,The competitive advantage of nation - Free Press* ISBN 0-684-84147-9. New York, pp 1 – 56, 1990

https://anre.ro/despre/statutul-si-rolul/ - ANRE: Statutul si rolul. 15.05.2024

https://www.brm.ro/wp-content/uploads/2020/09/Procedura-de-tranzactionare-a-produselor-pe-termen-mediu-si-lung-11.09.2020-1.pdf - Procedura de organizare si functionare a pietei produselor standardizate pe termen mediu și lung administrata de societate. 15.05.2024

https://www.ccifer.ro/actualites/n/news/le-marche-de-lenergie-en-roumanie.html - The Energy Market in Romania, CCI Française en Roumanie. 15.05.2024

https://ibn.idsi.md/sites/default/files/imag_file/148_152_Competitive%20potential%20of%20the%20enterprises.pdf - Competitive potential of the enterprises. 15.05.2024

https://lege5.ro/Gratuit/gm3tqojwgeya/ordinul-nr-143-2020-privind-obligatia-de-a-oferta-gaze-natural - ANRE: Ordinul nr. 143/2020 privind obligatia de a oferta gaze naturale pe pietele centralizate a producatorilor de gaze naturale a caror productie anuala realizata in anul anterior depaseste 3.000.000 MWh. 15.05.2024

https://legislatie.just.ro/Public/DetaliiDocument/139677 - Legea energiei electrice si a gazelor naturale nr. 123/2012. 15.05.2024

https://www.enssib.fr/bibliotheque-numerique/documents/66455-produire-un-rapport-d-activite-pourquoi-comment.pdf - Produire un rapport d'activité, pp 11-13. 15.05.2024

https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX%3A32014R1348- Regulamentul de punere in aplicare (UE) nr. 1348/2014 al Comisiei din 17 decembrie 2014 privind raportarea de date, pentru punerea in aplicare a articolului 8 alineatele (2) si (6) din Regulamentul (UE) nr. 1227/2011 al Parlamentului European si al Consiliului privind integritatea si transparenta pietei angro de energie Text cu relevanta pentru SEE. 15.05.2024

https://www.opcom.ro/uploads/doc/PCCB/GP/Procedura%20PCSU_02.2018_obs_clean.pdf. - Procedura operationala de desfasurare a licitatiilor. 15.05.2024

https://wedc-knowledge.lboro.ac.uk/resources/booklets/G009FR-La-redaction-d-un-rapport-le-guide-online.pdf - La redaction d'un rapport le guide online, WEDC, Loughborough University, ISBN 978 1 911252 00 9, Loughborough, 2016. 15.05.2024

REVOLUTIONIZING ARCHITECTURE: THE INTEGRATION OF 3D PRINTING TECHNOLOGY, VR EXPERIENCES, AIA AND VIDEO GAMES IN ARCHITECTURE

Daria Ștefania ISTRATE[1]

Ana Mihaela ISTRATE[2]

Abstract

STEM fields, comprising Science, Technology, Engineering and Mathematics have been considered for many decades the core preoccupation of male specialists. Yet, beginning with the Second World War, when women started working in military facilities and factories, and took over lots of the masculine responsibilities, but also as a result of the publicity that many women in STEM received due to the simplification of access to communication, these disciplines have increased in interest also because they are associated with the level of inventiveness of a society, but more importantly, with the level of growth in social and economic fields.

Once big academic institutions have included in their curriculum gender studies and entrepreneurship for women, more and more girls have started considering the field of STEM as a possible future career. In the case of Romania, career path for women in STEM has been subjected to stereotyping, caused by the cultural and political values that Romania had before the fall of Communism. But today, due to access to information, access to international systems of education and scholarship programs dedicated to women in STEM, more and more girls are engaged in the field, with spectacular results.

The present study is an analysis of the most recent statistical data related to women's engagement in the field of STEM, certain comparisons with other countries from the Central and Eastern European region and in the final part offers a set of predictions about the future of girls in STEM.

Keywords: STEM education, gender studies, entrepreneurship for women, Hofstede's Dimensions of Culture, role models for girls, gender disparities

JEL Classification: I24, J16

[1] Student, "Mihai Viteazul" National College, Ploiesti, Romania, stefy.istrate@gmail.com

[2] PhD, associate professor, Romanian-American University, Bucharest, Romania, ana.mihaela.istrate@rau.ro

1. Introduction

Today STEM education represents an important subject for societies, all over the world. Yet we have to admit that women are still underrepresented in many fields of activity, UNESCO suggesting that only 28% of women represent the percentage of the world researchers. Thus, the problem of education in STEM is the primary cause of this misrepresentation.

On the other hand, many cultures of the world are still putting a lot of pressure on women's role in the STEM field, as they are the subject of stereotyping, mockery and even bullying. Role models are still few, and they have a very small impact on the aspirations of young women. The only environment where changes have been made are in the field off entrepreneurship, in business field women being more and more appreciated for their values and there are efforts to put a more equal sign between male and female revenues, in order to achieve a more equal gender distribution of incomes.

There are different initiatives made by governments, non-governmental institutions, and entities at different levels, promoting girls and women in STEM. We can give the example of a robotics competition, that was developed in the United States of America twenty-five years ago, First Tech Challenge, that promoted the inclusion of girls in the field of STEM, and more specifically robotics, but there are also different educational programs like #Women in STEM, #Girls Who Code, #Black Girls Code, etc., that are all aimed at the same outcome: to encourage girls to pursue careers in the field of STEM.

The present article starts with a brief cultural overview of the most important causes of the gender discrepancies between cultures. It relies on Geert Hofstede's Country Comparison Tool, an instrument for the analysis of differences and similarities between cultures, according to the six indexes that the Dutch sociologist developed in the 1970s: Power Distance Index, Individualism vs. Collectivism Index, Motivation towards Achievement and Success, Long Term orientation and Indulgence vs. Restraint.

It then continues with a set of statistical data related to the number of girls and women in STEM fields, at European level, offered by the European Institute for Gender Equality, explaining the position that Romania has in this respect, while in the final part it offers a set of possible paths to follow for the younger members of the generation that is preparing now to pursue a career in the field of STEM.

2. Cultural overview on gender discrepancies

According to Geert Hoftsede's Dimensions of Culture, when comparing Romania, the second lowest European country on gender equality and Netherlands, which is the most gender equal country in Europe, we realize the magnitude of the difference. [1]

Hofstede's Dimensions of Culture represent a framework that the Dutch sociologist developed in order to better understand the differences that exist between cultures, with the particular focus on the labor force all over the world, but also help us understand our own behaviors, beliefs and values, that can be exploited in a positive way in the work environment.

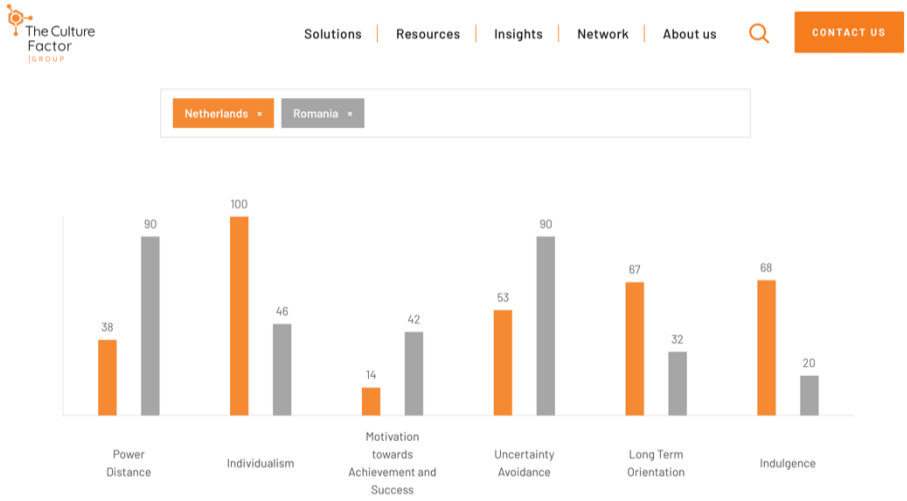


Figure 1. Hofstede's Country Comparison Tool – Netherlands vs. Romania[3]

Romania scores 90 on the Power Distance Index, compared to the low score of 38 in Netherlands. This means that in Romania hierarchies continue to be important, as they are the direct result of a totalitarian political system that lasted in Romania after the Second World War and until the fall of communism in 1989.

By comparison, Dutch employees are independent, hierarchy being perceived only as an element of convenience, and leaders are seen more as role models and coaches for their followers. Power is not centralized, as Dutch culture nurtures egalitarianism and does not agree with the concept of authority.

Egalitarianism is a concept in political philosophy, and in Netherlands it is deeply rooted in different cultural, social, and political factors. First of all, the historical background, the Dutch Revolt against Spanish ruling of the 16th century, which conducted to the establishment of the Dutch Republic, a period fostering independence, civic responsibility and a culture of tolerance. One of the cultural factors that favored Dutch egalitarianism was the welfare state, providing social security, healthcare and unemployment benefits, which all resulted in a reduction of economic disparities. [2]

With a mixed economy, blending free market capitalism and strong labor unions, fair wages are ensured, which promoted economic equality. Last but not least, the fact that the

[3] https://www.hofstede-insights.com/country-comparison-tool?countries=netherlands%2Cromania

Netherlands. Was one of the first members of the European Union since its inception, offered a set of regulations and policies that promoted social and economic cohesion.

Compared to Netherlands, Romania is still in a transition period, lasting for more than thirty-five years, after a long period of feudalism, followed by a period of communist totalitarianism, that had deeply influenced its social hierarchies and collective psyche. The transition to the market economy after 1989, brought a dramatic change, as well as social and economic challenges, contributing to the social stratification. Today we have huge gaps between the rich and the poor, with very thin middle class layer representation.

Romania's workplace is still hierarchical, with a clear distinction between management and employees. Traditional gender roles are more pronounced, although there are strong efforts to promote gender equality. Yet, in the system of education one can still find middle aged teachers in the class admitting gender inequality and accepting it per se.

In point of Individualism versus Collectivism, Romania and Netherlands exhibit again a huge difference. With a score of 100, the maximum, Netherlands is one of the most individualist societies in the world, with a high preference for distant social ties, where individuals take care of themselves and their immediate families. By comparison, Romania scores 46, which makes it a collectivistic country, with long term commitment to the group (family, extended family, and friends). Loyalty to the group is one of the most appreciated traits, overriding any other rule or regulation. In a collectivist society, all members of the group are expected to take responsibility, and at entrepreneurial level management is the management of the group.

On the Motivation towards achievement and success, Netherlands scores extremely low, just 14, while Romania scores 42, which means it is a "relatively consensus society". While in Netherlands the work-life balance is extremely important, in Romania we value the concept of *work in order to live*, with higher levels of solidarity and a value for equality, yet not as high as Netherlands exhibit.

With a high score of 90, on the Uncertainty Avoidance Index, compared to 53 in Netherlands, Romania maintains rigid codes of belief and behavior and is extremely intolerant to unusual ideas and beliefs. There is an emotional need for rules, time is perceived as money, punctuality is strictly appreciated, particularly at entrepreneurial level and security is highly appreciated as an element of individual motivation.

Romania scores 32 on the Long Term Orientation Index, compared to 67 in Netherlands, which means there is a difference between Romanian normative culture and the Dutch pragmatic nature. Romanians are deeply rooted in traditions and focus mainly on short term achievements, while in the Netherlands people easily adapt to changing conditions, thriftiness, and perseverance in achieving results being the main traits of the Dutch labor force.

Last but not least, on the Indulgence versus Restraint Index, Romania scores 20, while Netherlands has a score of 68, which means Romania is a culture of restraints, cynicism and pessimism being the main traits, versus joy of life and fun in the Dutch culture. Romania does not put very much emphasis on the value of leisure time, and people have the feeling that indulging themselves might be wrong. By comparison, Dutch people possess a positive attitude to life, with a tendency towards optimism and willingness to realize their desires.

This brief cultural analysis can help us understand different aspects related to the entrepreneurial environment of Romania, which is deeply rooted in totalitarianism, is based on cynicism and is a pessimistic society, where the role of women in top management positions is still not completely appreciated and valued.

3. An Overview of Gender Differences

Statistically speaking, 37,4% of all STEM (Science, Technology, Engineering and Mathematics) students will pursue an entrepreneurial career path after graduation, with more male students (61,38%) of all graduates focusing on developing their own start-up business, compared to 17,6% of female graduates that consider starting a business after graduation.

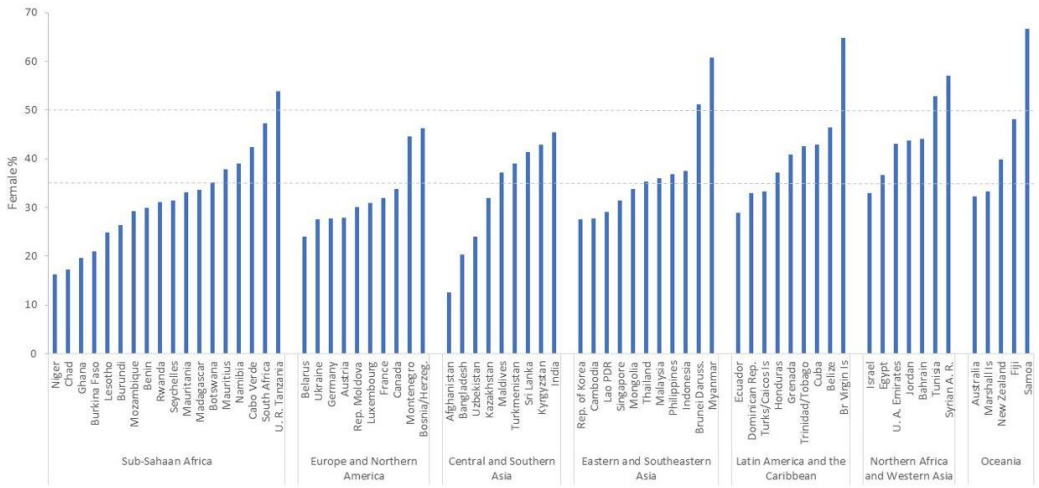


Figure 2. Share of STEM female graduates[4]

But the analysis of the gender differences within STEM fields highlighted a higher intention score for women who pan on becoming entrepreneurs, compared to men. “However, they present a negative score in human medicine, health, mathematics, and natural sciences. This

[4] https://world-education-blog.org/2024/04/25/new-uis-data-show-that-the-share-of-women-in-stem-graduates-stagnant-for-10-years/

result shows that women prefer to work in established organizations rather than become a traditional entrepreneur and start a venture". [3]

In the particular case of women, the entrepreneurial career results from the analysis of a series of factors modeling the intention, based on the theories of Ajzen (1991) and Linan&Chen (2009) which suggests that the career in STEM is impacted by the following factors:

Entrepreneurial self-efficacy
University climate
Social context
Entrepreneurial learning

The academic system is aware of the importance of the role of women in business field and that is why there are different support systems to accommodate and facilitate their engagement in all aspects of the business life. It also depends very much on the culture nurturing this environment. For example, in Romania, where the present study takes place, we have the highest number of female graduates in the field of STEM: natural sciences, mathematics, engineering, information technology, constructions, etc.

In 2021, in the European Union, higher education female graduates (level 5-8 according to the International Standard Classification of Education – ISCED) in the field of STEM represented 32,8% of all graduates, 0,3% more than the previous year (32,5%).

But of all STEM field female graduates, most of them come from Romania, with 42,5%, followed by Poland with 41,5%, Greece 40,09% and Italy with 339%.

At the opposite side of the scale, countries with the lowest representation of STEM female graduates in European Union are Belgium 27,4%, followed by Spain and Germany with 27,7% and Austria 26%.

Data from Romania places our country at the forefront of the percentage of female graduates in the field of STEM, compared to the overall number of higher education graduate sin these fields. These results prove a significant progress in engaging women in careers in the field of STEM. Yet, we as a society, need to continue to engage women more in efforts of ensuring a balanced representation in STEM fields, but also to harness the potential of diversity in the scientific and technological progress of the country.

Claudine Schmuck, in her book *Women in STEM Disciplines*, includes Central and Eastern Europe in the same poll of evolution with East Asia, highlighting the dramatic increase in the number of women graduating from STEM fields, which is 80% higher today than 10 years ago. "The fact that the perimeter of analysis used is not constant doesn't allow to consider in itself this growth as significant since it also reflects the fact that data is available for a larger number of countries, but it is interesting to observe that during the same period of time the total headcount of STEM graduates (men and women) has increased more slowly than that of female graduates (average 60% growth to compare with 80 %). Thus,

women share of all STEM graduates has increased from 43 to 48% (when health is included in the definition), but only from 30 to 34% when only EMC and science are included. Thus, the gender divide is not decreasing significantly in STEMS (EMC and science)". [4]

4. Gender Discrepancies in Romania and How Role Models Overcome the Barriers

The growing number of women who work in STEM fields, assure sustainable and equitable growth patterns in entrepreneurship and social entrepreneurship. These topics have been taken up in research, because of their importance in the employment of women across the world. Women nowadays show more interest in social entrepreneurship than men, which creates growth of positive social and environmental externalities. Overall, promoting female entrepreneurship creates an inclusive and diverse path, which others can approach by their example.

The gender gap still remains large in most economies, despite the potential benefits. This gap persists worldwide, with women owning one in three businesses, both in developed and developing economies. A fact that everybody knows is that women are misrepresented in economies with less income, where they do not have as many chances as men do to start a new business.

Research performed at international level has shown that the barriers women have to face in order to achieve their goals in entrepreneurship are two time more difficult than in the case of men. There are several obstacles that men do not have, and that is the reason why gender gaps in entrepreneurship form. Women have less access to capital, business opportunities, networking and are less likely to start their own business. But women cross these barriers, and one great example is the new concept, called "*mompreneurship*", which means entrepreneurship done by mothers, which shows women's capability to handle critical situations and also take care of their families. The solutions to this problem is to identify the key barriers and to promote female entrepreneurship until governments introduce policies which reduce these gender gaps.

This new concept is relevant for Romanian women, where constraints such as taking care of children or elderly prevent them from managing their own businesses. It is very important to develop an understanding of their experiences and realities, in order to lead by example and grow the numbers of women involved in business.

In the particular case of Romania, totalitarian way of thinking still persists in the mentality that women have to choose between family or career, or at least to make some sacrifices and choose a path that is more in favor of family responsibility, children's education and caregiving.

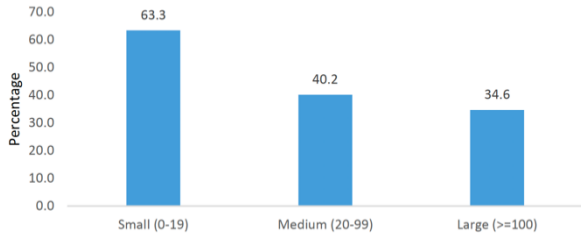


Figure 3. Percentage of female owners by company size[5]

As it can be seen from the statistics, there are fewer representatives of the big companies in Romania, compared to the small enterprises. That is why, especially for the young generation, role models are extremely important. In this respect we can mention a couple of successful women from different fields, bankers, mathematicians, economists, etc, that lead the next generation of female entrepreneurs:

Anca Vlad is the founding president of the Fildas-Catena Group, which includes one of the biggest chains of pharmacies in Romania. She is one of the most influential Romanian businesswomen, with vast experience in the pharmaceutical industry.

Veronica Savanciuc is the co-founding president of the Lowe Group Romania, one of the most important advertising agencies. Veronica is a prominent figure in the advertising and communication industry.

Mirela Iordan is the general director of Pfizer Romania, with an impressive career in the pharmaceutical industry, leading the operations of one of the biggest pharmaceutical companies in the Romanian market.

Cristina Bâtlan is known as the founder and owner of Musette, one of the most well-known footwear and fashion accessories brands in Romania, which has even managed to extend internationally.

Rucsandra Hurezeanu, known as the CEO of Ivatherm, is a doctor and entrepreneur in the field of dermato-cosmetics. Ivatherm is the first Romanian brand of dermato-cosmetics, which uses thermal water from Herculane.

Măriuca Talpeș, co-founder of Bitdefender, together with her husband, is one of the most successful global cyber security companies worldwide. She is a mathematician and CEO of a company that is engaged in the process of digitalization of education and production of modern interactive textbooks for public and private schools in Romania.

[5] World Bank Enterprise Survey 2018–2020, https://login.enterprisesurveys.org/content/sites/financeandprivatesector/en/library.html.

5. Conclusions

The world in which we live, which is becoming more and more digitalized and where sooner or later human activity will be replaced by artificial intelligence, still needs the direct input of specialists in the field of STEM, both men and women.

The gender gap existing in Romania between men and women in STEM fields, which has a direct impact on the number of women engaged in entrepreneurial activities, can be overcome by implementing different strategies to nurture a supportive environment. By developing targeted governmental policies to promote gender equality, like scholarships for girls in STEM, an inclusive curricula and anti-discrimination laws as well as developing mentorship programs in which successful businesswomen and generational leaders are involved, can be a solution towards an appropriate guidance for the younger generations. Yet, stereotypes still persist in the society and probably there is needed one more generation of successful girls and women in the field of STEM that need to provide outstanding success in order to change the mentalities.

References

[1] https://plato.stanford.edu/entries/egalitarianism/#PreDis – Definition of Egalitarianism. 30.05.2024

[2] https://www.hofstede-insights.com/country-comparison-tool?countries=netherlands%2Cromania – Netherlands vs. Romania Country Comparison Tool. 30.05.2024

[3] MARI, M., POGGESI, S., FOSS, L. – *Women's Entrepreneurship in STEM Disciplines. Issues and Perspectives* – ISBN 1431-1941, Springer Nature Switzerland. 2021. [13]

[4] SCHMUCK C. – *Women in STEM Disciplines* - ISBN 978-3-319-41648-8. Springer International Publishing. 2017 [14]

Bibliography

Author A-1, Author A-2, […] - *Paper Title (italic)* - Journal ISBN/ISSN. [page AA-BB]. [month-date] Year

[1] CLARK, J. – *Closing the Achievement Gap from an International Perspective. Transforming STEM for Effective Education* – ISBN 978-94-007-4356-4, Springer. 2014

MARI, M., POGGESI, S., FOSS, L. – *Women's Entrepreneurship in STEM Disciplines. Issues and Perspectives* – ISBN 1431-1941, Springer Nature Switzerland. 2021.

[2] ROBAYO-ABRIL, M, RUDE, B. – *The Gender Gap in Entrepreneurship in Romania. Background Study for the Romania Gender Assessment*, The World Bank Group, 2023

[3] SCHMUCK C. – *Women in STEM Disciplines* - ISBN 978-3-319-41648-8. Springer International Publishing. 2017

SCHNITTKR, J. – *Career Paths of Women Entrepreneurs in STEM – Challenges and*

SWE KINE, M., AREEPATTAMANNIL, S. – *STEAM Education. Theory and Practice* – ISBN – 979-3-030-04002-4. Springer Nature Switzerland. 2019

WOLVERTON, A., NAGAOKA, L., WOLVERTON, M. – *Breakin In. Women's Accounts of How Choices Shape STEM Careers* – ISBN 978-1-57922-428-8. Stylus Publishing. 2015

https://www.efainlacluj.ro/romania-are-cele-mai-multe-absolvente-din-ue-in-domeniile-stem-stiinte-naturale-matematica-inginerie-etc/. 28.05.2024

https://world-education-blog.org/2024/04/25/new-uis-data-show-that-the-share-of-women-in-stem-graduates-stagnant-for-10-years/. 28.05.2024

https://www.hofstede-insights.com/country-comparison-tool?countries=netherlands%2Cromania – Netherlands vs. Romania Country Comparison Tool. 30.05.2024

https://plato.stanford.edu/entries/egalitarianism/#PreDis – Definition of Egalitarianism. 30.05.2024

https://www.cnr-unesco.ro/ro/activitate/masa-rotunda-femei-in-stiinta-cercetare-dezvoltare-si-impact-societal. 30.05.2024

https://world-education-blog.org/2024/04/25/new-uis-data-show-that-the-share-of-women-in-stem-graduates-stagnant-for-10-years/. 30.05.2024

https://edition.pagesuite.com/html5/reader/production/default.aspx?pubname=&edid=0d078ba3-52ce-4fe6-8f1f-de9fdc314457. 30.05.2024

INVESTIGATION OF LANGUAGE LEARNING THROUGH USE OF ARTIFICIAL INTELLIGENCE AT UNIVERSITY LEVEL FOR BUSINESS STUDENTS

Gyongyver MĂDUȚA[1]

Abstract

There have been significant advancements in using artificial intelligence (AI) for language learning at the university level, especially for business students. AI technology enables personalized and adaptive learning experiences, allowing students to practice and improve their language skills at their own pace. One common application of AI in language learning is the use of chatbots and virtual tutors, which can engage students in conversations, provide instant feedback, and simulate real-life language usage. These AI-driven platforms often use natural language processing (NLP) algorithms to understand and generate human-like responses, enhancing the learning experience. Moreover, AI-powered language learning tools can analyze students' performance data to identify areas of improvement, tailor exercises to their specific needs, and track their progress over time. By leveraging machine learning algorithms, these platforms can continuously adapt to students' learning patterns and preferences, thereby optimizing their learning outcomes.

Keywords: AI language learning, immersive vr, educational gamification, experiential learning, online learning environments.

JEL Classification: I21, I23, I25, O33

1. Introduction

There have been notable scientific advancements in utilizing artificial intelligence (AI) for language learning at the university level, especially catering to business students. AI technology is revolutionizing the way languages are taught and acquired by offering personalized and adaptive learning experiences.

One significant development is the use of AI-powered chatbots and virtual tutors in language learning. These tools leverage natural language processing (NLP) algorithms to engage students in conversations, provide instant feedback, and simulate real-world language usage. By interacting with AI chatbots, students can practice their language skills in a more interactive and immersive manner. Wang et. al. [1] examined the effects of adaptive learning technologies, including AI-powered platforms, on students' motivation

[1] Assistant Professor PhD, Romanian-American University, gyongyver.maduta@rau.ro

and learning outcomes across various educational contexts. This provided insights into the effectiveness of personalized learning approaches in improving student engagement and achievement.

Furthermore, AI enables the creation of tailored learning experiences for individual students. Through machine learning algorithms, AI systems can analyze data on students' language proficiency, learning patterns, and preferences to customize exercises and materials accordingly [2]. This personalized approach helps students learn at their own pace and focus on areas where they need improvement.

Another area of advancement is the integration of AI-driven language learning platforms with virtual reality (VR) environments and interactive games [3]. These innovative tools provide students with practical scenarios and simulations that mimic real-life language usage in business contexts. By engaging with AI-powered VR environments and games, students can enhance their language skills in a more practical and experiential manner.

Overall, the integration of AI technologies in language learning at the university level for business students is shaping a more effective and engaging learning experience. These advancements not only enhance language proficiency but also equip students with the communication and cultural competencies necessary to succeed in today's globalized business environment.

Several AI-powered chatbots and virtual tutors exist for language learning, each employing various techniques and technologies to facilitate learning [4]. Here are some notable ones:

Duolingo: Duolingo offers a gamified language learning experience through its app, which employs AI algorithms to adapt to users' learning styles and provide personalized feedback. It uses techniques such as spaced repetition and natural language processing (NLP) to enhance learning efficiency.
Rosetta Stone: Rosetta Stone utilizes speech recognition technology to provide instant feedback on pronunciation and offers interactive lessons tailored to individual learners. It focuses on immersive language learning through visual and auditory cues.
Babbel: Babbel uses AI to personalize language lessons based on users' proficiency levels and learning goals. It emphasizes practical conversation skills and offers interactive exercises to reinforce learning.
Memrise: Memrise employs spaced repetition algorithms and multimedia content to help users memorize vocabulary effectively. It uses AI to adapt lessons based on learners' progress and performance.
Busuu: Busuu combines AI with a community of native speakers to offer personalized language learning experiences. It provides interactive exercises, writing practice, and live tutoring sessions with AI-powered feedback.
Lingvist: Lingvist utilizes AI algorithms to adapt language lessons to users' proficiency levels and learning pace. It focuses on vocabulary acquisition through contextual learning and spaced repetition.

HelloTalk: HelloTalk connects language learners with native speakers worldwide for language exchange. While not purely AI-powered, it incorporates AI features such as translation, pronunciation correction, and personalized learning recommendations.

These platforms typically leverage technologies such as NLP, machine learning, speech recognition, and adaptive algorithms to provide personalized learning experiences. They analyze user data, track progress, and offer tailored feedback to enhance language acquisition. Additionally, many of them incorporate interactive exercises, quizzes, and games to make learning more engaging and effective. AI has paved the way for innovative language learning methods, such as immersive virtual reality (VR) environments and interactive language games, which make learning more engaging and effective for business students. These technologies offer practical scenarios and real-world applications, enabling students to apply their language skills in relevant contexts.

2. Theoretical background

Research on the efficiency and effectiveness of AI-powered language learning platforms continues to evolve. To highlight some common themes and trends in the research up to present day, most studies have investigated the impact of personalized learning experiences facilitated by AI algorithms. Research suggests that adaptive learning systems, which tailor content and feedback to individual learners' needs and preferences, can lead to improved learning outcomes compared to traditional one-size-fits-all approaches [4].

Retention and Engagement: Studies have examined the role of AI-powered features such as spaced repetition, gamification, and interactive exercises in enhancing learner retention and engagement [5]. Findings indicate that incorporating these elements can increase motivation, participation, and knowledge retention among users.

Research has explored the effectiveness of AI-generated feedback in language learning contexts [6]. While automated feedback can provide immediate corrections and guidance to learners, studies suggest that it may not always be as accurate or nuanced as feedback from human instructors. However, advancements in natural language processing and machine learning continue to improve the quality and reliability of AI-generated feedback.

Errors and constructive criticism are an inevitable aspect of teaching and studying a foreign language. One way to characterize errors is as departures from the target language's norms. They highlight instances in which learners have improperly transferred rules from their first language to the foreign language or overgeneralized rules in the foreign language, hence illuminating trends in their development of interlanguage systems. Therefore, corrective feedback, which comes in a range of forms, serves as a warning to a language learner that their usage of the target language is inappropriate. Meta-linguistic information may or may not be included in corrective feedback, which can be explicit (e.g., "No, you should say does, not do") or implicit (e.g., "Yes, he goes to work every day").

Some research has compared the efficacy of different AI-powered language learning platforms and methodologies. Comparative studies have evaluated factors such as learning outcomes, user satisfaction, and engagement levels across various platforms to identify best practices and areas for improvement.

While many studies have demonstrated short-term benefits of AI-powered language learning interventions, there is growing interest in understanding their long-term effects. Research efforts are underway to assess whether skills acquired through AI-driven language learning platforms are retained over time [7] and how they transfer to real-world communication contexts.

It's worth noting that the field of AI-powered language learning is dynamic, with new research emerging regularly. Researchers are continually refining existing approaches, exploring novel technologies [8, 9], and investigating the intersection of AI with other fields such as linguistics, psychology, and education. For the latest developments and findings, I'd recommend consulting academic databases, conference proceedings, and research journals in the fields of language education, computer-assisted language learning (CALL), and artificial intelligence [10].

AI has the potential to greatly improve language teaching and student learning results in a number of ways:

1. Personalized Learning: To better meet the requirements of each student, AI algorithms can assess each student's strengths, weaknesses, learning preferences, and rate of learning. By ensuring that every student receives customized help, this customization promotes more successful learning outcomes.

2. Adaptive Learning Platforms: Equipped with artificial intelligence (AI), these platforms may modify the level of exercises and materials according to students' performance, guaranteeing that they are suitably challenged and involved. Additionally, these systems have the ability to give pupils instant feedback, enabling them to fix errors in real time.

3. Language Tutoring Bots: Artificial intelligence chatbots and online tutors can provide students more practice chances outside of the classroom. These chatbots may converse with pupils in an engaging manner, fix their pronunciation and grammar, and respond to their inquiries on grammar and vocabulary.

4. Natural Language Processing (NLP) Tools: NLP tools are capable of analyzing vast volumes of text to spot trends, typical mistakes, and areas in need of development. These findings may be used by educators to create more specialized lesson plans and provide students tailored feedback.

5. Language Assessment: By grading written assignments, assessing spoken language skills, and giving standardized examinations, AI can automate the assessment process. This

frees up teachers' time so they can concentrate on giving pupils insightful feedback and encouragement.

6. Language Generation: To enhance in-class learning, AI may provide language exercises, tests, and study guides. Learning may be made more effective and interesting by tailoring these resources to meet particular learning goals and student interests.

7. Multimodal Learning Experiences: By fusing text, audio, video, and interactive components, AI can provide multimodal learning experiences. Language learning applications, for instance, can offer textual explanations and visual assistance in addition to using speech recognition technology to assist students in honing their speaking and listening abilities.

AI-powered tools often incorporate interactive elements, gamification, and multimedia content, making learning more engaging for students [11]. Increased engagement can lead to higher levels of motivation and participation in language learning activities.

Personalized learning experiences tailored to individual student needs can lead to better retention of language concepts and skills [12]. When students receive targeted support and practice opportunities, they are more likely to retain and apply what they have learned. AI algorithms can identify gaps in students' knowledge and provide targeted interventions to address them. This personalized approach can accelerate students' language learning progress by focusing on areas where they need the most support.

With AI-powered language tutoring bots and adaptive learning platforms [13], students can receive immediate feedback on their grammar, pronunciation, and vocabulary usage. Continuous feedback and practice opportunities can lead to improvements in language proficiency over time. AI technologies can provide language instruction and support to students with diverse learning needs [13, 14], including those with disabilities or language-related challenges. By offering customizable learning experiences and adaptive materials, AI can help ensure that all students have access to high-quality language instruction.

Last but not least, AI can assist teachers in analyzing student data, identifying trends, and providing targeted support to individual students or groups [15]. By automating routine tasks such as grading and assessment, AI frees up teachers' time to focus on delivering personalized instruction and supporting students' unique learning needs. By integrating AI into language instruction from an early age, students can develop positive learning habits and attitudes towards language learning. The engaging and interactive nature of AI-powered language learning tools can encourage students to continue practicing and improving their language skills outside of the classroom.

3. Methodology

For our research, we gathered a corpus of classroom exchanges that comprised 19 transcriptions of foreign language instruction in English given by seven different teachers, amounting to around 12 hours. The instructors for the 19 Bucharest classes were divided into four Romanian teachers who teach in secondary schools and have Romanian as their first language, two native Americans who teach in universities, and one British teacher who teaches in an English further education college. The use of diverse educators from three distinct nations instills trust in us that our findings stem from a well-rounded corpus, rather than the peculiarities of a particular teaching methodology.

The majority of pupils are native speakers of Romanian.

In our research, we also asked the instructors to list the competency levels of each student in their class (es). Table 1 provides further information.

Level	Class	Trainer	Focus on	Duration
Beginner (A2)	1	1	Meaning	40
Beginner (A2)	1	1	Forms	40
Beginner (A2)	2	2	Forms	40
Beginner (A2)	3	2	Meaning	30
Beginner (A2)	4	3	Forms	30
Beginner (A2)	5	3	Forms	50
Beginner (A2)	1	3	Forms	50
Beginner (A2)	2	3	Forms	40
Beginner (A2)	2	3	Forms	40
Intermediate	3	4	Forms	50
Intermediate	4	4	Meaning	50
Intermediate	5	5	Meaning	50
Intermediate	6	5	Meaning	50
Advanced	7	6	Meaning	30
Advanced	8	6	Meaning	30
Advanced	9	6	Meaning	30
Advanced	10	7	Meaning	30
Advanced	11	7	Meaning	30
Advanced	12	7	Meaning	30
Total	12	7		680 minutes

Table 1: Data from Classroom Transcription

We requested instructors to record and provide us samples of their regular classroom interactions in order to gather data. We didn't tell the teachers what to focus on, what kind of errors to make, or what kinds of comments they should provide each other. The

recordings from the classroom can be divided into two main categories: those that concentrate on meaning, or lessons where the students talk about a range of cultural topics in the foreign language, and those that concentrate on forms, or lessons about various grammar topics (such as pronominalization, syntactic structures, possessives, and passive verbs). For each class, each trainer used one or more AI-powered language learning platforms or tools that align with the language proficiency levels and learning objectives of the students. These included platforms like Duolingo, Rosetta Stone, or educational chatbots specifically designed for language learning.

The selected AI-powered tools were integrated into regular classroom instruction alongside traditional teaching methods. We ensured that teachers and students are provided with appropriate training and guidance on how to use the tools effectively.

We analyzed the recorded classroom interactions to observe how the integration of AI-powered tools influences teaching practices and student learning experiences.

Special attention was paid to factors such as:

Frequency and nature of teacher-student interactions facilitated by AI tools.
Types of feedback provided by teachers and the effectiveness of corrective feedback in improving student language proficiency.
Student engagement levels during AI-supported activities compared to traditional instruction.
Differences in classroom dynamics and instructional approaches between lessons focusing on meaning (e.g., cultural topics) and lessons focusing on forms (e.g., grammar topics).

4. Results

When professors openly appreciate a student's input and when they just go on to the next question, subject, or explanation are the things that we are interested in learning about. We are also interested in the ways that teachers respond to students in various contexts and the most common and efficient forms of corrective feedback.

We gathered qualitative feedback from teachers and students regarding their experiences with the AI-powered tools by soliciting opinions on the usability, effectiveness, and perceived benefits or limitations of integrating AI into language instruction.

By conducting this adapted experiment, we gained insights into the potential benefits and challenges of incorporating AI-powered language learning tools into classroom instruction and contribute to the broader understanding of AI's role in education.

A native English speaker annotated the classroom recordings with definitions for the labels. In order to analyze the data files using a range of XML tools, such as tools for querying and displaying the data, we employed an annotation tool to add XML tags to the data files. We concentrated on the labels for this study that will most clearly elicit differences in use of AI. We were especially curious about the following:

Type and frequency of mistakes in vocabulary, grammar, and pronunciation.
Kind and frequency of various feedback formats (repetition and rewording for both instructor an AI tool).
Kind and frequency of corrections for vocabulary, pronunciation, and grammatical mistakes.
Student utterance type in front of instructor versus AI.

We totalled 680 minutes of classroom data and 100 minutes of student AI tutorial data from the same target group.

We took into account how the two teaching modalities distributed the various mistake categories and corrective feedback (Table 2.). Many errors can be seen in the speech of certain students. However, in order to maintain a fluid discourse, we discovered that teachers usually only addressed one mistake at a time, whereas AI tools/tutors address each mistake individually.

Level	Grammar (%)	Vocabulary (%)	Pronounciation (%)
Beginner (A2)	47	6	47
Intermediate	60	33	7
Advanced	45	36	19
Total	50	20	30

Table 2. Classroom Error Frequency by Learning Level

The relative frequency of the various errors' kinds varies depending on the learner's level. Among the most common mistakes made by students of all levels are grammatical ones. Pronunciation mistakes are just as common among novices as grammar faults, with vocabulary errors ranking second among intermediate and advanced learners.

We did discover that the two modes differed in how they interacted and how they repaired themselves. Firstly, it appears that teachers are worried about keeping the discourse flowing in classes where the emphasis is on communication. Rather than promoting self-repair sequences that might break up the flow of the discourse, they strive to help pupils solve errors by giving them the goal forms. Second, teachers must rely on modeling strategies like recast and explicit correction since students who are not yet fluent in the target language might not have the information necessary to self-correct.

5. Conclusions

Initially, while it might seem that there were fewer mistakes in tutorial mode compared to classroom mode, considering that we have 680 minutes of classroom data and only 100 minutes of tutorial (AI) data, we find that errors were actually much less common in the classrooms. This could be because students have more opportunities to interact with the AI in the AI tutorial mode than in the classroom form.

If the AI-powered tools are engaging and interactive, the results show increased student participation and motivation during AI-supported activities compared to traditional instruction. This suggests that AI has the potential to enhance student engagement in language learning.

Students demonstrated improvements in language proficiency over time, as measured by assessments or evaluations, which indicates that the AI-powered tools are effective in facilitating language acquisition. This suggests that AI has a positive impact on student learning outcomes in the language classroom.

Corrective feedback's success is contingent upon the specific linguistic elements that require correction as well as the circumstances surrounding the feedback's delivery. We looked into two of these situations in our investigation. The first is the student's ability—or lack thereof—to react to the corrective input in a classroom context. We discovered that, in a non-negligible percentage of cases—such as when using multiple AI language tools—the teacher moved on to the next question, topic, or kind of feedback (clarification, explanation) without waiting for the student to respond or validate the feedback.

The results also show differences in teacher responses and instructional approaches when using AI tools compared to traditional methods. Some teachers embrace the technology and adapt their teaching strategies to integrate AI seamlessly into their instruction, while others encountered challenges or resistance to change.

AI tools provide timely and relevant feedback to students, which indicated that AI-supported instruction leads to more efficient corrective feedback and language development. This suggests that AI use in classroom instruction has the potential to enhance the feedback process in language learning environments.

Analysis of usage patterns and feedback from teachers and students also reveals insights into the usability and effectiveness of the AI tools. Teachers and students expressed preferences for certain features or functionalities of the tools based on their experiences, which could inform future development and implementation strategies.

Given the diverse background of teachers and students in the experiment, the results also highlight cultural and linguistic factors that influence the effectiveness of AI-supported language instruction.

The investigation revealed difficulties and restrictions related to AI application in language learning settings. These might be problems with usability, technological obstacles, or incorporating AI technologies into current teaching methods and curriculum. Optimizing the use of AI in education requires recognizing and resolving these issues.

The experiment's findings have wider ramifications for language instruction and acquisition. Overall, research in a number of fields, such as intelligent tutoring systems, intelligent computer assisted language learning, and second language acquisition, has enhanced our study methodology. These differing viewpoints raise broader concerns about

how AI systems, by providing additional opportunities for engagement and encouraging student-generated repair, might help to mitigate the constraints or drawbacks associated with the classroom approach of handling student errors. In addition to highlighting topics for more study and development in the field of AI-driven language teaching, they might provide guidance for judgments on the adoption and application of AI-powered technologies in educational settings.

It is necessary to do further research on how learners react to various feedback systems and if they alter their learning over time. Longitudinal trials are also necessary to confirm the trends shown in our investigations.

References

[1] WANG, Xiaoman, et al. *The Efficacy of Artificial Intelligence-Enabled Adaptive Learning Systems From 2010 to 2022 on Learner Outcomes: A Meta-Analysis*. Journal of Educational Computing Research, 07356331241240459, May 15 2024.

[2] MURTAZA, Mir, et al. *AI-based personalized e-learning systems: Issues, challenges, and solutions*. IEEE Access, 10: 81323-81342, 2022.

[3] DIVEKAR*, Rahul R., et al. *Foreign language acquisition via artificial intelligence and extended reality: design and evaluation*. Computer Assisted Language Learning, 35.9: 2332-2360, 2022.

[4] JINMING, D. U.; DANIEL, Ben Kei. *A Systematic Review of AI-Powered Chatbots in EFL Speaking Practice: Transforming Language Education*. Computers and Education: Artificial Intelligence, 100230, 2024.

[5] GULOMOVNA, Xatamova Ziyoda; SOLIDJONOV, Dilyorjon. *Effective Methods To Increase Vocabulary In English As A Second Language*. Journal of Universal Science Research, 1.6: 735-745, 2023.

[6] YANG, Hongzhi; GAO, Chuan; SHEN, Hui-zhong. *Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study*. Education and Information Technologies, 29.4: 3837-3858, 2024.

[7] JEON, Jaeho. *Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives*. Computer Assisted Language Learning, 37.1-2: 1-26, 2024.

[8] BORGER, Jessica G., et al. *Artificial intelligence takes center stage: exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education*. Immunology and Cell Biology, 101.10: 923-935, 2023.

[9] ANIS, Muneeba. *Leveraging Artificial Intelligence for Inclusive English Language Teaching: Strategies and Implications for Learner Diversity*. Journal of Multidisciplinary Educational Research, 12.6, 2023.

