ICCSM 2025

BOOK OF PROCEEDINGS

1ST INTERNATIONAL CONFERENCE COMPUTER SCIENCES AND MANAGEMENT

WHERE DIGITAL & BUSINESS BECOME HUMAN

26-27 June 2025 | Tirana, Albania

ISBN 9789928347123

DOI 10.37199/c41000300

CONFERENCE CHAIR

Assoc. Prof. Merita Toska, POLIS University

PARTNER UNIVERSITIES

POLIS University, Albania
Université Lyon 2, France
Università Telematica San Raffaele, Italy
University of Western Macedonia, Greece
International University of Sarajevo, Bosnia & Herzegovina
Mother Teresa University, North Macedonia
Gebze Technical University, Turkey
Public International Business College, Kosovo
Rochester Institute of Technology – RIT Global Campus, Kosovo
Co-PLAN, Institute for Habitat Development, Albania
AI Hub Albania, Albania
Luralux, Albania

ORGANISING COMMITTEE

Dr. Blerta Mjeda
Dr. Emiliano Mankolli
Msc. Sonila Murataj
Msc. Andia Vllamasi
Msc. Klejda Hallaci
Msc. Erilda Muka
Msc. Armela Reka

SCIENTIFIC COMMITTEE

Prof. Dr. Jérôme Darmont, Université Lumière Lyon 2 (France)
Prof. Dr. Lydia Coudroy de Lille, Université Lumière Lyon 2 (France)
Prof. Dr. Jim Walker, Université Lumière Lyon 2 (France)
Prof. Dr. Besnik Aliaj, POLIS University, (Albania)
Prof. Dr. Daniela Sica, San Raffaele Roma University, (Italy)
Prof. Dr. Stefania Supino, San Raffaele Roma University, (Italy)
Prof. Dr. Arbana Kadriu, South East European University (North Macedonia)
Prof. Dr. Ing. Lejla Abazi-Bexheti, South East European University (North Macedonia)
Prof. Dr. Yusuf Sinan Akgül, Gebze Technical University (Turkey)
Assoc. Prof. Dr. Galia Marinova, Technical University of Sofia (Bulgaria)
Assoc. Prof. Dr. Vasil Guliashki, Technical University of Sofia (Bulgaria)
Assoc. Prof. Mehmet Göktürk, Gebze Technical University (Turkey)
Assoc. Prof. Yakup Genç, Gebze Technical University (Turkey)
Assoc. Prof. Habil Kalkan, Gebze Technical University (Turkey)
Assoc. Prof. Dr. Godiva Rëmbeci, POLIS University (Albania)
Assoc. Prof. Dr. Xhimi Hysa, POLIS University (Albania)
Assoc. Prof. Dr. Merita Toska, POLIS University (Albania)
Assoc. Prof. Dr. Sotir Dhamo, POLIS University (Albania)
Dr. Gennaro Maione, San Raffaele Roma University, (Italy)
Dr. Nicola Capolupo, San Raffaele Roma University, (Italy)
Dr. Benedetta Esposito, San Raffaele Roma University, (Italy)
Dr. Venera Demukaj, Rochester Institute of Technology (Kosovo)
Dr. Emil Knezović, International University of Sarajevo (BiH)
Dr. Šejma Aydin, International University of Sarajevo (BiH)
Dr. Azra Bičo, International University of Sarajevo (BiH)
Dr. Šejma Aydin, International University of Sarajevo (BiH)
Dr. Azra Bičo, International University of Sarajevo (BiH)
Dr. Hamza Smajić, International University of Sarajevo (BiH)
Dr. Panagiotis Kyratsis, University of Western Macedonia (Greece)
Dr. Delina Ibrahimaj, Minister of State for Entrepreneurship and Business Climate (Albania)
Dr. Elona Karafili, POLIS University (Albania)
Dr. Emi Hoxholli, POLIS University (Albania)
Dr. Shefqet Suparaku, POLIS University (Albania)
Dr. Manjola Hoxha, POLIS University (Albania)
Dr. Elsa Toska, POLIS University (Albania)
Dr. Emiliano Mankolli, POLIS University (Albania)

Dr. Albina Toçilla, POLIS University (Albania)
Dr. Sonia Jojic, POLIS University (Albania)
Dr. Ilda Rusi, POLIS University (Albania)
Dr. Ledian Bregasi, POLIS University (Albania)
Dr. Klodjan Xhexhi, POLIS University (Albania)
Dr. Endri Duro, POLIS University (Albania)
Dr. Remijon Pronja, POLIS University (Albania)
Dr. Vjosë Latifi, International Business Collage Mitrovica (Kosovo)
Dr. Agron Hajdari, International Business Collage Mitrovica (Kosovo)

Table of Contents

01

INFLUENCER MARKETING AND HUMAN CAPITAL:
THE STRATEGIC ROLE OF EMPLOYEES IN THE FOOD INDUSTRY

DOI: 10.37199/c41000301

Gennaro MAIONE

San Raffaele Roma University, Rome, Italy
gennaro.maione@uniroma5.it
ORCID 0000-0002-9167-6369

Mario D'ARCO

University of Sannio, Benevento, Italy
madarco@unisannio.it
ORCID 0000-0001-9564-6960

Orlando TROISI

University of Salerno, Fisciano, Italy
otroisi@unisa.it
ORCID 0000-0002-4420-0130

Corrado DE CRISTOFARO

University of Salerno, Fisciano, Italy
cdecristofa@gmail.com
ORCID 0009-0005-9930-8359

Abstract

Influencer marketing is a strategy aimed at driving consumers' brand awareness and/or their purchasing decisions by leveraging the influence of opinion leaders (Brown & Hayes, 2008; Scott, 2015). Influencer marketing scholars have so far paid attention to influencers outside the company (social media personalities not affiliated with the company) and their ability to stimulate trust in consumers and guide their purchase intentions (Lou & Yuan, 2019), while limited attention has been paid to the managerial implications deriving from the ability of internal subjects to influence (internal influencers). Awareness of the effectiveness of the influencer marketing strategy has sparked debate in the food industry about the extent to which internal and external influencers can influence consumers. Previous research suggests that authenticity and credibility are key drivers of influencer effectiveness (Audrezet et al., 2020), making internal influencers a potential strategic *asset for brands undergoing digital transformation. The aim of this study, in line with the theory of the human mark (Kowalczyk & Pounders, 2016), is to analyse how consumers perceive employees versus external influencers in terms of credibility, authenticity, and emotional connection to the brand. This study will help fill a gap in the literature on influencer marketing.*

The methodology adopted for this study is a 2x2 between-subjects experimental design, which allows us to analyse the joint effects of influencer type and the value alignment between the brand and the influencer on three dependent variables: consumer trust, consumer engagement, and purchase intention. The results of this work, in line with the literature on the topic under analysis (van Driel & Dumitrica, 2021), reveal that internal influencers elicit greater consumer trust and heighten perceptions of authenticity, whereas external influencers generate higher levels of consumer engagement.

Keywords: Internal influencer, External influencer, Behavioural intention, Food marketing

I. INTRODUCTION

The widespread adoption of social media has made influencer marketing one of the most effective communication tools in the contemporary landscape. Several scholars have focused on the analysis of the effectiveness of the use of external influencers (e.g. content creators with a significant following) in terms of brand perception and consumer behaviour (Hudders et al., 2021; Lou & Yuan, 2019), but little on the potential of internal influencers (e.g. employees or company owners) as a category of brand promoters. Drawing on source credibility theory (Hovland & Weiss, 1951; Ohanian, 1990), this study investigates how influencer type (internal vs. external) affects consumer trust, post-engagement, and behavioural intention.

Trust is recognised as a key element in effective persuasion and consumer loyalty, especially in digital contexts, where the authenticity of content is often questioned (Kim & Kim, 2021). For further theoretical support, we draw on engagement theory (Hollebeek et al., 2014). Furthermore, relationship marketing theory (Morgan & Hunt, 1994) and authenticity theory (Wood et al., 2008; Pöyry et al., 2019) provide a solid framework to understand how trust and authenticity mediate and moderate the effectiveness of influencer content.