[10] IMA, Nur; JIHAD, Adriani. Computer-Assisted Language Learning: The Impact in Language Education. Vifada Journal of Education, 2.1: 29-35 , 2024.

[11] YAKHINA, R., et al. *AI-Powered Tools For Enhancing Engagement And Academic Performance Of Engineering Students*. In: INTED2024 Proceedings. IATED. p. 2990-2996, 2024.

[12] KASWAN, Kuldeep Singh; DHATTERWAL, Jagjit Singh; OJHA, Rudra Pratap. *AI in personalized learning.* In: Advances in Technological Innovations in Higher Education. CRC Press, p. 103-117, 2024.

[13] FOLGIERI, Raffaella, et al. *AI-Powered Personalised Learning Platforms For EFL Learning: Preliminary Results*. In: Proceedings of the 16th International Conference on Computer Supported Education. Science and Technology Publications. p. 1-7, 2024.

[14] DE LA VALL, Roxana Rebolledo Font; ARAYA, Fabián González. *Exploring the benefits and challenges of AI-language learning tools*. Int. J. Soc. Sci. Humanit. Invent, 10: 7569-7576, 2023.

[15] POKRIVCAKOVA, Silvia. *Preparing teachers for the application of AI-powered technologies in foreign language education*. Journal of Language and Cultural Education, 7.3: 135-153, 2019.

Bibliography

[1] ALHARBI, Wael. *AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools*. Education Research International, https://doi.org/10.1155/2023/4253331 , pp.1-15 February 2023.

[2] ALSHAIKHI, Turky; KHASAWNEH, Mohamad Ahmad Saleem. *Investigating the Benefits and Drawbacks of Selected AI-Powered Instruments Employed for Foreign Language Education*. Kurdish Studies, 12.2, pp. 2156-2169, 2024.

[3] DIVEKAR*, Rahul R., et al. *Foreign language acquisition via artificial intelligence and extended reality: design and evaluation*. Computer Assisted Language Learning, 35.9: 2332-2360, 2022.

[4] JINMING, D. U.; DANIEL, Ben Kei. *A Systematic Review of AI-Powered Chatbots in EFL Speaking Practice: Transforming Language Education*. Computers and Education: Artificial Intelligence, 100230, 2024.

[5] FOLGIERI, Raffaella, et al. AI-Powered Personalised Learning Platforms For EFL Learning: Preliminary Results. In: Proceedings of the 16th International Conference on Computer Supported Education. Science and Technology Publications. p. 1-7, 2024.

[6] GULOMOVNA, Xatamova Ziyoda; SOLIDJONOV, Dilyorjon. Effective Methods To Increase Vocabulary In English As A Second Language. Journal of Universal Science Research, 1.6: 735-745, 2023.

[7] HUANG, Xinyi, et al. Trends, research issues and applications of artificial intelligence in language education. Educational Technology & Society, 26.1: 112-131, 2023.

[8] STRASSER, Thomas. *Narrow AI-Powered Visualization Facilitation Tools in Foreign Language Learning: A Visual Approach Promoting Equal Opportunities in Foreign Language Grammar Teaching*. In: Facilitating Visual Socialities: Processes, Complications and Ethical Practices. Cham: Springer International Publishing,. p. 85-98 , 2023

TESTING COMMAND MODIFICATIONS TO GRAPH A VECTOR FIELD OVER A CONE IN 3D USING MAPLE

Jose MUJICA[1]

Ramon MATA-TOLEDO[2]

Abstract

Undoubtedly science researchers have benefited from the arrival of mathematical software by different manufacturers such as MATLAB (™), Maple (™), Mathematica (™) and the like. The fast reactivity of these software tools to changes in their parameters allows researchers to observe and study mathematical models in real time and speed up their research process. It is desirable that these wonderful tools be improved and applied to a wide range of new applications. However, one important aspect to consider is the difficulty of translating the traditional mathematical formulas as written in books or manuals into the appropriate syntax required by the mathematical software. In the present paper, the authors show the difficulties to graph a vector field over a cone in 3D using both, an outdated version of the mathematical software Maple (™) and its latest version. Although a solution was obtained the authors think that this solution is far from being optimal. This work includes details of the code modifications and the combination of instructions or commands to obtain the desired solution.

Keywords: maple software, plot a cone, vector field over a cone, 3d vector field

JEL Classification: C02

1. Introduction

One powerful resource in the research world is to test a combination of scientific processes such as that of visualizing a vector field over a surface. It is widely known that visualizing a combination of mathematical functions may help explain many physical events.

Tools that permit us to make multiple quick tests over on complex simulations of world nature are very appreciated in the scientific field.

Recently the authors, using the mathematical software Maple(™) [1], have been studying moving particles on a membrane hoping to find a mathematical model that shows the

[1] EE, AES Fellow, Escuela Superior de Audio y Acústica, Venezuela, jmujica@escuelasuperiordeaudio.com.ve

[2] Professor Emeritus of Computer Sciences, James Madison University, US, matatora@jmu.edu

relationship between the membranes' surface areas and the quantity and direction of the particles pushed by them in a gaseous medium. Usually, we find examples in textbooks about how a vector field passes through a surface but not what happens when the field itself starts from a surface.

One of the many inconveniences we have experienced when working with mathematical software is the waste of time reconciling the syntax of the mathematical expressions as written in books or the Internet with that of the required software. An example using Maple is the instruction shown in Figure 1 and its corresponding cone of Figure 2. Notice that the traditional cone equation is not reflected in the code of the figure.
One powerful resource in the research world is to test a combination of scientific processes such as the visualization of a surface with that of a vector field. It is widely known that visualizing a combination of mathematical functions may help explain many physical events.

Tools that permit us to make multiple quick tests over complex behavior of simulations of world nature are very appreciated in the scientific field.

Recently the authors have been studying moving particles on a membrane hoping to find a mathematical model that shows the relationship between the membranes' surface areas and the quantity and direction of the particles pushed by them in a gaseous medium. Usually, we find textbooks with examples about how a vector field passes through a surface but not what happens when the field itself starts from a surface.

One of the many inconveniences we have experienced when working with mathematical software is the waste of time reconciling the syntax of the mathematical expressions as written in books or the Internet and that of the required software. An example using Maple™ [1] is the instruction shown in Figure 1 and its corresponding cone of Figure 2. Notice that the traditional cone equation is not reflected in the code of the figure.

Maple code to plot a Cone

```
plots:-display(plot3d([rho, theta, (1/4)*Pi], rho = 0 .. 3*sec((1/4)*Pi), theta= -Pi .. Pi, coords = spherical));
```

44. Minimum distance to the origin Find the point closest to the origin on the curve of intersection of the plane $2y + 4z = 5$ and the cone $z^2 = 4x^2 + 4y^2$.

Figure 1 Cone equation as presented in a book. [2]

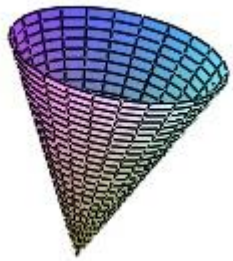


Figure 2 3D cone plotted on Maple

Let us consider another Maple software snippet using a different instruction to plot a cone anew. As seen, the classical cone equation is not reflected in the instruction syntax.

```
restart; x := r*cos(phi); y := r*sin(phi); plot3d([[x, y, r], [x, y, 4]], r = 0
.. 4, phi = 0 .. 2*Pi, style = surfacewireframe, color = grey, scaling =
constrained, labels = ["x", "y", z])
```

When using Maple, we experienced several difficulties when plotting a vector field over a surface. Fortunately, on the Internet we found a help file which after several trials and modifications allowed us to plot a vector field through a cone. This procedure is explained next.

1. PLOTTING A VECTOR FIELD ON A SURFACE WITH MAPLE.

The following code snippet is the one we found on the Internet. We also show the plot it produces. See figure 3.

```
with(plots): with(VectorCalculus):
SetCoordinates(cartesian[x, y, z]):
eqn := (x^2+3y^2)*exp(1/2*(-x^2-y^2)):
montagne := plot3d(eqn, x = -3..3, y = -3.6.. 3.6, shading = zgrayscale, grid = [300,300],
scaling = unconstrained):
normals := Gradient(z-eqn):
display(montagne,seq(seq(arrow([x, y, eqn],normals/2),colour = red),
x = -3..3),y =-3..3), scaling = constrained, axes = boxed)
```

```
> with(plots) : with(VectorCalculus) :
  SetCoordinates(cartesian[x, y, z]) :
   equal := (x^2 + 3*y^2) * exp(1/2*(-x^2-y^2)) :
   model := plot3d(equal, x = -3 .. 3, y = -3.6 .. 3.6, shading = xyz,
    grid = [300, 300], scaling = unconstrained) :
  normals := Gradient(z - equal) :
  display(model, seq(seq(arrow([x, y, equal], normals/2, colour = red),
    x = -3 ..3), y = -3 ..3), scaling = constrained);
```

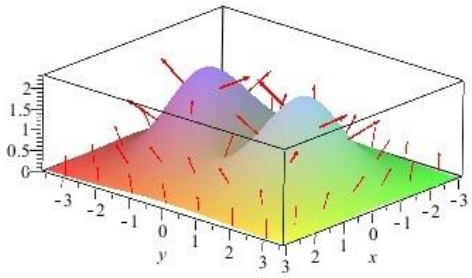


Figure 3 Vector field through a surface

When we tried to modify the equation of the previous snippet using the Cartesian expression of a cone to write it as shown in our reference book the software did not allow us to do so.

The authors also tried to modify the plot of vector field in 3D as produced by the Vector Calculus Tutorial in the Tool menu of Maple. This was not possible either. The next window, Figure 4, shows the only plot that we were able to obtain.

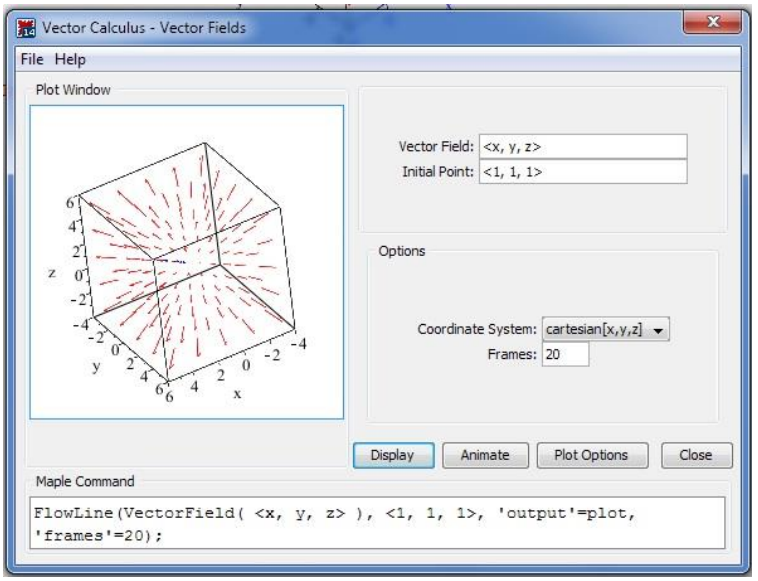


Figure 4 Vector Field on Maple using the Vector Calculus Tutorial

2 MODIFYING THE CODE SNIPPET

The following step illustrates the modifications implemented in the snippet of Figure 3.

Changing The Coordinates System and the Equation

The first try we did was to replace the equation of Section 1.

$$(x^2 + 3y^2) * e^{\left(\frac{1}{2}(-x^2-y^2)\right)}$$

By the Cartesian expression of the cone equation

$$x^2 + y^2 = z^2$$

After several trials we did not obtain the desired results. Therefore, we decided to change the Cartesian coordinates to spherical and proceed to write the conversion of the variables *x* and *y*

$x = r * \cos(phi)$ $y = r * \sin(phi)$

As a result, we obtained the cone's graph, but the vector field disappeared (See Figure 5).

```
with(plots): with(VectorCalculus):
SSetCoordinates('spherical'r, Φ, Θ):
x := r*cos(phi): y := r*sin(phi)):
model := plot3d([x, y, r],r = 0..4,phi = 0..2*Pi, shading = xyz, grid = [300,300], scaling =
unconstrained, axes = boxed):
normals := Gradient(z-model):
display(model, seq(seq(arrow([x, y,r*sin(phi)+r*cos(phi)],normals/4), colour = red), x =-
2..2,y =-2..2,scaling =constrained);
```

```
> with(plots) : with(VectorCalculus) :
  SSetCoordinates('spherical'r, ϕ, θ) :
  x := r*cos(phi) : y := r*sin(phi) :
  model := plot3d([x, y, r], r = 0 .. 4, phi = 0 .. 2·Pi, shading = xyz,
    grid = [300, 300], scaling = unconstrained, axes = boxed) :
  normals := Gradient(z − model) :
  display(model, seq(seq(arrow([x, y, r*sin(phi) + r·cos(phi)],
  normals/4, colour = red), x = −2 .. 2), y = −2 .. 2), scaling = constrained);
```

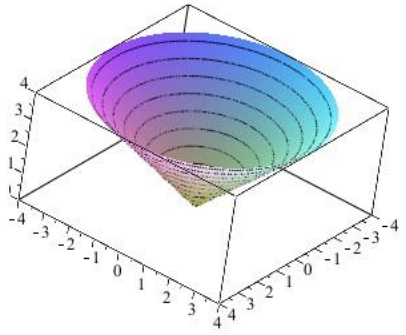


Figure 5 First try to plot a vector field on a cone

Changing The Limit of the Equation

The membrane we wanted to study is that of an electrodynamic loudspeaker with a cone shape. To do this we decided to change the limits of the equation from r = 0..4 to r = 1..4 of the cone to get a frustum (See Figure 6). We also set the third variable of the display function to 5 in lieu of the spherical expression (changes are shown in blue). The result we obtained, at this scale, is the representation of the vector field as a set of points.

```
with(plots): with(VectorCalculus):

SSetCoordinates('spherical'r, Φ, Θ):
x := r*cos(phi): y := r*sin(phi):
model := plot3d([x, y, r], r =1.. 4, phi = 0..2*Pi,shading = xyz,
grid = [300, 300],scaling = unconstrained, axes = boxed):
normals := Gradient(z-(model):
display(model,seq(seq(arrow([x, y, 5], normals/4), colour =
red), x = -2..2, y = -2..2),scaling = constrained)
```

```
> with(plots) : with(VectorCalculus) :
  SSetCoordinates('spherical'r, ϕ, θ) :
  x := r*cos(phi) : y := r*sin(phi) :
  model := plot3d([x, y, r], r = 1 .. 4, phi = 0 .. 2·Pi, shading = xyz,
    grid = [300, 300], scaling = unconstrained, axes = boxed) :
  normals := Gradient(z − model) :
  display(model, seq(seq(arrow([x, y, 5], normals/4, colour = red),
    x = −2 ..2), y = −2 ..2), scaling = constrained);
```

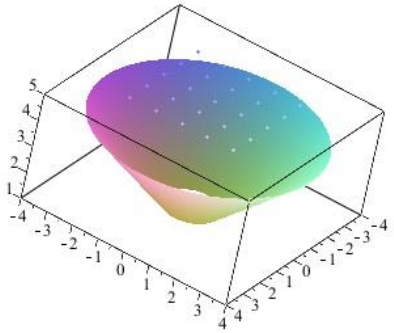


Figure 6 First view of the vector field over the cone

2.3 Obtaining a Visible Vector Field

The visualization of the vector field was obtained by increasing the variables x and y. This allowed us to obtain a wider view of the plot. Finally, we adjusted the plot by subtracting 2 from the axes x and y respectively. We also increased the third variable of the display function to 7 as shown in Figure 7. Likewise, all changes are shown in blue. However, the whirlpool vector field obtained did not show the direction that we expected.

```
with(plots): with(VectorCalculus):
SSetCoordinates('spherical'r, Φ, Θ):
x := r*cos(phi): y := r*sin(phi):
model := plot3d([x, y, r],r =1..4,phi = 0 ..2*Pi,shading = xyz,
grid = [300, 300],scaling = unconstrained, axes = boxed):
model1 :=fieldplot3d([4*x-2, 4*y-2,6], r = -4..4, phi =
-Pi..Pi,z =4..7):
normals := Gradient(z-model):
display(model,model1, seq(seq(arrow([x, y, 7], normals/4), colour=
red), x = -2..2), y = -2..2,scaling = constrained;)
```

```
> with(plots) : with(VectorCalculus) :
  SSetCoordinates('spherical'r, ϕ, θ) :
  x := r* cos(phi) : y := r* sin(phi) :
  model := plot3d([x, y, r], r = 1 .. 4, phi = 0 .. 2 · Pi, shading = xyz, grid = [300, 300],
  scaling = unconstrained, axes = boxed) :
  model1 := fieldplot3d([4 · x − 2, 4 · y − 2, 6], r = −4 ..4, phi = −Pi ..Pi, z = 4 ..7) :
  normals := Gradient(z − model) :
  display(model, model1, seq(seq(arrow([x, y, 7], normals/4, colour = red),
    x = −2 ..2), y = −2 ..2), scaling = constrained);
```

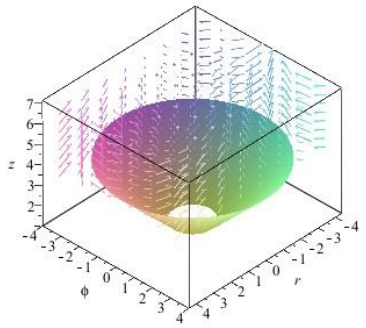


Figure 7 Whirlpool vector field over a cone

3 Looking for the Appropriate Vector Field

In our next step, we proceeded to look for a vector field with a unique origin point [4]. Fortunately, in the help file of the software was an example with the desired field's shape (See Figure 8).

```
> fieldplot3d([1, 0, 0], r = 0 .. 4, t = 0 .. (1/2)*Pi, p = 0 .. (1/2)*Pi, coords = spherical, axes = boxed);
```

```
> fieldplot3d([1, 0, 0], r = 0 ..4, t = 0 .. Pi/2, p = 0 .. Pi/2,
  coords = spherical, axes = boxed)
```

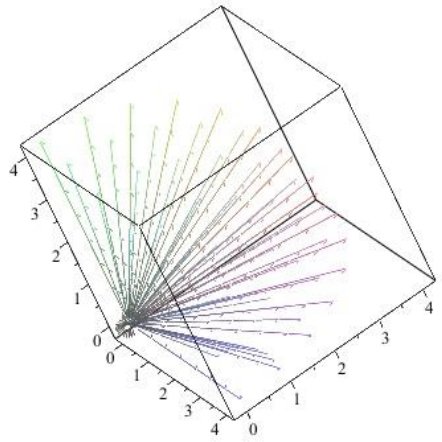


Figure 8 Vector field from the origin

3.1 First View of The Directed Vector Field Over the Cone

When we tried to integrate this command with the previous one, we did not know if these two were going to be compatible. Some adjustments were necessary to achieve the correct view. In this respect, in model 1, we changed the axes of Fieldplot3D by switching t and r to make it coincide with the direction of the cone (See Figure 9). Finally, we got the directed vector field over our cone, but still it did not have the correct width that we expected.

```
with(plots): with(VectorCalculus):
SSetCoordinates('spherical'r, Φ, Θ):
x := r*cos(phi):y := r*sin(phi))
model := plot3d([1.9*x,1.9*y,r-4], r = 1 .. 4*phi = 0 ..2* Pi, shading = xyz, grid = [300, 300], scaling = unconstrained, axes = boxed):
model1 := fieldplot3d([1, 0, 0], r = 0 .. Pi/1.5, t = 2 .. 7, p = 0 .. Pi/1.5), coords = spherical, axes = boxed, color = blue):
normals := Gradient(z-model):
display(model, model1, scaling = unconstrained);
```

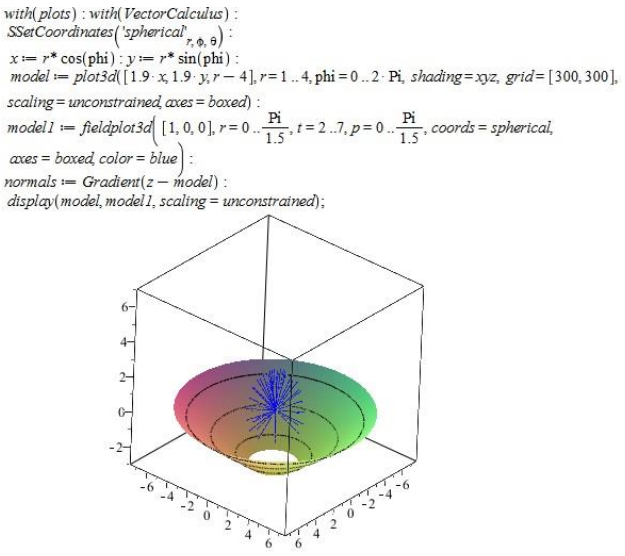


Figure 9 Directed Vector Field over The Cone

3.2 Final View of the Desired Vector Field Over the Cone

Based upon our partial success by obtaining a vector field coming from a cone loudspeaker membrane, we decided to modify the equations of the model and model1 and by tweaking the appropriate parameters we finally got the desired graphic as shown in Figure 10.

```
with(plots): with(VectorCalculus):
SSetCoordinates('spherical'r, Φ, Θ):
x := r*cos(phi)): y := r*sin(phi)):
model := plot3d([0.4*x,0.4*y,r/8],r =0.8..3, phi=0..2*Pi, shading = zgrayscale,grid = [300, 300], scaling = unconstrained, axes =boxed): model1:= fieldplot3d([1,0,0.5], r =0..Pi/2.8, t = 4..9,p = 0..Pi/2.8, coords = spherical,axes = boxed, color = blue):
normals := Gradient(z-model):
display(model, model1, scaling = unconstrained);
```

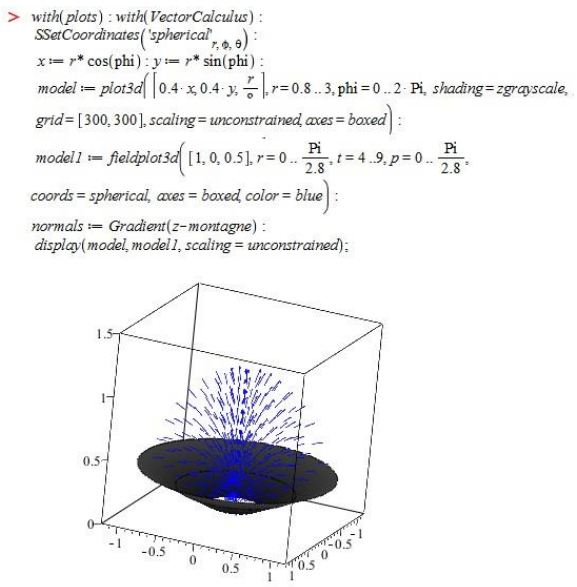


Figure 10 Our desired vector field over the cone

4. CONCLUSIONS

It is important to mention that the aim of this work is to encourage researchers and practitioners to employ powerful tools that improve their research without spending too much time learning how to use these tools. The learning curve to master the language and functionality of mathematical software may vary from weeks to months and, in some cases, years. For researchers mastering a tool is not the main objective of their endeavors. Tools are just an auxiliary complement to facilitate research and share the findings with the interested community.

It is common to encounter new auxiliary technologies and tools that can help us to teach our experiences, however, some of them may be easy to learn while others may have a long learning curve. Finding shortcuts that can speed up the learning process are always helpful and welcome.

This work's field of application extends beyond studying a loudspeaker membrane. For example, inverting the direction of the vector field could be used to study the tympani membrane or microwave parabolic antennas.

It is important to note that the Maple commands and codes written in this paper do not necessary match the exact syntax of Maple. What this means is that if you copy any piece of code and paste it on a worksheet, it may not work. This is because we exported them from Maple as plain text and then edited them manually.

References

[1] Maple Software plot. https://www.maplesoft.com/

[2] Thomas Jr. George, Calculus Early Transcendental, 13th Edition, Pearson, 2014

[3] Israel Robert, Plot a Vector Field on a Surface in Maple, (Emeritus Associate Professor of Mathematics at University of British Columbia.), 2016, Math Stack Exchange https://math.stackexchange.com/questions/2072082/plot-a-vector-field-on-a-surface-in-maple

[4] MarsdenJerrold and Tromba Anthony, Vector Calculus, W.H. Freeman and Company, California, 1976

[5] Zill, D. Advance Engineering Mathematics, 7th Edition, Jones & Barlett Learning, Burlington, MA, 2022.

Bibliography

Israel Robert, Plot a Vector Field on a Surface in Maple, (Emeritus Associate Professor of Mathematics at University of British Columbia.), 2016, Math Stack Exchange https://math.stackexchange.com/questions/2072082/plot-a-vector-field-on-a-surface-in-maple

Maple Software plot. https://www.maplesoft.com/

MarsdenJerrold and Tromba Anthony, Vector Calculus, W.H. Freeman and Company, California, 1976

Thomas Jr. George, Calculus Early Transcendental, 13th Edition, Pearson, US, 2014

Waterloo Maple Inc., Maple User Manual, Maplesoft, Canada, 2010

Zill, D. Advance Engineering Mathematics, 7th Edition, Jones & Barlett Learning, Burlington, MA, 2022.

THE USAGE OF 3D LASER SCANNING AND DIRECT DIGITAL MANUFACTURING FOR RESTORATION AND REPLICATION OF CULTURAL HERITAGE

Elena MUSEANU[1]

Maria Sarah VLĂDESCU[2]

Abstract

The integration of 3D laser scanning, and additive manufacturing has sparked a profound transformation in the realm of cultural heritage restoration and preservation. These cutting-edge technologies offer a blend of precision, non-invasiveness, and efficiency that has redefined the way artifacts and historical sites are documented and restored.

At the beginning of this revolution lies the ability of 3D laser scanning to meticulously capture detailed three-dimensional digital models of cultural objects and sites. This process is non-invasive, meaning it avoids physical contact with the artifacts, thereby minimizing the risk of damage during documentation. With such precision, conservators can create accurate replicas of damaged objects, preserving their historical authenticity while restoring them to their former glory.

The integration of additive manufacturing, commonly known as 3D printing, enables the efficient restoration of these objects. Utilizing the digital models obtained through laser scanning, conservators can recreate missing or deteriorated parts with unprecedented accuracy. This streamlined approach not only accelerates the restoration process but also ensures that the replicas seamlessly integrate with the original artifact. Beyond restoration, these technologies facilitate continuous monitoring and condition assessment, mitigating the risk of further deterioration. By digitally archiving cultural artifacts and sites, conservators can monitor changes over time, intervening promptly when necessary to preserve their integrity.

Through immersive experiences and interactive exhibits, individuals can explore historical sites and artifacts in ways previously unimaginable, fostering a deeper appreciation for our shared heritage.

Keywords: Cultural Heritage, 3D Laser Scanning, Additive Manufacturing, Restoration, Replication, Non-Invasive, Digital Models, Conservation Planning

JEL Classification: Z11

[1] PhD, associate professor, Romanian-American University, Bucharest, Romania, elena.museanu@rau.ro
[2] Student, "Mihai Viteazul" National College, Ploiesti, Romania

1. Introduction

Cultural heritage embodies the rich tapestry of history, identity, and achievements of human civilization, serving as a bridge connecting past, present, and future generations. These irreplaceable artifacts, however, are perpetually threatened by the ravages of time, the destructive forces of natural disasters, human conflict, and other factors that contribute to their gradual decay and destruction. Preservation and restoration of cultural heritage objects are essential to maintaining the historical continuity and educational value they offer. In recent years, advancements in technology, particularly in the fields of 3D laser scanning and additive manufacturing, have revolutionized the approach to heritage conservation. These cutting-edge technologies enable unprecedented precision and efficiency in the restoration and replication of damaged or destroyed artifacts, providing tools that were previously unimaginable. This paper delves into the transformative impact of 3D laser scanning and additive manufacturing on cultural heritage preservation, examining their roles, benefits, and implications for the future of this critical field. By leveraging these technological innovations, we can ensure that the cultural legacy of humanity is preserved and appreciated by future generations.

2. Historic Preservation

Historic preservation encompasses the practice of maintaining and safeguarding cultural heritage sites, buildings, artifacts, and other significant objects from the forces of deterioration or destruction. This multidisciplinary field integrates elements from archaeology, architecture, art history, and engineering, aiming to preserve the cultural treasures that hold immense educational and historical value for future generations.

Traditional preservation methods have often relied on physical restoration techniques, which, while effective, can be invasive and may alter the original materials of the artifacts, potentially compromising their authenticity. For example, restoring a historic building might involve replacing decayed wood or stone, which, although necessary, can lead to a loss of original craftsmanship and materials. Similarly, repairing ancient artifacts often involves cleaning or reassembling fragments, processes that risk damaging delicate items or obscuring historical details.

The introduction of 3D laser scanning and additive manufacturing technologies has marked a paradigm shift in the landscape of historic preservation. These technologies offer more precise and non-invasive alternatives, fundamentally changing how we approach the conservation and restoration of cultural heritage. 3D laser scanning captures detailed, high-resolution images of objects and structures, creating exact digital replicas without any physical contact. This non-invasive technique is particularly valuable for fragile artifacts that could be damaged by traditional methods.

Additive manufacturing, commonly known as 3D printing, uses these digital models to produce accurate physical replicas of artifacts. This technology allows conservators to create precise replacements for missing or damaged parts, seamlessly integrating them into the original items. These replicas can be made from materials that closely match the originals, ensuring that the restored artifacts retain their historical authenticity. For instance, a 3D-printed replica of a broken statue can be used to complete the original, preserving its aesthetic and cultural significance.

By employing these advanced methods, we can achieve a balance between maintaining the integrity of artifacts and utilizing modern techniques to ensure their longevity and accessibility for future study and appreciation. 3D laser scanning and additive manufacturing enable detailed documentation and precise restoration, enhancing the ability to study and understand historical objects and sites. Moreover, the digital records created through scanning provide invaluable data for future research and conservation efforts, ensuring that even if physical artifacts are lost or further damaged, their details are preserved. [1]

The integration of these technologies into historic preservation not only improves the accuracy and efficiency of restoration efforts but also expands the possibilities for education and public engagement. Digital models and 3D-printed replicas can be used in exhibitions, allowing more people to experience and learn from cultural heritage without risking damage to the original items. This accessibility fosters a greater appreciation for cultural heritage, encouraging efforts to protect and preserve these invaluable resources.

3. The Role of 3D Laser Scanning and Other Similar Technologies

3D laser scanning technology has emerged as a crucial tool in the preservation of cultural heritage, revolutionizing how we capture and conserve the intricate details of historical artifacts. This cutting-edge technology uses laser light to accurately measure and record the exact size, shape, and minute details of physical objects, creating highly detailed three-dimensional digital representations. These digital models are invaluable for various preservation efforts, including the documentation, restoration, and replication of cultural heritage objects.

The multifaceted role of 3D laser scanning in cultural heritage preservation is pivotal. It provides high-resolution data that serves as the foundation for both documentation and restoration efforts. By creating precise digital models, conservators gain a comprehensive understanding of the physical characteristics and conditions of artifacts. This detailed insight is crucial for developing informed and effective preservation strategies, ensuring that interventions are both accurate and minimally invasive.

Complementing 3D laser scanning is additive manufacturing, commonly known as 3D printing. This technology uses the detailed digital models generated by laser scanning to produce physical replicas of artifacts. The synergy between these technologies allows conservators to recreate exact copies of damaged or missing parts, which can then be

seamlessly integrated into the original artifacts. This process minimizes the intrusion on the original materials, preserving the artifact's authenticity and integrity. [2]

In addition to 3D laser scanning and additive manufacturing, other technologies such as photogrammetry and structured light scanning play significant roles in cultural heritage preservation. Photogrammetry involves taking multiple photographs of an object from different angles and using software to generate a 3D model. Structured light scanning projects a pattern of light onto an object and captures its shape by analyzing the distortion of the pattern. These technologies collectively enhance our ability to document and preserve cultural heritage objects with unprecedented accuracy and fidelity.

The integration of these advanced technologies ensures that the essence and authenticity of cultural artifacts are maintained. High-resolution digital models provide a permanent record of an artifact’s current state, which is essential for long-term preservation and future restoration efforts. Moreover, these technologies enable the creation of detailed replicas that can be used for educational purposes, exhibitions, and research, ensuring that the cultural and historical significance of these objects is accessible to a wider audience.

4. Precision and Detail

One of the most remarkable advantages of 3D laser scanning lies in its capability to capture intricate details with exceptional precision. Traditional methods of documentation and replication often fall short in accurately representing the complex geometries and surface textures of heritage objects, leading to potential loss of critical details. In contrast, 3D laser scanning produces highly detailed digital models with resolutions reaching down to a few micrometers. This extraordinary level of precision ensures that even the smallest and most delicate features are meticulously preserved, enabling restorations that faithfully honor the original craftsmanship.

The detailed digital models generated by 3D scanning are invaluable for conservators, providing a comprehensive reference for assessing the condition of artifacts, planning conservation strategies, and conducting virtual restorations. The ability to examine and manipulate these models in a digital environment allows for thorough analysis and planning without the risk of further damaging the physical objects. This capability is particularly crucial for fragile artifacts, where physical examination might pose significant risks. By leveraging these detailed digital representations, conservators can ensure that restoration efforts are accurate of the artifacts' original state. [3]

The precision of 3D laser scanning technology not only captures the visible aspects of artifacts but also their minute details that might otherwise be overlooked. Every scratch, inscription, and texture is recorded with unparalleled accuracy, creating a digital twin that serves as a permanent and detailed record of the artifact's current condition. This digital archive is invaluable for future reference, providing a baseline against which any changes or deteriorations can be measured over time.

Furthermore, the non-invasive nature of 3D laser scanning is particularly advantageous for delicate or deteriorating artifacts. Traditional methods often require direct contact, which can exacerbate damage or cause wear. 3D laser scanning, however, operates without physical touch, thereby eliminating the risk of further harm during the documentation process. This aspect of 3D laser scanning ensures that the preservation process itself does not contribute to the degradation of the artifacts.

The application of 3D laser scanning extends beyond mere documentation to include the planning and simulation of restoration processes. Conservators can use the digital models to test different restoration techniques in a virtual environment, predicting their outcomes before applying them to the actual artifacts. This approach minimizes the risk of irreversible mistakes and allows for more informed decision-making. By simulating various restoration scenarios, conservators can choose the most appropriate method that aligns with the artifact's historical and material integrity.

5. Non-Invasive

One of the most significant advantages of 3D laser scanning and similar technologies in cultural heritage preservation is their non-invasive nature. Traditional restoration techniques often necessitate physical contact with artifacts, which can pose substantial risks to fragile or deteriorated materials. This contact can lead to further damage, compromising the integrity of these invaluable objects. In stark contrast, 3D laser scanning technology involves no direct contact with the artifacts, dramatically reducing the risk of damage during the documentation process. This non-invasive approach is especially crucial for artifacts that are too delicate to endure traditional conservation methods, ensuring that their integrity is maintained throughout the restoration process.

Moreover, the non-invasive nature of 3D laser scanning allows for repeated scans over time, providing a dynamic means for conservators to monitor the condition of artifacts. This capability enables the detection of subtle changes or signs of deterioration without causing further impact on the objects. Such ongoing monitoring is essential for developing effective long-term preservation strategies, ensuring that interventions can be carefully planned and executed with minimal risk to the artifacts. By integrating non-invasive technologies into their practices, conservators can achieve a higher standard of care for cultural heritage objects, adeptly balancing the imperative of preservation with the necessity to minimize any potential harm. The ability to repeatedly scan artifacts over time provides conservators with a wealth of data that can be used to track and understand the progression of wear and deterioration. This continuous monitoring is invaluable for preemptive conservation measures, allowing conservators to address issues before they become critical. Additionally, the detailed and precise digital records created through 3D laser scanning serve as a permanent archive of the artifacts' current state. These digital models can be used for various purposes, including research, education, and virtual restoration, without risking damage to the original objects. [4]

In essence, the non-invasive nature of 3D laser scanning and similar technologies represents a transformative advancement in the field of cultural heritage preservation. By significantly reducing the risk of damage and enabling continuous monitoring and precise documentation, these technologies allow for a more meticulous and careful approach to conserving our shared cultural heritage. This balance between preservation and protection is crucial for ensuring that future generations can continue to study and appreciate these historical treasures. As such, the adoption of non-invasive technologies marks a crucial step forward in safeguarding the integrity and longevity of our cultural heritage.

6. Speed and Efficiency

Taking into consideration cultural heritage preservation, the integration of 3D laser scanning, and additive manufacturing has brought about a seismic shift, revolutionizing the pace and efficacy of conservation efforts. Gone are the days of laborious, time-consuming documentation and restoration methods, reliant on painstaking manual measurements and artisanal skills. With 3D laser scanning, the intricate details of artifacts can now be captured in a matter of hours, offering conservators a rapid and precise digital representation to guide their restoration endeavors.

Simultaneously, additive manufacturing has slashed the time required for producing physical replicas to a fraction of what traditional methods demanded. This accelerated restoration process empowers conservators to respond swiftly to preservation needs, whether stemming from sudden damage due to natural disasters or as part of long-term restoration projects. Moreover, the ability to swiftly generate accurate digital and physical models fosters collaboration among conservators, historians, and stakeholders, ensuring that preservation efforts are informed, coordinated, and timely.

By harnessing the speed and efficiency of these cutting-edge technologies, the field of cultural heritage preservation can operate with unprecedented effectiveness. Not only can urgent conservation needs be addressed promptly, but strategies for sustainable, long-term preservation can also be carefully planned and executed. In essence, the marriage of 3D laser scanning and additive manufacturing accelerates the journey towards safeguarding our shared cultural legacy, ensuring that it endures for generations to come.

7. Data Preservation

The role of 3D laser scanning extends far beyond physical restoration; it serves as a crucial tool for data preservation. Through the meticulous capture of digital models, 3D laser scanning immortalizes artifacts, providing comprehensive records of their current state. These digital replicas become invaluable resources for future research and conservation endeavors, safeguarding detailed information about cultural heritage objects even if the physical artifacts are lost or further damaged.

The creation of digital archives facilitated by 3D laser scanning revolutionizes access to cultural heritage. Stored and shared electronically, these digital records transcend geographical boundaries, allowing researchers, educators, and the public to study and appreciate artifacts remotely. This newfound accessibility fosters a deeper awareness and appreciation of cultural heritage, igniting efforts to protect and preserve these invaluable resources for future generations.

By creating detailed and accessible digital records, 3D laser scanning ensures that the knowledge and cultural significance of heritage objects endure beyond their physical form. These digital archives not only preserve the essence of cultural artifacts but also enable innovative approaches to engagement and research. In the digital age, data preservation through 3D laser scanning extends the life and impact of cultural artifacts, ensuring that their legacy persists for generations to come. [5]

Through the lens of 3D laser scanning, cultural heritage transcends physical boundaries, finding new life and relevance in the digital realm. As technology continues to evolve, the preservation of cultural heritage will be increasingly intertwined with digital innovation, ensuring that the stories and significance of our collective past endure for future exploration and appreciation.

8. Site Documentation

3D laser scanning is not limited to capturing the intricacies of individual artifacts; it also plays a vital role in documenting entire heritage sites comprehensively. This capability proves particularly invaluable for archaeological sites, historic buildings, and expansive landscapes, where traditional methods of documentation often fall short, presenting challenges and leaving gaps in understanding. Through the creation of detailed digital models, 3D laser scanning provides a thorough record of these sites, capturing their spatial relationships, structural features, and environmental context with unprecedented accuracy and fidelity.

These digital models serve as indispensable tools for site management and conservation planning. They enable conservators and researchers to conduct detailed analyses of site conditions, identify potential threats, and develop targeted preservation strategies. By providing a comprehensive understanding of the site's physical characteristics and vulnerabilities, 3D laser scanning empowers conservation efforts to be more effective and sustainable, ensuring the long-term protection of these invaluable cultural treasures.

Moreover, 3D laser scanning facilitates virtual reconstructions of heritage sites, offering researchers and the public immersive opportunities to explore and comprehend these sites in new and innovative ways. Through virtual reconstructions, individuals can experience the historical context and significance of these sites firsthand, fostering a deeper appreciation and understanding of our cultural heritage. This technology not only brings history to life but also democratizes access to heritage sites, allowing individuals from around the world to engage with and learn from these invaluable resources.

The benefits of 3D laser scanning extend beyond mere documentation; it plays a pivotal role in preserving heritage sites that are difficult to access or at risk of imminent damage. By capturing the complexity and context of these sites, 3D laser scanning ensures that their details are preserved for future generations, even in the face of ongoing threats and challenges. This technology serves as a safeguard against the loss of cultural heritage, preserving these sites in digital form for posterity.

9. Conservation Planning

The utilization of detailed data obtained through 3D laser scanning is indispensable in the meticulous planning of conservation endeavors. The accuracy of digital models furnished by this technology empowers conservators to conduct comprehensive assessments of both artifacts and sites. By scrutinizing these digital replicas, conservators can meticulously identify areas of deterioration, enabling them to develop targeted interventions that address specific issues without unnecessary disruption. This data-driven methodology ensures that conservation efforts are not only precise but also highly effective, as they are tailored to the exact needs of the artifact or site under consideration.

Furthermore, the ability to simulate various restoration techniques within a digital environment further enhances the decision-making process for conservators. By digitally testing different approaches, conservators can anticipate potential outcomes and mitigate the risk of unintended consequences. This is particularly crucial for complex restoration projects, where multiple factors must be carefully weighed to ensure the preservation of the artifact's integrity and historical significance. Through the use of 3D laser scanning, conservators can establish a robust foundation for conservation planning, enabling them to develop strategic and effective conservation plans that yield successful outcomes.

3D laser scanning revolutionizes the conservation planning process by providing conservators with invaluable tools and insights. By leveraging the detailed data and digital models generated through this technology, conservators can navigate the complexities of conservation projects with unparalleled precision and foresight. Ultimately, this enables them to safeguard the cultural heritage entrusted to their care in a manner that is both appropriate and successful, ensuring its preservation for future generations to cherish and appreciate.

10. Structural Analysis

3D laser scanning has emerged as a transformative tool in the realm of conservation planning, offering conservators unprecedented capabilities and insights. Through the meticulous capture of detailed data and the generation of digital models, this technology equips conservators with invaluable resources to navigate the intricate challenges inherent in conservation projects. By harnessing the power of 3D laser scanning, conservators can

achieve a level of precision and foresight that was previously unattainable, revolutionizing their approach to safeguarding cultural heritage.

At the heart of the revolution lies the ability of 3D laser scanning to produce highly accurate digital representations of cultural artifacts and sites. These digital models serve as dynamic reference points, capturing every nuance and intricacy of the object or structure being scanned. With such comprehensive data at their disposal, conservators are empowered to delve into the minutiae of conservation projects, meticulously analyzing each component and assessing its condition with unparalleled detail.

Moreover, the digital models generated by 3D laser scanning facilitate advanced simulation and analysis, enabling conservators to explore various conservation strategies virtually before implementation. Through simulated restoration scenarios, conservators can anticipate potential outcomes and refine their approaches with precision and foresight. This proactive approach minimizes the risk of unintended consequences and ensures that conservation efforts are both appropriate and successful.

Beyond its utility in individual conservation projects, 3D laser scanning also enhances collaboration and knowledge sharing within the conservation community. By creating digital archives of cultural artifacts and sites, this technology fosters a wealth of information that can be accessed and utilized by conservators worldwide. The ability to share digital models and data facilitates interdisciplinary collaboration and enables conservators to draw upon a global pool of expertise and resources, enriching their conservation efforts and expanding their collective knowledge base.

Ultimately, the integration of 3D laser scanning into the conservation planning process represents a monumental leap forward in the preservation of cultural heritage. By providing conservators with unprecedented tools and insights, this technology empowers them to navigate the complexities of conservation projects with precision and foresight. In doing so, it ensures that the cultural heritage entrusted to their care is safeguarded in a manner that is both responsible and effective, preserving it for future generations to cherish and appreciate. As 3D laser scanning continues to evolve and advance, its impact on the field of conservation planning will only grow, heralding a new era of preservation and stewardship for our shared cultural legacy.

11. Restoration and Replication

The restoration and replication of cultural heritage objects represents a paramount application of 3D laser scanning and additive manufacturing technologies, ushering in a new era of preservation and accessibility. These cutting-edge tools empower conservators to meticulously recreate accurate replicas of damaged or lost artifacts, thereby revitalizing their original appearance and functionality. Such a transformative capability proves indispensable, especially when faced with the daunting challenge of restoring artifacts that have endured partial destruction or are composed of materials that are no longer readily available.

Additive manufacturing, in particular, emerges as a pivotal component in this process, offering the versatility to produce replicas using materials that closely mimic the composition of the original artifacts. This ensures that the restored objects retain not only their visual authenticity but also their historical significance and cultural value. Whether the artifact is a centuries-old sculpture or a fragment of a once-grand monument, additive manufacturing enables conservators to breathe new life into these relics, preserving their essence for future generations to behold and appreciate.

The flexibility afforded by additive manufacturing allows for the creation of multiple replicas, serving diverse educational and research needs without diminishing the integrity of the originals. These replicas not only facilitate scholarly inquiries and hands-on learning experiences but also act as invaluable teaching aids, fostering a deeper understanding and appreciation of our shared cultural heritage. By democratizing access to these replicas, museums, educational institutions, and researchers can broaden the scope of cultural exploration and engagement, transcending geographical barriers and temporal constraints.

In cases where the original artifacts are too fragile or susceptible to damage, 3D-printed replicas emerge as a viable alternative for exhibition and study purposes. These replicas offer a tangible means of experiencing and interacting with cultural artifacts, bridging the gap between past and present while safeguarding the originals in controlled conservation environments. Whether displayed in museum exhibits or utilized for educational programs, these replicas ensure that the cultural and educational value of the artifacts remains palpable and accessible to all

In essence, the restoration and replication capabilities enabled by 3D laser scanning and additive manufacturing technologies signify a monumental leap forward in cultural heritage preservation. By harnessing the precision and versatility of these tools, conservators can breathe new life into damaged or lost artifacts, preserving their legacy for future generations to cherish and learn from. As we continue to refine and expand upon these technologies, the boundaries of cultural exploration and preservation will be pushed ever further, ensuring that the stories and treasures of our past endure for generations to come.

12. Preserving the Past and the Future - Examples and Counterexamples

The integration of 3D laser scanning and additive manufacturing into cultural heritage preservation marks a significant leap forward in the field. These cutting-edge technologies offer a myriad of benefits, enhancing not only the precision, efficiency, and non-invasiveness of restoration endeavors but also facilitating the creation of comprehensive digital archives. These archives serve as invaluable resources, guaranteeing the conservation of cultural heritage for posterity. By embracing these innovative tools, conservators can now safeguard and restore heritage objects with unprecedented accuracy, ensuring that the rich tapestry of human history remains intact.

Moreover, the utilization of digital technologies in heritage preservation holds promise for fostering greater public engagement and appreciation. Through virtual reconstructions and

digital archives, cultural heritage becomes more accessible to a global audience, transcending geographical boundaries and promoting widespread awareness and support for preservation initiatives. As these technologies continue to evolve and become more accessible, their role in safeguarding the stories and achievements of human civilization will undoubtedly become increasingly vital. By investing in and advancing these technologies, we can ensure that our cultural heritage is not only preserved but also celebrated for generations to come, enriching the collective human experience and fostering a deeper understanding of our shared past.

When considering the usage of 3D laser scanning and direct digital manufacturing (DDM) in the restoration and replication of cultural heritage, we can draw some insightful parallels through examples and counterexamples of good practices. These examples illustrate the benefits and challenges associated with the adoption of advanced technologies in preserving and replicating cultural artifacts.

The famous statues from Galleria dell'Accademia, Florence, Italy, underwent a comprehensive restoration project. 3D laser scanning was employed to create highly detailed digital models, which were used for analysis and to plan the restoration. The technology enabled restorers to understand the statue's condition in minute detail without physical contact, preserving the integrity of the original. Additionally, the digital model provides a permanent record that can be used for future studies and virtual exhibitions.

Fig. 1. Photographs from Galleria dell'Accademia, Florence, Italy[3]

On the other hand, an example of poor practices in the field of restoration is the direct corrosive action on the heritage object in favor of an almost instantaneous result, but with permanently damaging effects. The application of chemical compounds, both synthetic and organic, affects the object's ability to be maintained in an optimal state and also alters its specific characteristics. Moreover, manually created molds for the purpose of casting can damage the original, and may even completely destroy it.

[3] Personal collection

This Roman military pillar from the Carsium Museum, Hârșova, Romania was cleaned using organic solutions, resulting in minimal corrosive effects. Unfortunately, the milder and gentler the applied substance is to the object, the shorter the duration of its effectiveness.

Fig. 2. Carsium Museum, Hârșova, Romania[4]

13. Conclusions

The integration of 3D laser scanning, and additive manufacturing technologies marks a significant turning point in the realm of cultural heritage preservation. These cutting-edge innovations have redefined the approach to restoring and replicating damaged or destroyed cultural artifacts, offering a level of precision, efficiency, and non-invasiveness previously unattainable through traditional methods. By harnessing the power of these technologies, conservators and preservationists can now safeguard the invaluable treasures of our shared cultural heritage for future generations to cherish and learn from.

One of the most profound impacts of 3D laser scanning and additive manufacturing lies in their ability to provide unparalleled precision in the restoration process. Traditional methods often struggled to capture the intricate details and nuances of cultural artifacts, leading to compromises in accuracy and authenticity. However, with 3D laser scanning, conservators can now create highly detailed digital records of artifacts with resolutions down to a few micrometers. This level of precision ensures that even the smallest features are faithfully preserved, allowing for meticulous restorations that honor the original craftsmanship and historical significance of the objects.

Furthermore, the non-invasive nature of 3D laser scanning and additive manufacturing technologies represents a significant advancement in cultural heritage preservation. Unlike traditional restoration techniques that often require physical contact with artifacts, these

[4] Personal collection

innovative methods involve no direct interaction, minimizing the risk of damage or alteration to the original materials. This non-invasive approach is particularly crucial for fragile or deteriorated artifacts that cannot withstand traditional conservation methods, ensuring their integrity is preserved throughout the restoration process. 3D laser scanning and additive manufacturing enable the creation of detailed digital records and accurate physical replicas of cultural artifacts. These digital archives serve as comprehensive repositories of cultural heritage, capturing not only the physical characteristics but also the historical, cultural, and educational significance of the objects. By preserving this wealth of information in digital form, we can ensure that future generations have access to a rich and diverse cultural legacy, fostering greater appreciation and understanding of our collective heritage.

As we continue to refine and develop these tools, their impact on the preservation of cultural heritage will only continue to grow. By embracing these technological advancements, we can safeguard our shared cultural legacy and ensure that it remains accessible for all to appreciate and learn from, now and in the future. Through a combination of precision, efficiency, and non-invasiveness, 3D laser scanning and additive manufacturing are reshaping the landscape of heritage conservation, paving the way for new opportunities to protect and celebrate the treasures of human civilization

References

[1] LUHMANN T., ROBSON S., A KYKE S., BOEHM J. - *Close-Range Photogrammetry and 3D Imaging.* ISBN 978-3110302691. Walter de Gruyter. November 2013.

[2] JAMESON F. - *Archaeologies Of The Future. The desire called utopia and other science fictions.* ISBN 9781844675388. Verso Books. 2007.

[3] IOANNIDES M., FINK E., MOROPOULOU A., HAGEDORN-SAUPE M. - *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection.* ISBN 978-3-319-48495-2. Springer. November 2016.

[4] BOARDMAN C., BRYAN P. - *3D Laser Scanning for Heritage: Advice and Guidance on the Use of Laser Scanning in Archaeology and Architecture (Historic England Guidance).* ISBN 978-1848025219. Historic England. May 2018.

[5] KENNINGTON MCCARTHY J., BENJAMIN J., WINTON T., VAN DUIVENVOORDE W. - *3D Recording and Interpretation for Maritime Archaeology.* ISBN 978-3-030-03635-5. Springer. April 2019.

Bibliography

BOARDMAN C., BRYAN P. - *3D Laser Scanning for Heritage: Advice and Guidance on the Use of Laser Scanning in Archaeology and Architecture (Historic England Guidance).* ISBN 978-1848025219. Historic England. May 2018.

BANON, C., RASPALL, F. – *3D Printing Architecture Workflows, Applications and Trends*. ISBN 978-9-811-58387-2. Springer. 2021

IOANNIDES M., FINK E., MOROPOULOU A., HAGEDORN-SAUPE M. - *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection.* ISBN 978-3-319-48495-2. Springer. November 2016.

JAMESON F. - *Archaeologies Of The Future. The desire called utopia and other science fictions.* ISBN 9781844675388. Verso Books. 2007.

KENNINGTON MCCARTHY J., BENJAMIN J., WINTON T., VAN DUIVENVOORDE W. - *3D Recording and Interpretation for Maritime Archaeology.* ISBN 978-3-030-03635-5. Springer. April 2019.

LUHMANN T., ROBSON S., A KYKE S., BOEHM J. - *Close-Range Photogrammetry and 3D Imaging.* ISBN 978-3110302691. Walter de Gruyter. November 2013.

AI SOLUTIONS FOR SUSTAINABLE TOURISM MANAGEMENT: A COMPREHENSIVE REVIEW

Ioana Cristiana PATRICHI[1]

Abstract

The rapid growth of the tourism industry has brought significant economic benefits but has also raised concerns about its environmental and socio-cultural impacts. Sustainable tourism management is imperative to mitigate these negative effects and promote responsible tourism practices. In recent years, artificial intelligence (AI) has emerged as a powerful tool to address various challenges in sustainable tourism management. This paper presents a comprehensive review of AI solutions applied to sustainable tourism, encompassing resource management, environmental conservation, visitor management, and community engagement. Through an extensive analysis of existing literature, case studies, and implementation strategies, this review examines the effectiveness of AI technologies in optimizing tourism operations while minimizing environmental degradation and maximizing socio-economic benefits. Additionally, the paper discusses the ethical considerations, challenges, and future directions of integrating AI into sustainable tourism management practices.

Keywords: AI, artificial intelligence, sustainable tourism, tourism management, environmental conservation, visitor management, community engagement

JEL Classification: L83, O31, Q01

1. Introduction

In an era marked by rapid technological advancement, artificial intelligence (AI) stands at the forefront, revolutionizing industries and reshaping the way we perceive and interact with the world. AI, which replicates human intelligence through machine processes, has enormous potential to fuel innovation, streamline operations, and advance sustainability across multiple industries. One such industry ripe for transformation is tourism.

The tourism industry, often celebrated for its ability to connect cultures, foster economic growth, and promote environmental awareness, also grapples with significant challenges, particularly concerning sustainability. As global travel continues to surge, concerns regarding carbon emissions, over-tourism, and environmental degradation loom large.

[1] Lecturer PhD, Romanian-American University, ioana.cristiana.patrichi@rau.ro

However, amidst these challenges lies an opportunity for AI to emerge as a catalyst for positive change.

AI offers a plethora of tools and solutions that can empower the tourism sector to operate more sustainably. By leveraging AI-driven analytics, businesses can gain deeper insights into consumer behavior, preferences, and trends, enabling them to optimize resource allocation, reduce waste, and minimize environmental impact. Advanced data analytics can also facilitate the development of personalized travel experiences tailored to individual preferences, thereby promoting more responsible and meaningful tourism.

Furthermore, AI-driven systems have the capability to improve operational efficiency and optimize processes across the entire tourism value chain. From dynamic pricing algorithms that optimize revenue and resource utilization to intelligent logistics management systems that optimize transportation routes and minimize fuel consumption, AI holds the key to unlocking new levels of efficiency and sustainability.

Moreover, advancements in natural language processing and virtual assistants propelled by AI are transforming customer service and communication within the tourism sector. By providing travelers with real-time support, personalized recommendations, and interactive experiences, AI-powered chatbots and virtual assistants not only enhance customer satisfaction but also promote sustainable travel practices by disseminating relevant information and promoting eco-conscious behaviors.

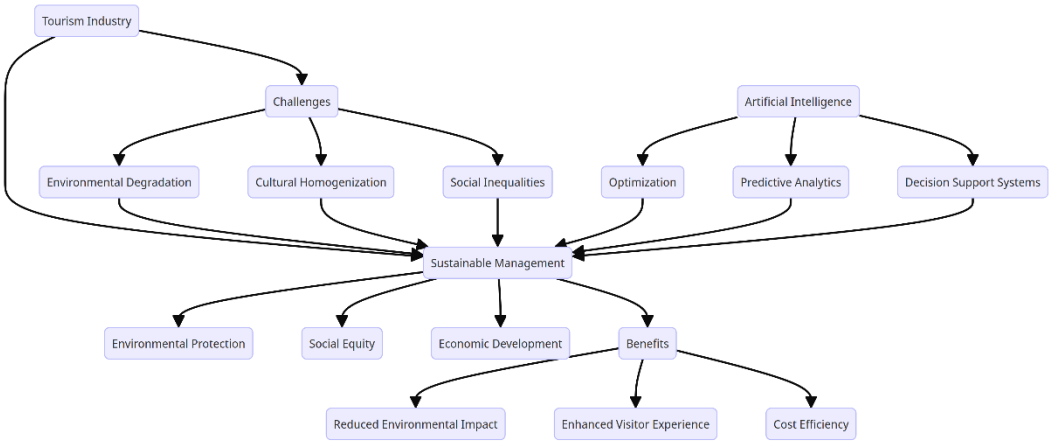


Figure 1. Interconnections Between Tourism, Sustainable Management, and AI Integration

In essence, AI has the potential to transform the tourism industry into a more sustainable and responsible sector, where economic prosperity coexists harmoniously with environmental stewardship and social well-being. By harnessing the power of AI-driven technologies, stakeholders within the tourism ecosystem can pioneer innovative solutions that preserve natural resources, protect cultural heritage, and foster authentic, enriching travel experiences for generations to come. As we embark on this journey towards a more

sustainable future, AI stands as a beacon of hope, guiding us towards a world where tourism serves as a force for good, enriching lives while safeguarding the planet.

The integration of AI technologies holds promise as a transformative force in revolutionizing how we conceive, plan, and manage sustainable tourism endeavors. Against this backdrop, this paper embarks on a comprehensive journey to unravel the multifaceted dimensions of AI solutions in the realm of sustainable tourism management. Through a meticulous examination of their applications, inherent advantages, persistent challenges, and future trajectories, this review seeks to illuminate the pivotal role of AI in charting a course towards a more sustainable and equitable tourism future.

2. Literature Review

The application of artificial intelligence (AI) in the domain of sustainable tourism management has gained increasing attention in recent years due to its potential to address complex challenges and optimize decision-making processes. This section provides an overview of existing literature on AI solutions for sustainable tourism management, focusing on key themes, methodologies, and findings. Artificial intelligence comprises a range of technologies designed to mimic human intelligence for problem-solving purposes (Bulchand, 2023; Tsaih, 2018); using AI technologies, businesses can automate routine tasks and enhance resource allocation (Garcia, 2023). According to Sousa (2024) the applications of AI in tourism and hospitality industry may take different forms, from customer service delivery robots in hotels and restaurants which can speed up some business processes and operations (Peric, 2021), to the use of technology at the check-in, check-out processes at the reception or conversational systems - chatbots and voice assistants (Bulchand, 2020). AI can also optimize energy consumption, water usage, and waste management in tourism facilities, thereby reducing operational costs and minimizing environmental footprints (Hussain, 2024).

Numerous studies have explored the use of AI technologies, such as machine learning and predictive analytics, in optimizing resource management in tourism destinations. For example, Li et al. (2019) developed a predictive model using machine learning algorithms to forecast tourist arrivals and optimize hotel room pricing strategies. The study demonstrated that AI-based approaches can improve revenue management and resource allocation in the hospitality sector, leading to enhanced sustainability and profitability. Also, AI-driven systems can work for efficient traffic management (Prahadeeswaran, 2023) and forecast and regulate the flow of tourists in popular destinations, preventing overcrowding (Milton, 2023).

Similarly, Wang et al. (2020) conducted a case study on the application of AI-driven smart energy systems in tourism destinations. By analyzing historical energy consumption data and weather patterns, the researchers developed an AI-powered energy optimization model

that reduced electricity costs and carbon emissions in hotels and resorts. The study highlights the potential of AI technologies to promote energy efficiency and environmental sustainability in the tourism sector.

In the context of environmental conservation, AI-enabled monitoring and surveillance systems have emerged as valuable tools for protecting natural ecosystems and biodiversity in tourism destinations. Zhang et al. (2021) utilized remote sensing technologies and machine learning algorithms to monitor deforestation activities in a protected area. The researchers demonstrated that AI-based systems could accurately detect illegal logging activities and facilitate timely intervention by conservation authorities, thereby mitigating environmental degradation and habitat loss.

Additionally, Liang et al. (2021) conducted a study on the use of AI-powered drones for wildlife conservation in ecotourism destinations. By deploying drones equipped with thermal imaging cameras and object recognition algorithms, the researchers monitored endangered species populations and identified poaching threats in real time. The findings underscore the potential of AI technologies to enhance conservation efforts and safeguard biodiversity in tourism hotspots.

The efficient management of tourist flows, and visitor experiences is essential for ensuring sustainable tourism development. AI-based analytics tools have been increasingly utilized to analyze visitor behavior patterns, preferences, and sentiments in tourism destinations. For instance, Kim et al. (2019) conducted a study on sentiment analysis of tourist reviews using natural language processing techniques. The researchers extracted valuable insights from online reviews to understand visitor perceptions, preferences, and satisfaction levels, enabling destination managers to tailor marketing strategies and enhance visitor experiences.

Similarly, Chen et al. (2020) developed an AI-powered recommendation system for personalized tourist itineraries. By analyzing historical travel data and user preferences, the researchers generated customized travel recommendations for tourists, optimizing itinerary planning and resource allocation. The study demonstrates the potential of AI technologies to improve the quality of tourist experiences and reduce overcrowding in popular destinations.

While AI solutions offer promising opportunities for sustainable tourism management, several challenges and ethical considerations need to be addressed. These include data privacy concerns, algorithmic biases, technological dependence, and socio-cultural impacts on local communities. Future research should focus on developing robust AI algorithms, enhancing data governance frameworks, and fostering interdisciplinary collaboration to ensure equitable and sustainable tourism development.

Overall, the literature reviewed in this section highlights the diverse applications and benefits of AI solutions for sustainable tourism management. By leveraging AI technologies, destination stakeholders can optimize resource utilization, improve visitor

experiences, and protect natural and cultural heritage, thereby contributing to the long-term sustainability of the tourism industry.

3. AI Applications in Sustainable Tourism Management

In the contemporary tourism industry, the incorporation of cutting-edge technologies has emerged as a necessity for the implementation of sustainable management practices. Among these technological advancements, Artificial Intelligence (AI) holds significant promise, presenting novel solutions to tackle intricate challenges inherent in sustainable tourism management.

1. Machine Learning (ML)

Machine learning, a branch of artificial intelligence, allows systems to learn and improve from experience without explicit programming. In sustainable tourism management, machine learning algorithms are employed to analyze vast amounts of data and identify patterns, trends, and insights that can inform decision-making and enhance sustainability efforts.

Applications:

Predictive Modeling: Machine learning models can forecast tourist arrivals, predict visitor behavior, and anticipate demand for tourism services, enabling destinations to optimize resource allocation and plan more effectively.

Recommendation Systems: ML algorithms power recommendation engines that suggest personalized travel itineraries, accommodation options, and tourist attractions based on individual preferences and past behavior.

Natural Resource Management: Machine learning is used to analyze environmental data, such as weather patterns, biodiversity metrics, and water quality measurements, to support conservation efforts and sustainable resource management in tourism destinations.

2. Natural Language Processing (NLP)

Natural language processing is a branch of AI that focuses on enabling computers to understand, interpret, and generate human language. In sustainable tourism management, NLP technologies are leveraged to analyze textual data from various sources, including social media, customer reviews, and online forums, to extract insights and sentiment related to tourism experiences and destinations.

Applications:

Sentiment Analysis: NLP algorithms analyze online reviews, social media posts, and customer feedback to gauge public sentiment towards tourism destinations, attractions, and experiences. This information can help destinations identify areas for improvement and address visitor concerns.

Chatbots and Virtual Assistants: NLP-powered chatbots and virtual assistants interact with tourists in natural language, answering queries, providing recommendations, and assisting with travel planning. These AI-driven assistants enhance customer service and engagement while reducing the workload on human staff.

Language Translation: NLP technologies facilitate language translation services for tourists, enabling them to communicate effectively with locals, access information in their native language, and navigate foreign destinations with ease.

3. Computer Vision

Computer vision is a field of AI that enables computers to interpret and analyze visual information from images and videos. In sustainable tourism management, computer vision technologies are utilized to analyze visual data captured by cameras, drones, and satellite imagery to monitor environmental conditions, assess tourist behavior, and enhance safety and security measures.

Applications:

Environmental Monitoring: Computer vision systems analyze satellite images and aerial photographs to track changes in land use, deforestation, coastal erosion, and other environmental indicators, providing valuable insights for conservation and sustainable land management.

Crowd Monitoring: Computer vision algorithms analyze video feeds from surveillance cameras to monitor crowd density, detect congestion hotspots, and optimize crowd management strategies to ensure visitor safety and enhance the tourist experience.

Security Surveillance: Computer vision technologies enable the automated detection of security threats and suspicious activities in tourist areas, helping authorities respond swiftly to incidents and maintain a safe and secure environment for visitors.

4. Predictive Analytics

Predictive analytics involves using statistical techniques and machine learning algorithms to analyze historical data and make predictions about future events or trends. In sustainable tourism management, predictive analytics is applied to anticipate tourist demand, optimize resource allocation, and mitigate environmental impacts.

Applications:

Demand Forecasting: Predictive analytics models analyze historical booking data, tourist arrivals, and market trends to forecast future demand for tourism services, accommodation, and attractions. These insights enable destinations to adjust pricing, capacity, and marketing strategies to meet demand fluctuations and optimize revenue.

Environmental Impact Assessment: Predictive analytics is used to assess the potential environmental impacts of tourism activities, such as carbon emissions, waste generation, and water consumption. By modeling different scenarios and outcomes, destinations can identify sustainable practices and policies to minimize negative environmental effects.

Risk Management: Predictive analytics helps tourism stakeholders identify and mitigate risks associated with natural disasters, climate change, and other disruptive events. By analyzing historical data and risk factors, destinations can develop contingency plans, enhance resilience, and ensure the safety and well-being of tourists and residents alike.

4. AI Solutions Driving Sustainable Tourism: Insights from the Hotel Industry and Destination Management

The hotel industry utilizes AI applications in sustainable tourism management to minimize environmental impact, optimize resource utilization, and promote responsible tourism practices. In a survey conducted in 2023 among hotel chains globally, artificial intelligence (AI) emerged as the primary area for anticipated innovations over the following two years. Approximately 86 percent of participants acknowledged this trend, while 74 percent of respondents highlighted energy management as another significant area of focus (Figure 2.).

For example, in the hospitality industry AI-powered systems analyze real-time data from sensors and smart meters to optimize energy consumption within hotel properties. These systems can adjust heating, ventilation, and air conditioning (HVAC) systems, lighting, and other energy-consuming devices based on occupancy levels, weather conditions, and time of day. By optimizing energy usage, hotels can reduce their carbon footprint and lower operational costs.

Also, AI algorithms analyze historical data to identify patterns and trends in waste generation within hotel properties. By predicting peak times of waste production and identifying areas of inefficiency, hotels can implement more effective waste management strategies. This may include recycling programs, composting initiatives, and the use of AI-powered sorting systems to minimize landfill waste.

And finally, AI technologies monitor water usage patterns within hotels and identify opportunities for conservation. By analyzing data from sensors and meters, AI systems can detect leaks, identify inefficiencies in water distribution systems, and optimize irrigation practices in landscaping. Additionally, AI-powered algorithms can adjust water flow rates in guest rooms and public areas to minimize waste without compromising guest comfort.

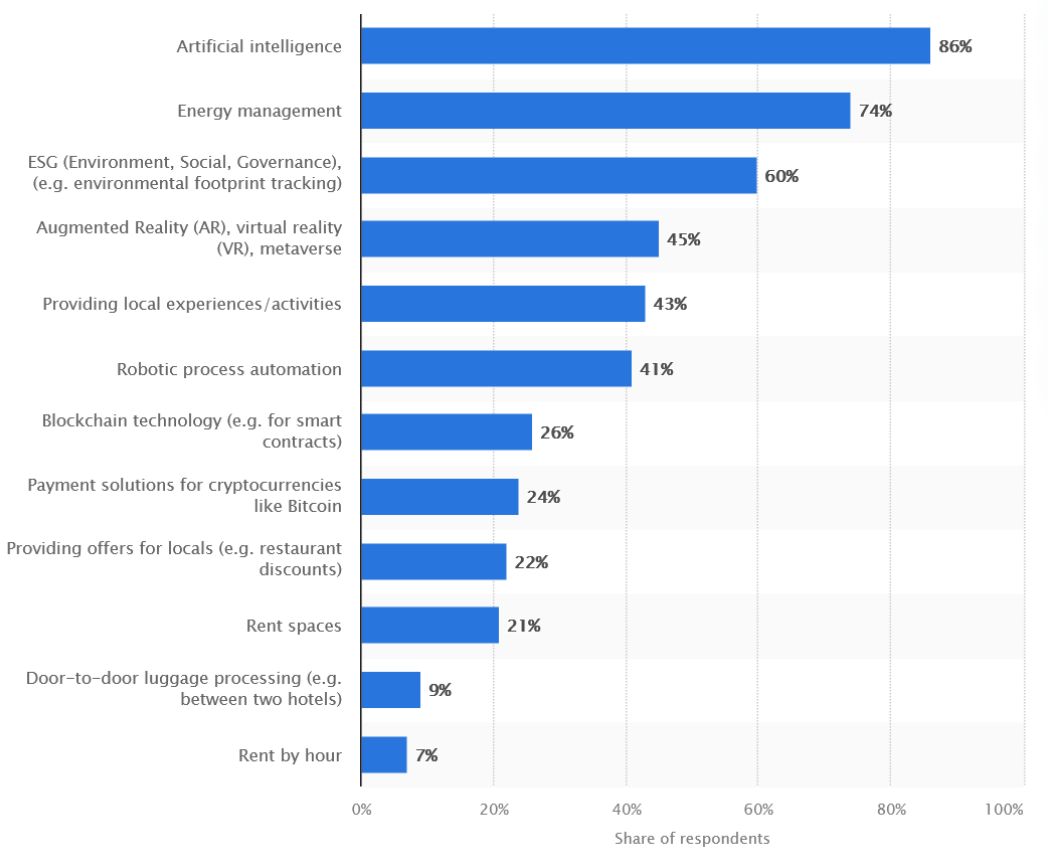


Figure 2. Hotel chains' areas with the most innovations arising in the next two years worldwide in 2023[2]

When it comes to destination management AI can also play a major role in promoting sustainable development. Hence, several crucial factors are pivotal for the effective implementation of AI, including:

Resource Management:

AI technologies such as machine learning and data analytics are being used to optimize resource allocation and utilization in tourism destinations. Predictive models can forecast tourist demand, enabling authorities to better manage infrastructure, transportation, and natural resources. For example, smart energy systems powered by AI algorithms can optimize energy consumption in hotels and resorts, reducing carbon emissions and operating costs.

Environmental Conservation:

[2] Source: Statista, https://www.statista.com/statistics/1453944/hotel-chains-most-innovative-services-worldwide/, 2024.

AI-based systems play a crucial role in monitoring and protecting natural ecosystems in tourism destinations. By integrating remote sensing technologies with machine learning algorithms, it is possible to monitor biodiversity, deforestation, and habitat degradation in real-time. Intelligent conservation drones equipped with AI-powered cameras can identify illegal activities such as poaching and deforestation, facilitating timely intervention by conservation authorities.

Visitor Management:

AI-powered analytics tools are transforming visitor management practices in tourism destinations, enhancing the visitor experience while minimizing overcrowding and congestion. Smart destination management systems use data from various sources, including mobile devices, social media, and sensors, to analyze visitor behavior patterns and preferences. This information enables destination managers to implement targeted marketing campaigns, dynamic pricing strategies, and crowd control measures to ensure a sustainable and enjoyable visitor experience.

Community Engagement

AI technologies are being employed to promote community participation and empowerment in tourism planning and decision-making processes. Residents can use social media sentiment analysis and online forums to express their thoughts, preferences, and suggestions about tourism development. AI-driven virtual reality platforms can facilitate community consultations and public engagement activities, fostering a sense of ownership and inclusivity among local stakeholders.

5. Discussion on How the AI solutions can be improved

Improving AI solutions for sustainable tourism management involves addressing several key areas to enhance their effectiveness, efficiency, and ethical considerations. Here are some points for discussion:

Data Quality and Availability:

Enhanced Data Collection: Improving data collection methods to gather high-quality, real-time data on environmental, social, and economic aspects relevant to sustainable tourism.

Data Sharing and Collaboration: Encouraging collaboration among stakeholders to share data and insights, facilitating the development of more comprehensive AI models and solutions.

Algorithmic Transparency and Interpretability:

Explainable AI: Developing AI models that provide transparent explanations for their decisions and recommendations, enhancing trust and accountability in the decision-making process.

Interdisciplinary Research: Integrating expertise from diverse fields such as computer science, environmental science, and social sciences to ensure AI solutions consider multi-dimensional impacts and trade-offs.

Ethical Considerations:

Ethical AI Frameworks: Establishing ethical guidelines and frameworks for the development and deployment of AI solutions in sustainable tourism management, addressing issues such as fairness, accountability, and privacy.

Community Engagement: Involving local communities and indigenous peoples in the design and implementation of AI solutions to ensure their perspectives and interests are represented and respected.

Scalability and Adaptability:

Modular and Scalable Solutions: Designing AI solutions that can be easily scaled and adapted to different contexts and locations, considering the diversity of tourism destinations and their specific sustainability challenges.

Continuous Learning and Improvement: Implementing feedback mechanisms to allow AI models to continuously learn and adapt based on new data and evolving sustainability priorities.

Human-AI Collaboration:

Augmented Intelligence: Fostering collaboration between AI systems and human experts to leverage the strengths of both, combining AI's computational power with human creativity, intuition, and contextual understanding.

User-Centered Design: Designing AI interfaces and tools with usability and user experience in mind, ensuring they are intuitive and accessible to a wide range of stakeholders, including policymakers, tourism operators, and local communities.

Policy and Regulatory Frameworks:

Regulatory Oversight: Establishing regulatory frameworks to govern the development and deployment of AI solutions in sustainable tourism, balancing innovation with the protection of human rights, environmental integrity, and cultural heritage.

Incentive Mechanisms: Implementing incentives and rewards for businesses and organizations that adopt and implement AI solutions that contribute to sustainable tourism goals, encouraging widespread adoption and investment in this area.

Advancing AI solutions for sustainable tourism management demands a concerted effort to address key facets spanning data quality, algorithmic transparency, ethical considerations, scalability, human-AI collaboration, and regulatory frameworks. Elevating data collection standards and fostering collaboration among stakeholders can bolster the foundation of AI models. Emphasizing explainability and interdisciplinary research ensures decisions align

with ethical principles and encompass diverse perspectives. Modular design principles and continuous learning mechanisms enable adaptability to varied contexts and evolving priorities. Human-AI collaboration harnesses the complementary strengths of both, while user-centered design enhances accessibility and usability. Finally, robust policy frameworks and incentive mechanisms are pivotal in guiding responsible AI deployment and fostering widespread adoption, thereby contributing to the overarching goal of sustainable tourism management.

6. Conclusion

Despite their potential benefits, AI solutions for sustainable tourism management face several multifaceted challenges and ethical considerations. These include intricate issues surrounding data privacy, the inherent biases embedded within algorithms, the growing dependency on advanced technologies, and the profound socio-cultural impacts on local communities. Additionally, the implementation of AI technologies demands substantial investment in sophisticated infrastructure, comprehensive capacity building, and meticulous stakeholder engagement.

Nonetheless, the tangible benefits of integrating AI into sustainable tourism management are compelling and multifarious. For instance:

Enhanced Decision-Making: AI algorithms possess the capacity to analyze vast and complex datasets, generating nuanced insights and sophisticated recommendations for strategic tourism management. This enables stakeholders to make data-driven decisions that harmonize economic growth with environmental stewardship.

Improved Resource Efficiency: AI-driven systems can precisely optimize resource allocation, streamline energy consumption, and innovate waste management practices. This leads to significant cost savings and mitigates the environmental impact of tourism operations.

Enhanced Visitor Experience: AI technologies facilitate personalized recommendations, implement dynamic pricing models, and deploy advanced crowd management solutions. These enhancements elevate the visitor experience by tailoring services to individual preferences and mitigating the adverse effects of overcrowding in popular destinations.

Strengthened Conservation Efforts: AI enables real-time and predictive monitoring of environmental indicators, supporting preemptive and adaptive management strategies to safeguard biodiversity and preserve natural habitats. This technology-driven approach bolsters long-term conservation objectives and enhances ecological resilience.

To fully harness the transformative potential of AI in sustainable tourism management, it is imperative to prioritize interdisciplinary research, foster collaborative partnerships, and promote comprehensive knowledge sharing among academic institutions, industry leaders, governmental bodies, and civil society organizations. Future research endeavors should

focus on refining AI algorithms to ensure robustness, advancing data governance frameworks to protect privacy and promote transparency, and critically examining the socio-cultural implications of AI deployment to foster inclusivity and equity in tourism development.

AI solutions present a myriad of opportunities to revolutionize sustainable tourism management practices. They empower destination stakeholders to optimize resource utilization, enhance the visitor experience, and protect both natural and cultural heritage. However, surmounting the associated challenges and addressing the ethical considerations intrinsic to AI implementation are crucial for ensuring that tourism development is socially inclusive, environmentally responsible, and economically viable over the long term. By fostering innovation, interdisciplinary collaboration, and ethical stewardship, stakeholders can leverage the capabilities of AI to cultivate a more sustainable, resilient, and equitable tourism industry.

References

[1] Bulchand-Gidumal, J., Secin, E., O'Connor, P., & Buhalis, D., *Artificial intelligence's impact on hospitality and tourism marketing: Exploring key themes and addressing challenges.* Current Issues in Tourism, https://doi.org/10.1080/13683500.2023.2229480, (2023).

[2] Tsaih, Rua-Huan, and Chih Chun Hsu. *"Artificial intelligence in smart tourism: A conceptual framework*." (2018).

[3] García-Madurga, M.-Á.; Grilló-Méndez, A.-J., *Artificial Intelligence in the Tourism Industry: An Overview of Reviews*. *Adm. Sci. 13*, 172. https://doi.org/10.3390/admsci13080172, 2023.

[4] Sousa, Ana & Pais, Sónia & Viana, Ana, *Applications of Artificial Intelligence in Tourism and Hospitality: A Systematic Literature Review*. 10.1007/978-3-031-44131-8_29, (2024).

[5] Perić, M.; Vitezić, V. *Tourism Getting Back to Life after COVID-19: Can Artificial Intelligence Help? Societies 11,* 115. https://doi.org/10.3390/soc11040115, 2021.

[6] Bulchand-Gidumal, Jacques, *Impact of Artificial Intelligence in Travel, Tourism, and Hospitality*, 10.1007/978-3-030-05324-6_110-1, 2020.

[7] Hussain, Khalid & Arsalan, Huzaifa, *Sustainable Tourism Management: Leveraging AI for Enhanced Customer Satisfaction,* 10.13140/RG.2.2.36272.57608, (2024).

[8] Prahadeeswaran, R. "A Comprehensive Review: The Convergence of Artificial Intelligence and Tourism." *International Journal for Multidimensional Research Perspectives* 1.2 (2023): 12-24.

[9] Milton, T & Philosophers, Chandigarh. *Artificial Intelligence in Tourism-A review of Trends Opportunities and Challenges*. 1. 01-11, IJMRP \ Vol. 1 \ Issue 2 \October 2023 www.chandigarhphilosophers.com ISSN:2584-0029 (online) (2023).

[10] Wang, Wei & Kumar, Neeraj & Chen, Jun-xin & Gong, Zhiguo & Kong, Xiangjie & Wei, Wei & Gao, Honghao. (2020). Realizing the Potential of Internet of Things for Smart Tourism with 5G and AI. IEEE Network. PP. 1-7. 10.1109/MNET.011.2000250.

[11] Zhang J., Wang Z., Bai L., Song G., Tao J. and Chen L., "*Deforestation Detection Based on U-Net and LSTM in Optical Satellite Remote Sensing Images," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS,* Brussels, Belgium, 2021, pp. 3753-3756, doi: 10.1109/IGARSS47720.2021.9554689.

[12] Kim, K., Park, O., Barr, J. and Yun, H., *"Tourists' shifting perceptions of UNESCO heritage sites: lessons from Jeju Island-South Korea"*, Tourism Review, Vol. 74 No. 1, pp. 20-29. https://doi.org/10.1108/TR-09-2017-0140, 2019.

[13] Chen, Lei, et al. "Personalized itinerary recommendation: Deep and collaborative learning with textual information." *Expert Systems with Applications* 144 (2020): 113070.

Bibliography

Kannan, Ragul & Philosophers, Chandigarh. *Revolutionizing the Tourism Industry through Artificial Intelligence: A Comprehensive Review of AI Integration, Impact on Customer Experience, Operational Efficiency, and Future Trends*, IJMRP \ Vol. 2 \ Issue 2 \February 2024 www.chandigarhphilosophers.com ISSN:2584-0029 (online) (2024).

Law, R., Li, G., Fong, D.K.C. and Han, X. (2019), "Tourism demand forecasting: a deep learning approach", Annals of Tourism Research, Vol. 75, pp. 410-423, doi: 10.1016/j.annals.2019.01.014.

Loureiro, S.M.C.; Nascimento, J. Shaping a View on the Influence of Technologies on Sustainable Tourism. *Sustainability* **2021**, *13*, 12691. https://doi.org/10.3390/su132212691

Majid, G. M., Tussyadiah, I., Kim, Y. R., & Pal, A. (2023). Intelligent automation for sustainable tourism: a systematic review. *Journal of Sustainable Tourism, 31*(11), 2421–2440. https://doi.org/10.1080/09669582.2023.2246681

Nannelli, M., Capone, F., & Lazzeretti, L. (2023). Artificial intelligence in hospitality and tourism. State of the art and future research avenues. *European Planning Studies, 31*(7), 1325–1344. https://doi.org/10.1080/09654313.2023.2180321

Shaolong Sun, Yunjie Wei, Kwok-Leung Tsui, Shouyang Wang, Forecasting tourist arrivals with machine learning and internet search index, Tourism Management, Volume 70, 2019,

Pages 1-10, ISSN 0261-5177, https://doi.org/10.1016/j.tourman.2018.07.010.

https://www.statista.com/statistics/1453944/hotel-chains-most-innovative-services-worldwide/, Statista oficial website, may 2024.

TRAFFIC-SIGN RECOGNITION

Marcel PRODAN[1]

Gabriel DOROBANȚU[2]

Narcis IONIȚĂ[3]

Mihai-Lucian VONCILĂ[4]

Nicolae TARBĂ[5]

Costin-Anton BOIANGIU[6]

Nicolae GOGA[7]

Abstract

Traffic-sign recognition is critical for vehicle safety applications, especially as self-driving cars become a reality. This paper proposes a solution based on existing approaches, utilizing deep learning and computer vision preprocessing to create a real-time algorithm that addresses the limitations of previous methods. The proposed algorithm aims to overcome as many drawbacks as possible and serve as a core component of advanced driver-assistance systems (ADAS). The proposed method is evaluated using the German Traffic Sign Recognition Benchmark (GTSRB) and the Belgium Traffic Sign Dataset (BTSD). This study concludes with a fully functional pipeline that can inspire the development of driving assistants and advance the future of self-driving cars.

Keywords: traffic-sign recognition, detection, neural networks, deep learning

JEL Classification: C45, C88

[1] PhD Student, Faculty of Automatic Control and Computers, University "Politehnica" of Bucharest, Romania, e-mail: marcoprod@gmail.com

[2] Student, Faculty of Automatic Control and Computers, University "Politehnica" of Bucharest, Romania, e-mail: gabriel.dorobantu@stud.mec.upb.ro

[3] Student, Faculty of Automatic Control and Computers, University "Politehnica" of Bucharest, Romania, e-mail: narcis.ionita2511@stud.fsa.upb.ro

[4] PhD Student, Faculty of Automatic Control and Computers, University "Politehnica" of Bucharest, Romania, e-mail: mihai_lucian.voncila@stud.acs.upb.ro

[5] PhD Student, Faculty of Automatic Control and Computers, University "Politehnica" of Bucharest, Romania, e-mail: nicolae.tarba@upb.ro

[6] Professor, Faculty of Automatic Control and Computers, University "Politehnica" of Bucharest, Romania, e-mail: costin.boiangiu@cs.pub.ro

[7] Professor, Faculty of Engineering in Foreign Languages, University "Politehnica" of Bucharest, Romania, e-mail: nicu.goga@upb.ro

1. Introduction

According to the World Health Organization's statistic on road traffic injuries [1], 1.3 million people die annually due to road crashes, with 93% of these crashes occurring in less developed countries where older cars with less equipped safety technology are in circulation. Additionally, between 20 to 50 million people get injured each year, leading to a 3% cost of a country's gross domestic product. This highlights the importance of detection and recognition of road signs in the real world. The European New Car Assessment Programme (Euro NCAP) has made it mandatory for all new cars sold in the EU to be equipped with this type of technology soon [2], as they place great value on car safety and have conducted surveys and safety campaigns regarding ADAS, stating that cars of the future need "readable" roads [3]. Detecting different signs in varying weather, daytime and road conditions is seen as a challenge, as current advanced driver assistance systems only have a defined subset of possible signs. Unfortunately, no comprehensive unbiased comparison of sign detecting algorithms has been implemented, and the slow development of this feature may be attributable to a lack of a big freely available benchmark data set.

The recognition process can be divided into two steps: detection and classification. When comparing the two, detection takes precedence because state-of-the-art classification systems can only compete with humans at best. Therefore, classification can be regarded as solved, at least for the time being [4][5]. While most of the attention of sign detection is on particular shapes, such as rectangles and circles, and the type of the sign (speed limit traffic signs), an additional system should be put in place when not focused on a single type of road signs. In most cases, the system uses a color-based segmentation, followed by a recognition stage. However, a large training database with plenty of road signs is required for this approach to work. To minimize the size of the database and get the intended learning process results, the color-based segmentation can be replaced by a combination of color and shape detection.

Key Takeaways:

Global Road Traffic Statistics: Annually, 1.3 million deaths and 20-50 million injuries from road crashes, mostly in less developed countries with older, less safe cars.
Economic Impact: Road accidents cost up to 3% of a country's GDP.
Importance of Road Sign Detection: Essential for reducing accidents; European Union mandates advanced detection technology in new cars.
Challenges in Sign Detection: Varying weather, daylight, road conditions, and limited sign recognition in current systems.
Lack of Comprehensive Studies: No extensive unbiased comparison of sign detection algorithms; hindered by lack of a large, free benchmark dataset.
Recognition Process: Involves detection and classification, with detection being more crucial as classification technology matches human performance.

Focus on Sign Shapes: Current systems primarily recognize shapes like rectangles and circles, specifically for speed limit signs.
Proposed Improvement: Combining color and shape detection for a more effective and compact training database.