This study aims to investigate whether internal influencers generate higher levels of trust than external influencers, and whether such trust, in turn, drives engagement and purchase intentions. We also investigate whether perceived authenticity moderates these relationships and, if so, whether such moderation enhances or attenuates the observed effects, depending on how authentic the influencer appears to the audience. At the heart of this investigation is a broader question, with significant implications for human resource management (HRM): can employees serve a dual role, both as productive team members and as effective brand ambassadors?

Employee-led influencer marketing represents a cost-effective alternative to traditional influencer campaigns while simultaneously boosting employee morale, brand pride and organisational cohesion.

II. THEORETICAL BACKGROUND AND HYPOTHESES DEVELOPMENT

II.1 Food Influencer Marketing

Influencer marketing attracts the attention of many scholars who, with their analyses, offer multifaceted perspectives on the impact of culinary marketing (Mainolfi et al., 2022; Lee et al., 2021) not only on consumer purchase decisions and loyalty, but also on the trust that is built over time (Lou & Yuan, 2019; Vrontis et al., 2021). Much remains to be explored regarding the criteria for selecting influencers and the alignment of brand messages with influencers' narratives.

Audrezet et al. (2020) highlighted the complexity of storytelling in marketing and identified strategies that merge personal narratives and promotional content. For example, food influencers often use inviting descriptions and share behind-the-scenes anecdotes about meals or personal experiences in preparing them. This approach may evoke emotions and encourage audiences to leave positive reviews or to spontaneously share the content, thereby increasing uptake and positive sentiment.

For Weng et al. (2022), changes in TikTok and Instagram algorithms are increasingly influencing the sustainability of users' attention span. For brands that need to maintain an image of authenticity and align with their audience's values, there is an urgent need to embrace new perspectives on culinary heritage and narratives about the origins of ingredients (Misra et al., 2024; Anjos et al., 2022).

These findings suggest an evolution in public preferences, increasingly oriented towards authentic and humanised relationships with brands rather than impersonal and sophisticated forms of advertising.

Teams frequently engage with multiple external influencers at the same time to increase brand awareness (Campbell & Farrel, 2020; Hudders et al., 2021). However, this practice can put the firm at risk of significant divergence in the messages that external influencers share (Childers et al., 2019), and this misalignment could erode trust among viewers who may suspect that monetisation strategies are overshadowing the culinary artist's passion. On the contrary, a stronger connection is more likely if marketing managers carefully select all collaborators (Janssen & Rudeloff, 2025).

It follows that strategic HRM policies could play a decisive role, as employees or business owners can be highly effective digital promoters, able to communicate the brand's ideals and enhance intangible assets that attract current and potential customers. Despite the lack of a cohesive core of studies on the subject under analysis, some recent scientific articles suggest that internal brand ambassadors can strongly influence employee morale and represent concrete examples of the so-called corporate pride, which leads to greater economic loyalty towards the company (Li et al., 2024).

The use of internal influencers, in addition to representing a low-cost form of marketing (Ivanov et al., 2024), has the advantage of making personal stories seem authentic, and the brand merges with their real presence on social media (Seçilmis et al., 2025)

II.2 Hypothesis development

Consistent with the Source Credibility Theory (Hovland et al., 1953; Hovland & Weiss, 1951), the reliability of the source of a message represents a crucial factor in communication (Ohanian, 1990). For this reason, we believe that reporting the use of internal or external influencers can influence the success of the influencer marketing campaign. Some studies have highlighted greater credibility for internal influencers compared to external ones (Seçilmis et al., 2025), while others, while recognising the greater popularity of external influencers, note that their sponsorships can generate greater scepticism among consumers (Audrezet et al., 2020). Based on these premises, we can advance the following research hypothesis:

H1: Posts shared by internal influencers will generate higher levels of consumer trust than those shared by external influencers.

According to the influencer marketing literature, purchase intentions of sponsored products or services also depend on the type of source from which the social media post originates (Kumar et al., 2023). For this reason, we hypothesise that:

H2: Source type (internal influencer vs. external influencer) has a positive effect on individuals' behavioural intention.

Several studies in the literature highlight that the type of source also influences consumer engagement with posts published by influencers (Giakoumaki & Kreppa, 2020). According to De Vries & Carlson (2014), external influencers can generate higher levels of engagement thanks to an entertainment-oriented communication style. For this reason, we hypothesise that:

H3: Posts shared by external influencers will generate higher levels of engagement than those shared by internal influencers.

Lou et al. (2019) and Sardar et al. (2024) suggest that engagement with influencer posts positively affects consumers' purchase intention for sponsored products or services. Therefore, let's assume that:

H4: Consumer engagement positively affects individuals' purchase intention.

Kim & Kim (2021) highlight that trust facilitates transactions and relationships between consumers in digital environments. In influencer marketing studies, trust refers to individuals' disposition to rely on content shared by influencers (Schultz, 2025) and is considered an antecedent of engagement (Connolly et al., 2023; Vohra & Bhardwaj, 2019). Therefore, let's assume that:

H5: Trust positively affects individuals' engagement with the influencer's post.

Trust can enhance the effectiveness of the influencer's message and foster the behavioural intention to acquire the recommended products or services (Kim & Kim, 2021). Therefore, we propose:

H6: Trust positively affects individuals' purchase intention.

Recently, perceived authenticity, which can be defined as "being true to oneself" (Wood et al., 2008), has emerged as a central element in influencer marketing (Mucundorfeanu et al., 2025; Audrezet et al., 2020). According to Pöyry et al. (2019), authenticity can strengthen the bond between the influencer and their followers.

An analysis of the current literature shows that Dondapati & Dehury (2024) examined the role of authenticity as a moderating factor in the relationship between influencer type and outcomes. Considering these studies, it is hypothesised that perceived authenticity may act as a moderating variable. Consequently, the following hypotheses were formulated:

H7a: Perceived authenticity positively moderates the relationship between source type (internal vs. external influencer) and trust.

H7b: Perceived authenticity positively moderates the relationship between source type (internal vs. external influencer) and engagement with the post.

H7c: Perceived authenticity positively moderates the relationship between source type (internal vs. external influencer) and behavioural intention (purchase from a pizza shop)

III. METHODOLOGY

To achieve the research objective, a 2x1 between-subjects experimental design was adopted; participants were randomly assigned to view a social post promoting a pizzeria, published by either an internal or an external influencer.

The aim was to measure behavioural purchase intention, considering consumer trust and engagement as mediating variables and perceived authenticity as a moderating variable. To avoid familiarity with the brand or the influencer, the two scenarios were created using a fictitious pizzeria and invented names for the influencers (Li et al., 2025; Mo and Wang, 2025).

The experiment was conducted in April 2024 via the Qualtrics online platform. The crowdsourcing platform Prolific (Palan & Schitter, 2018) was used to recruit participants, the United States was selected as the reference country, and the number of participants was set at 300.

To verify that participants correctly identified the post's source (internal vs. external influencer), they were asked to agree or disagree with the statement: "An external influencer created this post?"

Participants were then asked to fill out a questionnaire to assess levels of trust, engagement, perceived authenticity, and behavioural purchase intention. To ensure the validity of the measurements, the questionnaire was constructed by referring to already validated scales (Netemeyer et al., 2003; DeVellis & Thorpe, 2021).

Trust was measured through four items from Kennedy et al. (2001) and Kim & Kim (2021) and rated on a Likert scale from 1 ("strongly disagree") to 7 ("strongly agree"), consumer engagement was assessed with four items from De Vries and Carlson (2014) and Giakoumaki and Krepapa (2020), also with a Likert scale from 1 ("strongly disagree") to 7 ("strongly agree"). Perceived authenticity was measured through five items adapted from Mucundorfeanu et al. (2025) and Russell and Rasolofoarison (2017) with a Likert scale from 1 ("strongly disagree") to 7 ("strongly agree"), behavioral purchase intention was assessed with three items from Taylor et al. (1975) and Kim & Kim (2021), using a Likert scale from 1 ("absolutely unlikely") to 7 ("absolutely probable").