2. Related work

Early methods in traffic sign detection utilized a set of rules that im- posed restrictions on color and shape, as well as required signs to appear only in certain regions of an image. These regions are considered candidates, which are then recognized based on a template matching method using other images. Michael Shneier's article [6] on road sign detection is an example of this approach, where his algorithm performed sufficiently to be used and considered in real-time. However, it only addressed warning signs and a few regulatory signs, and its performance suffered when faced with blurry or affected images. Advancements in machine learning have led to the emergence of new approaches in traffic sign recognition. Many articles have proposed different methods that utilize support-vector machines (SVM) or convolutional networks (CNN). For instance, David Soendoro and Iping Supriana's article proposed an SVM method for classifying binary images with localized traffic signs, which resulted from a color-based method with CIELab + hue [7]. A more recent approach used a CNN with fewer parameters, smaller models, and easier training, achieving high accuracy of close to 97%, better than a classical convolutional network [8].

A new and innovative approach to traffic sign recognition is a CNN method that uses GPGPU [9] and Nvidia's latest solution in the automotive industry for autonomous vehicles, called Nvidia DRIVE [10]. This method focuses on solving severe illumination problems regarding low light or wide variance of light-like reflection, in images captured from real-world.

Method	Accuracy
CNN with 3 Spatial Transformers [17]	99.71%
Committee of CNNs [15]	99.46%
COSFIRE filters for object recognition [18]	98.97%
Human Performance [5]	98.84%
Multi-Scale CNNs [16]	98.31%

Method	Accuracy
Random Forests [19]	96.14%
LDA on HOG 2 [5]	95.68%

Table 1. Performance of various methods in the IJCNN2011 Competition

Traditional methods have also made use of a wide range of handcrafted techniques, such as distinctive shapes and colors, such as HOG [11][12] or SIFT [13][14], for classification with machine learning models like SVM, tree classifiers, and boosting. Several traffic sign recognition contests have been established to encourage participation from researchers in a variety of disciplines. The objective of one of these contests, GTSRB, was to produce a report on comparing learning algorithms for traffic sign recognition [5].

By integrating many deep convolutional neural network columns and preprocessing input images into as many little blocks as possible, IDSIA [15] was able to obtain an error rate of 0.54%. COSFIRE, on the other hand, used local and global characteristics with multi-scale CNNs to achieve an error rate of 1.03% [16]. The GTSRB dataset comprises images where the traffic sign takes up a significant amount of the image. However, in the actual world, categorizing images where the traffic sign only takes up a tiny portion of the traffic scene is more significant and should be the primary focus of studies.

Key Takeaways:

Early Traffic Sign Detection Methods: Utilized rule-based systems focusing on color, shape, and specific image regions; relied on template matching for recognition.
Limitations of Early Methods: Targeted mainly warning and regulatory signs; struggled with blurry or affected images.
Advancements with Machine Learning: Introduction of methods using Support-Vector Machines (SVM) and Convolutional Networks (CNN) for improved recognition.
Specific Machine Learning Approaches:
SVM for classifying binary images of localized traffic signs.
CNN with fewer parameters and easier training, achieving up to 97% accuracy.
Innovative Approaches:
CNN methods using GPGPU and Nvidia DRIVE for addressing severe illumination issues in images.
Performance of Various Methods (as per IJCNN2011 Competition):
CNN with Spatial Transformers: 99.71% accuracy.
Committee of CNNs: 99.46% accuracy.

Other methods like COSFIRE filters and Human Performance range between 98.31% to 98.97% accuracy.
Traditional Methods: Employed handcrafted techniques (e.g., HOG, SIFT) combined with machine learning models like SVM for classification.
Traffic Sign Recognition Contests: Encouraged research in diverse disciplines; focused on comparing learning algorithms.
Recent Developments in Deep Learning: IDSIA's integration of multiple CNN columns and preprocessing into small blocks achieved 0.54% error rate; COSFIRE combined local/global characteristics with CNNs for 1.03% error rate.
Significance of Real-World Application: Emphasizes the need for effective categorization in scenarios where traffic signs occupy a small portion of the scene.

3. Beyond State-of-the-Art

Most of the datasets used in state-of-the-art algorithms are focused on one type of sign or have images that heavily emphasize the area in which the sign is located. The goal of this work is to begin by training the proposed algorithm on a traffic sign-focused dataset and then expand it to detect signs in panoramic images. Many challenges associated with traffic sign recognition stem from datasets that are too focused on specific targets rather than real-life situations. Identifying traffic signs in a time-efficient manner is another challenge to be explored, considering that the solution may be implemented on an image capturing device, requiring a timely and relevant outcome. Additionally, the user experience aspect of the problem will be examined. As previously indicated, displaying the result in a non-distracting manner poses a challenge that has yet to be addressed. This entails not only presenting the result appropriately but also implementing a selection process to determine the display worthiness of the data. Although the objectives may seem formidable, the focus is on exploring them and laying the foundation for the next generation of traffic sign recognition software.

Key Takeaways:

The datasets used in current state-of-the-art algorithms typically focus on one type of sign or the sign's location.
The proposed work aims to train an algorithm on a traffic sign-focused dataset, then extend it to panoramic image sign detection.
Existing challenges include datasets being too narrow and not reflecting real-life situations.
Timely and efficient traffic sign recognition is crucial, especially for implementation in image capturing devices.
User experience is a key consideration, particularly in displaying results in a non-distracting manner.
A process for determining the display worthiness of data needs to be established.

The project seeks to explore these challenges and lay groundwork for advanced traffic sign recognition software.

4. Proof of concept

In the initial prototype of the application, a CNN model was developed using the GTSRB dataset [20]. To accomplish this, the Python programming language was utilized along with the NumPy module for mathematical calculations, OpenCV for image processing, and the Tensorflow module and Keras API3 for neural networks and deep learning support. The Scikit-learn library was also applied for streamlined training and testing of the machine learning model.

4.1. Preliminary architecture

The GTSRB dataset is imported and loaded into the application by using the images and their respective labels, which are defined as X and y in the code. The Convolutional Neural Network (CNN) will efficiently process each image because the input images are quite small (resized to 30x30 pixels). The dataset is divided into two models, namely the training model and the test model. The shape of the X train will be *(62734, 30, 30, 3)*, where the first number represents the number of images on which the model is trained. Similarly, the shape of the X test will be *(15684, 30, 30, 3)*, where the first number represents the number of images that are being tested, and the next two numbers represent the size of an image. The last number is the number of color channels, which is 3 for the RGB model.

The model is built using the Keras Sequential function, which allows for the layer-by-layer creation of the model. The layers that will be used are *Conv2D* layers, *MaxPool2D* layers, *Dropout* layers, and *Flatten* layers. *Conv2D* layers use the input images as 2D matrices, *MaxPool2D* layers down-sample the images, Dropout layers ignore random neurons based on a rate to enhance model training, and Flatten layers create a connection between a Conv2D layer and a Dense layer, with the latter being the output layer in the case of neural networks. The activation method used in *Conv2D layers* and some *Dense layers* (except the last one) will be *ReLU*, which stands for Rectified Linear Activation and has been proven to work well in neural networks. The last Dense layer will use the *softmax* activation to transpose the results into probabilities, with 43 nodes, representing each possible class outcome that describes road signs.

The model is built using the *Adam* optimizer to change the learning rate during training. To avoid the manual encoding of the y variable is used the loss parameter sparse categorical cross-entropy. The fit function is used to train the model, where the epochs parameter specifies the number of times the model runs through the data.

To classify a traffic sign, it is necessary to detect the sign first. The application's functionality is illustrated in a pipeline diagram presented in Figure 1.

Figure 1. Pipeline diagram of proposed solution implementation

The proposed scenario is straightforward: a vehicle is equipped with a camera that records the road and captures data for detecting and recognizing traffic signs. The system then displays warnings, alerts, and informative notifications to the driver about their surroundings. Every frame of the video input goes through the detection method, and the regions that are found are then verified before being supplied to the CNN-based recognition classifier. The recognized signs are filtered using the confidence level, and some other parameters are analyzed to determine the appropriate feedback message for the driver. Two approaches were attempted to detect the signs. The first approach used the Fast R-CNN method, which required extensive training time and did not yield conclusive or satisfactory results. The second approach utilized Maximally Stable Extremal Regions (MSER), which required preprocessing of the input image.

Figure 2. First line - blue mask and red mask; Second line - enhanced merged mask and original image with sign detected in a bounding box; In middle - cropped output traffic sign

In the MSER method, detection is split by color to target red or blue traffic signs. To achieve the best results, two separate routines are used for each color. Firstly, to obtain the red mask, the image is processed by performing contrast normalization over each channel and normalizing the red channel intensity. A binary thresholding with a threshold value near the

maximum intensity value is then performed to obtain the red mask. Secondly, for the blue mask, contrast enhancement of the original image is done, and it is then converted to the HSV color model for easy segmentation of the blue color. A lower and upper limit of the blue mask is defined to extract the blue area, refer to Figure 2. The resulting red mask and blue mask are merged using a bitwise operation, and the merged mask is dilated to enhance its features. The MSER method is applied to detect the regions of presumed road signs using this mask. Finally, some of the ROIs are discarded by specifying a minimum area, ensuring that the bounding boxes detected are square-like, and checking for any intersecting boxes. If two boxes have a large enough intersection, they are united. The final cropped ROIs are then passed on to the recognition algorithm, as shown in the middle of Figure 2.

4.2. Preliminary results

The model was trained for 3 epochs, and various input and compiling parameters were experimented with. In terms of CNN loss function, sparse categorical cross-entropy has 96.31% accuracy if the color model is RBG and varies to 97.69% for BGR and 97.72% for Grayscale.

For categorical cross-entropy, accuracy varies from 97.19% in RGB to 98.61% for BGR and 99.05% for Grayscale.

The conclusion for this step can be that the best-performing model, which used grayscale input images and categorical cross-entropy loss, achieved an accuracy of 99.05% and took 4 seconds less than the second most accurate model that used OpenCV's BGR model. This result was in line with our expectations based on previous studies, given that the grayscale model has only one-color channel, which explains its faster processing time. The BGR model, surprisingly, performed similarly to the grayscale model, while the RGB model performed the worst.

Based on the preliminary traffic sign detection findings, it can be deduced that additional inputs characterizing the state of the road are necessary for maximum performance under difficult lighting conditions. These inputs can be obtained either by integrating the application with the car's rain and light sensors, or through a Bluetooth connection with the driver's smartphone device that can access the internet and provide weather data. These inputs would be fed into the pipeline along with the camera input, allowing for improved detection of traffic signs in various lighting conditions.

Key Takeaways:

Developed a CNN model for traffic sign recognition using Python, NumPy, OpenCV, TensorFlow, Keras, and Scikit-learn.
Utilized the GTSRB dataset, processing images resized to 30x30 pixels.

The CNN architecture includes Conv2D, MaxPool2D, Dropout, and Flatten layers, with ReLU and softmax activations.
Implemented the Adam optimizer and sparse categorical cross-entropy for training.
Detection of traffic signs is performed first, followed by recognition using the CNN model.
Two detection methods tried: Fast R-CNN (ineffective) and MSER (effective, with color-based segmentation for red and blue signs).
Training the model for 3 epochs showed varying accuracies: highest with grayscale images (99.05%) and categorical cross-entropy loss.
Proposed integrating additional inputs from car sensors or smartphones for improved performance in different lighting conditions.

5. Implementation versus State-of-the-Art

In this section, a comparison will be made between the implementation presented and the results obtained in previous competitions using a given dataset. Specifically, the focus will be on the German Traffic Sign Detection Benchmark and the German Traffic Sign Recognition Benchmark.

5.1. Detection

A benchmark for single-image identification models in the areas of pattern recognition, computer vision, and image-based driver assistance is the German Traffic Sign Detection Benchmark. The dataset consists of 900 photos, separated into three categories that are suitable for multiple detection algorithms with diverse attributes. There are 600 training images and 300 evaluation images.

Algorithm	Prohibitive	Danger	Mandatory
HOG	91.3%	90.7%	69.2%
Hough-like	55.3%	65.1%	34.7%
Voila-Jones	98.8%	74.6%	67.3%

Table 2. The detection rate of all the preliminary algorithms [21]

The Viola-Jones detector, the HOG feature approach, and a model-based technique called Hough-like voting were the first three baseline techniques that were independently trained. Among the pool of independent approaches in the selected category, it was discovered that the Viola-Jones approach had the highest detection rate, with the HOG classifier also

performing well. However, for necessary (blue circular) and danger signals, both techniques' performance declined (red triangular). Because they make use of higher-order shape information, the model-based and HOG approaches were better able to manage this challenge.

The German Traffic Sign Detection Benchmark led researchers to the preliminary conclusion that traditional general-purpose detectors worked admirably and outperformed a cutting-edge model-based technique. Performance on unique subsets, like the obligatory signs, was still insufficient for industrial applications. Teams from all around the world responded with innovative solutions after being given a task in response to this.

Among these teams, seven distinguished themselves by achieving perfect results in at least one category. Their methods were more efficient than previous approaches and showed promising results for improving traffic sign detection in challenging scenarios.

Team	Prohibitive	Danger	Mandatory
wgy@HIT501	100%	99.91%	100%
visics	100%	100%	96.98%
LITS1	1000%	98.85%	92%
BolognaCVLab	99.98%	98.72%	95.76%
NII-UIT	98.11%	-	86.97%
wff	-	99.78%	97.62%
milan	-	96.55%	96%

Table 3. Competition Ranking by Area-Under-Curve (Average Overlap) [21]

One issue that is noticeable is how difficult it is to see the required traffic signs. The fact that they are put close to the ground, where they are vulnerable to deterioration or vandalism, and their color - blue, which appear to be difficult to identify in surrounding environment, can also be contributed to this.

Method	Correct recognition rate
Committee of CNNs	99.46%

Human Performance	98.84%
Sermanet	98.31%
Random Forests	96.14%

Table 4. Competition Ranking by Area-Under-Curve (Average Overlap) [21]

5.2. Recognition

The International Joint Conference on Neural Networks hosts the German Traffic Sign Recognition Benchmark, a multi-class, single-image classification task. The dataset is big and is thought to be a lifelike database, with over 40 classifications and over 50,000 photos in total. In the GTSRB final competition stage, four teams stood out from the crowd. Many updated algorithms were presented during the GTSRB competition, and a comparative evaluation of the traffic sign recognition performance of humans and cutting-edge machine learning algorithms was made. The GTSRB concluded that, while the human performance trial achieved a close accuracy of 99.22%, it was exceeded by the effective machine learning strategy - a group of CNNs with a 99.46% right classification rate. Convolutional neural networks, unlike classical computer vision, are capable of learning task-specific features from raw data. Finding the best *ConvNet* architecture for a given task, on the other hand, is mostly empirical.

The proposed approach, which uses categorical cross-entropy loss on a grayscale color model and has a correct recognition rate of 99.05%, is as close to the human performance experiment as the latter got to the machine learning approach. Therefore, the difference between our approach and the human performance experiment is only 0.17%. Comparing these results with those from the GTSRB final ranking would place this solution in second place, 0.21% ahead of third place, and 0.41% below first place.

Key Takeaways:

Comparison of Detection Techniques: The German Traffic Sign Detection Benchmark compared initial techniques like Viola-Jones, HOG, and Hough-like voting for traffic sign detection. Viola-Jones had the highest detection rate, especially for prohibitive signs.
Performance of Detection Algorithms: While traditional general-purpose detectors performed well, their efficacy was limited for certain signs, like mandatory ones. Innovative solutions from international teams showed improvement, with several achieving nearly perfect detection rates in specific categories.
Challenges in Sign Detection: Difficulty in detecting mandatory signs due to factors like placement, deterioration, and color (blue) was noted.

Recognition Benchmark Findings: The German Traffic Sign Recognition Benchmark at the International Joint Conference on Neural Networks highlighted the superiority of machine learning algorithms, particularly Convolutional Neural Networks (CNNs), over human performance in traffic sign recognition.
Algorithmic Advancements: A new approach using categorical cross-entropy loss on a grayscale color model nearly matched human performance in recognition tasks, placing it competitively in the GTSRB final rankings.
Empirical Nature of ConvNet Architecture: Finding the most effective ConvNet architecture for specific tasks remains largely empirical, despite their demonstrated effectiveness in learning task-specific features from raw data.

6. Fine-tuning for the proposed solution

After conducting testing and considering various options, a decision was made to replace the MSER method used for detection. This decision was based on the low accuracy rate of around 60% and the issue of many unwanted regions being identified as candidates due to different lighting scenarios. To address this, the proposed solution utilized a deep neural network (DNN) solution, as it holds great potential for this field. One such system is You Only Look Once (YOLO), a real-time object identification system capable of analyzing images at 30 frames per second (FPS) and even up to 45 frames per second with CUDA acceleration [22]. There are multiple versions of YOLO available, with YOLOv3 being the most suitable for our needs as it prioritizes accuracy over speed compared to YOLOv4. This decision was made with the intention of utilizing GPGPU acceleration later to improve speed of detection.

6.1. Improved architecture

The YOLO system is a real-time object detection solution that is promising for traffic sign detection. After considering different options, the MSER method was replaced due to its low accuracy of detection, around 60%, and the generation of many unwanted regions as candidate objects. The Fast R-CNN also did not prove to be satisfactory, so a deep neural network (DNN) solution was pursued. The YOLOv3 system was found to be the best fit for aimed purpose, with four classes needed for the GTSDB dataset in YOLO format, including prohibitory, danger, mandatory, and other. All 43 subclasses were grouped under each of these classes, from speed limit signs to priorities and directions signs. Pretrained YOLO uses the Darknet architecture and comes with 80 classes from the COCO dataset, but our specific needs required the adjustment to the four classes mentioned above. The model was trained for 8000 iterations, achieving an accuracy of 97.20% after around 8 hours of training, which was a marked improvement over the MSER implementation. However, iterating twice through the dataset for each epoch increased accuracy to 99.11%, though this came at the cost of a longer training time. During detection and recognition, the average

data loss was approximately 3.69%, primarily due to different light conditions or camera artifacts such as motion blur and out-of-focus signs. The performance in low light conditions was found to be quite good, with the data loss, in this case, being less significant.

6.2. Unifying models

The initial testing phase involved applying the application on the im- ages from the GTSDB dataset. A unified pipeline was achieved by loading the saved YOLO weights, the YOLO configuration file, and the custom CNN trained model. The YOLO model selects a set of ROIs for each image, and candidates with the best confidence scores and above a threshold are retained for recognition. Based on the maximum probability prediction using softmax, which returns distributed probabilities, a bounding box with the class description and confidence percentage is displayed on the image for the recognized sign. Figure 3 shows the output of the demonstrator application, where traffic signs detected by the YOLO trained model are outlined in a rectangle shape with a text description that consists of the class prediction and its confidence level of the recognition model, displayed in various colors. The detection and recognition rates are optimal when the distance between the road sign and the camera capturing is between 10 to 20 meters. The pipeline can also process a streamed video input in real time for traffic sign recognition. Based on specific parameters, such as display time or accuracy variation, and the output of the recognition or the class type output, different outcome cases can be generated. For example, the time the sign is displayed on the screen of a car, or how or when the driver is notified about the road sign or conditions. A graphical user interface using the Python library tkinter5 can be created for users to test, where based on a video input, a traffic sign is detected and classified to create a notification. Users can provide feedback on which notifications are useful or not and under what circumstances.

Figure 3. Applications detection and recognition results on the GTSDB dataset

Recorded videos from a dashboard camera were used in different scenarios and weather conditions to further test the application. The processes of identification of traffic signs are run in real-time on the videos, not precomputed, to gain experience on the application's behavior, which can be useful for future adjustments.

6.3. Bug tracking and code refactoring

The encountered issues pertained to the misinterpretation of output files and conversion type errors. For instance, during a detection and recognition session, the OpenCV library experienced breakage due to a frame being unreadable or because the last frame was not detected, causing the loop to keep reading none type variables. To address these bugs, new case statements were introduced that either broke the loop or passed to another frame if one existed. Another issue that arose was related to Python's OpenCV library not detecting the CUDA toolkit, which was necessary for obtaining real-time performance for the demonstrator application. Configuring OpenCV, CUDA, and cuDNN (CUDA DNN) for Windows can be quite challenging. The workaround for this problem involved installing the OpenCV library using cmake, which, unfortunately, lacks a precompiled version for Windows, and the compilation itself is a time-consuming task.

Key Takeaways:

Replacement of MSER Method: The MSER method was replaced due to its low accuracy (around 60%) and problems identifying unwanted regions under different lighting conditions.
Adoption of YOLO System: A deep neural network (DNN) solution using YOLO (You Only Look Once) was adopted, capable of processing images at 30-45 FPS with CUDA acceleration.
YOLOv3 Selection: YOLOv3, prioritizing accuracy over speed, was chosen for traffic sign detection, and later enhanced with GPGPU acceleration for faster detection.
Improvements in Architecture: YOLOv3 showed a significant improvement in accuracy (97.20% to 99.11%) for traffic sign detection, with data loss primarily due to lighting and camera artifacts.
Unifying Models for Application: The system unified YOLO weights, configuration, and a custom CNN model, achieving optimal detection and recognition at 10-20 meters distance.
Real-Time Video Processing: The system can process streamed video in real-time for traffic sign recognition, with a user interface for testing and feedback using tkinter5.
Testing with Dashboard Camera: The application was further tested in various weather conditions using dashboard camera videos to observe real-time performance.
Bug Tracking and Refactoring: Addressed issues included misinterpretation of output files and conversion errors, with solutions like introducing new case statements and configuring OpenCV with CUDA for Windows.

8. Conclusions

Precise results were achieved in the recognition stage as shown are presented in the section Related work. The categorical cross-entropy loss function yielded the highest precision, specifically with a value of 99.11% when applied to the grayscale color mode. By comparing this paper's results with those from GTSRB IJCNN as presented in Table 4, it can be concluded that this solution could have been a contender for first place in the competition. The precision achieved was near human precision levels, and any recognition software with over 98% precision is considered suitable for further development and implementation into Level 2+ systems, which do not drive the vehicle but instead provide driver support.

Source	Precision
Best GTSRB Machine Learning Algorithm	99.46%
GTSRB Human Precision	99.22%

Table 5. The final precision comparison

Further development would involve identifying areas where the software performs poorly and devising solutions for these areas, such as accounting for road and weather conditions, as well as driver behavior. To integrate the software into a vehicle, it could be embedded into the automobile's core system with surrounding sensors, or as an accessory such as a Bluetooth dashboard camera that is linked to the driver's mobile phone and serves as a display endpoint.

8.1. Future work

The perfected application will provide options for both online and offline users. The offline version will work seamlessly with the proposed algorithm implementation, while the online version will be served through a web API for handling detection and recognition tasks.

In terms of workflow, the client device will primarily serve as an image capture tool, allowing users to capture images that will be sent to the server via a post request, with the image attached. The server will be responsible for executing the prediction stage, which will culminate in the transmission of a post request back to the client application, containing

the prediction results. Depending on the application, the results may be accessed via a dictionary and displayed in various ways, such as using images, sound effects, or other formats.

Key Takeaways:

High Precision Achieved: The research yielded a precision of 99.11% using the categorical cross-entropy loss function in grayscale mode, comparable to top results in the GTSRB IJCNN competition.
Near Human-Level Precision: Achieved precision was close to human levels, indicating suitability for Level 2+ systems offering driver support.
Future Development Focus: Identify and improve areas of software weakness, considering road/weather conditions and driver behavior, with integration options including core system embedding or Bluetooth dashboard cameras linked to mobile phones.
Offline and Online Application Versions: The perfected application will offer both offline (with algorithm implementation) and online (via web API) versions for detection and recognition tasks.
Client-Server Workflow: The client device captures images, sends them to the server, which processes and returns prediction results. The results can be displayed in various formats such as images or sound effects.

References

[1] https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries - Road traffic injuries, 2021, 02.01.2023

[2] https://www.euroncap.com/en/press-media/press-releases/cars-of-the-future-need-readable-roads-now - *Press Release: Cars of the future need "readable" roads now*, 2013, 02.11.2022

[3] Procedure 2018/0145/cod, 2019. 02.11.2022

[4] Jan Salmen Johannes STALLKAMP, Marc SCHLIPSING, Christian IGEL - *Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition* Neural Networks, 32:323–332, 2012

[5] Dan CIREȘAN, Ueli MEIER, Jonathan MASCI, Jürgen SCHMIDHUBER - *A committee of neural networks for traffic sign classification* - 2011 International Joint Conference on Neural Networks, pages 1918–1921, 2011

[6] Michael SHNEIER - *Road sign detection and recognition*, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=823519. 02.11.2022

[7] David SOENDORO, Iping SUPRIANA - *Traffic sign recognition with color-based method, shape-arc estimation and SVM* - Proceedings of the 2011 International Conference on Electrical Engineering and Informatics. IEEE, July 2011.

[8] Wenhui LI, Daihui LI, Shangyou ZENG - *Traffic sign recognition with a small convolutional neural network* - IOP Conference Series: Materials Science and Engineering, 688(4):044034, Nov 2019

[9] Kwangyong LIM, Yongwon HONG, Yeongwoo CHOI, Hyeran BYUN - *Real-time traf- fic sign recognition based on a general purpose GPU and deep-learning* - PLoS One, 12(3):e0173317, March 2017

[10] https://www.nvidia.com/en-us/self-driving-cars - Self-driving cars. 01.11.2022.

[11] Zhiyong HUANG, Yuanlong YU, Jason GU, Huaping LIU - *An efficient method for traffic sign recognition based on extreme learning machine* - IEEE Transactions on Cybernetics, 47(4):920–933, 2017

[12] Yi YANG, Hengliang LUO, Huarong XU, Fuchao WU -*Towards real-time traffic sign detection and classification* - IEEE Transactions on Intelligent Transportation Systems, 17(7):2022–2031, 2016

[13] Fatin ZAKLOUTA, Bogdan STANCIULESCU - *Real-time traffic-sign recognition using tree classifiers* - IEEE Transactions on Intelligent Transportation Systems, 13(4):1507–1514, 2012

[14] Jack GREENHALGH, Majid MIRMEHDI - *Real-time detection and recognition of road traffic signs* - IEEE Transactions on Intelligent Transportation Systems, 13(4):1498–1506, 2012

[15] Dan CIREȘAN, Ueli MEIER, Jonathan MASCI, Jürgen SCHMIDHUBER - *Multi-column deep neural network for traffic sign classification. Neural Networks*, 32:333–338, 2012

[16] Pierre SERMANET, Yann LECUN - *Traffic sign recognition with multi-scale convolutional networks* - 2011 International Joint Conference on Neural Networks, pages 2809–2813, 2011

[17] Álvaro ARCOS-GARCÍA, Juan A. ÁLVAREZ-GARCÍA, Luis M. SORIA-MORILLO - *CNN with 3 spatial transformers*. Neural Networks, 32, 2012

[18] Baris GECER, George AZZOPARDI, Nicolai PETKOV - *Color-blob-based cosfire filters for object recognition* - Image and Vision Computing, 57:165–174, 2017

[19] Fatin ZAKLOUTA, Bogdan STANCIULESCU, Omar HAMDOUN - *Traffic sign classification using k-d trees and random forests* - 2011 International Joint Conference on Neural Networks, pages 2151–2155, 2011

[20] https://benchmark.ini.rub.de - German traffic sign benchmarks, 2019. 10.11.2022

[21] Sebastian HOUBEN, Johannes STALLKAMP, Jan SALMEN, Marc SCHLIPSING, Christian IGEL - *Detection of traffic signs in real-world images: The German traffic sign detection benchmark* - 2013 International Joint Conference on Neural Networks (IJCNN), pages 1–8, 2013

[22] https://pjreddie.com/darknet/yolo - Yolo: Real-time object detection. 26.01.2023.

Bibliography

Álvaro ARCOS-GARCÍA, Juan A. ÁLVAREZ-GARCÍA, Luis M. SORIA-MORILLO - *CNN with 3 spatial transformers*. Neural Networks, 32, 2012

Baris GECER, George AZZOPARDI, Nicolai PETKOV - *Color-blob-based cosfire filters for object recognition* - Image and Vision Computing, 57:165–174, 2017

Dan CIREȘAN, Ueli MEIER, Jonathan MASCI, Jürgen SCHMIDHUBER - *A committee of neural networks for traffic sign classification* - 2011 International Joint Conference on Neural Networks, pages 1918–1921, 2011

Dan CIREȘAN, Ueli MEIER, Jonathan MASCI, Jürgen SCHMIDHUBER - *Multi-column deep neural network for traffic sign classification. Neural Networks*, 32:333–338, 2012

David SOENDORO, Iping SUPRIANA - *Traffic sign recognition with color-based method, shape-arc estimation and SVM* - Proceedings of the 2011 International Conference on Electrical Engineering and Informatics. IEEE, July 2011.

Fatin ZAKLOUTA, Bogdan STANCIULESCU, Omar HAMDOUN - *Traffic sign classification using k-d trees and random forests* - 2011 International Joint Conference on Neural Networks, pages 2151–2155, 2011

Fatin ZAKLOUTA, Bogdan STANCIULESCU - *Real-time traffic-sign recognition using tree classifiers* - IEEE Transactions on Intelligent Transportation Systems, 13(4):1507–1514, 2012

François CHOLLET – *Deep Learning with Python* – Manning Publishing, November 2017, ISBN 9781617294433

Ian GOODFELLOW, Yoshua BENGIO, Aaron COURVILLE - *Deep Learning* – MIT Press, 2016

Jack GREENHALGH, Majid MIRMEHDI - *Real-time detection and recognition of road traffic signs* - IEEE Transactions on Intelligent Transportation Systems, 13(4):1498–1506, 2012

Jan Salmen Johannes STALLKAMP, Marc SCHLIPSING, Christian IGEL - *Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition* Neural Networks, 32:323–332, 2012

Kwangyong LIM, Yongwon HONG, Yeongwoo CHOI, Hyeran BYUN - *Real-time traffic sign recognition based on a general purpose GPU and deep-learning* - PLoS One, 12(3):e0173317, March 2017

Michael SHNEIER - *Road sign detection and recognition*, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=823519. 02.11.2022

Nadra BEN ROMDHANE, Hazar MLIKI, Mohamed HAMMAMI - *An improved traffic signs recognition and tracking method for driver assistance system* - 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), Okayama, Japan, 2016

Neel BHATT, Pratiksha LALDAS, Vivian Brian LOBO - *A Real-Time Traffic Sign Detection and Recognition System on Hybrid Dataset using CNN* - 2022 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2022

Njayou YOUSSOUF - *Traffic sign classification using CNN and detection using faster-RCNN and YOLOV4* - Heliyon, Volume 8, Issue 12, 2022, ISSN 2405-8440

Pierre SERMANET, Yann LECUN - *Traffic sign recognition with multi-scale convolutional networks* - 2011 International Joint Conference on Neural Networks, pages 2809–2813, 2011

Preeti BAILKE, Kunjal AGRAWAL*, Traffic Sign Classification Using CNN* – IJRASET, ISSN 2321-9653, 2022

Sebastian HOUBEN, Johannes STALLKAMP, Jan SALMEN, Marc SCHLIPSING, Christian IGEL - *Detection of traffic signs in real-world images: The German traffic sign detection benchmark* - 2013 International Joint Conference on Neural Networks (IJCNN), pages 1–8, 2013

Yi YANG, Hengliang LUO, Huarong XU, Fuchao WU -*Towards real-time traffic sign detection and classification* - IEEE Transactions on Intelligent Transportation Systems, 17(7):2022–2031, 2016

Ying SUN, Pingshu GE, Dequan LIU - *Traffic Sign Detection and Recognition Based on Convolutional Neural Network* - 2019 Chinese Automation Congress (CAC), Hangzhou, China, 2019

Wenhui LI, Daihui LI, Shangyou ZENG - *Traffic sign recognition with a small convolutional neural network* - IOP Conference Series: Materials Science and Engineering, 688(4):044034, Nov 2019

Zhiyong HUANG, Yuanlong YU, Jason GU, Huaping LIU - *An efficient method for traffic sign recognition based on extreme learning machine* - IEEE Transactions on Cybernetics, 47(4):920–933, 2017

https://benchmark.ini.rub.de - German traffic sign benchmarks, 2019. 10.11.2022

https://www.euroncap.com/en/press-media/press-releases/cars-of-the-future-need-readable-roads-now - *Press Release: Cars of the future need "readable" roads now*, 2013, 02.11.2022

https://www.nvidia.com/en-us/self-driving-cars - Self-driving cars. 01.11.2022.

https://pjreddie.com/darknet/yolo - Yolo: Real-time object detection. 26.01.2023.

https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries - Road traffic injuries, 2021, 02.01.2023

AGILE TRANSFORMATIONS IN THE DIGITAL BUSINESS ECOSYSTEM, AND INCREASED IT ALIGNMENT IN THIS CONTEXT

Ioan-Matei PURCĂREA[1]

Abstract

We are witnessing significant and transformative shifts in digital business ecosystem, actors' connection, the importance of valorizing digital technologies and strategic partnerships, as well of encouraging software developers' ability and motivation to use Gen AI that need to be treated as a broad business priority. There is no doubt about today's agile transformations and increased IT alignment, as well about remarkable movements like from improving IT projects' results to going beyond improvements to IT, based on business and IT together working. It is more and more challenging for companies to investigate digital capabilities' assessment in the context of proceeding along the growth path, and to accelerate cloud efforts being more focused on return on investment. Also, there is an impressive continuity of the debates regarding the significant potential of Gen AI technologies, where the need of discovering the right answer to where you sit in the value chain, better understanding the importance of focusing on the user, learning from your entire surrounding environment.

Keywords: Digital Business Ecosystem, Agile Transformations, Increased IT Alignment, Gen AI Technologies, Cloud Challenge, Digital Capabilities

JEL Classification: D83, L81, M31, O31, O33

1. Introduction

There is a real challenge today for businesses to become smarter benefiting from valorizing continuously advancing disruptive technologies, by integrating both disruptive technology use and digital thinking across a whole business, including at the level of the strategic alliance in its supply chain, while promoting a customer-centric culture, improving not only its services and experiences delivery, and its effectiveness, but also its long-term relationships, ensuring customer satisfaction, loyalty, renewed enthusiasm and advocacy. Recent research findings highlighted that business and IT must work together based on an agile program, the implemented culture change stimulating value creation supported by IT integration and improved IT projects' results depending directly on what technology leadership is creating for aligned teams. It was also recently revealed that in their effort to advance on the paths to growth and productivity businesses are required to even go beyond

[1] PhD, Romanian-American University, Romania, matei.purcarea@rau.ro

improvements to IT towards higher-value cloud, being known why a successful business must be cloud based.

Recent research by Greu [1], underlined the great expectations from the complexity science, as well as the fine linkage between ICT and Information Society, suggesting to expand and refine knowledge. Also, Chen et al. [2] invited researchers to a more in-depth analysis of platformization, our economic and social life being reshaped, for instance, by digital platforms. On the other hand, the President of the Romanian Academy Section of Science and Information Technology made recently reference to the convergence between man and computer in the elaboration of co-decisions [3].

And as there is always more to learn and to research, let us recall the words of Diamandis and Kotler [4], who focused on the impact of converging technologies not only on business and industries, but also on our lives: "New business models are no longer forces for stability and security. To compete in today's accelerated climate, these models are designed for speed and agility. Most importantly, none of this is in any danger of slowing down… No one really understands the impact AI will have on retail… AI makes retail cheaper, faster, and more efficient, touching everything from customer service to product delivery. It also redefines the shopping experience, making it frictionless and – once we allow AI to make purchases for us – ultimately invisible."

2. Digital business ecosystem, actors' connection, the importance of valorizing digital technologies and strategic partnerships, as well of encouraging software developers' ability and motivation to use Gen AI that need to be treated as a broad business priority

2.1 Digital business ecosystems' strategic functions and reinvention of customer engagement and experience within the resilient supply chains'

After methodically reviewing studies published in top information systems and management literature, Heim et al. [5] identified sharing information (like data or data analyses) between the connected actors as one of the digital business ecosystems' strategic functions (a capability, role, or task to achieve the desired change) taken by intermediaries and needing to be better understood for effective interactions management within multilateral sourcing relationships. They gave a typical example of the digital platforms' (such as e-commerce platforms Amazon, eBay, and Shopify) partnerships, highlighting how important is in today's quickly-advancing digital business ecosystems to anticipate and mitigate (through strategic planning and operational resilience) the specific risks (associated with the simultaneity of some actors' relevant status of gate opener and gatekeeper) in customer relationship management (CRM). And within the context in which actors' connection is the intermediaries' primary function (all other functions, including sharing information, building on the primary function), the intermediaries' gained control

(depending on the strategic functions fulfilled by each one) may be the initiator for encouraging needed innovation.

On the other hand, there are rigorous research papers signaling the imperative of better understanding the rising digital platforms (seen as multi-sided markets transforming nearly every industry) and ecosystems, the impact of digital platform innovation being considered significant [6]. It was highlighted, among other aspects, how new and adaptable means for inter-organizational relations can be encouraged by these digital platforms, being necessary to deepen the analysis of how to architect these above-mentioned interrelationships. We see here a fine linkage with some research approaches regarding the circumstances surrounding digital transformation considering the identified stages of digital transformation (digitization, digitalization, and digital transformation) and the suggested growth strategies to successfully transform digitally, as well as digital transformation context-specific manifestations [7]; [8].

Resilience is considered by Boston Consulting Group (BCG) experts today's new priority in rethinking supply chains, the needed improved visibility, agility, operational efficiency, and customer engagement coming from valorizing digital technologies and strategic partnerships ensuring shared innovation efforts and actionable insights while accessing target markets [9]. In the same time, it is essential to promote sustainability and ethics, adapting proactively to the dynamic global landscape.

Real-time integration with supply chain management (SCM) and e-commerce are important retailers' technology requirements, and consequently CRM software systems (like Salesforce, Oracle CX, Microsoft Dynamics 365, SAP CRM) need to fit in with retail companies' imperatives such as consumer insights and segmentation, customer experience management (CXM), retail marketing (considering precision campaigns) and selling (considering clienteling as customer service technique), ensuring including social listening and engagement tools across channels [10]. According to that approach, retail companies can benefit of an improved process of collecting customer data (pulling insights from it, accordingly), as well as of an increased customer lifetime value (CLV), moving consumers' loyalty from products to their brands.

2.2 Artificial intelligence, human cognition, collective intelligence, and customers perceptions of AI-enabled interactions

Focusing on key differences between artificial intelligence (AI) and human cognition, Felin and Holweg [11] drew attention to asymmetric beliefs' importance, and theory-based causal logic, by going beyond AI's data and prediction-orientation view. According to Taylor [12], when designing AI applications it is good to start by being aware of the need to enhance collective intelligence (CI) through AI.

With regard to the way in which software development is transformed by the generative artificial intelligence AI (Gen AI), BCG experts [13] brought to our attention software companies' need to treat Gen AI introduction within their organizational transformation so as to exploit it. What presupposes to overcome difficulties standing in the way of its adoption, and not only to choose the right descriptions of users' interaction with systems in order to reach their goals, but also to encourage developers' ability and motivation to use Gen AI, and being continuously value focused.

Analyzing how e-commerce future is gaining speed, retail industry experts [14] underlined retailers' pressure to ensure more profound customers relationships and e-commerce platforms' profitability based on adequate experiments (with logistics and value chains, as well as right price points). A retailers' competitive advantage will be created by experimenting with real-time deal comparison and overall search results' improvements with the help of Gen AI, where the impact on e-commerce given by Google's zero-click searches on (without clicking on a website, just answering directly the query on the search engine results page) will be significant. It is also predicted that the secondhand marketplaces (together with the rental platforms and resale) are going to become increasingly important within the context of rising customers' living expenses and sustainable purchasing habits. Recently, for instance, May 21 was declared by eBay (based in San Jose, California) as the first annual Recommerce Day, being already supported by first legislative steps to possibly preserve this new idea so as to be protected and respected, promoting re-commerce shopping behavior [15]. It is also worth mentioning that on the basis of people search Google changed in May 2024 the way of displaying its content type filters [16].

According to a SEO expert presented very recently by MarketingSherpa as having a worldwide experience [17], performance in SEO (within digital content creation) can become greater by strategically applying Gen AI expertise. It is considered crucial for SEO experts to become very skilled and experienced at using safe Gen AI, productivity being intensified with the help of AI (used in compliance with data protection laws, ethical considerations etc.), improving both content quality (including for preliminary research and idea generation), and quality control.

In general, as shown by McKinsey experts [18], the scope of the transformational value possible to be generated by Gen AI can be caught by exploiting it at scale, consequently companies' chief information officers (CIOs being in close relationships with business unit leaders) needing to act on facts to ensure business guidance, based on effective end-to-end automation. Companies must treat Gen AI as a broad business priority, embedding it into their business with the help of the built teams (having roles and skills like: DataOps, Site reliability engineer, DevOps engineer, Cloud architect, Solution/data architect, Platform owner, Full-stack developer, Data scientist, Data engineer) progressing beyond the information technology (IT) function. In the same time, as demonstrated by Singla et al. [19], in the repetitive and collaborative process of working with Gen AI providers it is

important to take into account not only to continuously share updates and discuss challenges, but also to align on priorities across technology providers, both building trust and fostering collaboration so as to open the maximum potential of Gen AI Singla et al. gave an interesting example of a partnership between a luxury retail company (that shared its product information catalog and details concerning customer preferences and behaviors' nuances) and a Gen AI provider (ensuring relevant prompts' engineering and necessary domain-specific data interpretation etc.), from this strategic partnership resulting personalized product recommendation system being remarkable (99% accuracy). These McKinsey experts also brought to our attention how essential for a scalable Gen AI ecosystem creation it is to ensure interoperability between models and components.

As in today's shifting business management landscape AI is increasingly utilized in the digital content by technology brands, Mushtaq and Kethuda [20] investigated how AI integration (as cutting-edge technology) into brands' marketing activities is perceived by customers (considering convenience, value, risk of AI adoption, and product/service quality as key indicators) so as to influence their purchase intentions, changing their behaviors. The insights provided by this research findings were seen as leading to changing business practices, ensuring technology brand differentiation and credibility.

Prior research of Purcărea et al. [21] has shown the need of improving consumers' information and education concerning AI interactions' benefits, within the context in which consumers' perceptions of AI was the input construct in the employed quantitative research using SEM (see figure 1 below). It was also revealed, among other aspects, retailers' need of becoming increasingly aware of the whole range of AI benefits for digital marketing to successful use AI-enabled interactions with consumers.

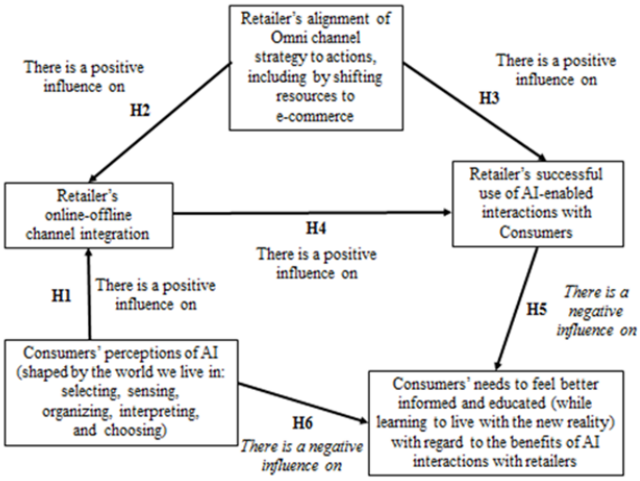


Figure 1. The theoretical research model[2]

[2] Source: Theodor PURCĂREA, Valeriu IOAN-FRANC, Ştefan-Alexandru IONESCU and Ioan Matei PURCĂREA - The Profound Nature of Linkage Between the Impact of the Use of Artificial Intelligence in Retail on Buying and Consumer Behavior and Consumers' Perceptions of Artificial Intelligence on the Path to

Since the middle of the past decade, we emphasized marketers' need to master the new rules of digitization and to think proactively about their tasks regarding potential disruptors [22]. As Jim Lecinski said, "In our experience at Google, marketers who move with speed, make their messages highly relevant, and use data (it beats opinion!) are best-positioned for success with today's buyers and modern media vehicles" [23]. More recently [24], we signaled marketers' challenge to successfully drive consumers' engagement, including by considering the metaverse-related activities. On the other hand, let's remember what Neil Perkin said, "With data, technology and machine learning changing marketing practice in such fundamental ways it is essential for agile marketers to remain focused on how they can bring technology and human capabilities together in ways that compound benefit and impact" [25].

3. Agile transformations, increased IT alignment, digital capabilities' assessment and cloud efforts' acceleration

3.1 Making agile principles work, improving leadership as a whole

According to BCG experts [26], to make agile principles work technology leaders must implement a culture change in which, for instance, value creation is supported immediately by the driven IT programs, and IT designer and developer teams can focus on results of notable significance, multiplying program, and processing a great number of units of information in a given amount of time. Their recent research findings revealed significant aspects, such as: chief information officers (CIO)s, chief technology officers (CTOs), and IT departments' heads have not been included in their companies' process of general decision-making when a technology implementation is initiated or were not successful in communicating in a direct way what they can expect from an initiated tech implementation; at most companies, improving IT projects' results will depend directly on what technology leadership is creating for teams; IT projects' success rates improve when proper valid tracking mechanisms are in place for them; deep coordination between technology and business counterparts (the party that a company is contracting with or a third party) will be required to implement GenAI. It is worth mentioning that BCG experts' investigation was focused on in-house software development, mainly custom programs and applications tailored specifically (for example, targeting marketing automation).

After the Agile Manifesto introduction 23 years ago, when radical changes were introduced in the new software development process [27], we were witnessing an increasing adoption of agile approaches while trying to comprehend the role of agile – both as-a-tool (in the contingency perspective), and as-a-culture (in the configuration perspective) – a not only in

the Next Normal. Amfiteatru Economic, 23(56), pp.9-32. p-ISSN 1582-9146, e-ISSN 2247–9104. February 2021.

driving innovation and digital transformation, but also in improving leadership as a whole [28]. Within the last-mentioned framework, it was launched finally an invitation to continuous research so as to better light on connections of agile with management and organization theories. And as in agile transformations tensions are rising, it is very important to not forget the tensions identified by Strode et al. [29], considering, of course, each organizations' chosen transformation focus (operational change, strategical change or cultural change).

Research findings by Al Jabri et al. [30] revealed how enterprise agility is improved by increased IT alignment (the last one being achieved by company's transforming and seizing), while enterprise agility is intensified by company's transforming, sensing and seizing capabilities. On the other hand, the link between enterprise agility and transforming was found as being moderated by digital transformation (DT).

3.2 Investigating digital capabilities' assessment in the context of proceeding along the growth path, and accelerating cloud efforts being more focused on ROI

Based on prior literature and focused on assessing digital capabilities (in connection with their desired DT goals) in the DT process, Baiyere et al. [31] presented a so-called MIND (Management, Infrastructure, Networking/Sourcing, and Development) framework as a result of a several years lasting design science research project, this framework including both a MIND Canvas, and a MIND Process. A balanced scorecard (as the adapted BSC and providing strategic utility), a Status Map (giving a brief idea of what the positioning of an organization's digital capability is like) and a Capability Matrix (incorporating the digital capability concepts into the above-mentioned framework) were the instruments forming a MIND toolkit. Research authors invited for future research going beyond DT, for instance to consider investigating digital capabilities' assessment in the context of proceeding along the growth path.

As explained recently by McKinsey Global Institute [32], only from the growth of productivity (as a measure of economic performance considering output per unit of input) in time can come a higher standard of living, if we think, for instance, about the in the nineties seen dot-com boom created by the ascending digitization and other forces, supply chains' modernization of supply chains by the very large stores like Walmart (the U.S. biggest retail store chain) etc. According to this approach, productivity will be impacted by the advances in AI, being expected significant progress.

Recent research by Betley et al. [33] revealed that in their effort to advance on the paths to growth and productivity (being behind North American and Asian companies) European companies are required finally to capture overall value from cloud, going beyond improvements to IT towards higher-value cloud use cases in their business operations, avoiding that the obtained value remains of limited use and smaller than other similar results. Consequently, European companies are seen as needing to accelerate their cloud

efforts being more focused on return on investment (ROI), investing in the necessary capabilities to capture the above-mentioned overall value from cloud based on a clear view compared to the current situation (the last one being characterized as follows: the variance by sector of the thresholds to determine cloud adoption scale, retailers, for example, requiring comparatively more workloads on cloud; a notable portion of the technology estate of the European companies adopting cloud are continuing to keep over 20 percent of their activity on-premises).

Let us recall how priorly Arora et al. [34] brought to our attention that the cloud investment-and-return equation could be transformed by Gen AI. They made their comments on the basis of the findings in McKinsey report, showing, for example, various aspects, such as: where cloud's total value lies etc.; how many percentage points of incremental ROI can be added by Gen AI to cloud programs and through which key benefits.

In an interview given to McKinsey Digital [35], the senior vice president and divisional CIO for life, annuities, and distribution at Lincoln Financial Group (an insurance company), Rob Klaczak, talked about their decision to transition to a cloud-based system, revealing significant aspects from their journey, such as: how not only effective communication and information flow, but also collective engagement was ensured on the basis of the hub-and-spoke model, according to this model, the hub (company's core team) being connected to the spokes (company's various parts of the business and IT); essential for navigating their security complexities were two things: first, to ensure early engagement with their chief information security officer (CISO), and second to promote a strong relationship with the above-mentioned CISO's team; one of their focus was on optimizing for cost efficiency in a cloud-based consumption-driven structure (to ensure both valorization of necessary technologies for storage, security, analytics, and AI, as well as their customers' charging only based on their resource usage).

4. A great frustration and a great challenge for IT departments. Forging business links to IT treated as a strategic capability

Trust among companies can be increased with the help of technology, and that on the basis of reliable data sharing systems, transparency for all involved parties being provided by modern software [36]. That is very important at the level of a supply chain, for instance, taking into account companies' executives concern about their suppliers and customers' way of using shared data. To avoid data sharing risks can be brought into play AI joint model trained by a joint venture or a third party (benefitting this way from the value of sharing data), an emerging alternative being federated learning (the training being initialized independently by each participant).

According to McKinsey experts [37], IT departments has been frustrated for years because of all the time question regarding why some technology transformations achieve the desired result while others fail, but here was also identified an easy-to-understand opportunity for growth. Their research findings demonstrate that technology transformation initiatives (such as: a) data capabilities, sharing a tech-backed business strategy, and modernizing infrastructure; b) data capabilities, redesigning tech-organization, and transforming talent management and sourcing partnerships; c) shaping a tech-backed business strategy, and delivery model and core engineering principles) had a significant effect (as follows): a) on enhanced IT infrastructure; b) shaped a tech-backed business strategy; c) redesigned tech-organization and operating model. To valorize their technology foundations as much as possible companies need to be based (beyond investment in foundational improvements and ensuring alignment between technology and business stakeholders with regard to building products) on strategic alignment and have a strong business case, being aware of the followings: the difficulty of encouraging others (employees, shareholders etc.) to support their vision of modernizing legacy technology; only IT organization and operating model transformation (despite being essential for companies to benefit as much as possible from modernized technology) it is not sufficient enough to achieve better performance, being recommendable to lean on product- and platform centric operating models focused on the end-user experience.

There is no doubt that in the face of continuous increasing competition and adaptation under pressure it is essential to have that technology capability to achieve strategic goals, forging business links to IT treated as a strategic capability making the business to exploit the technology so as to deliver the expected value [38]. Consequently, business and IT must work together based on an agile program, the alignment being ensured across all levels (top management/board level, business process implementation, IT governance, and technology platform).

5. Conclusions

At the end of March this year, President of the Romanian Academy Section of Science and Information Technology, Academician Florin Gheorghe Filip, spoke in an interview about the history of the term digital humanism, the DIGHUM initiative and the Vienna Manifesto launched at the Technical University of Vienna in 2019 and other similar ideas, as well as about important contributions to the definition of the concept [3]. He also mentioned, among other aspects (such as ChatGPT), the more recent trend called dataism (as opposite to the one mentioned before), documenting and emphasizing the continuity of the debates, including in relation to “a great turning point” in human evolution, and the convergence between man and computer in the elaboration of co-decisions etc. In the end he referred to young people as the ones who will live in the world that they think and prepare.

Coming back to the Digital Humanism Initiative [39] it should be signalized that on the occasion of an online-only event (also live-streamed) taking place on June 11, 2024, Gordon Burtch (from Questrom School of Business, Boston University, USA) will explain the large language models' (LLMs) impact on online knowledge communities, starting from the significant potential of Gen AI technologies (mainly LLMs like ChatGPT) to impact both participation and content production in the above-mentioned communities. He will provide significant aspects based on their comparative data analysis at Stack Overflow and Reddit developer communities [40].

It is worth adding within this context that also in March this year took place an interesting podcast episode of McKinsey on Building Products [41] on software product management and engineering exploration, in which the chief product officer (CPO) of Reddit, Pali Bhat (known for being the very first to do product innovation for various large tech platforms, including Google), highlighted, among other aspects, the followings: only the highest-quality authentic conversations move towards the Reddit top; discovering the right answer to where you sit in the value chain is the next step after understanding how an AI and ML pipeline would work; how important is to focus on the user, learning from your entire surrounding environment, but avoiding to be distracted with insignificant details, learning from mistakes' magnitude and treating them as learning opportunities.

Also, it is useful to remember within this context that at the end of last year McKinsey experts [42], explained why a successful business must be cloud based according to their new research (paraphrasing American psychologist William Schutz), ensuring the foundation (enabling a lot of automation – by thinking about infrastructure as code, and entire systems as code – the cloud being built for), then everything else becoming more efficient, and cloud benefits coming from both enabling the business use cases, and allowing to work a lot faster to company's application developers and engineers based on updating skills (without forgetting how expensive are skill sets). Of course, if a company establish multiple foundations (compared to the above-mentioned one foundation), it is the situation (resulting from failing to effectively ensuring control) so-called cloud sprawl (known from many sources as the uncontrolled proliferation of an organization's cloud instances, services or providers etc.).

There is no doubt, we have great expectations from the complexity science, from the optimized working together of business and integrated and aligned IT, better understanding and leveraging platformization, disruptive technologies, and digital innovation, and last but not least better living in our world that we think and prepare.

References

[1] Victor GREU - *Improving Information and Communication Technology to Innovate the Paradigms of World Progress, Part 1*. Romanian Distribution Committee Magazine, vol. 15(1), pp. 16-28, ISSN 2069-0134. March 2024.

[2] Julie CHEN, Niels VAN DOORN, Rafael GROHMANN, Andrea POLLIO, Cheryll SORIANO - *Introducing Platforms & Society*. Platforms & Society Volume 1: 1–2, eISSN 29768624, ISSN 29768624. 2024. https://journals.sagepub.com/doi/10.1177/29768624241235492

[3] Cătălin MOSOIA - *The future must be carefully thought out and prepared*. Romanian Academy, March 26, 2024. Retrieved from: https://acad.ro/institutia/media_florin_filip.html

[4] Peter H. Diamandis, Steven Kotler - *The Future Is Faster Than You Think: How Converging Technologies Are Transforming Business, Industries, and Our Lives* - [pages 87, 100]. ISBN 1982109661, 9781982109660. Publisher Simon and Schuster. [January 28] 2020.

[5] Sophie HEIM, Rebekka SCHINDLER, Tianyu GE, Holger WITTGE - *And Here, All Threads Run Together: Understanding the Strategic Functions of Intermediaries in Digital Business Ecosystems*. European Conference of Information Systems, ECIS 2024 Proceedings. 2., pp. 1-2, 8, June 2024. Retrieved from: https://aisel.aisnet.org/ecis2024/is_governance/track21_is_govern/2

[6] Mark DE REUVER, Carsten SØRENSEN, Rahul C. BASOLE - *The digital platform: a research agenda*. Journal of Information Technology 33, 124–135. eISSN 14664437, ISSN 02683962. June 1, 2018. https://doi.org/10.1057/s41265-016-0033-3

[7] Peter C. VERHOEF, Thijs BROEKHUIZEN, Yakov BART, Abhi BHATTACHARYA, John Qi DONGA, Nicolai FABIAN, Michael HAENLEIN - *Digital transformation: A multidisciplinary reflection and research agenda*. Journal of Business Research, 122: 889–901, ISSN 0148-2963. 02 November 2019. https://doi.org/10.1016/j.jbusres.2019.09.022

[8] Manuel MÜHLBURGER and Barbara KRUMAY - *Towards a context-sensitive conceptualisation of digital transformation*. Journal of Information Technology, Vol. 0(0) 1–16, Special Issue: Digital Transformation as Geo-Political, Organizational and Technological Nexus, 124–135. eISSN 14664437, ISSN 02683962. November 23, 2023. https://journals.sagepub.com/doi/10.1177/02683962231219514

[9] Nikolaus LANG - *Imagine This... The End of Free Trade*. BCG, International Trade, May 13, 2024. Retrieved from: https://www.bcg.com/featured-insights/podcasts/imagine-this/end-of-free-trade?

[10] Chuck SCHAEFFER - *How to Choose the Best CRM Software for your Retail Company*. CustomerThink, February 15, 2021. Retrieved from: https://customerthink.com/how-to-choose-the-best-crm-software-for-your-retail-company/?

[11] Teppo FELIN, Matthias HOLWEG - *Theory Is All You Need: AI, Human Cognition, and Decision Making*. SSRN, pp. 1-46, February 24, 2024. http://dx.doi.org/10.2139/ssrn.4737265

[12] Jacob TAYLOR - *The most important question when designing AI*. The Brookings Institution, pp. 1-6. May 20, 2024. Retrieved from: https://www.brookings.edu/articles/the-most-important-question-when-designing-ai/?

[13] Pranay AHLAWAT, Julie BEDARD, Sankalp DAMANI, Ben FELDMAN, Elizabeth LUCERO, Abhishek SAMDARIA and Geoffrey SIPPERLY - *The Art of Scaling GenAI in Software*. BCG, Technology, Media, and Telecommunications, May 17, 2024. Retrieved from: https://www.bcg.com/publications/2024/the-art-of-scaling-genai-in-software?

[14] Guy ELLIOTT, Sudip MAZUMDER - *8 Trends Accelerating the Future of E-Commerce in 2024*. Publicis Sapient, May 2024. Retrieved from: https://www.publicissapient.com/insights/future-ecommerce-trends?

[15] Andrew Adam NEWMAN - *Why eBay is declaring the first Recommerce Day*. Retail Brew, May 21, 2024. Retrieved from: https://www.retailbrew.com/stories/2024/05/20/why-ebay-is-declaring-the-first-recommerce-day?

[16] Bobby JACK - *Google Offers a New Way to Show Results: Here's How You Find It*. MakeUseOf, May 22, 2024. Retrieved from: https://www.makeuseof.com/google-offers-a-new-way-to-show-results-heres-how-you-find-it/?

[17] Andrei IUNISOV - *Safely Leveraging AI in SEO: Best practices and Google compliance tips*. MarketingSherpa, May 23rd, 2024. Retrieved from: https://sherpablog.marketingsherpa.com/search-marketing/ai-seo/?

[18] Aamer BAIG, Douglas MERRILL, Danesha MEAD and Stephen XU - *Moving past gen AI's honeymoon phase: Seven hard truths for CIOs to get from pilot to scale*. McKinsey Digital, May 13, 2024. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/moving-past-gen-ais-honeymoon-phase-seven-hard-truths-for-cios-to-get-from-pilot-to-scale

[19] Alex SINGLA, Alexander SUKHAREVSKY, Ben ELLENCWEIG, Guilherme CRUZ, Carlo PALERMO - *Strategic alliances for gen AI: How to build them and make them work*. McKinsey Quarterly, May 22, 2024. Retrieved from: https://www.mckinsey.com/capabilities/quantumblack/our-insights/strategic-alliances-for-gen-ai-how-to-build-them-and-make-them-work?

[20] Tahir MUSHTAQ, Onder KETHUDA - *The effect of AI adoption on customer perceptions and purchase intentions in the changing business management landscape*, in Advances in Management and Innovation, Conference proceedings, Cardiff School of Management, pp. 128-134, 21 - 22 May 2024. Retrieved from: https://scholar.google.com/scholar_url?url=https://figshare.cardiffmet.ac.uk/articles/conference_contribution/Advances_in_Management_and_Innovation_Conference_Proceedings_2024/25860580/1/files/46439878.pdf%23page%3D122&hl=en&sa=X&d=17946673746450609&ei=xBFRZpfBHe2q6rQPysWIiAI&scisig=AFWwaebP1ls0hZHY0fNP32iq_dWS&oi=scholaralrt&hist=IoUu45oAAAAJ:12317441535909106215:AFWwaeZPDkveuOCMEVM5Guld6DwZ&html=&pos=6&folt=rel&fols=

[21] Theodor PURCĂREA, Valeriu IOAN-FRANC, Ştefan-Alexandru IONESCU and Ioan Matei PURCĂREA - *The Profound Nature of Linkage Between the Impact of the Use of Artificial Intelligence in Retail on Buying and Consumer Behavior and Consumers' Perceptions of Artificial Intelligence on the Path to the Next Normal*. Amfiteatru Economic, 23(56), pp.9-32. p-ISSN 1582-9146, e-ISSN 2247–9104. February 2021.

[22] Costel Iliuta NEGRICEA & Ioan Matei PURCAREA - *Digital Marketing, Digital Disruption*, and the New Rules of Digitization. Holistic Marketing Management, vol. 5(1), pp. 11-17. ISSN 2247-1189. March 2015.

[23] Neil T. BENDLE, Paul W. FARRIS, Phillip E. PFEIFE, David J. REIBSTEIN - *Marketing Metrics* – p. XV. ISBN 10: 0-13-408596-5, 13: 978-0-13-408596-8. Publisher Pearson FT Press. August 27, 2015.

[24] Ioan Matei PURCAREA - *Digital Customer Engagement, CX Future and the Advance of the Metaverse-Related Activities Including in E-Commerce Retail*. Holistic Marketing Management, vol. 12(3), pp. 13-21. ISSN 2247-1189. October 2022.

[25] Neil PERKIN - *Agile Marketing: Unlock Adaptive and Data-driven Marketing for Long-term Success* – p. 12. ISBN-10: 1398605107, ISBN-13: 978-1398605107. Publisher Kogan Page. April 26, 2022.

[26] Silvio PALUMBO, Benjamin REHBERG, and Haoran LI - *Software Projects Don't Have to Be Late, Costly, and Irrelevant*. BCG, April 30, 2024. Retrieved from: https://www.bcg.com/publications/2024/software-projects-dont-have-to-be-late-costly-and-irrelevant

[27] Mattia BIANCHI, Giacomo MARZI, Marina DABIĆ - *Guest editorial: agile beyond software—in search of flexibility in a wide range of innovation projects and industries*. IEEE Transactions on Engineering Management, 69:3454–3458, pp. 1-9. ISSN 1558-0040. 03 November 2022. https://doi.org/10.1109/TEM.2022.3206408

[28] Stefano MAGISTRETTI & Daniel TRABUCCHI - *Agile-as-a-tool and agile-as-a-culture: a comprehensive review of agile approaches adopting contingency and*

configuration theories. Review of Managerial Science. ISSN 18636683, 18636691. 09 March 2024. https://doi.org/10.1007/s11846-024-00745-1

[29] Diane E. STRODE, Helen SHARP, Leonor BARROCA, Peggy GREGORY, Katie TAYLOR - *Tensions in organizations transforming to agility*. IEEE Transactions on Engineering Management, 69(6), pp. 3572 – 3583. ISSN 1558-0040. 28 April 2022. https://doi.org/10.1109/TEM.2022.3160415

[30] Madhad Ali AL JABRI, Samira SHALOH, Nour SHAKHOOR, Mohamed Yacine HADDOUD, Bader Yousef OBEIDAT - *The Impact of Dynamic Capability on Enterprise Agility: The Intervening Roles of Digital Transformation and IT Alignment*. Journal of Open Innovation: Technology, Market, and Complexity. ISSN 2199-8531. 4 April 2024. https://doi.org/10.1016/j.joitmc.2024.100266

[31] Abayomi BAIYERE, Hannu SALMELA, Harri NIEMINEN, Tomi KANKAINEN - *Assessing digital capabilities for digital transformation—The MIND framework*. Information Systems Journal, pp. 1–33. ISSN 13501917, 13652575. 7 March 2024. https://onlinelibrary.wiley.com/doi/pdf/10.1111/isj.12519

[32] Chris BRADLEY, Jan MISCHKE, Marc CANAL - *A new McKinsey look at the role of productivity in sustainable growth*. McKinsey Global Institute (MGI) report "Investing in productivity growth", April 23, 2024. Retrieved from: https://www.mckinsey.com/about-us/new-at-mckinsey-blog/new-mgi-research-will-we-invest-in-productivity-to-thrive?

[33] Bernardo BETLEY, Hana DIB, Bjørnar JENSEN, and Bernhard MÜHLREITER - *The state of cloud computing in Europe: Increasing adoption, low returns, huge potential*. McKinsey Digital, pp. 1-12, April 2, 2024. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-state-of-cloud-computing-in-europe-increasing-adoption-low-returns-huge-potential

[34] Chhavi ARORA, Will FORREST, Mark GU and James KAPLAN - *In search of cloud value: Can generative AI transform cloud ROI?* McKinsey Digital, Report, pp. 1-7, November 15, 2023. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/in-search-of-cloud-value-can-generative-ai-transform-cloud-roi?

[35] Rob KLACZAK, Sanjay KUMAR, Colin GUNTER, Sanjay KANIYAR, Yash RAJYAGURU and Mayukh SAMAJDER - *From legacy to cloud: Lessons from the trenches*. McKinsey Digital, Interview, pp. 1-7, March 11, 2024. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/from-legacy-to-cloud-lessons-from-the-trenches?

[36] François CANDELON, Guillaume SAJUST DE BERGUES, David ZULUAGA MARTÍNEZ, Harsha Chandra SHEKAR, and Marcos AGUIAR - *The Benefits of Data Sharing Now Outweigh the Risks*. BCG Henderson Institute, Boston Consulting Group's strategy think tank, pp. 1-8, April 29, 2024. Retrieved from:

https://www.bcg.com/publications/2024/the-benefits-of-data-sharing-now-outweigh-the-risks

[37] Anusha DHASARATHY, Pranav HIMATSINGKA, Naufal KHAN and Thomas ELSNER - *Five key questions to get a tech transformation right*. McKinsey Digital, Survey, pp. 1-12, July 14, 2023. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/five-key-questions-to-get-a-tech-transformation-right?

[38] Oliver BOSSERT and Björn MÜNSTERMANN - *Business's 'It's not my problem' IT problem*. McKinsey Technology, March 2023, pp. 1-8, March 30, 2023. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/businesss-its-not-my-problem-it-problem

[39] CAIML (Center for Artificial Intelligence and Machine Learning) - *The Digital Humanism Initiative*. DIGHUM, 26-05-2024. Retrieved from: https://caiml.org/dighum/

[40] CAIML. Lecture Series on Digital Humanism - *Generative AI Degrades Online Communities*. The Digital Humanism Initiative, DIGHUM, 26-05-2024. Retrieved from: https://caiml.org/dighum/dighum-lectures/gordon-burtch-generative-ai-degrades-online-communities-2024-06-11/

[41] Chandra GNANASAMBANDAM and Rikki SINGH - *Pali Bhat on building intelligent products*. McKinsey & Company, Technology, Media &Telecommunications, March 29, 2024. Retrieved from: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/pali-bhat-on-building-intelligent-products?

[42] Mark GU and James KAPLAN - *Getting ahead in the cloud*. McKinsey Digital, pp. 1-13, November 30, 2023. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/getting-ahead-in-the-cloud?

Bibliography

AHLAWAT P., BEDARD J., DAMANI S., FELDMAN B., LUCERO E., SAMDARIA A. and SIPPERLY G. - *The Art of Scaling GenAI in Software*. BCG, Technology, Media, and Telecommunications, May 17, 2024. Retrieved from: https://www.bcg.com/publications/2024/the-art-of-scaling-genai-in-software?

AL JABRI M. A., SHALOH S., SHAKHOOR N., HADDOUD M. Y. and OBEIDAT B. Y. - *The Impact of Dynamic Capability on Enterprise Agility: The Intervening Roles of Digital Transformation and IT Alignment*. Journal of Open Innovation: Technology, Market, and Complexity. ISSN 2199-8531. 4 April 2024. https://doi.org/10.1016/j.joitmc.2024.100266

ARORA C., FORREST W., GU M. and KAPLAN J. - *In search of cloud value: Can generative AI transform cloud ROI?*. McKinsey Digital, Report, pp. 1-7, November 15, 2023. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/in-search-of-cloud-value-can-generative-ai-transform-cloud-roi?

BAIG A., MERRILL D., SINHA M., MEAD D. and XU S. - *Moving past gen AI's honeymoon phase: Seven hard truths for CIOs to get from pilot to scale*. McKinsey Digital, May 13, 2024. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/moving-past-gen-ais-honeymoon-phase-seven-hard-truths-for-cios-to-get-from-pilot-to-scale

BAIYERE A., SALMELA H., NIEMINEN H., KANKAINEN T. - *Assessing digital capabilities for digital transformation—The MIND framework*. Information Systems Journal, pp. 1–33. ISSN 13501917, 13652575. 7 March 2024. https://onlinelibrary.wiley.com/doi/pdf/10.1111/isj.12519

BENDLE N. T., FARRIS P. W., PFEIFE P. E., REIBSTEIN D. J. - *Marketing Metrics* – p. XV. ISBN 10: 0-13-408596-5, 13: 978-0-13-408596-8. Publisher Pearson FT Press. August 27, 2015.

BETLEY B., DIB H., JENSEN B. and MÜHLREITER B. - *The state of cloud computing in Europe: Increasing adoption, low returns, huge potential*. McKinsey Digital, pp. 1-12, April 2, 2024. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-state-of-cloud-computing-in-europe-increasing-adoption-low-returns-huge-potential

BIANCHI M., MARZI G., DABIĆ M. - *Guest editorial: agile beyond software—in search of flexibility in a wide range of innovation projects and industries*. IEEE Transactions on Engineering Management, 69:3454–3458, pp. 1-9. ISSN 1558-0040. 03 November 2022. https://doi.org/10.1109/TEM.2022.3206408

BOSSERT O. and MÜNSTERMANN B. - *Business's 'It's not my problem' IT problem*. McKinsey Technology, March 2023, pp. 1-8, March 30, 2023. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/businesss-its-not-my-problem-it-problem

BRADLEY C., MISCHKE J., CANAL M. - *A new McKinsey look at the role of productivity in sustainable growth*. McKinsey Global Institute (MGI) report "Investing in productivity growth", April 23, 2024. Retrieved from: https://www.mckinsey.com/about-us/new-at-mckinsey-blog/new-mgi-research-will-we-invest-in-productivity-to-thrive?

CAIML (Center for Artificial Intelligence and Machine Learning) - *The Digital Humanism Initiative*. DIGHUM, 26-05-2024. Retrieved from: https://caiml.org/dighum/

CANDELON F., SAJUST DE BERGUES G., ZULUAGA MARTÍNEZ D., SHEKAR H. C. and AGUIAR M. - *The Benefits of Data Sharing Now Outweigh the Risks*. BCG

Henderson Institute, Boston Consulting Group's strategy think tank, pp. 1-8, April 29, 2024. Retrieved from: https://www.bcg.com/publications/2024/the-benefits-of-data-sharing-now-outweigh-the-risks

CHEN J., VAN DOORN N., GROHMANN R., POLLIO, A., SORIANO C. - *Introducing Platforms & Society*. Platforms & Society Volume 1: 1–2, eISSN 29768624, ISSN 29768624. 2024. https://journals.sagepub.com/doi/10.1177/29768624241235492

DE REUVER M., SØRENSEN C. BASOLE R. C. - *The digital platform: a research agenda*. Journal of Information Technology 33, 124–135. eISSN 14664437, ISSN 02683962. June 1, 2018. https://doi.org/10.1057/s41265-016-0033-3

DHASARATHY A., HIMATSINGKA P., KHAN N. and ELSNER T. - *Five key questions to get a tech transformation right*. McKinsey Digital, Survey, pp. 1-12, July 14, 2023. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/five-key-questions-to-get-a-tech-transformation-right?

DIAMANDIS P. H., KOTLER S. - *The Future Is Faster Than You Think: How Converging Technologies Are Transforming Business, Industries, and Our Lives* - [pages 87, 100]. ISBN 1982109661, 9781982109660. Publisher Simon and Schuster. [January 28] 2020.

ELLIOTT G. and MAZUMDER S. - 8 Trends Accelerating the Future of E-Commerce in 2024. Publicis Sapient, May 2024. Retrieved from: https://www.publicissapient.com/insights/future-ecommerce-trends?

FELIN T., HOLWEG M. - *Theory Is All You Need: AI, Human Cognition, and Decision Making*. SSRN, pp. 1-46, February 24, 2024. http://dx.doi.org/10.2139/ssrn.4737265

GNANASAMBANDAM C. and SINGH R. - *Pali Bhat on building intelligent products*. McKinsey & Company, Technology, Media &Telecommunications, March 29, 2024. Retrieved from: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/pali-bhat-on-building-intelligent-products?

GREU V. - *Improving Information and Communication Technology to Innovate the Paradigms of World Progress, Part 1*. Romanian Distribution Committee Magazine, vol. 15(1), pp. 16-28, ISSN 2069-0134. March 2024

GU M. and KAPLAN J. - *Getting ahead in the cloud*. McKinsey Digital, pp. 1-13, November 30, 2023. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/getting-ahead-in-the-cloud?

HEIM S., SCHINDLER, R., TIANYU G., WITTGES H. - *And Here, All Threads Run Together: Understanding the Strategic Functions of Intermediaries in Digital Business Ecosystems*. European Conference of Information Systems, ECIS 2024 Proceedings. 2., pp. 1-2, 8, June 2024. Retrieved from: https://aisel.aisnet.org/ecis2024/is_governance/track21_is_govern/2

IUNISOV A. - *Safely Leveraging AI in SEO: Best practices and Google compliance tips*. MarketingSherpa, May 23rd, 2024. Retrieved from: https://sherpablog.marketingsherpa.com/search-marketing/ai-seo/?

JACK B. - *Google Offers a New Way to Show Results: Here's How You Find It*. MakeUseOf, May 22, 2024. Retrieved from: https://www.makeuseof.com/google-offers-a-new-way-to-show-results-heres-how-you-find-it/?

KLACZAK R., KUMAR S., GUNTER C., KANIYAR S., RAJYAGURU Y. and SAMAJDER, M. - *From legacy to cloud: Lessons from the trenches*. McKinsey Digital, Interview, pp. 1-7, March 11, 2024. Retrieved from: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/from-legacy-to-cloud-lessons-from-the-trenches?

LANG N. - *Imagine This... The End of Free Trade*. BCG, International Trade, May 13, 2024. Retrieved from: https://www.bcg.com/featured-insights/podcasts/imagine-this/end-of-free-trade?

CAIML. Lecture Series on Digital Humanism - *Generative AI Degrades Online Communities*. The Digital Humanism Initiative, DIGHUM, 26-05-2024. Retrieved from: https://caiml.org/dighum/dighum-lectures/gordon-burtch-generative-ai-degrades-online-communities-2024-06-11/

MAGISTRETTI S. & TRABUCCHI D. - *Agile-as-a-tool and agile-as-a-culture: a comprehensive review of agile approaches adopting contingency and configuration theories*. Review of Managerial Science. ISSN 18636683, 18636691. 09 March 2024. https://doi.org/10.1007/s11846-024-00745-1

MOSOIA C. - *The future must be carefully thought out and prepared*. Romanian Academy, March 26, 2024. Retrieved from: https://acad.ro/institutia/media_florin_filip.html

MÜHLBURGER M. and KRUMAY B. - *Towards a context-sensitive conceptualisation of digital transformation*. Journal of Information Technology, Vol. 0(0) 1–16, Special Issue: Digital Transformation as Geo-Political, Organizational and Technological Nexus, 124–135. eISSN 14664437, ISSN 02683962. November 23, 2023. https://journals.sagepub.com/doi/10.1177/02683962231219514

MUSHTAQ T., KETHUDA O. - *The effect of AI adoption on customer perceptions and purchase intentions in the changing business management landscape*, in Advances in Management and Innovation, Conference proceedings, Cardiff School of Management, pp. 128-134, 21 - 22 May 2024. Retrieved from: https://scholar.google.com/scholar_url?url=https://figshare.cardiffmet.ac.uk/articles/conference_contribution/Advances_in_Management_and_Innovation_Conference_Proceedings_2024/25860580/1/files/46439878.pdf%23page%3D122&hl=en&sa=X&d=17946673746450609&ei=xBFRZpfBHe2q6rQPysWIiAI&scisig=AFWwaebP1ls0hZHY0fNP32iq_d

WS&oi=scholaralrt&hist=IoUu45oAAAAJ:12317441535909106215:AFWwaeZPDkveuOCMEVM5Guld6DwZ&html=&pos=6&folt=rel&fols=

NEGRICEA C. I. & PURCAREA I. M. - *Digital Marketing, Digital Disruption, and the New Rules of Digitization*. Holistic Marketing Management, vol. 5(1), pp. 11-17. ISSN 2247-1189. March 2015.

NEWMAN A. A. - *Why eBay is declaring the first Recommerce Day*. Retail Brew, May 21, 2024. Retrieved from: https://www.retailbrew.com/stories/2024/05/20/why-ebay-is-declaring-the-first-recommerce-day?

PALUMBO S., REHBERG B. and LI H. - *Software Projects Don't Have to Be Late, Costly, and Irrelevant*. BCG, April 30, 2024. Retrieved from: https://www.bcg.com/publications/2024/software-projects-dont-have-to-be-late-costly-and-irrelevant

PERKIN, N. - *Agile Marketing: Unlock Adaptive and Data-driven Marketing for Long-term Success* – p. 12. ISBN-10: 1398605107, ISBN-13: 978-1398605107. Publisher Kogan Page. April 26, 2022.

PURCAREA I. M. - *Digital Customer Engagement, CX Future and the Advance of the Metaverse-Related Activities Including in E-Commerce Retail*. Holistic Marketing Management, vol. 12(3), pp. 13-21. ISSN 2247-1189. October 2022.

SCHAEFFER, C. - *How to Choose the Best CRM Software for your Retail Company*. CustomerThink, February 15, 2021. Retrieved from: https://customerthink.com/how-to-choose-the-best-crm-software-for-your-retail-company/?

SINGLA A., SUKHAREVSKY A., ELLENCWEIG B., CRUZ G., PALERMO C. and ABRAHAM J. C. - *Strategic alliances for gen AI: How to build them and make them work*. McKinsey Quarterly, May 22, 2024. Retrieved from: https://www.mckinsey.com/capabilities/quantumblack/our-insights/strategic-alliances-for-gen-ai-how-to-build-them-and-make-them-work?

STRODE D. E., SHARP H., BARROCA L., GREGORY P. and TAYLOR K. - *Tensions in organizations transforming to agility*. IEEE Transactions on Engineering Management, 69(6), pp. 3572 – 3583. ISSN 1558-0040. 28 April 2022. https://doi.org/10.1109/TEM.2022.3160415

TAYLOR J. - *The most important question when designing AI*. The Brookings Institution, pp. 1-6. May 20, 2024. Retrieved from: https://www.brookings.edu/articles/the-most-important-question-when-designing-ai/?

VERHOEF P.C., BROEKHUIZEN T., BART Y., BHATTACHARYA A., DONG J. Q., FABIAN N., HAENLEIN M. - *Digital transformation: A multidisciplinary reflection and research agenda*. Journal of Business Research, 122: 889–901. ISSN 0148-2963. 02 November 2019. https://doi.org/10.1016/j.jbusres.2019.09.022

EXPLORING COMPRESSION STRATEGIES FOR LARGE LANGUAGE MODELS TOWARDS EFFICIENT ARTIFICIAL INTELLIGENCE IMPLEMENTATIONS

Doinița ȘENDRE[1]

Dana-Mihaela PETROŞANU[2]

Alexandru PÎRJAN[3]

Abstract

The rapid advancements of Artificial Intelligence (AI) technologies, particularly Large Language Models (LLMs), have brought and accelerated significant innovations across various domains. Regardless of their widespread usefulness, the scalability of LLMs poses considerable challenges, primarily due to their substantial demands on computational and energy resources. This article explores the importance of developing and applying effective compression techniques to mitigate these numerous challenges. Techniques such as pruning, quantization, and knowledge distillation are analyzed for their potential to decrease a LLM's size and its associated computational demands, while striving to maintain performance integrity. Each technique inherently presents unique trade-offs between model efficiency and accuracy, requiring a nuanced understanding of their applications. We have made an in-depth analysis into the complexities of implementing these techniques, highlighting the balance required between performance and compression, along with the complex process of customization to specific LLM architectures. The article further analyzes the very important validation and testing phases that are much needed for ensuring that compressed models perform adequately in real-world applications. We have also considered the future adaptability of compression techniques to evolving AI models and architectures. The conducted study emphasizes the ongoing need for innovative research in model compression in order to make AI technologies more sustainable and accessible across various sectors, thereby expanding their potential benefits while addressing the limitations and risks associated with their deployment.

Keywords: Large Language Models, Model Compression, Pruning Techniques, Quantization, Knowledge Distillation, Computational Efficiency, Neural Network Optimization, Artificial Intelligence Scalability

[1] PhD Professor, School of Management-Marketing, Romanian-American University, 1B, Expozitiei Blvd., district 1, code 012101, Bucharest, Romania, doinita.sendre@rau.ro

[2] PhD Lecturer, Department of Mathematics-Informatics, National University of Science and Technology Politehnica Bucharest, 313, Splaiul Independentei, district 6, code 060042, Bucharest, Romania, dana.petrosanu@upb.ro

[3] PhD Hab. Professor, School of Computer Science for Business Management, Romanian-American University, 1B, Expozitiei Blvd., district 1, code 012101, Bucharest, Romania, alexandru.pirjan@rau.ro

JEL Classification: O3, O33, O34, O35, O36

1. Introduction

In the contemporary landscape of AI, LLMs have emerged as very important tools, driving innovations across a multitude of domains. These models, which include prominent examples such as Generative Pre-trained Transformer (GPT) and Bidirectional Encoder Representations from Transformers (BERT), leverage vast datasets to understand and generate human-like text, offering capabilities that extend into Natural Language Processing (NLP), Machine Learning (ML), and beyond. The applications of LLMs are numerous, covering various areas such as automated content generation [1–3], real-time language translation [4–7], sentiment analysis [8–11], and even aiding in medicines discovery [4,12–15], neurosciences [16], geology [17] or in legal document analysis [18–20]. This wide-ranging applicability emphasizes the models' growing importance in both academic research and industry.

Nevertheless, the scalability of LLMs presents significant challenges. As the size and complexity of these models increase, so do their demands for computational and energy resources. Training state-of-the-art LLMs often requires extensive hardware setups, including multiple high-end Graphics Processing Units (GPUs) or Tensor Processing Units (TPUs), which are cost-prohibitive and also raise environmental concerns due to their high energy consumption. The intensive computational requirements can limit the accessibility of cutting-edge AI technologies, particularly for researchers and organizations with limited resources [1,3,13,21,22].

Given these constraints, the development and application of efficient compression techniques for LLMs assume critical importance. Compression techniques aim to reduce the size of neural networks without significantly compromising their performance. Techniques such as knowledge distillation, quantization, and pruning are employed to create lighter models that retain the efficacy of their larger counterparts while being more economical and environmentally sustainable. The implementation of such compression techniques is not devoid of challenges. Compressed models often face trade-offs between size, speed, and accuracy. While a smaller model size may result in faster computation times and lower energy usage, it might also lead to a decrease in the model's ability to generalize across tasks or maintain the same level of accuracy as the original model. Conversely, maintaining high accuracy can limit the degree of achievable compression. In addition, the process of model compression can be complex and requires careful tuning and validation to ensure that the reduced model still adheres to the performance standards necessary for practical applications.

Despite these challenges, the advantages of model compression are significant, offering a pathway towards more sustainable and accessible AI technologies. As this field progresses, understanding the nuances of various compression techniques and their impact on model performance will be essential. This article aims to make an in-depth analysis into these aspects, presenting an overview of the current methodologies, challenges, and potential future directions in the compression of LLMs.

2. Research methodology

The research methodology section of this study plays a very important role in establishing the analytical approach used to understand compression strategies for LLMs. Given the field's rapid growth and the increasing focus on computational efficiency, this study aims to identify and analyze relevant scientific literature through a well-defined and systematic process. The chosen methodology ensures comprehensive coverage of the relevant scientific literature research and addresses important issues concerning compression techniques in AI.

The Clarivate Web of Science (WoS) database has been selected for its extensive indexing of high-quality scientific literature across multiple disciplines. WoS provides a curated database of peer-reviewed articles, ensuring that the retrieved scientific works maintain the highest standards of academic rigor. The usage of the WoS database has enabled us to draw from a reliable source that is widely recognized in academic circles for its authority and breadth, thereby enhancing the credibility of the research findings.

Consequently, the selected database has been chosen for this study due to its comprehensive and authoritative collection of scientific literature across various disciplines. This choice has offered us curated indexing, allowing access to high-quality, reliable information. By employing the query "TS=((LLM* OR LARGE LANGUAGE MODEL*) AND (AI OR ARTIFICIAL INTELLIGENCE OR MACHINE LEARNING) AND (SIZE COMPRES* OR SIZE REDUC*))", we have ensured the retrieval of relevant scientific articles that explicitly discuss scientifically compression methods applied to LLMs.

The rationale for using WoS extends beyond its exhaustive scope to its advanced search capabilities. The specific query structure uses truncation and Boolean operators, ensuring inclusivity by capturing all forms and variations of keywords related to LLMs and their compression, within the context of AI. By employing the TS field (Topic Search), the search identifies these terms across article titles, abstracts, and keywords, enhancing the likelihood of retrieving relevant research. The inclusion of multiple terms and logical connectors ensures that the search is neither too broad nor too narrow, focusing on specific scientific discussions around LLM compression.

Further filtering the search results to include only scientific research articles, while excluding review articles and conference proceedings, has been necessary for ensuring the relevance and rigor of the findings. Research articles present primary research, providing original data, insights, and methodologies that directly contribute to the understanding and advancement of model compression in LLMs. This type of content is foundational, offering empirical evidence that other forms of literature may not provide in its final finished form. Excluding review articles helps minimize the bias in the selection process, ensuring that the analysis relies on new and innovative research contributions.

Similarly, excluding conference proceedings was required as these papers often represent preliminary findings. While conference proceedings are valuable, they may lack the depth and methodological transparency of fully developed research articles. This further ensures that the collected scientific literature pool for our conducted study consists of in-depth, thoroughly vetted studies that meet high academic standards, providing a solid foundation for advancing knowledge in this domain.

3. Analysis of trends over time and the main research areas in the scientific literature regarding the compressing techniques of LLMs

The exploration of trends over time in the publication of scientific literature related to the compression of LLMs reveals valuable insights into the evolving interests and advancements within this important area of AI research.

By examining the distribution of publications from 2020 throughout to April 20, 2024, we can discern shifts in focus, emerging themes, and the overall growth of the field. There has been a significant rise in publications from 2020 through 2022, peaking in 2022 with 16 publications, a slight decrease in 2023 and an early count of 7 publications for 2024 as of 20-April-2024. This suggests a growing interest and development in the field of model compression techniques.

Starting with 2020, the scientific community began to place an increased emphasis on the scalability and efficiency of LLMs [5,23–25]. The year saw a modest number of publications, totaling four, which reflects the emerging stage of awareness and technological development concerning the compression of these complex models. As LLMs like GPT-3 and others began to demonstrate potent capabilities in various domains, ranging from NLP to automated content generation, the computational and environmental costs associated with these models started to draw significant attention (Figure 1).