All factors showed Cronbach's alphas above 0.7 and composite reliabilities above 0.7 (Hair et al., 2010), indicating the measurement model's reliability.

IV. RESULTS

IV.1 Sample Characteristics

The study participants ranged in age from 19 to 74 years, with a mean age of 39.06 years (standard deviation = 12.83). Regarding gender, 60% of the participants identified as female, 38.3% as male, and 1.7% selected the "other" option.

66.7% had a bachelor's degree, 21.3% had a high school diploma, and 12% had a master's degree. Regarding employment, 72% of participants were employed, 10.3% were unemployed, 6.3% were students, 6% were retired, and 5.3% were housewives. Monthly income was distributed across several bands, with the largest group comprising 15% of participants reporting earnings of more than $5,000 per month.

The participants' nationality was predominantly American (89.7%), with smaller percentages from other origins, including Pakistani, Hispanic, Chinese, and Indian.

IV.2 Results of the Structural Equation Model

The effectiveness of the experimental manipulation (Hauser et al., 2018) was assessed using an independent-samples t-test in IBM SPSS 25. The results showed that participants exposed to the post published by an internal influencer (Group 1) and those exposed to the post published by an external influencer (Group 2) reported significantly different levels of recognition of visual stimuli (internal influencer group: M = 2.50; standard deviation = 1.085; external influencer group: M = 5.81; standard deviation = 0.870; $t(298) = -29.174$; $p < 0.001$).

PROCESS Model 85 was used to conduct a moderated mediation analysis. In this analysis, influencer type was treated as the independent variable (X), while trust and engagement served as sequential mediators (M1 and M2). Perceived authenticity was included as a moderating variable (W), and behavioural purchase intention was the independent variable (Y).

The results of the study confirm hypothesis H1: internal influencers generated significantly higher levels of trust than external influencers ($t(1, 1) = -0.8596$; $p<0.001$), and internal sources are perceived as more credible.

H2 was not supported: The direct effect of source type on behavioural intention was not statistically significant (χ^2 $\beta= 0.0250$; $p = 0.7861$), suggesting that source type alone does not determine purchase intention.

H3 was confirmed: external influencers elicited significantly higher levels of engagement than internal influencers ($t(36) = 0.3898$; $p = 0.0025$). It should be noted that such engagement did not translate into an increase in purchase intention, which is why H4 was not supported: engagement did not represent a significant predictor of behavioural intention $\beta\beta = -0.0199$; $p = 0.6309$).

H5 was rejected: trust did not significantly predict engagement $\beta = -0.0354$; $p = 0.6028$), indicating that in this context, trust and engagement operate as independent processes.

H6 was confirmed: trust had a significant positive effect on purchase intention ($\chi^2 = 0.2097$; $p < 0.001$), reaffirming its central role in influencing consumer behaviour.

H7a was supported: perceived authenticity significantly moderated the relationship between source type and trust, with an interaction coefficient (β) of 0.3377 ($p < 0.001$). Furthermore, the indirect effect of source type on purchase intention, mediated by trust, increased at higher levels of perceived authenticity, with a coefficient $\beta = -0.1803$ at the mean level of authenticity. The moderated mediation index was found to be - 0.0806, with a 95% confidence interval of [-0.1504; -0.0275].

H7b was not confirmed: perceived authenticity did not significantly moderate the effect of source type on engagement with the post (interaction $\beta = 0.0656$; $p = 0.5751$), suggesting that engagement is more likely influenced by the content format or the influencer's personality than by perceived sincerity.

H7c was not supported, as authenticity did not moderate the effect of source type on purchase intention through the engagement pathway. The serial mediation path (type → trust → engagement → intention) was also not significant.

V. DISCUSSIONS

The results of this study show that the type of source (internal vs. external influencer) did not exert a direct effect on behavioural purchase intention, contrasting with existing literature that suggests a positive relationship between source type and individual intention to purchase sponsored products or services (Kumar et al., 2023; Sardar et al., 2024). This result can be explained by considering the mediating roles of trust and engagement, which are more determinants of behavioural outcomes.

Internal influencers significantly increased consumer trust, a finding consistent with studies by Seçilmis et al. (2025) and Ivanov et al. (2024), which found that internal figures (e.g., employees) are perceived as more credible and transparent due to their direct affiliation with the brand.

External influencers elicited higher levels of engagement, in line with research by De Vries & Carlson (2014), who demonstrated that content generated and disseminated by external influencers stimulates greater user interaction on social media (De Vries & Carlson, 2014).

In this study, engagement was not found to be a significant predictor of purchase intention. This finding is not consistent with previous studies (Lou et al., 2019; Sardar et al., 2024). The reason could be the greater superficiality with which it was observed in the present analysis (e.g., a 'like').

Perceived authenticity significantly affected the relationship between source type and trust; for example, internal influencers were more effective at generating trust when perceived as authentic. This finding is consistent with previous studies (Pöyry et al., 2019). It should be noted that authenticity did not significantly affect the relationship between source type and post engagement. This suggests that engagement may be more influenced by aesthetic factors rather than influencer credibility (Wang et al., 2022).

The results of the study show that trust is more important than engagement in determining consumer behaviour. Consumers seem to be more attracted to an authentic voice from within the organisation.

The results suggest that authenticity is more persuasive when it is associated with a direct connection to the brand. Even if an external influencer appears sincere and honest, the lack of direct affiliation with the company can limit the trust-building effect of that authenticity.

These findings highlight the importance of relational alignment between influencers and promoted brands, in line with previous studies (Audrezet et al., 2020), and contribute to the literature by showing that internal authenticity has a stronger impact on consumer trust than perceived authenticity in external content creators.

V.1 Theoretical and Practical Implications

The findings of this work suggest multiple theoretical and managerial implications. The study, thanks to the integration of source credibility theory (Hovland & Weiss, 1951; Ohanian, 1990), engagement theory (Holebeek et al., 2014), commitment-trust theory in relationship marketing (Morgan & Hunt, 1994), and authenticity theory (Wood et al., 2008; Pöyry et al., 2019), presents a comprehensive model that highlights the complex interaction between influencer type, trust, engagement, authenticity and consumer behavioral intention.

The results of this work, demonstrating that internal influencers are perceived as more trustworthy than external content creators, reinforce the source credibility theory. These results are consistent with the studies of Seçilmis et al. (2025) and Ivanov et al. (2024). However, the results show that trust alone is not sufficient; its effectiveness depends on the level of perceived authenticity in the context.

The results of this work confirm that external influencers tend to generate greater engagement in posts, thus contributing to engagement theory (De Vries & Carlson, 2014; Leung et al., 2022). Contrary to Lou et al. (2019), this study found that engagement does not significantly predict behavioural intention, questioning the assumption that all forms of engagement have the same value.

This work confirms the central tenet of commitment-trust theory (Morgan & Hunt, 1994), that trust is a key mediator influencing purchase intention. Trust, more than engagement, underpins behavioural outcomes in digital brand-consumer interactions (Kim & Kim, 2021). The study empirically supports the causal sequence proposed by recent trust-based digital marketing models (Connolly et al., 2023), in which trust precedes and reinforces consumer willingness.

The results reveal that authenticity strengthens the effect of internal influencers on trust, which in turn affects consumer intention. These results are consistent with the studies by Pöyry et al. (2019) and Mucundorpeanu et al. (2025), which found that authenticity is central to the effective communication of influencers.

A final piece that emerges is the contextuality of authenticity, which is not only a personal trait of the influencer but is more persuasive when it aligns with the corporate identity or the role of an employee.

Regarding managerial implications, this study, which highlights the strategic value of internal influencers, suggests that professionals should not rely exclusively on external influencers to gain visibility but should invest in employee-driven content strategies to build deeper, more sustainable relationships with consumers.