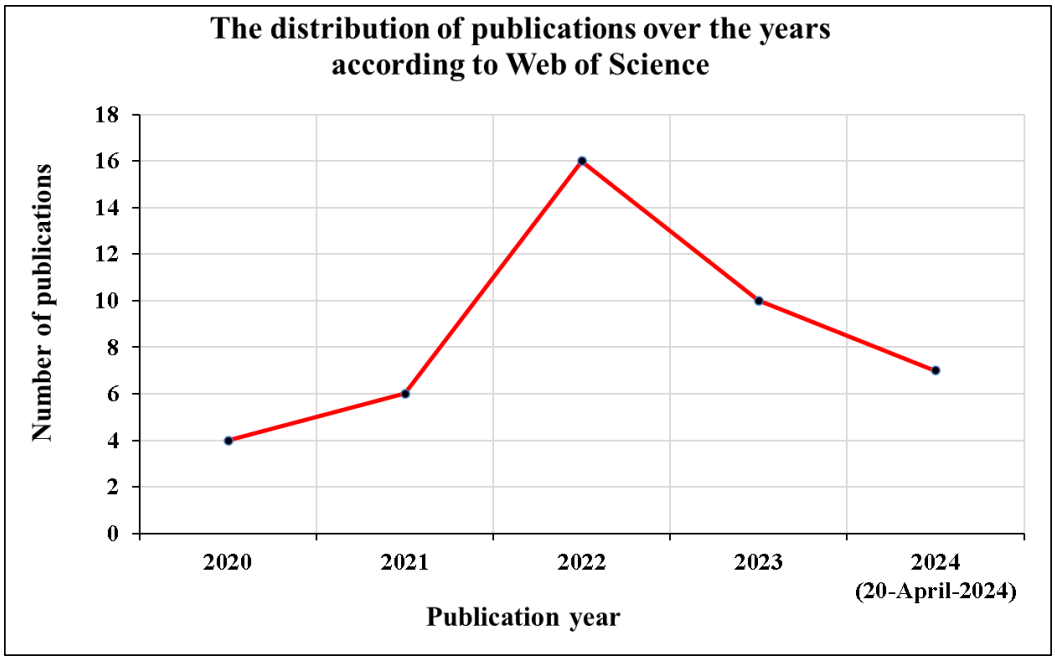


Figure 1. Exploration of trends over time in the publication of scientific literature related to the compression of LLMs[4]

In 2021, the publication count slightly increased to six [26–31]. This rise coincides with a broader comprehension within the AI research community and industry regarding the practical limitations imposed by the massive size and resource demands of state-of-the-art LLMs. The year 2021 saw enhancements in compression techniques such as knowledge distillation and quantization, tailored to mitigate these limitations. The slight increase in publications could be attributed to the consolidation of earlier findings and the initiation of more focused research projects aiming to refine and apply these emerging compression methods more effectively.

The year 2022 marked a significant peak with 16 publications, showcasing a robust interest and concerted effort in tackling the challenges associated with LLMs [2,4,9–11,15,32–41]. This surge can be interpreted as a response to the very important need for more sustainable AI practices, as the AI field grapples with the dual challenges of advancing technology and reducing its carbon footprint. During this period, research studies likely expanded into exploring the efficiency of individual compression techniques and especially their hybrid forms, such as integrating pruning with quantization, to achieve even greater reductions in model size and computational overhead.

In 2023, there was a notable decrease in publications to ten [12,13,18,42–48]. This drop might reflect a phase where the research community began to fit in the rapid advancements

[4] Source: The figure was devised based on the official data retrieved from Clarivate Web of Science in April 2024.

made in the previous years, shifting its focus from pioneering new methods to optimizing and validating existing techniques. It is also plausible that as some of the most accessible problems were being solved, the challenges became more complex, requiring longer cycles of research and development to achieve breakthroughs at the same pace as before.

Moving towards the partial data available for 2024, with seven publications recorded by April 20, the continuing interest in this area is clear, although with a publication rate that suggests a stabilization, or a slight decrease compared to the high mark in 2022 [1,3,21,22,49–51]. This trend could indicate several scenarios like a maturation of the field where major innovations become rarer and more incremental or a shift in focus towards other emerging areas of AI that require foundational research or simply the cyclical nature of research funding and publication outputs.

Overall, the trajectory of publications from 2020 to 2024 emphasizes a significant and growing recognition of the importance of developing effective compression techniques for LLMs. This trend is driven by the ongoing need to make AI technologies more accessible and sustainable, especially as these models find broader applications across industries and sectors. The data also suggests an increasing complexity in tackling the numerous challenges of compressing LLMs, reflecting deeper collaborations across computational and applied sciences.

As we look to the future, it is very important for the research community to continue promoting innovations in this space, particularly as the deployment environments for AI become more diverse and demanding. The adaptability of compression techniques to new model architectures and the integration of AI systems into edge devices and mobile platforms will likely be key areas of focus. Furthermore, as AI continues to integrate into more aspects of everyday life, ensuring the efficiency and sustainability of these systems will remain a major concern, driving ongoing research and interest in the field of LLM compression.

In the following, the research areas involved in the development and application of compression techniques for LLMs have been analyzed. The Computer Science field dominates the research areas with 41 publications, highlighting the central role of this field in the development and application of compression techniques for LLMs [2,5–13,18,21,23–27,29,31–33,36–38,41,42,45–47,49–60]. Engineering follows with 20 publications, indicating significant interdisciplinary work involving practical and technical aspects of implementing these techniques [1,4,6,10,18,27,30,32,33,35,42,43,47,49,51,52,56,60–62]. Physics [1,10,30,31,45,59], Telecommunications [6,27,32,33,47], and Materials Science [1,30,48] show fewer contributions, but emphasize the multi-disciplinary approach involving fundamental principles, data transmission, and possibly the materials used in computational hardware for AI. These findings suggest a robust and interdisciplinary effort in refining AI model efficiency, with a strong concentration in Computer Science. The trends reflect the academic and practical importance of this research area, along with

potential shifts in focus or emerging subfields within the broader AI research community (Figure 2).

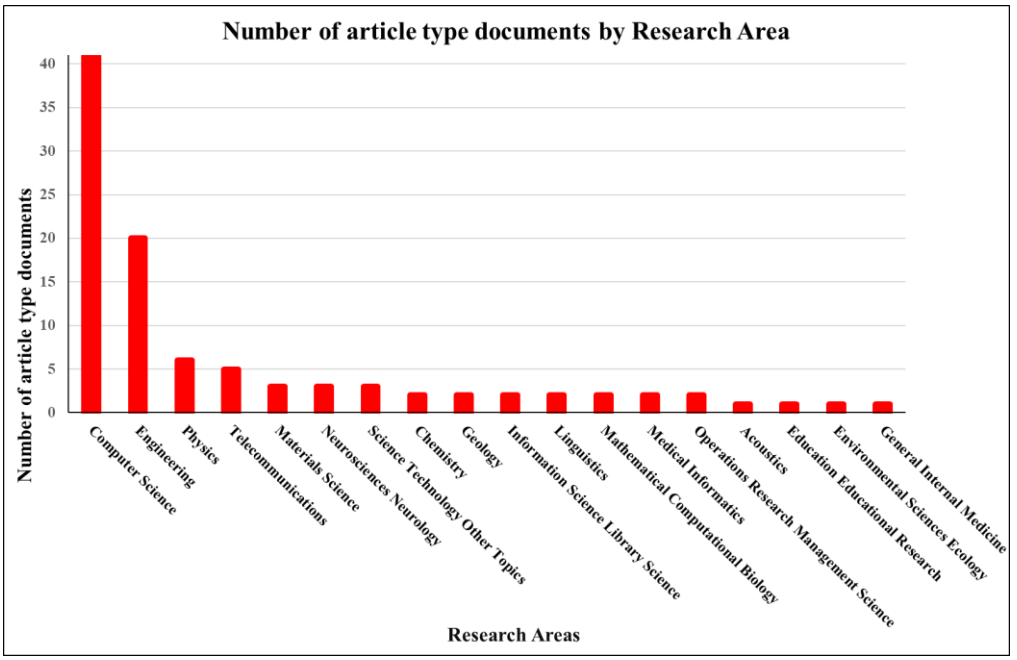


Figure 2. Research areas involved in the development and application of compression techniques for LLMs [5]

In order to obtain a comprehensive analysis of the correlation between different research areas involved in the development and application of compression techniques for LLMs, it is important to study the complex interactions and interdisciplinary efforts that characterize this field. This detailed exploration will highlight the predominant trends, collaborations across disciplines and highlight potential gaps along with opportunities for future research.

The rapid advancement of AI, particularly in the domain of LLMs such as GPT [1] and BERT [10,46,50], has required the exploration and implementation of various compression techniques to make these models more accessible and sustainable. The central challenge consists of reducing the computational and energy demands of these models without significantly compromising their performance. The interdisciplinary nature of this challenge has brought together experts from Computer Science, Engineering, Physics, Telecommunications, and Materials Science, each contributing unique perspectives and methodologies.

Computer Science is of extreme importance for research in AI model compression. It provides the theoretical frameworks, algorithms, and software implementations necessary

[5] Source: The figure was devised based on the official data retrieved from Clarivate Web of Science in April 2024.

for developing effective compression techniques such as pruning, quantization, and knowledge distillation. The field's dominance in the literature, accounting for over 68% of the publications, is indicative of its central role. Researchers in Computer Science work on algorithmic modifications and optimizations that can significantly reduce the size and computational complexity of LLMs. Pruning techniques are developed to remove redundant weights from neural networks, and quantization methods are applied to reduce the precision of the numerical values used in models, thereby decreasing the memory requirements and accelerating computation.

Engineering, with a substantial 33% of the publications, primarily focuses on the practical application and implementation of these compression techniques. Engineering research often bridges the gap between theoretical Computer Science models and real-world applications, addressing challenges related to hardware design, software-hardware integration, and the scalability of AI systems. Engineers work on adapting compression techniques to be compatible with existing and emerging hardware platforms, ensuring that compressed models are theoretically effective and practically viable. Engineering research might explore structured pruning techniques that are more amenable to conventional hardware architectures, therefore enhancing the efficiency of matrix operations that are extremely important for deploying models on general-purpose GPUs and on other accelerators.

Physics and Telecommunications contribute to a lesser extent, representing 10% and 8.33% of the publications, respectively. Nevertheless, their contributions are very important for understanding and improving the physical and network constraints associated with deploying AI models. Physics research might focus on the thermodynamic and quantum properties of materials used in hardware that supports AI computations, potentially leading to innovations in energy-efficient computing architectures. Meanwhile, Telecommunications research addresses the data transmission aspects, necessary for deploying AI models in distributed systems and for real-time applications such as automated content generation and real-time language translation.

Materials Science, though only accounting for 5% of the publications, plays a very important role in the development of new materials that can enhance the performance and efficiency of computational hardware. Research in this area might explore novel semiconductor materials or advanced manufacturing techniques that can be used to build more efficient GPUs and TPUs, which are needed for training and running large-scale AI models.

The correlation between these research areas can be seen in the collaborative efforts that aim to address the complex challenges posed by LLMs. The integration of Computer Science and Engineering is evident in the development of hardware-aware algorithms where compression techniques are tailored to the specific capabilities and limitations of the hardware used to run the models. Similarly, the collaboration between Materials Science

and Physics can lead to breakthroughs in hardware technology, such as the development of energy-efficient neural network processors that could further enhance the viability of compressed models. These interdisciplinary interactions promote innovation, ensure that the solutions developed are robust and applicable in a variety of settings. Advancements in Telecommunications research can enhance the deployment capabilities of AI models by improving the efficiency of data transfer across networks, which is very important for applications like cloud-based AI services and mobile AI applications [63].

Despite the robust collaborative efforts, there are gaps and challenges that need to be addressed to further advance the field of AI model compression. One of the significant challenges is the potential loss of accuracy and model generalizability due to compression. While compression techniques aim to minimize the impact on performance, there is often a trade-off between model size and its ability to perform complex tasks. Future research studies need to focus on developing compression techniques that can maintain high accuracy while achieving substantial reductions in model size and computational requirements.

Another area for potential improvement is the adaptability of compression techniques to new AI architectures and algorithms. As AI continues to evolve, with new models and approaches being developed at a rapid pace, compression techniques also need to be adaptable to these changes. This requires ongoing research and development to ensure that the techniques are effective for current models and also for future AI systems. The complexity of implementing these compression techniques also poses a significant challenge. Effective compression requires a deep understanding of both the architecture of the model and of the underlying algorithms. Customizing compression strategies to fit a specific model without losing essential functionalities demands extensive experimentation, complex engineering, and iterative tuning. This process is further complicated by the need for rigorous validation and testing to ensure that the compressed models perform adequately in real-world applications. Validation involves extensive testing against diverse datasets to identify any potential degradation in performance that may have been introduced during the compression process. This is extremely important for maintaining the trust and reliability of AI technologies in sensitive applications such as healthcare, finance, and autonomous driving.

Furthermore, the integration of emerging technologies such as federated learning and edge computing with model compression techniques could open new avenues for deploying AI in decentralized and privacy-preserving manners. These technologies allow for AI models to be trained and operated directly on user devices, reducing the need for data transmission and central processing, thereby enhancing user privacy and system efficiency. Conversely, this integration presents unique challenges, including the need for models that are compressed, robust enough to handle variable data environments and the computational capabilities of edge devices.

The field also faces ethical considerations, particularly in terms of bias and fairness. Compressed models, by necessity, simplify the representations learned by larger models, which could lead to the amplification of biases present in the training data. Ensuring that compression techniques do not exacerbate these biases requires careful attention to the design of both the model and the training process. This includes implementing strategies for bias detection and mitigation during both the training and compression phases.

Interdisciplinary collaboration will be of utmost importance for overcoming these challenges. As illustrated by the current distribution of research efforts, no single field can address all aspects of model compression alone. Collaborative projects that bring together experts from Computer Science, Engineering, Physics, Materials Science, and Telecommunications can leverage the strengths of each discipline to develop more comprehensive and effective solutions. Combining the theoretical insights from Computer Science with practical implementations from Engineering and cutting-edge materials from Physics can lead to the development of next-generation AI systems that are both powerful and efficient.

Future research studies should also focus on creating standardized frameworks and tools for implementing and evaluating model compression techniques. Such frameworks would help unify the efforts across different research areas and facilitate the sharing of best practices and benchmarks. This could accelerate the development of new compression methods and their adoption in industry and academia.

The compression of LLMs is a dynamic field that covers multiple disciplines, each contributing valuable and very important insights and technologies to address the challenges associated with these advanced AI systems. The ongoing collaboration and integration of diverse research areas are essential for advancing the state of the art in AI model compression. By continuing to encourage these interdisciplinary efforts and by addressing the technical, practical, and ethical challenges head-on, the research community can ensure that AI technologies become more sustainable, efficient, and accessible. This will expand the potential applications of AI and will also ensure that it is deployed in a manner that is beneficial and equitable for all sectors of society. The continual evolution of AI demands a proactive approach to research and development in model compression, making it an exciting and critical area of study in the expanding field of AI.

4. Challenges associated with the successful application of compression techniques to LLMs

In the following, we make an in-depth analysis into the main challenges associated with the successful application of compression techniques to LLMs. This analysis provides an overview of the complexities and main obstacles that researchers and practitioners face

when striving to reduce the computational and resource demands of these complex AI systems without compromising their performance.

One of the foremost challenges in compressing LLMs is maintaining a delicate balance between the model's size, speed, and accuracy. Compression techniques such as pruning, quantization, and knowledge distillation aim to reduce the physical size of the neural networks and the resources they require. Nevertheless, each method introduces potential trade-offs [24,42,49].

4.1. Pruning compression technique

Pruning is a compression technique applied to LLMs aimed at reducing the model's size and computational demands by eliminating redundant or non-critical parameters. It involves eliminating weights or neurons that contribute least to the model outputs, but it can also inadvertently remove elements that are necessary for certain tasks, leading to a loss in model generalizability or performance on specific benchmarks.

This technique is very important for making neural networks more efficient, especially in scenarios where computational resources are constrained, or costs need to be minimized. The concept of pruning stems from the observation that not all weights in a neural network contribute equally to its performance, suggesting that some can be removed with minimal impact on the model's efficacy.

Pruning can be broadly categorized into two types: structured and unstructured. Unstructured pruning involves the removal of individual weights across the network's matrices, leading to sparse connectivity between neurons. This type of pruning is highly flexible and can result in significant model size reduction. Nonetheless, it requires specialized software and hardware that can efficiently handle sparse matrices to realize computational speedups [64–66].

Structured pruning, on the other hand, removes entire rows, columns, or filters from matrices, leading to a reduced complexity in the network’s architecture. This form of pruning is more amenable to conventional hardware as it maintains the dense matrix structures necessary for optimized GPU utilization.

The implementation of pruning typically follows a three-step process:

I. Training: The neural network is first fully trained to learn the complex patterns in the data.

II. Removal (actual pruning): After training, weights that contribute the least to the output (often those with the smallest magnitudes) are identified and removed. This process can be iterative, involving re-training the network several times to refine which weights are pruned.

III. Fine-tuning: Once pruning is complete, the network undergoes additional training or fine-tuning with the remaining weights to recover any loss in performance due to the pruning process.

Various algorithms and criteria can be used to determine which weights to prune, such as magnitude-based pruning, where weights below a certain threshold are removed, and gradient-based pruning, which considers the sensitivity of the output to changes in each weight. One of the most evident benefits of pruning is the reduction in model size. By removing non-essential weights, the model becomes lighter, which saves storage space and also reduces the bandwidth needed for deploying the model in distributed systems or on edge devices.

Pruned models often require fewer computational resources. This is particularly beneficial in resource-limited environments where reducing the number of operations per inference can lead to faster response times and a lower power consumption. Structured pruning aligns well with existing hardware architectures, potentially increasing the efficiency of matrix operations. This can be particularly advantageous when deploying models on general-purpose GPUs or other accelerators that benefit from dense matrix operations.

The primary drawback of pruning is the potential reduction in model accuracy. Pruning important weights, even if they appear insignificant, can impair the model's ability to generalize from the training data to real-world applications. This requires a careful balance between the degree of pruning and the maintenance of model performance. Determining the optimal strategy for pruning is not a trivial aspect. It requires extensive experimentation with different pruning levels, methods, and fine-tuning cycles, which can be time-consuming and computationally expensive. Moreover, the criteria for pruning must be carefully chosen to avoid removing weights critical for certain tasks. The effectiveness of pruning is highly dependent on the quality of the initial model training. Poorly trained models might retain redundant weights while also lacking sufficient diversity in the weights that contribute to critical decision-making processes within the model.

Pruning presents a viable method for compressing LLMs by reducing unnecessary complexities and enhancing computational efficiency. While it offers considerable advantages in terms of model size and operational speed, it also poses challenges, including potential losses in accuracy and the complexity of its implementation. As the field of AI continues to evolve, further research and development are necessary to refine pruning techniques, ensuring they can reduce resource demands without significantly compromising the performance and adaptability of LLMs. This ongoing advancement will be extremely important in making AI technologies more accessible and sustainable, catering to a broader range of applications and environments.

4.2. Quantization compression technique

Quantization is a very important compression technique applied to LLMs to reduce the computational resources required for their operation while attempting to maintain acceptable levels of accuracy and performance. While this significantly decreases the size and increases the processing speed, it can lead to quantization errors, where the approximation of values causes a drop in accuracy. Quantization, in the context of ML and specifically in LLMs, refers to the process of reducing the precision of the numerical values used in a model. Traditionally, neural networks use floating-point arithmetic to perform calculations. These calculations, while accurate, are computationally expensive and resource intensive. Quantization addresses this aspect by approximating these floating-point numbers into lower-bit representations, typically using integer formats.

The basic principle behind quantization is the mapping of a continuous set of values (like those represented in floating-point) into a discrete set of values (like integers). This mapping reduces the memory requirements and speeds up the computation as integer operations are generally faster and more power-efficient on modern computing hardware than floating-point operations. Quantization can be broadly categorized [21,67,68] into three types:

I. Post-Training Quantization: This technique is applied after a model has been fully trained. The weights and activations, which are originally in floating-point, are converted into a lower-bit format. The main advantage of post-training quantization is its simplicity and ease of implementation as it does not require retraining the model.

II. Quantization-Aware Training (QAT): This method integrates quantization into the training process itself. By simulating the effects of quantization during training, QAT helps the model adjust its parameters to minimize the loss in accuracy that typically occurs when quantization is applied post-training.

III. Dynamic Quantization: This technique primarily quantizes the activations based on their distribution in real-time as they vary from one input to another. It is typically applied at the inference stage and is particularly useful for models where activation ranges can vary significantly.

Quantization incorporates various methodologies for mapping and representing values in a compact form. In uniform quantization, the range between the smallest and largest values is divided evenly, making this method straightforward and well-suited for hardware implementation due to its simplicity. Conversely, non-uniform quantization employs techniques such as logarithmic scaling, where the intervals between quantized values vary, enhancing fidelity particularly in regions near zero, where precision is often most important. Additionally, scalar and vector quantization approaches treat data differently, namely scalar quantization is processing each value independently, while vector quantization handles groups of values collectively, based on their overall distribution.

The application of quantization to LLMs brings significant advantages, primarily reducing the model's size through the use of fewer bits to represent each weight and activation, therefore diminishing the storage requirements. This compression also enables faster inference speeds as computations with lower-bit values are quicker on specialized hardware, a key aspect for applications demanding real-time processing. Furthermore, the reduced computational complexity leads to lower energy consumption, which is very important for models deployed on energy-constrained devices such as mobile phones and embedded systems.

One must take into account that quantization is not without its challenges. The reduction in bit precision can result in accuracy loss, particularly in complex tasks where high precision is essential. Moreover, some quantization techniques may depend on specific hardware capabilities to achieve computational benefits, limiting their usefulness in environments lacking such support. Implementing quantization also adds complexity, even if post-training quantization is relatively straightforward, quantization-aware training (QAT) necessitates adjustments to the training protocol, potentially introducing additional overhead in tuning and validation. Additionally, the effectiveness of quantization can vary significantly with different model architectures, requiring careful evaluation to understand its impact on each unique case.

Quantization presents a viable method for compressing LLMs, offering significant benefits in terms of reduced model size, increased processing speed, and enhanced energy efficiency. Nevertheless, the trade-offs in terms of potential accuracy loss and the complexities involved in its implementation must be carefully managed. Future research is needed to develop more advanced quantization techniques that can minimize accuracy loss while maximizing computational efficiency. This ongoing development is necessary for the broader adoption and application of LLMs in resource-constrained environments, contributing to the advancement of accessible and sustainable AI technologies. In the context of achieving LLM compression by means of quantization, it is important to highlight the extremely important balance between efficiency gains and potential pitfalls. As the field of AI continues to evolve, quantization will play an essential role in enabling the deployment of advanced neural networks in diverse and challenging real-world applications.

4.3. Knowledge Distillation compression technique

Knowledge distillation (KD) is a model compression technique that has gained significant attention in the field of AI, specifically in the context of LLMs. This technique involves transferring knowledge from a larger, often more cumbersome model (referred to as the "teacher") to a smaller, more efficient model (referred to as the "student"). The predominant goal of KD is to enable the student model to perform at par with the teacher model while requiring less computational power and memory, thereby making the deployment of AI

technologies more feasible in resource-constrained environments. The challenge consists in ensuring that the student model captures the nuanced understanding of the teacher model without needing the same computational resources. These techniques, while effective in reducing the size and computational load, must be applied judiciously to avoid undermining the model's ability to perform its intended tasks accurately.

KD operates on the premise that a complex model, which has been extensively trained and has a deep understanding of the data, can impart this knowledge to a simpler model. The process involves two main stages: the training of the teacher model and the distillation phase where the student learns from the teacher. The teacher model is typically a fully trained, high-capacity model that achieves high performance on the tasks for which it is designed. This model's depth and complexity allow it to capture subtle patterns in large datasets, making it an effective but resource-intensive solution.

During distillation, the student model is trained to predict the hard target labels of the training data and also to mimic the output distributions (soft targets) provided by the teacher model [69]. Soft targets are the probabilities or logits produced by the teacher for each class, which carry more information per example than hard labels. An example of this aspect is given by the fact that while a hard label might indicate the correct translation of a sentence, the soft targets could reveal how closely other potential translations compare, according to the teacher's understanding. The distillation loss, typically a form of cross-entropy between the soft targets of the teacher and the outputs of the student, guides the student training. This loss is often combined with the traditional hard target loss, balancing learning from the teacher and adhering to the ground truth [70].

A common method to enhance the effectiveness of KD is temperature scaling. This involves modifying the "softmax" function used during training by introducing a temperature parameter that controls the smoothness of the output probability distribution [71]. A higher temperature results in a softer probability distribution over classes, which provides more informative gradients for the student model during training. Beyond basic temperature scaling, custom strategies may involve adjusting the layers of the student that receive guidance from the teacher or altering the representation forms that the student should learn [69]. Some approaches focus on distilling intermediate representations (features) instead of just output probabilities [20].

One of the primary advantages of KD is the reduction in the size and computational requirements of the student model. This allows the deployment of complex AI models on devices with limited hardware capabilities, such as mobile phones and embedded devices. Despite its smaller size, a well-distilled student model can achieve performance close to that of the teacher model, making this technique particularly valuable for applications where performance cannot be compromised. By learning from the soft probabilities, student models often generalize better to new data compared to training from scratch or from hard

labels alone. This is because the soft labels encode additional information about the relationships between different classes [72].

The success of the student model heavily relies on the quality of the teacher model. A poorly trained teacher model can mislead the student, resulting in worse performance than a model trained directly from the data. The process of KD can be complex, involving careful tuning of the temperature parameter and of the distillation loss. Finding the right balance between learning from soft and hard targets requires extensive experimentation. Although the student model is lighter, the overall training time including the teacher's training can be substantial. Additionally, the resources required for training the teacher model are significant [73].

KD is particularly useful in scenarios where deploying large models is not feasible. It has been successfully applied in NLP tasks like machine translation [4–6], sentiment analysis [8,10], and question-answering [3] systems. Future research in KD is likely to focus on improving the efficiency of the distillation process, developing more robust student models that can outperform their teacher models, and extending the applicability of this technique to newer and more complex model architectures.

KD stands out as a promising technique for model compression. Its ability to transfer deep knowledge from large, resource-intensive models to more manageable counterparts without significant loss in performance is a major advantage. As AI technologies continue to evolve, optimizing and refining KD will be of extreme importance for advancing the practical deployment of AI systems, particularly in environments where resources are constrained.

The broader implications of KD extend beyond just model size and computational efficiency. By enabling powerful models to be compressed into more manageable forms, KD opens up new possibilities for AI applications in areas that were previously considered impractical due to hardware limitations. This includes real-time applications on mobile devices, such as live language translation and advanced on-device AI assistance, which can now benefit from deep learning insights without connectivity or high-power consumption. In addition, the democratization of access to advanced AI technologies through techniques like KD can help bridge the gap between well-funded, large-scale research institutions and smaller organizations or startups. This could offer equal opportunities and promote innovation across various sectors by making cutting-edge AI tools more accessible and less expensive to deploy.

One must also carefully take into consideration the challenges and limitations of KD. The reliance on a high-quality teacher model means that any inherent biases or errors in the teacher model are likely to be transferred to the student model. This could perpetuate or even amplify undesirable characteristics unless carefully managed. Additionally, the complexity of the distillation process itself may pose barriers to its widespread adoption, as it requires significant expertise and resources to implement effectively. Technical

challenges include the need for careful calibration of the distillation parameters, such as the temperature of the "softmax" function and the balance between different components of the loss function [73]. These parameters can significantly influence the effectiveness of the distillation and require detailed empirical evaluation to optimize.

Future studies might explore automated methods for optimizing these parameters or developing more adaptive distillation techniques that can dynamically adjust based on the student model's performance. Additionally, extending the concept of distillation beyond the teacher-student framework to include multiple teachers or collaborative distillation processes could offer new ways to enhance model performance and efficiency. As AI continues to advance, the role of KD is likely to grow, particularly in the development of AI models that are both powerful and practical for everyday applications.

The exploration of hybrid models that combine KD with other compression techniques such as pruning and quantization could yield even more efficient and robust AI systems. Furthermore, integrating KD into the lifecycle of AI development, from training through deployment, could help in continuously refining models in a resource-efficient manner. The integration of KD with emerging technologies like federated learning, where models are trained across multiple decentralized devices while keeping all the training data local, could further enhance privacy and scalability. This represents a significant step forward in creating AI systems that are both powerful and privacy-preserving.

KD is a powerful tool for model compression that offers significant benefits, including reduced model size, retained performance, and enhanced generalization. Nonetheless, its successful implementation requires careful consideration of various technical and ethical factors. Continued research and development in this area are essential to fully attain its potential and address the ongoing challenges. As the field progresses, KD will play an increasingly important role in making advanced AI technologies more accessible and sustainable, contributing significantly to the advancement of both the science and application of AI.

4.4. Complexity of implementation and trade-offs in model usefulness

The process of implementing compression techniques is inherently complex. It requires a deep understanding of both the architecture of the model and the underlying algorithms. Each LLM has unique characteristics based on its training data, structure, and intended use cases. Customizing compression strategies to fit a specific model without losing essential functionalities demands extensive experimentation, complex engineering, and iterative tuning.

The appropriate level of pruning or the best quantization scheme can vary widely between models. Developers must conduct numerous trials to identify the most effective parameters,

which can be a time-consuming and resource-intensive process. Additionally, each iteration must be rigorously tested to ensure that the compressed model still meets the performance criteria necessary for its application.

Compressed models often face significant trade-offs in their usefulness. A smaller model size generally leads to faster computation times and lower energy usage, which are beneficial for deploying models on edge devices or in environments where computing resources are limited. Nevertheless, these advantages may come at the cost of reduced accuracy or a diminished ability to generalize across different tasks and datasets. A model compressed for efficient translation might struggle with the subtleties of the language that are extremely important for tasks like sentiment analysis or legal document interpretation. Ensuring that a model remains versatile across various applications while being compressed is a significant challenge.

The endeavor to compress LLMs involves an array of complex procedures that require an in-depth comprehension of model architectures and underlying algorithms while also demanding careful customization to harmonize with the model's intrinsic characteristics without undermining its core functionalities. This complexity emerges from the inherent diversity and characteristics of LLMs, which are shaped by their training data, structural configurations, and the specific applications for which they are intended. As such, the implementation of compression techniques is a complex challenge that requires precise engineering, extensive experimentation, and iterative refinement to ensure that the integrity and effectiveness of the model are preserved.

The first step in the compression of a LLM is a thorough understanding of its architecture. LLMs, such as GPT and BERT, are built on complex neural network architectures that involve multiple layers of processing units, each responsible for understanding different aspects of the input data. These models employ mechanisms like attention and transformer architectures, which allow them to handle vast amounts of data and capture complex patterns in language.

Understanding these mechanisms is necessary for effective compression because each component of the architecture plays a specific role in the model's learning and inference processes. The attention mechanism in transformers facilitates the model's ability to focus on relevant parts of the input data, enhancing its understanding and generation of language. Compressing such a model without a detailed understanding of these components could lead to significant losses in functionality and performance, as critical aspects of the model’s capability to process and generate language might be inadvertently diminished.

Given the unique characteristics of each LLM, developing a one-size-fits-all compression approach is impractical. Instead, compression strategies must be tailored to fit specific models. This customization involves modifying existing compression techniques, such as

pruning, quantization, and KD, to align with the model's architecture and the requirements of its application domain.

The choice between structured and unstructured pruning depends on the specific model and the computational resources available. Structured pruning, which involves removing entire neurons or layers, may be suitable for models where computational efficiency is a priority and can be aligned with hardware that benefits from dense matrix operations. Conversely, unstructured pruning, which targets individual weights for removal, might be preferred when minimal impact on model performance is most important, and the available hardware can efficiently handle sparse matrices.

The process of implementing compression techniques is inherently iterative. It often requires multiple cycles of compression, testing, and tuning to find the optimal balance between model size, speed, and accuracy. Each iteration involves applying a compression technique, evaluating the model's performance on a set of benchmarks, and adjusting the parameters of the compression technique based on the outcomes.

This iterative process is very important because it allows for the gradual refinement of the compression strategy, minimizing the risk of degrading the model's performance. An example of this aspect is given by the fact that during the pruning process, an initial round of weight removal might show minimal impact on performance. Subsequent rounds might gradually increase the amount of pruning, with continuous monitoring to ensure that the performance does not fall below acceptable thresholds.

Implementing compression techniques requires advanced engineering skills and effective management of computational resources. Engineers must be adept at both software development and ML, with a deep understanding of how changes to the model’s architecture and parameters affect its behavior. This dual expertise is necessary to modify the model efficiently and to implement the compression techniques without introducing bugs or errors that could lead to unexpected behavior.

Additionally, managing computational resources effectively is extremely important, especially in environments where these resources are limited. Compression techniques can reduce the computational load of LLMs, making it feasible to deploy them on lower-end hardware or in resource-constrained environments. Nevertheless, achieving these reductions without excessive expenditure on computational resources during the compression process itself requires careful planning and management.

Despite the progress in model compression techniques, numerous challenges remain. The complexity of implementation is not merely a technical barrier but also a strategic one, involving decisions about which aspects of a model are essential for its intended use and how those aspects can be preserved during compression. Future studies in this area are likely to focus on developing more automated and adaptive compression techniques that can dynamically adjust to the model's performance during the compression process.

Advances in ML, such as reinforcement learning and meta-learning, could potentially be harnessed to automate the selection and tuning of compression parameters, reducing the need for manual intervention and making the process more efficient.

In addition, as AI and ML continue to evolve, the adaptability of compression techniques to new model architectures and training paradigms will be determinant factors. This includes the ability to compress models that use emerging techniques such as few-shot learning, unsupervised learning, or transfer learning, in situations where traditional compression approaches may not be directly applicable. The compression of LLMs represents a significant technical endeavor that holds the potential to transform the scalability and applicability of AI technologies, particularly in environments constrained by computational resources or hardware capabilities. The intricacies involved in achieving effective compression without compromising the functional integrity of these models require a disciplined approach to understanding, customizing, and iteratively refining the technologies used.

4.5. Validation and testing

An important part of implementing compression techniques consists in the validation phase, where the performance of the compressed model is rigorously assessed against benchmarks that are representative of real-world tasks. This validation is more than a one-time event, is a continuous process that ensures the model remains robust and performs well under different conditions and datasets. In order to attain an effective validation, diverse datasets that cover possible scenarios that the model may encounter in practical applications are used. This extensive testing helps identify any potential degradation in performance or in functionality that might have been introduced during the compression process. Additionally, it provides insights into how the model performs on various tasks, which is very important for understanding the trade-offs made during compression .

Researchers must consider that as new types of neural network architectures are developed, they may introduce different characteristics or sensitivities that need to be considered in the compression process. Architectures that use mechanisms beyond attention, such as those incorporating dynamic neural networks or capsule networks, often require novel approaches to compression that can accommodate their unique properties [74–77]. The complexity of compressing LLMs also necessitates collaboration across various disciplines within both academia and industry. This multidisciplinary approach brings together expertise from areas such as ML, software engineering, hardware design, and application development. Such collaborations can accelerate the refinement of compression techniques and help bridge the gap between theoretical research and practical application.

Collaborative efforts can also facilitate the sharing of best practices, tools, and resources, making it easier for smaller organizations or individuals to adopt and benefit from

compressed models. By democratizing access to advanced AI technologies, the broader AI community can drive innovation and application across a wider array of sectors. As compression techniques become more advanced and widely implemented, it is necessary to consider the ethical and social implications of deploying compressed models. These considerations include the transparency of model behaviors, the fairness of their outputs, and their accessibility to various user groups. Addressing these issues requires careful design of the compression process to ensure that it does not inadvertently introduce biases or reduce the model's ability to handle diverse data inputs fairly. It also involves developing guidelines and standards for the responsible use of compressed models, particularly in sensitive areas such as healthcare, law enforcement, and financial services.

The compression of LLMs is a complex, dynamic, and critically important area of research within AI. The successful implementation of compression techniques requires a deep understanding of the underlying technologies, a commitment to rigorous testing and validation, and a proactive approach to adapting these methods to new developments in the field. As these techniques continue to evolve, they promise to make AI more accessible and sustainable, thereby expanding the potential for these technologies to benefit society. The ongoing research, collaboration, and ethical consideration will be of extreme importance for obtaining these benefits while mitigating the risks associated with AI deployment in diverse environments.

5. Conclusions

The field of AI and ML is rapidly evolving, with new models and techniques being developed continuously. Ensuring that compression techniques remain effective as models evolve is a significant challenge. Compressed models must be robust, effective with current technologies and also adaptable to future developments. Researchers and developers must anticipate changes in hardware capabilities, data availability, and model architectures. This requires a forward-thinking approach to compression, where techniques are designed to optimize current models and adapt to next-generation AI technologies.

The compression of LLMs presents a complex array of challenges that cover technical, practical, and strategic dimensions. Balancing performance trade-offs, managing complex implementation processes, ensuring robust validation, and future-proofing technologies are all critical to the success of these initiatives. As the field progresses, developing more advanced, efficient, and adaptable compression techniques will be essential for making AI technologies more sustainable and accessible to a broader range of users and applications. This ongoing effort will require an intensive collaboration between researchers, engineers, and industry stakeholders in order to overcome these challenges and harness the full potential of compressed LLMs.

Acknowledgment

"The authors would like to express their gratitude for the logistics support received from the Center of Research, Consultancy and Training in Economic Informatics and Information Technology RAU-INFORTIS of the Romanian-American University and the Center for Computational Science and Machine Intelligence of the Romanian-American University."

References

[1] Tuan, N.T.; Moore, P.; Thanh, D.H. V; Pham, H. V A Generative Artificial Intelligence Using Multilingual Large Language Models for ChatGPT Applications. APPLIED SCIENCES-BASEL 2024, 14, doi:10.3390/app14073036.

[2] Lin, J.F.; Liu, Y.L.; Cleland-Huang, J. Information Retrieval versus Deep Learning Approaches for Generating Traceability Links in Bilingual Projects. Empir Softw Eng 2022, 27, doi:10.1007/s10664-021-10050-0.

[3] Zou, Z.; Mubin, O.; Alnajjar, F.; Ali, L. A Pilot Study of Measuring Emotional Response and Perception of LLM-Generated Questionnaire and Human-Generated Questionnaires. Sci Rep 2024, 14, doi:10.1038/s41598-024-53255-1.

[4] An, B.; Long, C.J. Paraphrase Based Data Augmentation For Chinese-English Medical Machine Translation. JOURNAL OF ELECTRONICS & INFORMATION TECHNOLOGY 2022, 44, 118–126, doi:10.11999/JEIT210926.

[5] Vázquez, R.; Raganato, A.; Creutz, M.; Tiedemann, J. A Systematic Study of Inner-Attention-Based Sentence Representations in Multilingual Neural Machine Translation. COMPUTATIONAL LINGUISTICS 2020, 46, 387–424, doi:10.1162/coli_a_00377.

[6] Zhu, X.N.; Yang, M.Y.; Zhao, T.J.; Zhu, C.H. Minimum Bayes-Risk Phrase Table Pruning for Pivot-Based Machine Translation in Internet of Things. IEEE ACCESS 2018, 6, 55754–55764, doi:10.1109/ACCESS.2018.2872773.

[7] Kim, K.; Park, E.J.; Shin, J.H.; Kwon, O.W.; Kim, Y.K. Divergence-Based Fine Pruning of Phrase-Based Statistical Translation Model. Comput Speech Lang 2017, 41, 146–160, doi:10.1016/j.csl.2016.06.006.

[8] Boiy, E.; Moens, M.F. A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts. Inf Retr Boston 2009, 12, 526–558, doi:10.1007/s10791-008-9070-z.

[9] Karthik, E.; Sethukarasi, T. A Centered Convolutional Restricted Boltzmann Machine Optimized by Hybrid Atom Search Arithmetic Optimization Algorithm for Sentimental Analysis. Neural Process Lett 2022, 54, 4123–4151, doi:10.1007/s11063-022-10797-7.

[10] Catelli, R.; Pelosi, S.; Esposito, M. Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian. Electronics (Basel) 2022, 11, doi:10.3390/electronics11030374.

[11] Nedjah, N.; Santos, I.; Mourelle, L.D. Sentiment Analysis Using Convolutional Neural Network via Word Embeddings. Evol Intell 2022, 15, 2295–2319, doi:10.1007/s12065-019-00227-4.

[12] Buonocore, T.M.; Crema, C.; Redolfi, A.; Bellazzi, R.; Parimbelli, E. Localizing In-Domain Adaptation of Transformer-Based Biomedical Language Models. J Biomed Inform 2023, 144, doi:10.1016/j.jbi.2023.104431.

[13] Tan, R.; Lin, Q.; Low, G.H.; Lin, R.X.; Goh, T.C.; Chang, C.C.E.; Lee, F.F.; Chan, W.Y.; Tan, W.C.; Tey, H.J.; et al. Inferring Cancer Disease Response from Radiology Reports Using Large Language Models with Data Augmentation and Prompting. JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION 2023, 30, 1657–1664, doi:10.1093/jamia/ocad133.

[14] Gronsbell, J.; Minnier, J.; Yu, S.; Liao, K.; Cai, T.X. Automated Feature Selection of Predictors in Electronic Medical Records Data. Biometrics 2019, 75, 268–277, doi:10.1111/biom.12987.

[15] Shoulson, I.; Arbatti, L.; Hosamath, A.; Eberly, S.W.; Oakes, D. Longitudinal Cohort Study of Verbatim-Reported Postural Instability Symptoms as Outcomes for Online Parkinson's Disease Trials. J Parkinsons Dis 2022, 12, 1969–1978, doi:10.3233/JPD-223274.

[16] Scheffer, L.K. Analysis Tools for Large Connectomes. Front Neural Circuits 2018, 12, doi:10.3389/fncir.2018.00085.

[17] Brenowitz, N.D.; Bretherton, C.S. Prognostic Validation of a Neural Network Unified Physics Parameterization. Geophys Res Lett 2018, 45, 6289–6298, doi:10.1029/2018GL078510.

[18] Abdelhakim, M.; Liu, B.Q.; Sun, C.G. Ar-PuFi: A Short-Text Dataset to Identify the Offensive Messages towards Public Figures in the Arabian Community. Expert Syst Appl 2023, 233, doi:10.1016/j.eswa.2023.120888.

[19] Petroșanu, D.M.; Pîrjan, A.; Tăbușcă, A. Tracing the Influence of Large Language Models across the Most Impactful Scientific Works. Electronics (Switzerland) 2023, 12.

[20] Kim, M.-Y.; Rabelo, J.; Babiker, H.K.B.; Rahman, M.A.; Goebel, R. Legal Information Retrieval and Entailment Using Transformer-Based Approaches. The Review of Socionetwork Strategies 2024, doi:10.1007/s12626-023-00153-z.

[21] Cho, M.S.; Vahid, K.A.; Fu, Q.C.; Adya, S.; Mundo, C.C.D.; Rastegari, M.; Naik, D.; Zatloukal, P. EDKM: An Efficient and Accurate Train-Time Weight Clustering for Large

Language Models. IEEE COMPUTER ARCHITECTURE LETTERS 2024, 23, 37–40, doi:10.1109/LCA.2024.3363492.

[22] Schonfeld, E.; Pant, A.; Shah, A.R.Y.; Sadeghzadeh, S.; Pangal, D.; Rodrigues, A.; Yoo, K.; Marianayagam, N.; Haider, G.; Veeravagu, A. Evaluating Computer Vision, Large Language, and Genome-Wide Association Models in a Limited Sized Patient Cohort for Pre-Operative Risk Stratification in Adult Spinal Deformity Surgery. J Clin Med 2024, 13, doi:10.3390/jcm13030656.