This leads to the need for a strategic change in the selection criteria of influencers, not focusing exclusively on the number of followers or entertainment value, companies should also evaluate the alignment of the influencer with the brand and their ability to convey authentic narratives integrated with corporate values.

Another element that emerges from the findings is the link between marketing and HRM. Using employees as brand ambassadors has been shown to strengthen organisational pride and commitment (Li et al., 2024), especially when they are free to share content that reflects their authentic work experience. This opens a new frontier for human resource (HR) departments, which can collaborate with marketing to identify and enhance employees best suited to represent the brand online.

VI. CONCLUSIONS

This work contributes to studies on food marketing on social media by offering an analysis of the psychological mechanisms through which the interaction among source type, trust, post engagement, and perceived authenticity influences consumers' behavioural intentions.

However, it is worth highlighting that the present work is not without limitations. It should be noted that: 1) the use of fictitious pizzerias and influencers may reduce the ecological validity of the results (Reis and Judd, 2000); 2) the generalizability of the results is limited by the fact that the study focused on a single industry (restaurants) and a single country (United States); 3) self-reported measures of behavioral intention, although common in digital marketing research, do not always translate into actual purchasing behaviors. Intentions do not reliably predict actions in the real world due to factors such as social desirability bias, memory errors, or situational constraints (Chandon et al., 2005; Podsakoff et al., 2003).

Future research can extend this work in several ways. First, it could be interesting to replicate the study with real influencers and brands to improve ecological validity and measure actual consumer behaviour. Conducting comparative studies across different cultural markets could help understand how cultural values shape perceptions of influencer authenticity and trustworthiness. Longitudinal studies could monitor the development of trust, engagement, and purchase intention over time following repeated exposure to internal or external influencers. Finally, it may be appropriate to investigate the impact of internal influencers on HRM, focusing on their organisational impact and considering factors such as employee morale, brand loyalty, and alignment with the corporate identity.

REFERENCES

Anjos, C. J. F., Marques, S. and Dias, A. (2022), "The impact of Instagram influencer marketing in the restaurant industry", *International Journal of Service Science, Management, Engineering, and Technology* (IJSSMET), Vol. 13 No. 1, pp. 1–20, https://doi.org/10.4018/IJSSMET.297496

Audrezet, A., de Kerviler, G. and Moulard, J. (2020), "Authenticity under threat: When social media influencers need to go beyond self–presentation", *Journal of Business Research*, Vol. 117, pp. 557–569, https://doi.org/10.1016/j.jbusres.2018.07.008

Brown, D., & Hayes, N. (2008). *Influencer marketing*: Who Influences Your Customers? Oxford, UK: Butterworth-Heinemann.

Campbell, C. and Farrell, J. R. (2020), "More than meets the eye: The functional components underlying influencer marketing", *Business Horizons*, Vol. 64 No. 4, 469–479, https://doi.org/10.1016/j.bushor.2020.03.003

Chandon, P., Morwitz, V. G. and Reinartz, W. J. (2005), "Do intentions really predict behavior? Self-generated validity effects in survey research", *Journal of Marketing*, Vol. 69 No. 2, pp. 1–14. https://doi.org/10.1509/jmkg.69.2.1.60755

Childers, C. C., Lemon, L. L. and Hoy, M. G. (2019), "#Sponsored #Ad: Agency perspective on influencer marketing campaigns", *Journal of Current Issues and Research in Advertising*, Vol. 40 No. 3, pp. 258–274, https://doi.org/10.1080/10641734.2018.1521113

Connolly, R., Sanchez, O. P., Compeau, D. and Tacco, F. M. D. S. (2023), "Understanding engagement in online health communities: a trust-based perspective", *Journal of the Association for Information Systems*, Vol. 24 No. 2, pp. 345–378, https://aisel.aisnet.org/jais/vol24/iss2/8

De Vries, N. J. and Carlson, J. (2014), "Examining the drivers and brand performance implications of customer engagement with brands in the social media environment", *Journal of Brand Management*, Vol. 21, 495–515, https://doi.org/10.1057/bm.2014.18

DeVellis, R. F. and Thorpe, C. T. (2021), "Scale development: Theory and applications", Sage Publications.

Dondapati, A. and Dehury, R. K. (2024), "Virtual vs. Human influencers: The battle for consumer hearts and minds", *Computers in Human Behavior: Artificial Humans*, Vol. 2 No. 1, 100059, https://doi.org/10.1016/j.chbah.2024.100059

Giakoumaki, C. and Krepapa, A. (2020), "Brand engagement in self-concept and consumer engagement in social media: The role of the source", *Psychology & Marketing*, Vol. 37 No. 3, pp. 457–465. https://doi.org/10.1002/mar.21312

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. and Tatham, R. L. (2010), "Multivariate data analysis" (7th Ed.). Pearson Prentice Hall.

Hauser, D. J., Ellsworth, P. C. and Gonzalez, R. (2018), "Are manipulation checks necessary?", *Frontiers in Psychology*, Vol. 9, pp. 1–10. https://doi.org/10.3389/fpsyg.2018.00998

Hollebeek, L. D., Glynn, M. S. and Brodie, R. J. (2014), "Consumer brand engagement in social media: Conceptualization, scale development, and validation", *Journal of Interactive Marketing*, Vol. 28, pp. 149–165, https://doi.org/10.1016/j.intmar.2013.12.002

Hovland, C. I. and Weiss, W. (1951), "The influence of source credibility on communication effectiveness", Public Opinion Quarterly, Vol. 15 No. 4, 635–650.

Hovland, C.I., Janis, I.L. and Kelley, H.H. (1953), "Communication and persuasion", Yale University Press.

Hudders, L., De Jans, S. and De Veirman, M. (2021), "The commercialization of social media stars: A literature review and conceptual framework on the strategic use of social media influencers", *International Journal of Advertising*, Vol. 40 No. 3, pp. 327–375, https://doi.org/10.1080/02650487.2020.1836925

Ivanov, A., Hashim, M. J. and Tacheva, Z. (2024), "From Hype to High Bids: How Internal Influencers Shape Engagement and Performance in C2C Auctions", available at SSRN: http://dx.doi.org/10.2139/ssrn.5080086

Janssen, H. and Rudeloff, C. (2025), "Exploring the role of influencers in shaping employer brands: A comparative study of corporate and third-party influencers", *Corporate Reputation Review*, Vol. 28 No. 1, pp. 57-70. https://doi.org/10.1057/s41299-023-00177-9

Kennedy, M. S., Ferrell, L. K. and LeClair, D. T. (2001), "Consumers' trust of salesperson and manufacturer: An empirical study", *Journal of Business Research*, Vol. 51 No. 1, pp. 73–86. https://doi.org/10.1016/S0148-2963(99)00039-9

Kim, D. Y. and Kim, H. Y. (2021), Trust me, trust me not: A nuanced view of influencer marketing on social media", *Journal of Business Research*, Vol. 134, 223–232, https://doi.org/10.1016/j.jbusres.2021.05.024

Kowalczyk, C. M., & Pounders, K. (2016). Transforming celebrities through social media: The role of authenticity and emotional attachment. Journal of Product & Brand Management, 25(4), 345-356.

Kumar, R., Kaushal, S. K. and Kumar, K. (2023), "Does source credibility matter in promoting sustainable consumption? Developing an integrated model", *Social Responsibility Journal*, Vol. 19 No. 7, pp. 1320–1347, https://doi.org/10.1108/SRJ-07-2021-0257

Lee, P. Y., Koseoglu, M. A., Qi, L., Liu, E. C. and King, B. (2021), "The sway of influencer marketing: Evidence from a restaurant group", *International Journal of Hospitality Management*, Vol. 98, article n. 103022, https://doi.org/10.1016/j.ijhm.2021.103022

Li, G., Nan, G., Li, D. and Tan, Y. (2024), "Cooperating or Cultivating: How Do Online Retailers Operate Live Influencer Marketing Channels?", available at SSRN: https://ssrn.com/abstract=4915669

Li, Y., Zhang, C. and Zhou, M. (2025), "The impact of open branding on consumer brand attitudes: the moderating role of power distance belief", *Journal of Research in Interactive Marketing*, Vol. 19 No. 3, 426–441, https://doi.org/10.1108/JRIM-03-2024-0133

Lou, C. and Yuan, S. (2019), "Influencer marketing: How message value and credibility affect consumer trust of branded content on social media", *Journal of Interactive Advertising*, Vol. 19 No. 1, 58–73. https://doi.org/10.1080/15252019.2018.1533501

Lou, C., Tan, S. S. and Chen, X. (2019), "Investigating consumer engagement with influencer-vs. brand-promoted ads: The roles of source and disclosure", *Journal of Interactive Advertising*, Vol. 19 No. 3, pp. 169-186, https://doi.org/10.1080/15252019.2019.1667928

Mainolfi, G., Marino, V. and Resciniti, R. (2022), "Not just food: Exploring the influence of food blog engagement on intention to taste and to visit", *British Food Journal*, Vol. 124 No. 2, 430–461, https://doi.org/10.1108/BFJ-04-2021-0400

Misra, A., Dinh, T. D. and Ewe, S. Y. (2024), "The more followers the better? The impact of food influencers on consumer behaviour in the social media context", *British Food Journal*, Vol. 126 No. 12, pp. 4018–4035, https://doi.org/10.1108/BFJ-01-2024-0096

Mo, T. and Wang, W. (2025), "The virtual new or the real old? The effect of temporal alignment between influencer virtuality and brand heritage narration on consumers' luxury consumption", *Psychology & Marketing*, Vol. 42 No. 2, pp. 470–492, https://doi.org/10.1002/mar.22132

Morgan, R. M. and Hunt, S. D. (1994), "The commitment-trust theory of relationship marketing", *Journal of Marketing,* Vol. 58 No. 3, pp. 20–38, https://doi.org/10.1177/002224299405800302

Mucundorfeanu, M., Balaban, D. C. and Mauer, M. (2025), "Exploring the effectiveness of digital manipulation disclosures for Instagram posts on source credibility and authenticity of social media influencers", *International Journal of Advertising*, Vol. 44 No. 1, pp. 131–163, https://doi.org/10.1080/02650487.2024.2381973

Netemeyer, R. G., Bearden, W. O. and Sharma, S. (2003), "Scaling procedures: Issues and applications", Sage Publications.

Ohanian, R. (1990), "Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness", *Journal of Advertising*, Vol. 19 No. 3, pp. 39–52, https://doi.org/10.1080/00913367.1990.10673191

Palan, S. and Schitter, C. (2018), "Prolific.ac—A subject pool for online experiments", *Journal of Behavioral and Experimental Finance*, Vol. 17, 22–27, https://doi.org/10.1016/j.jbef.2017.12.004

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y. and Podsakoff, N. P. (2003), "Common method biases in behavioral research: A critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 No. 5, pp. 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Pöyry, E. I., Pelkonen, M., Naumanen, E. and Laaksonen, S.- M. (2019), "A call for authenticity: Audience responses to social media influencer endorsements in strategic communication", *International Journal of Strategic Communication*, Vol. 13 No. 4, pp. 336–351, https://doi.org/10.1080/1553118X.2019.1609965

Reis, H. T. and Judd, C. M. (Eds.). (2000), "Handbook of research methods in social and personality psychology", Cambridge University Press.

Russell, C. A. and Rasolofoarison, D. (2017), "Uncovering the power of natural endorsements: A comparison with celebrity-endorsed advertising and product placements", *International Journal of Advertising*, Vol. 36 No. 5, pp. 761–778, https://doi.org/10.1080/02650487.2017.1348033

Sardar, S., Tata, S. V. and Sarkar, S. (2024), "Examining the influence of source factors and content characteristics of influencers' post on consumer engagement and purchase intention: A moderated analysis", *Journal of Retailing and Consumer Services*, Vol. 79, 103888, https://doi.org/10.1016/j.jretconser.2024.103888

Scott, D. M. (2015). *The new rules of marketing and PR: How to use social media, online video, mobile applications, blogs, news releases, and viral marketing to reach buyers directly*. John Wiley & Sons.

Taylor, J. W., Houlahan, J. J. and Gabriel, A. C. (1975), "The purchase intention question in new product development: a field test", *Journal of Marketing*, Vol. 39 No. 1, pp. 90–92, https://doi.org/10.1177/002224297503900119

Seçilmiş, C., Kiliç, İ., Sari, Y. and Şenel, E. (2025), "The role of influencer marketing in the rapid rise of Nusr-et: the cognitive response theory", *Journal of Hospitality and Tourism Insights*, Vol. 8 No. 1, pp. 62–79, https://doi.org/10.1108/JHTI-10-2023-0773

Van Driel, L., & Dumitrica, D. (2021). Selling brands while staying "authentic": The professionalization of Instagram influencers. Convergence, 27(1), 66-84.

Vrontis, D., Makrides, A. and Christofi, M. (2021), "Social media influencer marketing: A systematic review, integrative framework and future research agenda", *International Journal of Consumer Studies*, Vol. 45 No. 4, pp. 617–644, https://doi.org/10.1111/ijcs.12647

Wang, C. H., Sher, S. T. H., Salman, I., Janek, K. and Chung, C. F. (2022), ""TikTok made me do it": teenagers' perception and use of food content on TikTok", In *Proceedings of the 21st annual ACM interaction design and children conference* (pp. 458–463).

Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M. and Joseph, S. (2008), "The authentic personality: A theoretical and empirical conceptualization and the development of the "authenticity scale"", *Journal of Counseling Psychology*, Vol. 55 No. 3, pp. 385–399, https://doi.org/10.1037/0022-0167.55.3.385

02

RECONFIGURING WORK IN THE AGRIFOOD CHAIN: PROFILING EMPLOYABILITY SKILLS VIA BIG DATA AND TRANSFORMER-BASED LANGUAGE MODELS

DOI: 10.37199/c41000302

Francesco SMALDONE

G. Marconi University of Rome (Rome, Italy)
f.smaldone@unimarconi.it
ORCID 0000-0001-7581-2810

Benedetta ESPOSITO

San Raffaele Roma University (Rome, Italy)
benedetta.esposito@uniroma5.it
ORCID 0000-0002-4476-3339

Stefania SUPINO

San Raffaele Roma University (Rome, Italy)
stefania.supino@uniroma5.it

Daniela SICA

San Raffaele Roma University (Rome, Italy)
daniela.sica@uniroma5.it
ORCID 0000-0002-3466-3844

Abstract

Shifting towards a circular economy requires a comprehensive and systemic overhaul of job functions and the associated skill sets, ensuring alignment with the core principles of resource efficiency, waste reduction, and sustainable value creation. This shift transcends traditional industrial boundaries, calling for the seamless integration of circular strategies across a wide range of economic sectors. Among these, the agri-food sector stands out as a pivotal arena for change due to its multifaceted role.

Realising this transformation demands a workforce equipped with highly specialised and flexible competencies capable of driving innovation, implementing circular economy solutions, and effectively coordinating circular strategies throughout the entire food supply chain. Such capabilities *encompass a robust understanding of interdisciplinary methods, advanced technological applications, and the ability to navigate complex sustainability challenges.*

Despite the recognised importance of the agri-food sector in advancing circularity on a broader scale, previous research has largely treated circular economy-related competencies or agri-food-specific skills as isolated areas of inquiry, often failing to address the full supply chain in a cohesive manner. Moreover, while an expanding body of literature examines the skills needed for circular economy adoption across various industries, to the best of our knowledge, no study has systematically and holistically mapped the competencies essential for accelerating circularity in the agri-food system.

Accordingly, this study addresses that gap by applying an advanced topic modelling and semantic analysis framework to a curated dataset of 7,943 job advertisements collected between June and October 2024 from leading recruitment platforms and corporate career pages across the United States. The analysis focuses exclusively on job postings within the agrifood domain, aiming to uncover the most salient employability skills and their latent thematic groupings.