[23] Dong, L.B.; Guo, Q.M.; Wu, W.L.; Satpute, M.N. A Semantic Relatedness Preserved Subset Extraction Method for Language Corpora Based on Pseudo-Boolean Optimization. Theor Comput Sci 2020, 836, 65–75, doi:10.1016/j.tcs.2020.07.020.

[24] Choudhary, T.; Mishra, V.; Goswami, A.; Sarangapani, J. A Comprehensive Survey on Model Compression and Acceleration. Artif Intell Rev 2020, 53, 5113–5155, doi:10.1007/s10462-020-09816-7.

[25] Holtzen, S.; Van den Broeck, G.; Millstein, T. Scaling Exact Inference for Discrete Probabilistic Programs. PROCEEDINGS OF THE ACM ON PROGRAMMING LANGUAGES-PACMPL 2020, 4, doi:10.1145/3428208.

[26] Rajpal, D.; Garg, A.R.; Mahela, O.P.; Alhelou, H.H.; Siano, P. A Fusion-Based Hybrid-Feature Approach for Recognition of Unconstrained Offline Handwritten Hindi Characters. Future Internet 2021, 13, doi:10.3390/fi13090239.

[27] Kumar, V.; Recupero, D.R.; Riboni, D.; Helaoui, R. Ensembling Classical Machine Learning and Deep Learning Approaches for Morbidity Identification From Clinical Notes. IEEE ACCESS 2021, 9, 7107–7126, doi:10.1109/ACCESS.2020.3043221.

[28] McNorgan, C. The Connectivity Fingerprints of Highly-Skilled and Disordered Reading Persist Across Cognitive Domains. Front Comput Neurosci 2021, 15, doi:10.3389/fncom.2021.590093.

[29] Eronen, J.; Ptaszynski, M.; Masui, F.; Smywinski-Pohl, A.; Leliwa, G.; Wroczynski, M. Improving Classifier Training Efficiency for Automatic Cyberbullying Detection with Feature Density. Inf Process Manag 2021, 58, doi:10.1016/j.ipm.2021.102616.

[30] Li, X.; Luo, D.H.; Cheng, Y.; Wong, K.Y.; Hung, K. Identifying the Primary Odor Perception Descriptors by Multi-Output Linear Regression Models. APPLIED SCIENCES-BASEL 2021, 11, doi:10.3390/app11083320.

[31] Hoppe, N.; Winter, J.M.; Adami, S.; Adams, N.A. ALPACA - a Level-Set Based Sharp-Interface Multiresolution Solver for Conservation Laws. Comput Phys Commun 2022, 272, doi:10.1016/j.cpc.2021.108246.

[32] Wang, H.Z.; Qu, Z.H.; Zhou, Q.H.; Zhang, H.B.; Luo, B.Y.; Xu, W.C.; Guo, S.; Li, R.X. A Comprehensive Survey on Training Acceleration for Large Machine Learning Models in IoT. IEEE Internet Things J 2022, 9, 939–963, doi:10.1109/JIOT.2021.3111624.

[33] Bae, J.; Cheon, B.D.; Kim, H.Y. Pro-Attention: Efficient Probability Distribution Matching-Based Attention Through Feature Space Conversion. IEEE ACCESS 2022, 10, 131192–131201, doi:10.1109/ACCESS.2022.3229055.

[34] Chen, Y.C. Effects of Technology-Enhanced Language Learning on Reducing EFL Learners' Public Speaking Anxiety. Comput Assist Lang Learn 2022, doi:10.1080/09588221.2022.2055083.

[35] Naser, M.Z. Do We Need Exotic Models? Engineering Metrics to Enable Green Machine Learning from Tackling Accuracy-Energy Trade-Offs. J Clean Prod 2023, 382, doi:10.1016/j.jclepro.2022.135334.

[36] Pinto, J.P.; Viana, P.; Teixeira, I.; Andrade, M. Improving Word Embeddings in Portuguese: Increasing Accuracy While Reducing the Size of the Corpus. PeerJ Comput Sci 2022, 8, doi:10.7717/peerj-cs.964.

[37] Tan, J.H.; Tan, Y.H.; Chan, C.S.; Chuah, J.H. ACORT: A Compact Object Relation Transformer for Parameter Efficient Image Captioning. Neurocomputing 2022, 482, 60–72, doi:10.1016/j.neucom.2022.01.081.

[38] Tang, X.Y.; Zeng, S.; Yu, F.; Yu, W.; Sheng, Z.Y.; Kang, Z. Self-Supervised Anomaly Pattern Detection for Large Scale Industrial Data. Neurocomputing 2023, 515, 1–12, doi:10.1016/j.neucom.2022.09.069.

[39] Chen, X.; Wang, J.; Gomes, J.; Dubovik, O.; Yang, P.; Saito, M. Analytical Prediction of Scattering Properties of Spheroidal Dust Particles With Machine Learning. Geophys Res Lett 2022, 49, doi:10.1029/2021GL097548.

[40] Romero, R.W.A.; Viallon, M.; Spaltenstein, J.; Petrusca, L.; Bernard, O.; Belle, L.; Clarysse, P.; Croisille, P. CMRSegTools: An Open-Source Software Enabling Reproducible Research in Segmentation of Acute Myocardial Infarct in CMR Images. PLoS One 2022, 17, doi:10.1371/journal.pone.0274491.

[41] Ali, A.; Pinciroli, R.; Yan, F.; Smirni, E. Optimizing Inference Serving on Serverless Platforms. PROCEEDINGS OF THE VLDB ENDOWMENT 2022, 15, 2071–2084, doi:10.14778/3547305.3547313.

[42] Kashikar, P.; Sentieys, O.; Sinha, S. Lossless Neural Network Model Compression Through Exponent Sharing. IEEE Trans Very Large Scale Integr VLSI Syst 2023, 31, 1816–1825, doi:10.1109/TVLSI.2023.3307607.

[43] Zhang, J.F.; Zhang, Z.Y. Machine Learning Hardware Design for Efficiency, Flexibility, and Scalability [Feature]. IEEE CIRCUITS AND SYSTEMS MAGAZINE 2023, 23, 35–53, doi:10.1109/MCAS.2023.3302390.

[44] Woods, L.T.; Rana, Z.A. Constraints on Optimising Encoder-Only Transformers for Modelling Sign Language with Human Pose Estimation Keypoint Data. J Imaging 2023, 9, doi:10.3390/jimaging9110238.

[45] Bati, A.; Bryngelson, S.H. RoseNNa: A Performant, Portable Library for Neural Network Inference with Application to Computational Fluid Dynamics. Comput Phys Commun 2024, 296, doi:10.1016/j.cpc.2023.109052.

[46] Bhanage, D.A.; Pawar, A. V Robust Analysis of IT Infrastructure's Log Data with BERT Language Model. INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS 2023, 14, 705–714.

[47] Li, D.C.; Maulana, M.R.; Chou, L.D. NNSplit-SOREN: Supporting the Model Implementation of Large Neural Networks in a Programmable Data Plane. COMPUTER NETWORKS 2023, 222, doi:10.1016/j.comnet.2022.109537.

[48] Tüzemen, E.S.; Yüksek, A.G.; Demir, I.; Horoz, S.; Altuntas, I. Modeling of Temperature-Dependent Photoluminescence of GaN Epilayer by Artificial Neural Network. JOURNAL OF THE AUSTRALIAN CERAMIC SOCIETY 2023, 59, 1145–1159, doi:10.1007/s41779-023-00911-w.

[49] Zhu, J.; Wang, L.Y.; Han, X.; Liu, A.M.; Xie, T. Safety and Performance, Why Not Both? Bi-Objective Optimized Model Compression Against Heterogeneous Attacks Toward AI Software Deployment. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING 2024, 50, 376–390, doi:10.1109/TSE.2023.3348515.

[50] Pujar, S.; Zheng, Y.H.; Buratti, L.; Lewis, B.; Chen, Y.C.; Laredo, J.; Morari, A.; Epstein, E.; Lin, T.N.; Yang, B.; et al. Analyzing Source Code Vulnerabilities in the D2A Dataset with ML Ensembles and C-BERT. Empir Softw Eng 2024, 29, doi:10.1007/s10664-023-10405-9.

[51] Wu, R.H.; Zhu, X.Y.; Chen, J.S.; Liu, S.; Zheng, T.Y.; Liu, X.; An, H. SWattention: Designing Fast and Memory-Efficient Attention for a New Sunway Supercomputer. JOURNAL OF SUPERCOMPUTING 2024, doi:10.1007/s11227-024-05890-8.

[52] Chung, J.; Shin, T.; Yang, J.S. Simplifying Deep Neural Networks for FPGA-Like Neuromorphic Systems. IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS 2019, 38, 2032–2042, doi:10.1109/TCAD.2018.2877016.

[53] Huang, S.P.; Jin, L.W.; Xue, K.N.; Fang, Y. Online Primal-Dual Learning for a Data-Dependent Multi-Kernel Combination Model with Multiclass Visual Categorization Applications. Inf Sci (N Y) 2015, 320, 75–100, doi:10.1016/j.ins.2015.05.009.

[54] Ablimit, M.; Kawahara, T.; Hamdulla, A. Lexicon Optimization Based on Discriminative Learning for Automatic Speech Recognition of Agglutinative Language. Speech Commun 2014, 60, 78–87, doi:10.1016/j.specom.2013.09.011.

[55] Zahera, H.M.; El-Sisi, A.B. Accelerating Training Process in Logistic Regression Model Using OpenCL Framework. INTERNATIONAL JOURNAL OF GRID AND HIGH PERFORMANCE COMPUTING 2017, 9, 34–45, doi:10.4018/IJGHPC.2017070103.

[56] Tiwari, A.K.; Shreevastava, S.; Som, T.; Shukla, K.K. Tolerance-Based Intuitionistic Fuzzy-Rough Set Approach for Attribute Reduction. Expert Syst Appl 2018, 101, 205–212, doi:10.1016/j.eswa.2018.02.009.

[57] Manojlovic, I.; Svenda, G.; Erdeljan, A.; Gavric, M. Time Series Grouping Algorithm for Load Pattern Recognition. Comput Ind 2019, 111, 140–147, doi:10.1016/j.compind.2019.07.009.

[58] Nguyen, P.T.; Le, A.C.; Ho, T.B.; Nguyen, V.H. Vietnamese Treebank Construction and Entropy-Based Error Detection. Lang Resour Eval 2015, 49, 487–519, doi:10.1007/s10579-015-9308-5.

[59] Antcheva, I.; Ballintijn, M.; Bellenot, B.; Biskup, M.; Brun, R.; Buncic, N.; Canal, P.; Casadei, D.; Couet, O.; Fine, V.; et al. ROOT - A C++ Framework for Petabyte Data Storage, Statistical Analysis and Visualization. Comput Phys Commun 2009, 180, 2499–2512, doi:10.1016/j.cpc.2009.08.005.

[60] Ittiphalin, M.; Arnonkijpanich, B.; Pathumnakul, S. An Artificial Intelligence Model to Estimate the Fat Addition Ratio for the Mixing Process in the Animal Feed Industry. J Intell Manuf 2017, 28, 219–228, doi:10.1007/s10845-014-0972-x.

[61] Tapiador-Morales, R.; Linares-Barranco, A.; Jimenez-Fernandez, A.; Jimenez-Moreno, G. Neuromorphic LIF Row-by-Row Multiconvolution Processor for FPGA. IEEE Trans Biomed Circuits Syst 2019, 13, 159–169, doi:10.1109/TBCAS.2018.2880012.

[62] Abed, S.; Khalil, Y.; Modhaffar, M.; Ahmad, I. High-Performance Low-Power Approximate Wallace Tree Multiplier. INTERNATIONAL JOURNAL OF CIRCUIT THEORY AND APPLICATIONS 2018, 46, 2334–2348, doi:10.1002/cta.2540.

[63] Tăbuscă, A.; Tăbuscă, S.-M. Impact of 5G Technology in Global Economy. Cybersecurity and Legal Issues. Journal of Information Systems & Operations Management 2020, 13, 177–189.

[64] Moayed, H.; Mansoori, E.G.; Moosavi, M.R. An Efficient Pruning Method for Subgraph Matching in Large-Scale Graphs. Journal of Supercomputing 2023, 79, 10511–10532, doi:10.1007/s11227-023-05061-1.

[65] Cinquin, O. ChIP-GPT: A Managed Large Language Model for Robust Data Extraction from Biomedical Database Records. Brief Bioinform 2024, 25, doi:10.1093/bib/bbad535.

[66] Yan, H.; Liu, Y.; Jin, L.; Bai, X. The Development, application, and Future of LLM Similar to ChatGPT. Journal of Image and Graphics 2023, 28, 2749–2762, doi:10.11834/jig.230536.

[67] Li, W.; Hu, A.; Xu, N.; He, G. Quantization and Hardware Architecture Co-Design for Matrix-Vector Multiplications of Large Language Models. IEEE Transactions on Circuits and Systems I: Regular Papers 2024, doi:10.1109/TCSI.2024.3350661.

[68] Luo, Y.; Wei, Z.; Xu, G.; Li, Z.; Xie, Y.; Yin, Y. Enhancing E-Commerce Chatbots with Falcon-7B and 16-Bit Full Quantization. Journal of Theory and Practice of Engineering Science 2024, 4, 52–57, doi:10.53469/jtpes.2024.04(02).08.

[69] Tan, C.; Liu, J. Improving Knowledge Distillation With a Customized Teacher. IEEE Trans Neural Netw Learn Syst 2024, 35, 2290–2299, doi:10.1109/TNNLS.2022.3189680.

[70] Jiang, Y.; Weng, J.; Zhang, X.; Yang, Z.; Hu, W. A CNN-Based Born-Again TSK Fuzzy Classifier Integrating Soft Label Information and Knowledge Distillation. IEEE Transactions on Fuzzy Systems 2023, 31, 1843–1854, doi:10.1109/TFUZZ.2022.3215566.

[71] Yang, Z.; Zhang, Y.; Sui, D.; Ju, Y.; Zhao, J.; Liu, K. Explanation Guided Knowledge Distillation for Pre-Trained Language Model Compression. ACM Transactions on Asian and Low-Resource Language Information Processing 2024, 23, doi:10.1145/3639364.

[72] Du, Y.; Niu, J.; Wang, Y.; Jin, X. Multi-Stage Knowledge Distillation for Sequential Recommendation with Interest Knowledge. Inf Sci (N Y) 2024, 654, doi:10.1016/j.ins.2023.119841.

[73] Rao, J.; Meng, X.; Ding, L.; Qi, S.; Liu, X.; Zhang, M.; Tao, D. Parameter-Efficient and Student-Friendly Knowledge Distillation. IEEE Trans Multimedia 2024, 26, 4230–4241, doi:10.1109/TMM.2023.3321480.

[74] Arnes, J.I.; Horsch, A. Schema-Based Priming of Large Language Model for Data Object Validation Compliance. SSRN Electronic Journal 2023, doi:10.2139/ssrn.4453361.

[75] Sandholm, T.E.; Mukherjee, S.; Huberman, B.A. Randomness Is All You Need: Semantic Traversal of Problem-Solution Spaces with Large Language Models. SSRN Electronic Journal 2024, doi:10.2139/ssrn.4721407.

[76] Yang, J.; Wang, Y. Toward Auto-Modeling of Formal Verification for NextG Protocols: A Multimodal Cross- and Self-Attention Large Language Model Approach. IEEE Access 2024, 12, 27858–27869, doi:10.1109/ACCESS.2024.3366803.

[77] Insuasti, J.; Roa, F.; Zapata-Jaramillo, C.M. Computers' Interpretations of Knowledge Representation Using Pre-Conceptual Schemas: An Approach Based on the BERT and Llama 2-Chat Models. Big Data and Cognitive Computing 2023, 7, doi:10.3390/bdcc7040182.

Bibliography

Abdelhakim, M.; Liu, B.Q.; Sun, C.G. Ar-PuFi: A Short-Text Dataset to Identify the Offensive Messages towards Public Figures in the Arabian Community. Expert Syst Appl 2023, 233, doi:10.1016/j.eswa.2023.120888.

Abed, S.; Khalil, Y.; Modhaffar, M.; Ahmad, I. High-Performance Low-Power Approximate Wallace Tree Multiplier. INTERNATIONAL JOURNAL OF CIRCUIT THEORY AND APPLICATIONS 2018, 46, 2334–2348, doi:10.1002/cta.2540.

Ablimit, M.; Kawahara, T.; Hamdulla, A. Lexicon Optimization Based on Discriminative Learning for Automatic Speech Recognition of Agglutinative Language. Speech Commun 2014, 60, 78–87, doi:10.1016/j.specom.2013.09.011.

Ali, A.; Pinciroli, R.; Yan, F.; Smirni, E. Optimizing Inference Serving on Serverless Platforms. PROCEEDINGS OF THE VLDB ENDOWMENT 2022, 15, 2071–2084, doi:10.14778/3547305.3547313.

An, B.; Long, C.J. Paraphrase Based Data Augmentation For Chinese-English Medical Machine Translation. JOURNAL OF ELECTRONICS & INFORMATION TECHNOLOGY 2022, 44, 118–126, doi:10.11999/JEIT210926.

Antcheva, I.; Ballintijn, M.; Bellenot, B.; Biskup, M.; Brun, R.; Buncic, N.; Canal, P.; Casadei, D.; Couet, O.; Fine, V.; et al. ROOT - A C++ Framework for Petabyte Data Storage, Statistical Analysis and Visualization. Comput Phys Commun 2009, 180, 2499–2512, doi:10.1016/j.cpc.2009.08.005.

Arnes, J.I.; Horsch, A. Schema-Based Priming of Large Language Model for Data Object Validation Compliance. SSRN Electronic Journal 2023, doi:10.2139/ssrn.4453361.

Bae, J.; Cheon, B.D.; Kim, H.Y. Pro-Attention: Efficient Probability Distribution Matching-Based Attention Through Feature Space Conversion. IEEE ACCESS 2022, 10, 131192–131201, doi:10.1109/ACCESS.2022.3229055.

Bati, A.; Bryngelson, S.H. RoseNNa: A Performant, Portable Library for Neural Network Inference with Application to Computational Fluid Dynamics. Comput Phys Commun 2024, 296, doi:10.1016/j.cpc.2023.109052.

Bhanage, D.A.; Pawar, A. V Robust Analysis of IT Infrastructure's Log Data with BERT Language Model. INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS 2023, 14, 705–714.

Boiy, E.; Moens, M.F. A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts. Inf Retr Boston 2009, 12, 526–558, doi:10.1007/s10791-008-9070-z.

Brenowitz, N.D.; Bretherton, C.S. Prognostic Validation of a Neural Network Unified Physics Parameterization. Geophys Res Lett 2018, 45, 6289–6298, doi:10.1029/2018GL078510.

Buonocore, T.M.; Crema, C.; Redolfi, A.; Bellazzi, R.; Parimbelli, E. Localizing In-Domain Adaptation of Transformer-Based Biomedical Language Models. J Biomed Inform 2023, 144, doi:10.1016/j.jbi.2023.104431.

Buttrick, N. Studying Large Language Models as Compression Algorithms for Human Culture. Trends Cogn Sci 2024, 28, 187–189.

Catelli, R.; Pelosi, S.; Esposito, M. Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian. Electronics (Basel) 2022, 11, doi:10.3390/electronics11030374.

Chen, X.; Wang, J.; Gomes, J.; Dubovik, O.; Yang, P.; Saito, M. Analytical Prediction of Scattering Properties of Spheroidal Dust Particles With Machine Learning. Geophys Res Lett 2022, 49, doi:10.1029/2021GL097548.

Chen, Y.C. Effects of Technology-Enhanced Language Learning on Reducing EFL Learners' Public Speaking Anxiety. Comput Assist Lang Learn 2022, doi:10.1080/09588221.2022.2055083.

Cho, M.S.; Vahid, K.A.; Fu, Q.C.; Adya, S.; Mundo, C.C.D.; Rastegari, M.; Naik, D.; Zatloukal, P. EDKM: An Efficient and Accurate Train-Time Weight Clustering for Large Language Models. IEEE COMPUTER ARCHITECTURE LETTERS 2024, 23, 37–40, doi:10.1109/LCA.2024.3363492.

Choudhary, T.; Mishra, V.; Goswami, A.; Sarangapani, J. A Comprehensive Survey on Model Compression and Acceleration. Artif Intell Rev 2020, 53, 5113–5155, doi:10.1007/s10462-020-09816-7.

Chung, J.; Shin, T.; Yang, J.S. Simplifying Deep Neural Networks for FPGA-Like Neuromorphic Systems. IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS 2019, 38, 2032–2042, doi:10.1109/TCAD.2018.2877016.

Cinquin, O. ChIP-GPT: A Managed Large Language Model for Robust Data Extraction from Biomedical Database Records. Brief Bioinform 2024, 25, doi:10.1093/bib/bbad535.

Dong, G.; Chen, W. Blockwise Compression of Transformer-Based Models without Retraining. Neural Networks 2024, 171, 423–428, doi:10.1016/j.neunet.2023.12.001.

Dong, L.B.; Guo, Q.M.; Wu, W.L.; Satpute, M.N. A Semantic Relatedness Preserved Subset Extraction Method for Language Corpora Based on Pseudo-Boolean Optimization. Theor Comput Sci 2020, 836, 65–75, doi:10.1016/j.tcs.2020.07.020.

Du, Y.; Niu, J.; Wang, Y.; Jin, X. Multi-Stage Knowledge Distillation for Sequential Recommendation with Interest Knowledge. Inf Sci (N Y) 2024, 654, doi:10.1016/j.ins.2023.119841.

Eronen, J.; Ptaszynski, M.; Masui, F.; Smywinski-Pohl, A.; Leliwa, G.; Wroczynski, M. Improving Classifier Training Efficiency for Automatic Cyberbullying Detection with Feature Density. Inf Process Manag 2021, 58, doi:10.1016/j.ipm.2021.102616.

Gronsbell, J.; Minnier, J.; Yu, S.; Liao, K.; Cai, T.X. Automated Feature Selection of Predictors in Electronic Medical Records Data. Biometrics 2019, 75, 268–277, doi:10.1111/biom.12987.

Holtzen, S.; Van den Broeck, G.; Millstein, T. Scaling Exact Inference for Discrete Probabilistic Programs. PROCEEDINGS OF THE ACM ON PROGRAMMING LANGUAGES-PACMPL 2020, 4, doi:10.1145/3428208.

Hoppe, N.; Winter, J.M.; Adami, S.; Adams, N.A. ALPACA - a Level-Set Based Sharp-Interface Multiresolution Solver for Conservation Laws. Comput Phys Commun 2022, 272, doi:10.1016/j.cpc.2021.108246.

Huang, S.P.; Jin, L.W.; Xue, K.N.; Fang, Y. Online Primal-Dual Learning for a Data-Dependent Multi-Kernel Combination Model with Multiclass Visual Categorization Applications. Inf Sci (N Y) 2015, 320, 75–100, doi:10.1016/j.ins.2015.05.009.

Insuasti, J.; Roa, F.; Zapata-Jaramillo, C.M. Computers' Interpretations of Knowledge Representation Using Pre-Conceptual Schemas: An Approach Based on the BERT and Llama 2-Chat Models. Big Data and Cognitive Computing 2023, 7, doi:10.3390/bdcc7040182.

Ittiphalin, M.; Arnonkijpanich, B.; Pathumnakul, S. An Artificial Intelligence Model to Estimate the Fat Addition Ratio for the Mixing Process in the Animal Feed Industry. J Intell Manuf 2017, 28, 219–228, doi:10.1007/s10845-014-0972-x.

Jiang, Y.; Weng, J.; Zhang, X.; Yang, Z.; Hu, W. A CNN-Based Born-Again TSK Fuzzy Classifier Integrating Soft Label Information and Knowledge Distillation. IEEE Transactions on Fuzzy Systems 2023, 31, 1843–1854, doi:10.1109/TFUZZ.2022.3215566.

Kane, M.J.; King, C.; Esserman, D.; Latham, N.K.; Greene, E.J.; Ganz, D.A. A Compressed Large Language Model Embedding Dataset of ICD 10 CM Descriptions. BMC Bioinformatics 2023, 24, doi:10.1186/s12859-023-05597-2.

Karthik, E.; Sethukarasi, T. A Centered Convolutional Restricted Boltzmann Machine Optimized by Hybrid Atom Search Arithmetic Optimization Algorithm for Sentimental Analysis. Neural Process Lett 2022, 54, 4123–4151, doi:10.1007/s11063-022-10797-7.

Kashikar, P.; Sentieys, O.; Sinha, S. Lossless Neural Network Model Compression Through Exponent Sharing. IEEE Trans Very Large Scale Integr VLSI Syst 2023, 31, 1816–1825, doi:10.1109/TVLSI.2023.3307607.

Kim, K.; Park, E.J.; Shin, J.H.; Kwon, O.W.; Kim, Y.K. Divergence-Based Fine Pruning of Phrase-Based Statistical Translation Model. Comput Speech Lang 2017, 41, 146–160, doi:10.1016/j.csl.2016.06.006.

Kim, M.-Y.; Rabelo, J.; Babiker, H.K.B.; Rahman, M.A.; Goebel, R. Legal Information Retrieval and Entailment Using Transformer-Based Approaches. The Review of Socionetwork Strategies 2024, doi:10.1007/s12626-023-00153-z.

Kumar, V.; Recupero, D.R.; Riboni, D.; Helaoui, R. Ensembling Classical Machine Learning and Deep Learning Approaches for Morbidity Identification From Clinical Notes. IEEE ACCESS 2021, 9, 7107–7126, doi:10.1109/ACCESS.2020.3043221.

Li, D.C.; Maulana, M.R.; Chou, L.D. NNSplit-SOREN: Supporting the Model Implementation of Large Neural Networks in a Programmable Data Plane. COMPUTER NETWORKS 2023, 222, doi:10.1016/j.comnet.2022.109537.

Li, W.; Hu, A.; Xu, N.; He, G. Quantization and Hardware Architecture Co-Design for Matrix-Vector Multiplications of Large Language Models. IEEE Transactions on Circuits and Systems I: Regular Papers 2024, doi:10.1109/TCSI.2024.3350661.

Li, X.; Luo, D.H.; Cheng, Y.; Wong, K.Y.; Hung, K. Identifying the Primary Odor Perception Descriptors by Multi-Output Linear Regression Models. APPLIED SCIENCES-BASEL 2021, 11, doi:10.3390/app11083320.

Lin, J.F.; Liu, Y.L.; Cleland-Huang, J. Information Retrieval versus Deep Learning Approaches for Generating Traceability Links in Bilingual Projects. Empir Softw Eng 2022, 27, doi:10.1007/s10664-021-10050-0.

Lin, Y.J.; Chen, K.Y.; Kao, H.Y. LAD: Layer-Wise Adaptive Distillation for BERT Model Compression. Sensors 2023, 23, doi:10.3390/s23031483.

Luo, Y.; Wei, Z.; Xu, G.; Li, Z.; Xie, Y.; Yin, Y. Enhancing E-Commerce Chatbots with Falcon-7B and 16-Bit Full Quantization. Journal of Theory and Practice of Engineering Science 2024, 4, 52–57, doi:10.53469/jtpes.2024.04(02).08.

Manojlovic, I.; Svenda, G.; Erdeljan, A.; Gavric, M. Time Series Grouping Algorithm for Load Pattern Recognition. Comput Ind 2019, 111, 140–147, doi:10.1016/j.compind.2019.07.009.

McNorgan, C. The Connectivity Fingerprints of Highly-Skilled and Disordered Reading Persist Across Cognitive Domains. Front Comput Neurosci 2021, 15, doi:10.3389/fncom.2021.590093.

Moayed, H.; Mansoori, E.G.; Moosavi, M.R. An Efficient Pruning Method for Subgraph Matching in Large-Scale Graphs. Journal of Supercomputing 2023, 79, 10511–10532, doi:10.1007/s11227-023-05061-1.

Naser, M.Z. Do We Need Exotic Models? Engineering Metrics to Enable Green Machine Learning from Tackling Accuracy-Energy Trade-Offs. J Clean Prod 2023, 382, doi:10.1016/j.jclepro.2022.135334.

Nedjah, N.; Santos, I.; Mourelle, L.D. Sentiment Analysis Using Convolutional Neural Network via Word Embeddings. Evol Intell 2022, 15, 2295–2319, doi:10.1007/s12065-019-00227-4.

Nguyen, P.T.; Le, A.C.; Ho, T.B.; Nguyen, V.H. Vietnamese Treebank Construction and Entropy-Based Error Detection. Lang Resour Eval 2015, 49, 487–519, doi:10.1007/s10579-015-9308-5.

Petroșanu, D.M.; Pîrjan, A.; Tăbușcă, A. Tracing the Influence of Large Language Models across the Most Impactful Scientific Works. Electronics (Switzerland) 2023, 12.

Pinto, J.P.; Viana, P.; Teixeira, I.; Andrade, M. Improving Word Embeddings in Portuguese: Increasing Accuracy While Reducing the Size of the Corpus. PeerJ Comput Sci 2022, 8, doi:10.7717/peerj-cs.964.

Pujar, S.; Zheng, Y.H.; Buratti, L.; Lewis, B.; Chen, Y.C.; Laredo, J.; Morari, A.; Epstein, E.; Lin, T.N.; Yang, B.; et al. Analyzing Source Code Vulnerabilities in the D2A Dataset with ML Ensembles and C-BERT. Empir Softw Eng 2024, 29, doi:10.1007/s10664-023-10405-9.

Qi, H.; Cao, J.; Chen, S.; Zhou, J. Compressing Recurrent Neural Network Models through Principal Component Analysis. Stat Interface 2023, 16, 397–407, doi:10.4310/22-SII727.

Rajpal, D.; Garg, A.R.; Mahela, O.P.; Alhelou, H.H.; Siano, P. A Fusion-Based Hybrid-Feature Approach for Recognition of Unconstrained Offline Handwritten Hindi Characters. Future Internet 2021, 13, doi:10.3390/fi13090239.

Rao, J.; Meng, X.; Ding, L.; Qi, S.; Liu, X.; Zhang, M.; Tao, D. Parameter-Efficient and Student-Friendly Knowledge Distillation. IEEE Trans Multimedia 2024, 26, 4230–4241, doi:10.1109/TMM.2023.3321480.

Romero, R.W.A.; Viallon, M.; Spaltenstein, J.; Petrusca, L.; Bernard, O.; Belle, L.; Clarysse, P.; Croisille, P. CMRSegTools: An Open-Source Software Enabling Reproducible Research in Segmentation of Acute Myocardial Infarct in CMR Images. PLoS One 2022, 17, doi:10.1371/journal.pone.0274491.

Sandholm, T.E.; Mukherjee, S.; Huberman, B.A. Randomness Is All You Need: Semantic Traversal of Problem-Solution Spaces with Large Language Models. SSRN Electronic Journal 2024, doi:10.2139/ssrn.4721407.

Scheffer, L.K. Analysis Tools for Large Connectomes. Front Neural Circuits 2018, 12, doi:10.3389/fncir.2018.00085.

Schonfeld, E.; Pant, A.; Shah, A.R.Y.; Sadeghzadeh, S.; Pangal, D.; Rodrigues, A.; Yoo, K.; Marianayagam, N.; Haider, G.; Veeravagu, A. Evaluating Computer Vision, Large Language, and Genome-Wide Association Models in a Limited Sized Patient Cohort for Pre-Operative Risk Stratification in Adult Spinal Deformity Surgery. J Clin Med 2024, 13, doi:10.3390/jcm13030656.

Shoulson, I.; Arbatti, L.; Hosamath, A.; Eberly, S.W.; Oakes, D. Longitudinal Cohort Study of Verbatim-Reported Postural Instability Symptoms as Outcomes for Online Parkinson's Disease Trials. J Parkinsons Dis 2022, 12, 1969–1978, doi:10.3233/JPD-223274.

Tăbuscă, A.; Tăbuscă, S.-M. Impact of 5G Technology in Global Economy. Cybersecurity and Legal Issues. Journal of Information Systems & Operations Management 2020, 13, 177–189.

Tan, C.; Liu, J. Improving Knowledge Distillation With a Customized Teacher. IEEE Trans Neural Netw Learn Syst 2024, 35, 2290–2299, doi:10.1109/TNNLS.2022.3189680.

Tan, J.H.; Tan, Y.H.; Chan, C.S.; Chuah, J.H. ACORT: A Compact Object Relation Transformer for Parameter Efficient Image Captioning. Neurocomputing 2022, 482, 60–72, doi:10.1016/j.neucom.2022.01.081.

Tan, R.; Lin, Q.; Low, G.H.; Lin, R.X.; Goh, T.C.; Chang, C.C.E.; Lee, F.F.; Chan, W.Y.; Tan, W.C.; Tey, H.J.; et al. Inferring Cancer Disease Response from Radiology Reports Using Large Language Models with Data Augmentation and Prompting. JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION 2023, 30, 1657–1664, doi:10.1093/jamia/ocad133.

Tang, X.Y.; Zeng, S.; Yu, F.; Yu, W.; Sheng, Z.Y.; Kang, Z. Self-Supervised Anomaly Pattern Detection for Large Scale Industrial Data. Neurocomputing 2023, 515, 1–12, doi:10.1016/j.neucom.2022.09.069.

Tapiador-Morales, R.; Linares-Barranco, A.; Jimenez-Fernandez, A.; Jimenez-Moreno, G. Neuromorphic LIF Row-by-Row Multiconvolution Processor for FPGA. IEEE Trans Biomed Circuits Syst 2019, 13, 159–169, doi:10.1109/TBCAS.2018.2880012.

Tiwari, A.K.; Shreevastava, S.; Som, T.; Shukla, K.K. Tolerance-Based Intuitionistic Fuzzy-Rough Set Approach for Attribute Reduction. Expert Syst Appl 2018, 101, 205–212, doi:10.1016/j.eswa.2018.02.009.

Tuan, N.T.; Moore, P.; Thanh, D.H. V; Pham, H. V A Generative Artificial Intelligence Using Multilingual Large Language Models for ChatGPT Applications. APPLIED SCIENCES-BASEL 2024, 14, doi:10.3390/app14073036.

Tüzemen, E.S.; Yüksek, A.G.; Demir, I.; Horoz, S.; Altuntas, I. Modeling of Temperature-Dependent Photoluminescence of GaN Epilayer by Artificial Neural Network. JOURNAL OF THE AUSTRALIAN CERAMIC SOCIETY 2023, 59, 1145–1159, doi:10.1007/s41779-023-00911-w.

Vázquez, R.; Raganato, A.; Creutz, M.; Tiedemann, J. A Systematic Study of Inner-Attention-Based Sentence Representations in Multilingual Neural Machine Translation. COMPUTATIONAL LINGUISTICS 2020, 46, 387–424, doi:10.1162/coli_a_00377.

Wang, H.Z.; Qu, Z.H.; Zhou, Q.H.; Zhang, H.B.; Luo, B.Y.; Xu, W.C.; Guo, S.; Li, R.X. A Comprehensive Survey on Training Acceleration for Large Machine Learning Models in IoT. IEEE Internet Things J 2022, 9, 939–963, doi:10.1109/JIOT.2021.3111624.

Woods, L.T.; Rana, Z.A. Constraints on Optimising Encoder-Only Transformers for Modelling Sign Language with Human Pose Estimation Keypoint Data. J Imaging 2023, 9, doi:10.3390/jimaging9110238.

Wu, R.H.; Zhu, X.Y.; Chen, J.S.; Liu, S.; Zheng, T.Y.; Liu, X.; An, H. SWattention: Designing Fast and Memory-Efficient Attention for a New Sunway Supercomputer. JOURNAL OF SUPERCOMPUTING 2024, doi:10.1007/s11227-024-05890-8.

Yan, H.; Liu, Y.; Jin, L.; Bai, X. The Development，application，and Future of LLM Similar to ChatGPT. Journal of Image and Graphics 2023, 28, 2749–2762, doi:10.11834/jig.230536.

Yang, J.; Wang, Y. Toward Auto-Modeling of Formal Verification for NextG Protocols: A Multimodal Cross- and Self-Attention Large Language Model Approach. IEEE Access 2024, 12, 27858–27869, doi:10.1109/ACCESS.2024.3366803.

Yang, Z.; Zhang, Y.; Sui, D.; Ju, Y.; Zhao, J.; Liu, K. Explanation Guided Knowledge Distillation for Pre-Trained Language Model Compression. ACM Transactions on Asian and Low-Resource Language Information Processing 2024, 23, doi:10.1145/3639364.

Zahera, H.M.; El-Sisi, A.B. Accelerating Training Process in Logistic Regression Model Using OpenCL Framework. INTERNATIONAL JOURNAL OF GRID AND HIGH PERFORMANCE COMPUTING 2017, 9, 34–45, doi:10.4018/IJGHPC.2017070103.

Zhang, J.F.; Zhang, Z.Y. Machine Learning Hardware Design for Efficiency, Flexibility, and Scalability [Feature]. IEEE CIRCUITS AND SYSTEMS MAGAZINE 2023, 23, 35–53, doi:10.1109/MCAS.2023.3302390.

Zhu, J.; Wang, L.Y.; Han, X.; Liu, A.M.; Xie, T. Safety and Performance, Why Not Both? Bi-Objective Optimized Model Compression Against Heterogeneous Attacks Toward AI Software Deployment. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING 2024, 50, 376–390, doi:10.1109/TSE.2023.3348515.

Zhu, X.N.; Yang, M.Y.; Zhao, T.J.; Zhu, C.H. Minimum Bayes-Risk Phrase Table Pruning for Pivot-Based Machine Translation in Internet of Things. IEEE ACCESS 2018, 6, 55754–55764, doi:10.1109/ACCESS.2018.2872773.

RECONSIDERATION OF LONG-TERM INTERNSHIPS AND FUTURE PROPOSALS

Masataka SUKEGAWA[1]

Kazuma HATSUSHIO[2]

Nobutaka SUZUKI[3]

Abstract

The purpose of this paper is to examine the benefits and challenges of long-term internships through a comparative analysis of overseas internships, overseas online internships, and domestic internships. The results of the analysis suggest that overseas internships have a positive impact on English language learning and cross-cultural understanding but may pose risks in terms of daily life; while overseas online internships can develop English language and technical skills, they also pose challenges in terms of cross-cultural understanding, maintaining motivation, distance from others, and communication environment; and domestic internships can develop behavior as a member of society and confirm overseas orientation.

Keywords: long-term internships, overseas on-site internships, overseas online internships, domestic on-site internships

JEL Classification: I21

1. Introduction

As internationalization progresses further, the demand for global human resources is increasing. On the other hand, instability in world affairs, such as the novel coronavirus pandemic (hereafter referred to as "COVID-19") and the Ukraine crisis, has led to restrictions on travel to and from overseas. This situation has resulted in restrictions such as the inability to conduct internships at companies located overseas (hereinafter referred to as "overseas on-site internships"). In response to this situation, internships with companies located overseas conducted online (hereinafter referred to as "overseas online internships") have been implemented by using remote meeting tools such as Zoom and chat tools such as Slack to conduct internships in cooperation with the overseas companies while

[1] Nagaoka University of Technology, Information and Management Systems Engineering, Japan, s203361@stn.nagaokaut.ac.jp

[2] Nagaoka University of Technology, Information and Management Systems Engineering, Japan, s203369@stn.nagaokaut.ac.jp

[3] Nagaoka University of Technology, Information and Management Systems Engineering, Japan, nsuzuki39@kjs.nagaokaut.ac.jp

the intern remains in Japan. However, there are still communication and other issues to be solved in overseas online internship programs, and overseas on-site internship programs also have some risk-related issues. In addition, the future of internships in Japan (hereafter referred to as "domestic internships"), also affected by the outbreak of the COVID-19 pandemic, needs to be examined. In this paper, we discuss the effects and issues of long-term, overseas, online, and domestic internships as research questions. A long-term internship (practical training) program at Nagaoka University of Technology is used as a case study.

The structure of this paper is as follows. In Part 2, the definition of internship types and previous studies are reviewed. Part 3 describes the research methodology and case selection. Part 4 summarizes the research results. Part 5 presents the results and discussion of the comparative case study analysis. Part 6 serves as the conclusion of this paper.

2. Review of Previous Studies

First, we will review the internship programs discussed in this paper, followed by a review of previous studies of overseas on-site internships and overseas online internships.

2.1 Internships

The basic recognition and promotion measures of internships are presented in the "Basic Approach to the Promotion of Internships and Other Measures to Support Student Career Development" [1], dated June 13, 2022, which is a revision of the three-ministry agreement of September 1997 (between the Ministry of Education, Culture, Sports, Science and Technology, Ministry of International Trade and Industry, and Ministry of Labor). According to the basic concept, by linking academic training at universities and social experiences, internships are expected to deepen academic training, stimulate motivation for learning, and foster vocational awareness, and are important initiatives that can be fully expected to have supportive effects on educational and career development. In particular, "industry-academia cooperative efforts to support students' career development" are organized into four types. Of these, there is Type 3 (internship for general abilities and specialized use), in which work experience is mandatory and the purpose is to "identify one's own abilities" and "obtain evaluation materials", and Type 4 (advanced specialized internship (trial)).

2.2 Length of Internship Period

According to the results of a survey conducted by a study group of the Ministry of Economy, Trade and Industry [2], a questionnaire survey of students indicates that graduate students in the sciences are the most likely to participate in long-term internships, with many of them

focusing on understanding the atmosphere of the workplace and the business of the company as their objectives. Furthermore, the longer the duration of internship, the greater the effect on academic behavior and the greater the interest in society, as interns were exposed to actual work and were influenced by their own aptitudes and strengths. Kofu et al. [3] argue that internships that allow participating students to gain practical work experience are highly effective as career education, not only in providing them with an understanding of their current abilities and the structure of society, but also in enhancing their desire to find employment and improve their own abilities. It is argued that internships are highly effective as career education.

Thus, from the viewpoint of the definitions of "internship" and "career education", we believe that long-term internships are desirable, and the subject of this study is long-term internships.

2.3 Overseas On-site Internships

Yamashita et al. [4] show the educational effects of overseas fieldwork in terms of cross-cultural understanding, cross-cultural communication skills, a sense of daily life, adaptability and experience, and language ability. Noguchi et al. [5] showed improvements not only in cross-cultural understanding, international awareness, and English proficiency, but also in research and learning-related items such as a spirit of challenge, an attitude toward learning what one does not know, teamwork, problem-solving skills, and knowledge application. Nagata et al. [6] noted that not only did they promote understanding of diversity and independent behavior, but also promoted the image of global engineers and overseas orientation, and that after an overseas internship, there was an impact on career paths, job hunting, and English language learning.

Tanaka et al. [7] conducted a questionnaire survey of students who participated in a debriefing session presented by participants in an overseas field study program. Of the 589 responses, 47% indicated that they "want to participate" and 51% indicated that they "do not want to participate". Reasons for wanting to participate included "I want to experience foreign culture and interact with people from other countries" and "I want to improve my English skills", while reasons for not wanting to participate included "I am concerned about living abroad (food, safety, etc.)", "The cost of traveling abroad is high", and "I am worried about my language skills". In the questionnaire regarding the need for English, approximately 400 respondents (about 70% of the total number of respondents) felt that English was necessary. Although they understood the necessity of English, more than half of the students did not want to participate in an overseas field study program.

The effect of overseas study programs is very pronounced in terms of future careers and personal futures, as well as English proficiency and cross-cultural understanding. However, previous studies have not dealt much with the risks associated with overseas study programs. In addition to the recent outbreak of COVID-19, the possibility of being caught

in the Ukraine crisis and other international crises, exposure to danger in developing countries, and the possibility of contracting diseases cannot be denied.

2.4 Overseas Online Internships

Tsujii et al. [8] conceived and implemented an online internship program to solve architectural design problems presented by social entrepreneurs in emerging countries. As a result, Tsujii et al. pointed out that the online internship had several advantages, such as the sequential participation of third-party experts and advisors online, in addition to students and local personnel, and the online internship seemed to have a flat relationship with diverse participants, including host entrepreneurs and professional intermediaries, without any differences in their positions. The online version of the program has advantages, such as the ability to communicate with participants in a face-to-face setting. However, it was also pointed out that in order to communicate online what is difficult to communicate only face-to-face, it is necessary to enhance the means of expression (e.g., video, VR, etc.).

Research Methods and Case Selection

3.1 Research Methods and Nagaoka University of Technology's Long-term Internships (Practical Training)

Case studies will be conducted to examine future long-term internships. Case studies [9] are suitable for dealing with forms of research questions such as "how", and are also suitable for exploratory studies such as this one.

The subject of this paper is students who participated in "practical training", a long-term internship program at Nagaoka University of Technology. The details of the on-the-job training and its effects have already been discussed by Iida [10], Maruyama [11], and Sugimoto [12] in "Engineering Education".

Practical training is a long-term internship program in which fourth-year undergraduate students who plan to continue on to graduate school at Nagaoka University of Technology are dispatched to companies and other institutions for practical training during a five-month period from October to February. The purpose of this program is to cultivate practical and technical sensibility, and it has been conducted since 1979, when the university first opened. About 300 students are dispatched every year, and as of the 2022-23 academic year, a total of 13,937 students had been dispatched, 994 of whom were sent overseas [13].

Before the COVID-19 outbreak, practical training was conducted in person both in Japan and overseas. However, in the 2020-21 and 2021-22 academic years, when the COVID-19 pandemic struck, a change from the previous method was unavoidable. In principle, in-country practical training was conducted face-to-face from November to February (a little

over three months), with measures such as shortening the period by one month and providing vaccinations and health observation. In some cases, the training was conducted remotely from universities or homes. On the other hand, overseas practical training was cancelled in 2020-21 due to overseas travel restrictions, and in 2021-22, a new form of internship was conducted in a fully remote online format.

Thus, internships at Nagaoka University of Technology used to be conducted by actually traveling overseas before the pandemic, but after the pandemic, the internship format was changed to an online one.

The long-term internship program at Nagaoka University of Technology was selected as a case study because it has both overseas on-site and overseas online types, and it has a long history and a proven track record.

For the research method, a comparative case analysis was conducted on three cases, and a case report was prepared and discussed.

The first case is an overseas on-site internship, where the internship is conducted face-to-face; the second case is an overseas online internship, where the internship is conducted online from within Japan, without traveling overseas; and the third case is an on-site internship done in Japan (a domestic company with a high ratio of overseas sales). In this case, it is necessary to compare the overseas on-site type and the overseas online type, and we decided that an international Japanese company would be preferable in this case. The criterion for an international company was defined as "having an overseas sales ratio of 50% or more".

The survey was conducted on Nagaoka University of Technology students and graduates who had participated in internships corresponding to the above three cases. The number of subjects surveyed was seven, and the details of the subjects are shown in Table 1, including the destination country, year of internship, and industry.

Type of internship	**Year of internship**	**Destination (Country)**	**Industry**	**Surveyed person**
Overseas on-site internship	2019	I (Philippines)	IT education industry	A
		II (Indonesia)	Instrumentation manufacturer	B
Overseas online internship	2021	III (Vietnam)	IT service industry	C
				D

		IV (Romania)	IT service industry	E
				F
Domestic on-site internship	2021	V (Japan)	Industrial material manufacturer	G

Table 1. Details of Surveyed Persons

Semi-structured interviews were conducted online with the internship participants corresponding to the above three cases for one hour each from June to July 2022, and each case was analyzed.

3.2 Questions

The interview questions were developed based on the studies by Kofu et al. [3], Inoue et al. [14], Tan et al. [15], and others. Specifically, the questions were about the company, the content of the internship, communication methods, advantages of the program, problems and issues, and the degree to which they would recommend the company to their juniors. Table 2 shows the questions asked.

No.	Questions
1	About the company to which you were dispatched
2	Content of internship
2-1	Please tell us why you chose the company to which you were dispatched.
2-2	What were your main tasks and practical training during your internship?
2-3	How did you communicate with the company?
2-4	How did you access company resources, security, etc.?
3	Looking back on the internship
3-1	What did you gain from participating in this internship program?
3-2	What were the benefits of this internship program?

3-3	What were your expectations for this internship and how did they differ from your expectations?
3-4	Please tell us about any difficulties or problems you encountered during this internship.
3-5	Please tell us about any issues or challenges you faced during this internship.
3-6	If the same program were offered next year under similar circumstances, to what extent would you recommend it to younger students? Please tell us why you would recommend this program. (Answer in a 5-point scale: 1. Not at all, 2. Not very much, 3. Undecided, 4. Somewhat, 5. Definitely)

Table 2. Questions

4. Survey Results

4.1 Overseas On-site Internships

Semi-structured interviews were conducted with two graduates of Nagaoka University of Technology who participated in the study before the COVID-19 pandemic.

In 2019, Mr. A received on-the-job training at an IT education company in the Philippines, and Mr. B received on-the-job training at an Indonesian manufacturer of measuring instruments. In their opinion, in addition to the basic working experience, they were able to experience cultural differences (religion, climate, food, etc.) through direct life experience (shopping, bus transportation, social gatherings, etc.) and understand different cultures. They also said that it was good that they were able to gain experience abroad without having to take a leave of absence, that they were able to have an experience similar to studying abroad, and that the range of work they were able to engage in expanded in the latter half of the program period, allowing them to do practical training that they had not planned for. However, they also had feelings of confusion and uneasiness in an unfamiliar environment, and they both felt unwell due to the unfamiliarity with the place of residence, water supply, food, etc., and the influence of the social situation. The reasons given were that they thought they would learn more from overseas than in Japan in terms of cross-cultural exchange and English communication, that they felt stronger both mentally and physically as a result of the program, that it would be a good experience for their future, and that being in a tough environment would help them to grow.

4.2 Overseas Online Internships

Due to the impact of COVID-19, the overseas internship program in the 2021-22 academic year was switched to an online internship program conducted from within Japan. The survey was conducted on students who participated in the overseas online internship

program. Four surveys were conducted, including two at Company C (Mr. C and Mr. D) and two at Ding (Mr. E and Mr. F).

The first two interns (Mr. C and Mr. D) were from Company C. Company C is an IT service provider in Vietnam. They commented that the internship provided them with technical learning (system development and back-end fields), experience in team development, a good opportunity to learn business English communication (English chat, daily reports in English, speaking and listening in web meetings), and a chance to participate in an internship without having to change their living environment more than other placements. In addition, they said that it was difficult to maintain motivation because the environment did not change as participation was always done from their laboratory or room (Mr. C), communication was difficult due to connectivity issues, there were no social gatherings or other events, and it was difficult to get a sense of distance from the people in charge at the company. As to whether they would recommend the program to their juniors, Mr. C answered "4. Somewhat" and Mr. D answered "3. Undecided". Mr. C's reasons were that speaking in person is the best way to improve English, that it is a good option for those who do not want to go there, and that overseas development experience is not only for finding a job but also for learning various skills as an engineer (application development, system overview, team development, etc.). Mr. D's reasons were that using English online and being able to work remotely were good experiences, but there were a few times when he found it difficult in terms of communication. He also said that because of the need for autonomy, those who cannot work independently should work on-site.

Next are two interns for Ding (Mr. E and Mr. F). Ding is an IT service provider in Romania. They mentioned that they gained technical knowledge (RPA and software development), confidence from being able to communicate with people from other countries, and understanding of cultural differences (Mr. E). Other opinions included the opportunity to communicate in English with people outside the company (lunch meetings with Romanian people, lectures by Romanian university professors, etc.), and the fact that they enjoyed development and were able to create their own rhythm, which kept their motivation up (Mr. E). However, they also pointed out that they felt uneasy because they had lost contact with the company at times, and that they had few contacts at the internship site, which made them anxious when they were working during the internship. As for Mr. F's reasons, although he experienced a variety of work, he did not feel a sense of fulfillment or accomplishment after the internship, and although he was attracted by the overseas aspect, he was more likely to recommend the program to his juniors if he could get more face-to-face experience in Japan.

In contrast with the other four interns, they were not assigned any work beyond what they had originally planned.

4.3 Domestic Internships

Lastly, we examined internships at domestic companies that have a high ratio of overseas sales. The number of cases studied is one, that is, Mr. G of Bosha. Bosha is a manufacturer of industrial goods, and the ratio of overseas sales exceeds 70%. The opinions obtained were as follows: "I could understand what the company is looking for as a member of society, I deepened my understanding of English (I can listen to English but cannot speak it, and I changed my awareness that it is important to communicate with others even if my pronunciation is not great)", "I was not the type to show interest in foreign countries, but I felt that foreign countries were closer to me (the hurdle was lowered)", and "I learned that I could work in a foreign country". He also had opportunities to communicate with people in overseas offices in English (he was allowed to participate in remote meetings with overseas business partners), overseas elements were incorporated in many places (morning assemblies were held half in English, meetings were held in English, and in some departments, employees studied English in the morning before starting work), and he felt closer to overseas countries (the hurdle was lowered), and the opportunity to interact with people from other departments, overseas branches, and other overseas locations while being involved in the management of events that were not originally planned to connect the various locations. However, he said that he did not have many opportunities to talk with people from overseas. He had a total of three opportunities to speak with people at remote meetings and overseas offices, and there were no English speakers in the department during his internship period (although English speakers were assigned to the department before and after his internship period).

The first question was whether he would recommend the program to his juniors "as a regular internship", and the second is "as an overseas internship program". The reason is that it is important to know the gap between one's position in the university and one's position in the company, because it is not always the same. Next, for "as an overseas internship program", the answer he was selected as "3. Undecided", the reason being that that if someone is interested in overseas, they should go to an overseas internship program, and that it is good as an opportunity to create a global viewpoint because it is difficult to develop a yearning for and interest in overseas countries in ordinary life.

5. Discussion

We have examined three cases so far. Based on these surveys, the items were derived and summarized in Table 3.

Items	Overseas on-site internships	Overseas online internships	Domestic internships
Understanding of English	Highly effective	Medium effect	Small effect
Cross-cultural understanding	Highly effective	Medium effect	Small effect

Feelings of anxiety	Yes	None	None
Maintaining motivation	Easy	Difficult	Easy
Distance from others	Easy	Difficult	Easy

Table 3. Survey Results

The following items were obtained: "understanding of English", "understanding of different cultures", "feelings of anxiety", "maintaining motivation", and "feelings of distance from others". The items are explained one by one.

"Understanding of English" was found to be effective to a certain extent, as in the previous studies [4] and [5]. Mr. G, who was dispatched to a domestic site, did not mention whether he was able to master English conversation, although the hurdle for doing so was lowered. The overseas online program is based on chat tools, while the on-site program is based on face-to-face conversation, so the improvement in speaking, listening, and writing skills may be skewed.

The results for "cross-cultural understanding" were similar to the previous studies [4], [5], and [6], and a certain effect was obtained. However, the effect was not observed in the domestic on-site program, and only one out of four participants in overseas online programs achieved intercultural understanding, indicating that additional verification of the effect is needed.

"Feelings of anxiety" varied from case to case. In the case of overseas on-site programs, anxiety about living conditions was a constant concern; participants had to live in a different culture from that of Japan, and had a sense of anxiety about climate, water, food, safety, and so on. On the other hand, anxiety in overseas online programs came from being unable to grasp the "sense of distance" from their counterparts and see what is going to happen next. Anxiety always follows because they do not know what is happening at the worksite, such as a lack of specific instructions, few actions on daily reports, and few meetings.

"Maintaining motivation" is largely due to the fact that the environment around them does not change. This is why it is not seen in overseas on-site and domestic on-site programs. Many opinions were raised regarding overseas online programs. However, it was not seen in some (Mr. E).

"Distance from others" is not a problem for overseas and domestic on-site programs, but is an issue for overseas online programs. If this issue is not addressed, it will lead directly to "feelings of anxiety".

The first difference is the scope of work assigned to the interns. Both overseas and domestic on-site internships gave the interns a variety of experiences beyond the scope of the internship program they had originally planned. The second difference is the tendency of

the answers to the questions in Table 2. On the other hand, many participants of overseas online programs tended to cite their work experience and knowledge in terms of skills.

The results of these analyses can be summarized in three cases.

The overseas on-site internships were highly effective in terms of English learning and cross-cultural understanding, and there were no problems in maintaining motivation and an understanding one's distance from others. On the other hand, this type of internship program was not suitable for the students who felt uneasy about their daily life.

Overseas online internships enable growth in English language skills and technical aspects. The ability to improve one's English ability and skills in their field of specialty without significantly changing their living environment is a highly attractive feature for students interested in going abroad. On the other hand, it is suggested that the interns have problems in "understanding different cultures", "maintaining motivation", and "feelings of distance from others". It is desirable for participants to actively engage in the internship in terms of "understanding different cultures" and "maintaining motivation". With regard to "distance from others", it is desirable for both the company and the participant to make active efforts. Therefore, it is important to confirm the acceptance system of the company. Examples include whether or not there are efforts to eliminate the sense of distance (e.g., social gatherings) and its track record of accepting internship participants. The communication environment should also be considered. In one case (Mr. C), when participating in monthly meetings, communication was not possible due to the sound quality of the other party, and no improvement was made until the end. Because this depends on the company participating in the internship, it is necessary to check the communication environment (remote meeting tools, chat tools, etc.) when selecting a company. When participating in an overseas online internship, it is advisable to get advice from previous participants or to participate in an internship with prior information from your laboratory if possible.

Domestic internships were suggested to deepen understanding of one's position as a member of society and what is expected of oneself. In addition, through the interview with Mr. G, there were scattered indications of an overseas orientation. Nagata et al. [6] confirmed this in overseas on-site programs, but the results of this survey indicate that it was also confirmed in domestic on-site programs.

6. Conclusion

In this paper, we examined the effects and issues of long-term internships as a research question. The number of online internships increased after COVID-19. The results of the analysis suggest that overseas on-site internships have a positive impact on English language learning and cross-cultural understanding, but pose risks in terms of daily life. While overseas online internship programs can develop English and technical skills, they also presents challenges in terms of cross-cultural understanding, maintaining motivation,

distance from others, and the communication environment. Domestic on-site internships have the potential to cultivate the behavior of a member of society and to confirm overseas orientation.

Acknowledgments

We would like to express our sincere gratitude to the company officials in the case studies for accepting our students as interns. We also thank the interns who participated in the interview surveys. This research was supported by the JSPS Grant-in-Aid for Scientific Research 20K05055.

References

[1] https://www.mext.go.jp/a_menu/koutou/sangaku2/20220610-mxt_ope01_01.pdf - Basic concept in promoting initiatives related to student career development support including internships. 05-03-2024

[2] https://www.mext.go.jp/content/20201209-mxt_gakushi01-000011520_01.pdf - Study group on the effect of long-term internships on connections between students and companies: Results of a questionnaire survey of students. 05-03-2024

[3] Kenji Kofu, Akifumi Okabe. *Internship ni yoru kougakukyouiku no kouka* (*Effect of Engineering Education by Internship)* - Journal of JSEE Vol. 55 No. 5, ISSN-L 1341-2167. 2011(in Japanese)

[4] Toru Yamashita, Somei Ohnuki , Naoyuki Funamizu. *Geneki gakusei oyobi syuusyokugo no syakaijin heno tuisekityousa karamiru kougakukeikaigaiinternship no kouka (Effect of Overseas Internship -Research on Both Students and Graduates)* - Journal of JSEE Vol. 64 No. 5. ISSN-L 1341-2167. 2016 (in Japanese)

[5] Toru Noguchi, Kozo Yoshikawa, Masato Nakamura. *Kougakukeidaigakuin ni okeru kaigaiinternshipkyouiku to sono kouka* (*Overseas Internship Education in Engineering Graduate Courses and Evaluation of Educational Effect)* - Journal of JSEE Vol. 56 No. 3, ISSN-L 1341-2167. 2008 (in Japanese)

[6] Hirokazu Nagata, Mayu Kashimura, Yoshihito Shirai.*Rikoukeigakusei wo taisyoutosita kaigaiinternship no kyouikukouka to sinnrosenntakuheno eikyo (Educational Effects of Overseas Internships for Science and Technology Majors)* - Journal of JAGCE Vol. 6 No. 1, ISSN-L 2188-3505. 2008 (in Japanese)

[7] Takanori Tanaka, Tatsumi Kameno. *Kaigaiinternship wo tujita kokusaigijutusyaikuseino torikumi* (*Educational Program for Training International Engineers through International Internships)* - Journal of JSCE Vol. 2 (p.91-96), ISSN-L1884-7781. 2010 (in Japanese)

[8] Hiroyuki Tsujii , Harumi Kagawa, Aewol Kim.*Shinkoukoku no syakaikadaikaiketu ni torikumu online・internship no kaihatu* (*Development of Online Internship to Solve Social Issues in Emerging Countries)* - Journal of JSEE Vol. 70 No. 3 (p. 48-53), ISSN-L 1341-2167. 2022 (in Japanese)

[9] Robert K. Yin. *Case Study Research: Design and Methods (2nd edition)* , Sage Publications, Inc. 1994, ISBN-13 978-0803956636.

[10] Seishi Iida. *Nagaokagijutukagakudaigaku no kyoikuseido ni okeru internship no yakuwari to kinou -genjo to tenbo-* (*Role and Function of Internship in the Education System of Nagaoka University of Technology)* - Journal of JSEE Vol. 51 No. 3 (p.18-21), ISSN-L 1341-2167. 2022 (in Japanese)

[11] Kyuichi Maruyama. *Tyokiinternship(jitumukunren)niokeru sanngakurennkeikyouiku no sinnten* (*Co-operated Education at Nagaoka University of Technology through Long Term Internship)* - Journal of JSEE Vol. 57 No. 4 (p.37-41), ISSN-L 1341-2167. 2009 (in Japanese)

[12] Mitsutaka Sugimoto.*Nagaokagijutukagakudaigaku no sotugyoukennkyuu tosite no jitumukunnrenn(tyouki internship)* (*Jitsumu-kunren (Long term internship) at Nagaoka University of Technology as Bachelor's Thesis)* - Journal of JSEE Vol. 60 No. 5 (p.75-789, ISSN-L 1341-2167. 2012 (in Japanese)

[13] https://www.nagaokaut.ac.jp/annai/koho/kankobutsu/integrated_report/index.files/202301.pdf - Nagaoka University of Technology: Integrated Report 2023, Nagaoka University of Technology, 01-12-2023 (in Japanese)

[14] Masahiro Inoue, Nobuhiro Ohe, Kazunori Mano.*Global Project Based Learning to Cyber-Physical kuukan heno kakudai -rikoukeideno onlinekokusaikyoudougakusyuu (COIL)no genjou to kongo* (*Global Project Based Learning and its Extension to Cyber-Physical Space: Current and Future of Cooperative Online International Learning (COIL) in Science and Engineering)* - Journal of JSEE Vol. 70 No. 3 (p.9-14), ISSN-L 1341-2167. 2022 (in Japanese)

[15] Wai Kian Tan, Minoru Umemoto. *International Industrial Internship: A Case Study from a Japanese Engineering University Perspective* - Education Sciences Vol. 11 No. 4, EISSN-L 2227-7102. 2021

Bibliography

Hirokazu Nagata, Mayu Kashimura, Yoshihito Shirai. *Rikoukeigakusei wo taisyoutosita kaigaiinternship no kyouikukouka to sinnrosenntakuheno eikyo (Educational Effects of Overseas Internships for Science and Technology Majors)* - Journal of JAGCE Vol. 6 No. 1, ISSN-L 2188-3505. 2008 (in Japanese)

Hiroyuki Itami. *How to Write a Creative Paper*, ISBN 4-641-07649-9. Yuhikaku Publishing Co., Ltd. [12-01] 2001

Hiroyuki Tsujii , Harumi Kagawa, Aewol Kim.*Shinkoukoku no syakaikadaikaiketu ni torikumu online・internship no kaihatu* (*Development of Online Internship to Solve Social Issues in Emerging Countries)* - Journal of JSEE Vol. 70 No. 3 (p. 48-53), ISSN-L 1341-2167. 2022 (in Japanese)

Kazuhisa Todayama. *Thesis Classroom*, ISBN 978-4-14-091194-5. NHK Publishing, Inc. [08-30] 2012

Kenji Kofu, Akifumi Okabe. *Internship ni yoru kougakukyouiku no kouka* (*Effect of Engineering Education by Internship)* - Journal of JSEE Vol. 55 No. 5, ISSN-L 1341-2167. 2011(in Japanese)

Kyuichi Maruyama. *Tyokiinternship(jitumukunren)niokeru sanngakurennkeikyouiku no sinnten* (*Co-operated Education at Nagaoka University of Technology through Long Term Internship)* - Journal of JSEE Vol. 57 No. 4 (p.37-41), ISSN-L 1341-2167. 2009 (in Japanese)

Masaaki Takane, *Methodology of Creation*, ISBN 978-4-06-145553-5. Kodansha Ltd. [09-18] 1979.

Masahiro Inoue, Nobuhiro Ohe, Kazunori Mano.*Global Project Based Learning to Cyber-Physical kuukan heno kakudai -rikoukeideno onlinekokusaikyoudougakusyuu (COIL)no genjou to kongo* (*Global Project Based Learning and its Extension to Cyber-Physical Space: Current and Future of Cooperative Online International Learning (COIL) in Science and Engineering)* - Journal of JSEE Vol. 70 No. 3 (p.9-14), ISSN-L 1341-2167. 2022 (in Japanese)

Mitsutaka Sugimoto.*Nagaokagijutukagakudaigaku no sotugyoukennkyuu tosite no jitumukunnrenn(tyouki internship)* (*Jitsumu-kunren (Long term internship) at Nagaoka University of Technology as Bachelor's Thesis)* - Journal of JSEE Vol. 60 No. 5 (p.75-789, ISSN-L 1341-2167. 2012 (in Japanese)

Nobuo Takahashi, *Understanding Business Management*, ISBN 978-4-62306-104-4. Minerva Shobo [10-30] 2011

Kiyohiro Oki. *Core Text: International Management*, ISBN 978-4-88384-266-7. Shinseisha [12-25] 2017

Robert K. Yin. *Case Study Research: Design and Methods (2nd edition)* , Sage Publications, Inc. 1994, ISBN-13 978-0803956636.

Takanori Tanaka, Tatsumi Kameno. *Educational Program for Training International Engineers through International Internships* - Journal of JSCE Vol. 2 (p.91-96), ISSN-L1884-7781. 2010

Toru Noguchi, Kozo Yoshikawa, Masato Nakamura. *Overseas Internship Education in Engineering Graduate Courses and Evaluation of Educational Effect* - Journal of JSEE Vol. 56 No. 3, ISSN-L 1341-2167. 2008

Toru Yamashita, Somei Ohnuki , Naoyuki Funamizu. *Effect of Overseas Internship - Research on Both Students and Graduates* - Journal of JSEE Vol. 64 No. 5. ISSN-L 1341-2167. 2016

Seishi Iida. *Role and Function of Internship in the Education System of Nagaoka University of Technology* - Journal of JSEE Vol. 51 No. 3 (p.18-21), ISSN-L 1341-2167. 2022

Wai Kian Tan, Minoru Umemoto. *International Industrial Internship: A Case Study from a Japanese Engineering University Perspective* - Education Sciences Vol. 11 No. 4, EISSN-L 2227-7102. 2021

https://www.mext.go.jp/a_menu/koutou/sangaku2/20220610-mxt_ope01_01.pdf - Basic concept in promoting initiatives related to student career development support including internships. 05-03-2024

https://www.mext.go.jp/content/20201209-mxt_gakushi01-000011520_01.pdf - Study group on the effect of long-term internships on connections between students and companies: Results of a questionnaire survey of students. 05-03-2024

https://www.nagaokaut.ac.jp/annai/koho/kankobutsu/integrated_report/index.files/202301.pdf - Nagaoka University of Technology: Integrated Report 2023, Nagaoka University of Technology. 01-12-2023

THE INTEGRATION OF THE LATEST TECHNOLOGICAL ADVANCEMENTS IN AGRICULTURE. WHAT ARE THEIR EXACT APPLICATIONS AND HOW DO THEY WORK?

Alexandru TĂBUȘCĂ[1]

Ștefan BĂICOIANU[2]

Abstract

Agriculture is not only about planting crops and raising animals, it is way more than that. Agriculture is a business which supports human life, provides raw materials and builds strong economies, playing an extremely important role in our livelihoods, even if we may not always notice that. Experts from this sector, however, say that the new global situation asks for a so called "revolution" in agriculture. Only in the last two decades, climate changes have been responsible for a productivity decrease of 21% in this business. Moreover, taking into consideration the pandemic and the political situation from the recent years, which led to a raise of the production costs, farmers and entrepreneurs are more and more worried about the way they are going to make their businesses profitable.

We should also mention that the worldwide food need is going to be 50% higher in 2050 than in the present. Fortunately, technological advancements come as a solution to these problems. Not only do they have an impact on the production costs, but they also help in reducing pollution around the world. A 2022 Deloitte study in collaboration with the Environmental Defense Fund revealed that the use of technology in agriculture can decrease with 9.8 gigatons the production of carbon dioxide-equivalent emissions (CO2e) between 2020 and 2050, as well as save up to 100 billion US dollars in costs to farmers by 2030. The agricultural technology, also known as "AgriTech", promises a more efficient use of equipment and an increase in crop yields, all of these while following a sustainable production plan, generally referred to as "precision agriculture".

Keywords: precision agriculture, artificial intelligence, internet of things, drones, robots, sustainability, optimization

JEL Classification: O13

1. Introduction

[1] PhD, Associate Professor, Romanian-American University, Romania, alex.tabusca@rau.ro

[2] Student, "Mihai Viteazul" National College, Ploiești, Romania, stefanbaicoianu24@gmail.com

Technological advances have always played an important role in the history of agriculture. In fact, we are going now through a period in which the future of our planet is shaped by these modern tools. This article explains in detail the application of the latest technologies in agriculture, describing in parallel how they work. Artificial Intelligence (AI) and the Internet of Things (IoT) are the two main components of AgriTech.

The Artificial Intelligence in Agriculture Market is "projected to grow from USD 1.7 billion in 2023 to USD 4.7 billion by 2028", reveals a report of MarketsandMarkets [1]. The integration of AI and IoT technologies transforms traditional farming practices into efficient and sustainable operations. From autonomous tractors to robots and drones, Artificial Intelligence and Internet of Things underlie all new machinery and systems used in agriculture.

2. Key applications

The Internet of Things connects devices and systems with processing ability with other computers, ensuring at the same time continuous data transfer. The information received over the Internet (or other communication networks) can be farther incorporated in a software programme. This is how AI and IoT add functionality to the modern tools used in agriculture (such as drones and robots). In short, a dataset, which for example can be formed of images or temperature readings over a week, is collected by sensors and cameras. Then, this information is analyzed by Artificial Intelligence algorithms, allowing farmers to make data-driven decisions regarding their crops. Next, we will delve deeper into how these modern agricultural machineries work and what is their exact application.

2.1. Drones

Drone technology plays a significant role in the agricultural sector, providing a unique view on sustainable farming. Drones are equipped with advanced cameras, furnishing high-resolution aerial images in order to monitor crop health. In this regard, some farmers already use satellite images, but their accuracy and precision are less exact. Modern drones are so precise that based on camera input, they are able to indicate specific location to the millimeter. It can be about a place where a plant disease was spotted, so that farmers can treat it in a very rapid and efficient way, before being too late and losing from the harvest. Or it can be about the security and management of a farm. For example, in this way, the use of machinery and equipment can be remotely monitored, saving both time and money, because the demand of security people is less.

Figure 1. An example of a camera-equipped agricultural drone[3]

Moreover, drones are used to spray pesticides and fertilizers. According to Tom Wolf's article on "Sprayers 101" website, this practice is widespread in south-east Asia, a great percentage of the agricultural areas being treated in such manner [2]. Using this technique, farmers improve crop yield by targeting specific areas and save chemical costs by applying no more solution than needed (which is almost impossible when referring to traditional spraying methods). Besides, by using a smaller quantity of pesticides, soil pollution is reduced and the nutritional quality of the crop is improved. This measure has a positive effect on human health, too. A study of the "World Health Organization" (WHO) from 2018 reveals that "about one-tenth of the world's population becomes ill every year from eating contaminated food" [3]. "Contaminated food intake is the main pathway for soil contaminants to the human body" informs research of "Science Communication Unit" in collaboration with "University of the West of England" [4].

Figure 2. Drone applying spray treatment[4]

[3] Source: https://www.croptracker.com/blog/drone-technology-in-agriculture.html
[4] Source: https://www.croptracker.com/blog/drone-technology-in-agriculture.html

In addition to pest detection, drones' camera vision technology provides information about plant health according to leaf color. Normalized Difference Vegetation Index (NDVI) is the metric used in this process. The principle behind this indicator is based on the spongy layers found on leaf backside. When healthy, leaves reflect a great amount of Near Infrared (NIR) light. At the opposite side, stressed and dead leaves reflect less NIR light. This is how NDVI imagery distinguishes healthy plants from sick ones.

Such information is essential for crop management because in this way farmers are able to identify irregularities and combat them. One of the reasons for plant stress is the lack of water. By getting NDVI feedback from drones, farmers are able to optimize water usage by irrigating only the problematic areas of the crop and not the entire land. We all know that these days our planet confronts water scarcity and it is crucial to use this resource carefully.

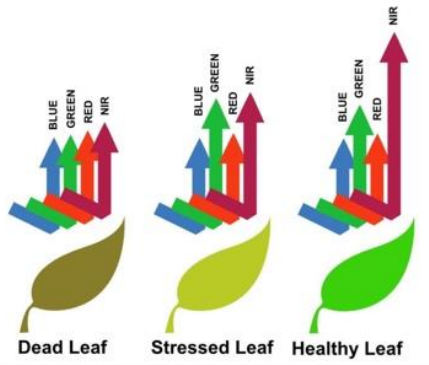


Figure 3. A representation of the light reflected by three types of leaves[5]

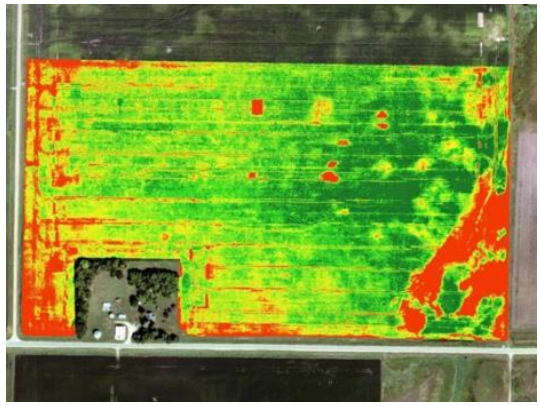


Figure 4. An example of NDVI imagery[6]

[5] Source: http://www.agasyst.com/portals/NDVI.html
[6] Source: http://www.agasyst.com/portals/NDVI.html

2.2. Robots

Robots represent a valuable tool within the agricultural industry, especially when talking about large farms, but not only. These robots are mainly designed to be able to harvest fruits and vegetables without damaging them. Being equipped with sensors and cameras (as in the case of drones), agricultural robots detect when crops are ready to be picked and with the help of an articulated arm and a transport unit for mobility, they successfully achieve their tasks. In fact, harvesting robots integrate in their design parts from both traditional six-axis robots and mobile robots.

Figure 5. An example of a harvesting robot[7]

Apart from harvesting, agricultural robots are used for planting and seeding tasks. In order to drive autonomously, robots use both computer vision technology and predefined GPS coordinates, so that they require no human intervention. Some of the widespread types of robots from this category include autonomous tractors and robotic arm planters. From drilling holes and depositing seeds at precise depths, to covering them with soil, planting robots are capable of achieving repetitive human tasks. Actually, when talking about this kind of actions, robots have higher accuracy and precision than people do, their operations leading to better crop yields. According to Markets and Markets, agricultural robots' market in the entire world was estimated at USD 13.5 billion in 2023 and is expected to reach USD 40.1 billion by 2028 [5].

3. Machine Learning in AgriTech

Machine Learning (ML) is a branch of Artificial Intelligence which enables computers to learn from a dataset and make predictions. ML models imitate the way humans learn, improving their accuracy once they have gathered a greater amount of data. The computer vision software used to develop drones and robots described in section 2 relies on Machine Learning algorithms, which are mostly created using *Python* programming language.

[7] Source: https://howtorobot.com/expert-insight/agricultural-robots

Along plant disease detection, ML is used in the agricultural sector to predict crop yield, which is an extremely important indicator. Depending on it, farmers adopt different strategies regarding land management. In this section we will focus on the ML algorithms behind crop yield prediction models.

First of all, according to the way that models are trained, Machine Learning is divided into two subcategories: supervised and unsupervised learning. Supervised learning is when machines receive labeled datasets, so that they do not have to independently identify any patterns within the information. Moreover, these algorithms analyze the relationship between input and output, so that they can make new predictions. For example, a supervised model might be used to predict crop yield based on specific parameters, such as soil type, temperature, precipitation or humidity. This category includes regression and decision tree algorithms. In contrast, unsupervised learning models are trained without any labels. Instead, machines group the information based on rules they create. Unsupervised learning algorithms are particularly suited for handling large amounts of data and establishing relationships between them, which is also the case of crop yield prediction. Artificial neural networks are used in this scope.

3.1. Artificial Neural Networks (ANNs)

Artificial Neural Networks are a more complex type of unsupervised learning algorithms, designed after the structure of a biological brain. ANNs are usually presented as systems of interconnected “neurons” (also known as “units”) divided into three types of layers. Data is received through input neurons, which is farther sent to the hidden layer, where patterns within the information are recognized. More complex networks may have more hidden layers. Finally, the “impulse” is transferred to the output neurons which return predictions as a response to the ANN input.

Links between units from different layers are called “synapses”. Each of them stores a “weight” attribute that determines the influence of the connection on the learning process (more specifically on the other artificial neurons). As in the case of the human brain, not all synapses have equal efficiencies (strengths).

Moreover, while biological neurons are stimulated by strong electrical impulses, ANNs are activated using mathematical functions which map the input to the output, as illustrated in figure 7. The learning technique behind ANNs is called “backpropagation” and consists of the modification of the “weight” feature according to the difference between predicted and actual outputs. It changes until the neural network reaches the minimal possible error rate.

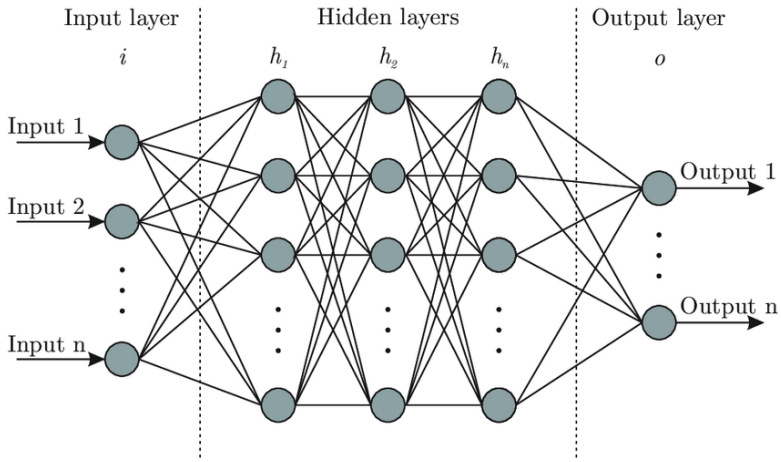


Figure 6. A representation of an Artificial Neural Network [8]

Artificial Network	Biological Network
Input layer	Dendrite
Neuron / Unit	Cell body
Weights	Synapses
Output layer	Axon

Table 1. An analogy between Artificial Neural Networks (ANNs) and Biological Neural Networks (BNNs)

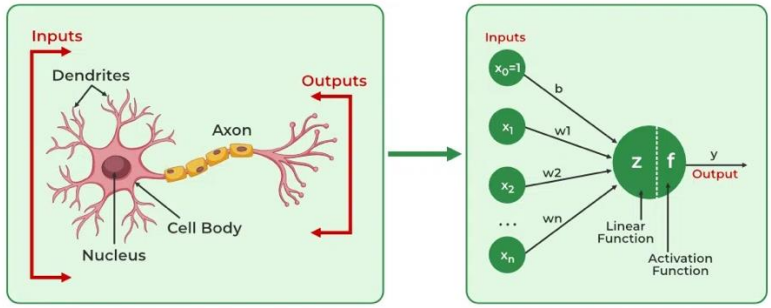


Figure 7. BNNs vs ANNs[9]

[8] Source: https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o_fig1_321259051

[9] Source: https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/

3.2. Regression models

Regression analysis relies on the relationship between input and output data by fitting a mathematical equation (which can be linear or polynomial). Regression can be represented in a cartesian plane, where the x-axis is the input feature, and the y-axis is the target value. For example, when referring to linear regression, the prediction line is fitted so that the sum of the squared vertical distances between line and data points is minimal (which is the definition of error). This technique is called "Least Squares Regression". Figure 9 shows an example of a crop yield prediction dataset. In reality it is made up of 28248 rows and 7 columns of information.

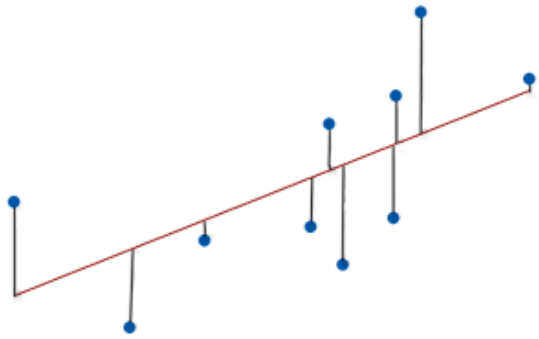


Figure 8. An illustration of a "Least Squares Regression" line[10]

	Area	Item	Year	hg/ha_yield	average_rain_fall_mm_per_year	pesticides_tonnes	avg_temp
1986	Bahamas	Cassava	1990	106667	NaN	484.59	25.74
1987	Bahamas	Maize	1990	15000	NaN	484.59	25.74
1988	Bahamas	Sweet potatoes	1990	38118	NaN	484.59	25.74
1989	Bahamas	Cassava	1991	100000	NaN	484.59	25.66
1990	Bahamas	Maize	1991	16667	NaN	484.59	25.66
1991	Bahamas	Sweet potatoes	1991	31385	NaN	484.59	25.66

Figure 9. A part of a dataset used for crop yield prediction[11]

3.3. Decision tree models

Decision tree algorithms use tree-like structures to make decisions and predictions. One of their greatest advantages is that they can handle high-dimensional data with good accuracy, which can be crucial when taking into consideration the scope of these models: crop yield prediction. They are similar to recursion functions, dividing input features into smaller and smaller subsets (based on their significance) until a dead end is reached. Each split of the

[10] Source: https://statisticsbyjim.com/regression/least-squares-regression-line/
[11] Source: https://www.javatpoint.com/crop-yield-prediction-using-machine-learning

tree is done by passing through a decision point (node), resulting in a new one or in a leaf which represents the outcome variable (the prediction). Connections between nodes are called "branches" and they are represented with arrows as in figure 10. When visualizing decision trees, branches are illustrated with responses to the decision rules such as *yes* and *no*. Final predictions are made by going through the entire tree, from the root node to a leaf.

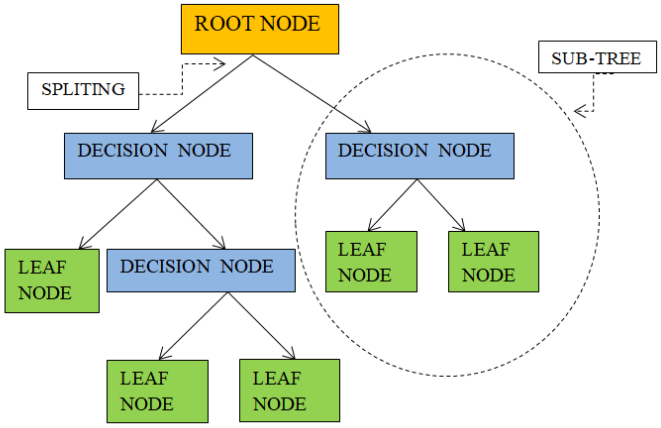


Figure 10. An example of a decision tree structure[12]

4. Conclusion

The integration of the latest technological advancements in agriculture is a "must", not a "should". Generally speaking, if we do not adapt to the current global situation and refuse to use the modern tools that we have at our disposal, the future of our planet is going to be worse and worse. The more is the case of agriculture, which is the second most polluting industry in the world, being directly responsible for almost 8.5% of total greenhouse gas emissions. Apart from the economic impact that tech innovations have on this sector, farmers and entrepreneurs should primarily focus on the environmental aspect that is in their hands. Even if buying such equipment may sound a bit expensive at the beginning (especially for small and midsize producers), there is nothing more important than a friendly environment which supports human life. Regarding this problem, governments should financially help farmers or directly provide them access to the latest technology. It is an extremely valuable tool for humanity and it would be a sin not to use it, especially when in the game is the future of our home, the Earth.

[12] Source: https://www.researchgate.net/figure/Decision-tree-structure_fig1_360443851

References

[1] https://www.marketsandmarkets.com/Market-Reports/ai-in-agriculture-market-159957009.html (the introduction section) – Global forecast to 2028 of Artificial Intelligence in Agriculture Market; last access: 26.05.2024

[2] https://sprayers101.com/challenges-drone/ (the introduction section) - A review about drone spraying; last access: 26.05.2024

[3], [4] https://openknowledge.fao.org/server/api/core/bitstreams/fe5df8d6-6b19-4def-bdc6-62886d824574/content/src/html/chapter-04-3.html (4.3.1 "Pathways for exposure" section) – Impact of soil pollution on human health; last access: 26.05.2024

[5] https://www.globenewswire.com/en/news-release/2023/08/16/2726443/0/en/Agriculture-Robots-Market-Size-Industry-Share-Trends-2028-MarketsandMarkets.html#:~:text=Agriculture%20robots%20market%20in%20terms,report%20published%20by%20MarketsandMarkets%E2%84%A2 - Global agriculture robots market size and trends to 2028; last access: 26.05.2024

Bibliography:

HENKE, C. – *Cultivating Science, Harvesting Power: Science and Industrial Agriculture in California.* ISBN 978-0-262-08373-7. August 2008

MUKHOPADHYAY, S. – *Smart Sensing Technology for Agriculture and Environmental Monitoring.* ISSN 1876-1100. February 2012

LI, D., ZHAO, C. – *Computer and Computing Technologies in Agriculture II, Volume 2*. ISSN 1571-5736/1861-2288. June 2009

https://www.cropin.com/iot-in-agriculture - Integration of IoT technology in agriculture; last access: 29.05.2024

https://www.forbes.com/sites/ganeskesari/2024/03/31/the-future-of-farming-ai-innovations-that-are-transforming-agriculture/ - AI innovations in the agricultural sector; last access: 29.05.2024

https://www.linkedin.com/pulse/ai-iot-powerful-duo-smart-farming-agriculture-clearspot-ai/ - AI and IoT in agriculture; last access: 29.05.2024

https://tolluncrewedsystems.com/blog/how-drones-are-used-in-agriculture/ - Use of drones in agriculture; last access: 29.05.2024

https://www.idtechex.com/en/research-report/agricultural-robots-and-drones-2017-2027-technologies-markets-players/525 - Agricultural robots and drones: market review and technology assessment; last access: 29.05.2024

https://howtorobot.com/expert-insight/harvesting-robots - An article about types of agricultural robots; last access: 29.05.2024

https://www.javatpoint.com/crop-yield-prediction-using-machine-learning - Machine Learning models used in crop yield prediction; last access: 29.05.2024

https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7c68a32212c1f86f535f4c1658ff68399d0a9ddd – Artificial Neural Networks (ANNs) for crop yield prediction; last access: 29.05.2024

JOURNAL OF INFORMATION SYSTEMS & OPERATIONS MANAGEMENT

VOL. 18 NO. 1 / 2024

ISSN: 1843-4711