The methodological approach is grounded in natural language processing and unsupervised machine learning. Job descriptions were pre-processed through lemmatisation, tokenisation, and stopword removal, followed by the generation of dense sentence embeddings using a transformer-based large language model. These embeddings served as the input to Bidirectional Encoder Representations from Transformers (BERTopic), a neural topic modeling algorithm that integrates semantic representations with dimensionality reduction and density-based clustering.

The findings confirm that employability in the agrifood sector is increasingly defined by a hybrid profile of competencies. The methodological contribution of this study lies in the integration of large language models (LLM)-based text mining with interpretable topic modeling and correlation logic, offering a replicable and scalable approach to skills profiling in sectors undergoing technological and environmental transformation.

Keywords: agri-food industry; big data; circular economy; employability; skills; text mining.

I. INTRODUCTION

The transition towards a circular economy constitutes a foundational reconfiguration of industrial and organisational systems. It involves a structural shift from the traditional linear model of *take-make-dispose* towards regenerative models of production and consumption that aim to close material loops, extend product lifecycles, and minimise ecological impact (Borrello et al., 2017; Esposito et al., 2025; Kirchherr et al., 2017). As a result, labour markets are undergoing a corresponding transformation, whereby employability is progressively defined by a new logic of sustainable value creation, ecological resilience, and innovation-led productivity (European Commission, 2020).

Within this evolving landscape, the agri-food sector is widely recognised as both a major contributor to environmental degradation and a potential enabler of circularity. Its global relevance, ecological embeddedness, and socio-economic complexity underscore its strategic position in the green transition (Burger et al., 2019). The sector's exposure to climate volatility, natural resource constraints, and regulatory pressures has intensified demand for workforce capabilities that transcend traditional agricultural knowledge and embrace sustainability as a core operational principle (Cedefop, 2012; Straub et al., 2023; Trivelli et al., 2019). Undeterred by the strategic importance of this sector, relatively few studies have attempted to systematically examine the employment implications of circular transition within the agri-food domain (Trevisan et al., 2024). Despite this urgency, the academic literature has yet to offer a comprehensive mapping of the employability requirements emerging across the agri-food supply chain in response to circular economy imperatives.

Research to date has primarily examined green skills in generalised settings or assessed agri-food competencies in isolation, thereby neglecting the integrative perspective required for systemic transformation (Merrifield, 2013; Smaldone et al., 2025a). Existing research has largely focused on either macro-level economic dynamics or firm-level innovation processes, with limited attention to how these changes alter the nature of work and the skillsets required across different phases of the agri-food value chain (Wu, 2025). The complexity of circular transitions demands not only technical proficiency but also interdisciplinary insight, adaptive capacity, and collaborative skillsets, dimensions often omitted from traditional workforce planning models (OECD, 2020; Trevisan et al., 2025).

This gap in the literature is particularly problematic when considering the increasingly hybrid nature of skill requirements in circular economy sectors. Today's professionals are expected to integrate technical expertise with cross-disciplinary thinking, digital fluency, and collaborative problem-solving abilities (De los Rios et al., 2017; Summerton et al., 2019). These transversal and adaptive skills become especially relevant in agri-food systems, where workers are called upon to navigate complex, interconnected processes spanning production, processing, logistics, and environmental compliance (Guyot Phung, 2019; Smaldone et al., 2025b).

Conventional approaches to identifying skills, reliant on surveys, classification frameworks, and qualitative foresight, risk underestimating the speed and heterogeneity of change in dynamic labour environments. By contrast, recent advances in data science and computational linguistics offer promising avenues for capturing real-time shifts in demand through large-scale text analytics of job advertisements (Smaldone et al., 2022; Smaldone et al., 2025a). These techniques enable the extraction of emergent competencies and latent thematic structures directly from employer-generated data, thereby facilitating a bottom-up, demand-led model of skill intelligence (Janssens et al., 2021; Sumter et al., 2021).

The present study leverages these methodological innovations to analyse a curated corpus of 7,943 job advertisements collected between June and October 2024 from leading U.S. recruitment platforms. The analysis applies transformer-based natural language processing and neural topic modelling to uncover the implicit logic through which agri-food employability is articulated and operationalised in a circular context. Emphasis is placed on identifying and categorising core, advanced, and transversal skills, and on clustering them into semantically coherent domains across the supply chain. In doing so, the study seeks to contribute both to empirical labour market intelligence and to the development of scalable methodologies for skill profiling in sectors undergoing ecological transition.

II. METHODOLOGY

This study adopts an advanced text mining approach that leverages natural language processing and machine learning to explore the evolving configuration of employability skills in the agri-food sector. The analysis is based on a curated dataset comprising 7,943 job advertisements collected between June and October 2024 from leading employment platforms and corporate career websites operating in the United States. The selection criteria targeted postings explicitly referring to agri-food-related functions, including agricultural production, food processing, quality control, logistics, environmental compliance, sustainability consultancy, and innovation in food systems. To prepare the textual data for semantic analysis, a multi-step preprocessing pipeline was implemented. Initially, all job descriptions were subjected to lowercasing, lemmatisation, tokenisation, and the removal of stopwords, punctuation, digits, and extraneous symbols. Furthermore, terms with extremely low or high document frequency were filtered to eliminate lexical noise and enhance analytical precision. These procedures ensured the semantic tractability of the corpus and facilitated the extraction of meaningful linguistic structures from unstructured text (Munzert et al., 2015; Silge et al., 2017).

Following preprocessing, the corpus was converted into dense semantic vectors using Sentence-BERT (SBERT), a transformer-based language model tailored for sentence-level embedding generation. SBERT fine-tunes the BERT architecture using a Siamese network structure, enabling high-performance similarity detection across short texts, a key capability for identifying latent relationships among job postings and the skill terms they contain (Reimers & Gurevych, 2019). These embeddings were used as input for topic modelling. To identify thematic clusters of co-occurring skill expressions, the BERTopic algorithm was employed. This neural topic modelling framework integrates SBERT embeddings with dimensionality reduction via UMAP and density-based clustering via HDBSCAN, allowing the model to autonomously infer the number and composition of topics based on the topological density of the embedded corpus (Grootendorst, 2022). Compared to traditional methods such as Latent Dirichlet Allocation, BERTopic is capable of retaining contextual nuance and offers a more accurate reconstruction of semantic fields within highly heterogeneous textual datasets. The resulting topic clusters were reviewed and interpreted as distinct thematic areas of skill demand. A keyword analysis was conducted to extract the most representative terms from each cluster, which were subsequently matched against internationally recognised skill taxonomies developed for green, digital, and transversal competencies (Cedefop, 2012; OECD, 2020).

II. RESULTS

The semantic analysis of the 7,943 agri-food-related job advertisements revealed a structured yet diverse demand for competencies reflecting both traditional operational functions and emerging circular economy priorities. A clear concentration of frequency was observed among a subset of core skills, with sustainability knowledge emerging as the most frequently required, appearing in 14.6% of all job postings. This skill was often associated with roles involving strategy implementation, reporting, and regulatory alignment, suggesting a widespread institutionalisation of sustainability principles across the agri-food value chain.

Closely following were quality assurance (12.8%) and food safety regulations (11.7%), which together reflect the sector’s ongoing preoccupation with compliance, traceability, and product integrity. These skills were particularly prevalent in job roles linked to processing, manufacturing, and distribution, aligning with both consumer safety mandates and international trade standards.

Supply chain coordination appeared in 10.3% of job postings, underscoring the sector’s growing reliance on logistics optimisation and integrated value chain planning. This demand aligns with the increasing complexity of circular logistics frameworks, particularly in managing reverse flows, waste segregation, and bio-based production. The skill of environmental compliance was cited in 8.9% of advertisements, further confirming the relevance of regulatory awareness and monitoring capabilities in the circular transition. Other frequently occurring competencies included technical reporting (7.5%), process optimisation (6.7%), and agricultural monitoring (5.9%), all of which support operational efficiency and resource-sensitive management practices. A cluster of specialised skills associated with circular strategies appeared in smaller but significant proportions. Circular production planning was found in 4.8% of postings, often connected to positions in innovation, sustainability engineering, or system design. Similarly, data interpretation (4.1%) emerged as an increasingly transversal requirement, not confined to analytical roles but distributed across managerial and field-level functions.

Less frequent yet notable were skills such as waste-reduction strategies (2.7%), resource-efficiency management (2.4%), and product innovation (3.2%). Their distribution indicates a gradual diffusion of circular-economy logic into operational practices, particularly among firms engaging with environmental certifications or eco-labelling schemes. Low-frequency skills included those of high specialisation, such as Life Cycle Assessment (1.4%), biowaste valorisation (1.2%), and eco-design (0.7%). While quantitatively marginal, the presence of these terms highlights an emerging niche in agri-food employment requiring advanced training and interdisciplinary expertise. Overall, the frequency distribution suggests that agri-food employability is currently defined by a hybrid profile in which traditional compliance and safety roles coexist with an expanding demand for system-level thinking and circular innovation. This dual structure reflects both the sector's regulatory inertia and its potential for sustainable transformation underpinned by data, technology, and design-oriented skills.

The application of BERTopic to the corpus yielded a three-topic solution that captured the latent semantic organisation of skill-related demands in the agri-food sector. Each topic reflected a coherent thematic nucleus and was categorised based on the prevailing nature of the competencies it encompassed, namely, hard, soft, or transversal skills. The largest topic, accounting for 61.4% of all job postings, was dominated by hard skills. These included competencies such as food safety management, HACCP implementation, quality assurance protocols, technical reporting, environmental compliance, and production process monitoring. Job roles associated with this topic were primarily situated in operational and regulatory domains, indicating the centrality of technical expertise and conformity to legal and quality standards in agri-food employability.

A second topic, comprising 22.3% of the advertisements, corresponded to transversal skills. The lexical composition of this cluster revealed high-frequency references to project management, sustainability coordination, supply chain logistics, data analysis, and cross-functional collaboration. Unlike the previous group, this topic cut across multiple organisational areas, reflecting the increasing demand for integrative, system-level roles that require interaction between technical, managerial, and sustainability-related functions. These skills often appeared in job descriptions that combined planning responsibilities with operational oversight and reporting duties. The third and smallest topic, representing 16.3% of the corpus, captured soft skills. Terms such as communication, teamwork, adaptability, flexibility, initiative, and problem-solving featured prominently in this cluster. Postings in this category frequently referenced collaborative environments, stakeholder engagement, and interpersonal dynamics, often in the context of project facilitation or client-facing functions. Despite its lower overall weight, the soft skill topic revealed a distinct communicative and behavioural component of employability, embedded particularly in positions that demanded responsiveness to changing operational contexts or multi-actor coordination.

The topic model confirmed a clear quantitative predominance of hard skills, followed by transversal and soft skill clusters, respectively. This tripartite distribution suggests a layered architecture of skill demand, in which technical specialisation remains foundational, while system integration and behavioural competence are emerging as essential complements in the evolving agri-food labour market.

III. DISCUSSIONS

The three-topic structure identified through BERTopic modelling offers important insights into the configuration of employability within the agri-food sector, confirming and expanding upon recent theoretical and empirical contributions to green labour market transformation. The quantitative predominance of hard skills, found in over 61% of analysed job postings, is aligned with prior observations that the agri-food industry remains highly regulated and operationally structured, relying on a workforce capable of ensuring compliance with increasingly stringent environmental and safety standards. The prominence of terms such as HACCP, food safety, and environmental monitoring echoes findings by Janssens et al. (2021), who described the circular transition in this sector as primarily "compliance-led", whereby firms engage with sustainability through institutionalised control mechanisms rather than through radical reorganisation of work processes (Esposito et al., 2025). This regulatory and technical orientation, while unsurprising, suggests a partial mismatch between the complexity of circular challenges and the linearity of skill expectations still prevailing across many segments of the industry. The modelled topic dedicated to transversal skills, which accounted for 22.3% of the corpus, provides evidence of an ongoing expansion in the semantic space of agri-food employability. Skills such as project management, sustainability reporting, supply chain integration, and cross-departmental coordination point to the growing demand for system-level professionals who are capable of mediating between technical knowledge and organisational adaptation. These findings resonate with the OECD's (2020) conceptualisation of transversal green skills as catalysts for innovation in traditionally conservative sectors, enabling the embedding of sustainability principles across structural silos (Guyot Phung, 2019).

Equally noteworthy is the emergence of soft skills as an autonomous cluster within the topic model, comprising 16.3% of the analysed postings. Although often underrepresented in technical literature, behavioural competencies such as teamwork, flexibility, initiative, and communication were frequently mentioned in association with leadership, problem-solving, and stakeholder engagement. This aligns with the argument advanced by Sumter et al. (2021), who emphasised the relevance of "sustainable work" frameworks in which soft skills serve not as auxiliary attributes but as central components of employability in socially complex and environmentally dynamic contexts (Smaldone et al., 2022). Moreover, the clear separation of soft skills from both hard and transversal clusters in the model indicates a shift in how employers articulate expectations, reinforcing the trend toward hybrid professional profiles that blend operational fluency with social and cognitive agility. Taken together, the distribution of the three topics reflects a labour market undergoing a staged but multidimensional transformation. The continued dominance of hard skills suggests that many employers still interpret circularity through the lens of compliance, technical optimisation, and regulatory adherence. At the same time, the significant presence of transversal competencies suggests a parallel evolution toward integrative roles that are structurally positioned to facilitate systemic transition. Soft skills, while quantitatively less dominant, appear to function as enabling traits for those tasked with navigating uncertainty, mediating institutional change, or fostering innovation from within. This layered configuration of skill demand is consistent with the thesis advanced by Smaldone et al. (2025) thesis that productivity in knowledge-intensive green sectors depends not merely on technical expertise but on the strategic orchestration of heterogeneous competencies. The tripartite architecture revealed by this study thus confirms the coexistence of distinct but interconnected domains of employability in the agri-food sector. While circularity as a principle has become increasingly visible in job advertisements, its operationalisation still relies heavily on traditional regulatory frameworks. At the same time, the gradual incorporation of transversal and soft dimensions signals the early emergence of more adaptive, holistic, and strategic workforce models. This suggests that the agri-food sector, despite its structural conservatism, is beginning to develop internal pathways for innovation-compatible labour practices, though these remain unevenly distributed and often under-recognised in recruitment language.

IV. CONCLUSIONS

This study has provided a data-driven reconstruction of employability dynamics in the agri-food sector under the lens of circular economy transition. Through the application of transformer-based topic modelling on a corpus of nearly 8,000 job advertisements, a tripartite semantic structure was identified, encompassing hard, transversal, and soft skills. The quantitative predominance of hard skills confirms the continued salience of technical and regulatory competencies in a sector shaped by safety, quality, and environmental standards. However, the meaningful presence of transversal and soft skill clusters indicates a gradual evolution toward more integrated and adaptive workforce profiles. The methodological contribution of the study lies in the combination of large language model embeddings and unsupervised clustering techniques, which allowed for the extraction of latent thematic structures from unstructured labour market data. This approach offers a replicable framework for semantic skill profiling that may be applied to other sectors undergoing technological, ecological, or normative transformation. In conceptual terms, findings suggest that employability within the agri-food system cannot be fully captured through static taxonomies or linear definitions of professional roles. Instead, it is increasingly shaped by hybrid logics that reflect the convergence of regulatory compliance, cross-functional coordination, and behavioural flexibility. This layered configuration reinforces the importance of viewing circular economy transitions not only as technical shifts but as organisational and human capital challenges. Future research may expand upon these insights by integrating temporal dynamics and longitudinal comparisons, or by linking job ad analysis with educational curricula, workplace performance indicators, or policy implementation outcomes. For workforce developers, educators, and institutional planners, the results underscore the need to design skill-building strategies that embrace complexity, promote versatility, and align with the evolving grammar of sustainability-oriented employment.

REFERENCES

Borrello, M., Caracciolo, F., Lombardi, A., Pascucci, S., & Cembalo, L. (2017). Consumers' perspective on circular economy strategy for reducing food waste. Sustainability, 9(1), 141.

Burger, M., Stavropoulos, S., Ramkumar, S., Dufourmont, J., & van Oort, F. (2019). The heterogeneous skill-base of circular economy employment. Research Policy, 48(1), 248–261.

CEDEFOP. (2012). Future skills supply and demand in Europe: Forecast 2012. Publications Office of the European Union.

De los Rios, I. C., & Charnley, F. J. (2017). Skills and capabilities for a sustainable and circular economy: The changing role of design. Journal of Cleaner Production, 160, 109–122.

Esposito, B., Smaldone, F., Raimo, N., Sica, D., & Vitolla, F. (2025). Enhancing circular economy disclosure towards Net Zero: The role of circular economy capabilities. International Journal of Productivity and Performance Management.

European Commission. (2020). A new Circular Economy Action Plan: For a cleaner and more competitive Europe. European Commission.

Grootendorst, M. (2022). BERTopic: Neural topic modeling with class-based TF-IDF. arXiv:2203.05794.

Guyot Phung, C. (2019). Implications of the circular economy and digital transition on skills and green jobs in the plastics industry. Field Actions Science Reports, Special Issue 19, 100–107.

Janssens, L., Kuppens, T., & Van Schoubroeck, S. (2021). Competences of the professional of the future in the circular economy: Evidence from the case of Limburg, Belgium. Journal of Cleaner Production, 281, 125365.

Kirchherr, J., Reike, D., & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. Resources, Conservation and Recycling, 127, 221–232.

Merrifield, A. (2013). The politics of the encounter: Urban theory and protest under planetary urbanization (Vol. 19). University of Georgia Press.

Munzert, S., Rubba, C., Meißner, P., & Nyhuis, D. (2015). Automated data collection with R: A practical guide to web scraping and text mining (Vol. 1). Wiley.

OECD. (2020). Green skills and innovation for inclusive growth. OECD Publishing.

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using siamese BERT-networks. arXiv:1908.10084.

Silge, J., Robinson, D., & Robinson, D. (2017). Text mining with R: A tidy approach (p. 194). O'Reilly.

Smaldone, F., Ippolito, A., Lagger, J., & Pellicano, M. (2022). Employability skills: Profiling data scientists in the digital labour market. European Management Journal, 40(5), 671–684.

Smaldone, F., Lagger, J., & Mainolfi, G. (2025a). Profiling employability skills for life cycle professionals amid the circular and digital economy. The International Journal of Life Cycle Assessment, 1–18.

Smaldone, F., Supino, S., Sica, D., & Esposito, B. (2025b). Uncovering employability skills for circular food systems: A big data approach. International Journal of Economic Behavior.

Straub, L., Hartley, K., Dyakonov, I., Gupta, H., van Vuuren, D., & Kirchherr, J. (2023). Employee skills for circular business model implementation: A taxonomy. Journal of Cleaner Production, 410(1), 137027.

Summerton, L., Clark, J. H., Hurst, G. A., Ball, P. D., Rylott, E. L., Carslaw, N., … & McElroy, C. R. (2019). Industry-informed workshops to develop graduate skill sets in the circular economy using systems thinking. Journal of Chemical Education, 96(12), 2959–2967.

Sumter, D., de Koning, J., Bakker, C., & Balkenende, R. (2021). Key competencies for design in a circular economy: Exploring gaps in design knowledge and skills for a circular economy. Sustainability, 13(2), 776.

Trevisan, A. H., Arcebi, F., Dukovska-Popovska, I., Mascarenhas, J., Terzi, S., & Sassanelli, C. (2024). A research agenda advancing circular economy skills. In International Conference on Interoperability for Enterprise Systems and Applications (pp. 207–220). Springer Nature Switzerland.

Trevisan, A. H., Boscarato, A., Acerbi, F., Terzi, S., & Sassanelli, C. (2025). Enhancing circular economy education and training for the manufacturing sector: A holistic skills framework. Journal of Environmental Management, 380, 124982.

Trivelli, L., Apicella, A., Chiarello, F., Rana, R., Fantoni, G., & Tarabella, A. (2019). From precision agriculture to Industry 4.0: Unveiling technological connections in the agrifood sector. British Food Journal, 121(8), 1730–1743.

Wu, Y. (2025). Identifying and prioritizing essential skills for the circular economy workforce. Journal of Environmental Management, 393, 127237.

03

USER-CENTERED DIGITAL PRODUCT DESIGN: A TRANSPORTATION-RELATED CASE STUDY

DOI: 10.37199/c41000303

George LYCHRIDAS

University of Western Macedonia (Kila Kozani, Greece)
giorgoslychridas@gmail.com
ORCID 0009-0002-6436-0286

Thomas LIOKOS

University of Western Macedonia (Kila Kozani, Greece)
thomasliokos@gmail.com
ORCID 0009-0006-7050-3289

Anastasios TZOTZIS

University of Western Macedonia (Kila Kozani, Greece)
a.tzotzis@uowm.gr
ORCID 0000-0002-6942-9636

Panagiotis KYRATSIS

University of Western Macedonia (Kila Kozani, Greece)
pkyratsis@uowm.gr
ORCID 0000-0001-6526-5622

Abstract

The term design is not limited to the appearance of the final product; it encompasses elements that justify its shape and functions. It is an escalating procedure that is heavily based on research to assess the problem properly, the use of an array of user-centred design tools, the evaluation of the proposed idea, and finally, the improvement of potential drawbacks. The above-mentioned procedure is crucial to the product design methodology, which is not linear but characterised by a repeated rhythm, providing the opportunity for more efficient products to emerge. In the present paper, a number of digital product design tools were used as a medium for developing a potential recreational concept vehicle. The project reflects a vision to design a vehicle whose design characteristics could correlate with Mediterranean tradition and summer holiday habits. As a result, it could be equipped with technical elements that facilitate motion on soft terrains. Furthermore, a series of creativity tools was used to transform the research data into design inspiration, to achieve optimal outcomes during the design process, and to prototype technologies/methods to provide insights into the form and functionality of the design solution.

Keywords: User-centred Design, Product Design, Digital Design Tools, CAD, Prototyping

I. INTRODUCTION

Design is a crucial factor in the development of new products, as it can positively affect a company's economic prosperity and corporate identity. Usually, the design team consists of several members who represent the creative section of a company to find innovative solutions. It is a dominant belief that the perception of a product's design is limited to its aesthetic characteristics. Still, it combines functionality and modern appearance (Roper et al., 2016). A successful product design can captivate and attract prospective consumers while also distinguishing itself from competitors.

Frequently, the term “good design” is heard, but it is rather difficult to identify with quantitative elements. Generally, design success can be evaluated in different ways, which depend on the person’s perspective (Roy Robin 2010). Furthermore, the uniqueness of a product is a major factor in its success. This characteristic stems not only from new design elements but also from older statements that function uniquely. More unique products can be produced when companies focus on developing effective methods to collect and document customers’ needs and, subsequently, to transform them into design requirements (Falahat et al., 2024).

To discover worthy new products, it is necessary to use effective methods and software systems throughout the design process. Traditionally, the role of the designer is balanced between their artistic background and technical knowledge. However, as of today, changes are driven by the emergence of new technologies that can more precisely present the characteristics of a product, such as augmented reality, before it enters the manufacturing process (Eric Lutters et al., 2014). The influence of the technological equipment has also emerged in the initial stages of the design process, as sketches are depicted digitally using pen displays. Simultaneously, the emergence of computational design, which relies on mathematical sequences, can yield variations of the original 3D model and changes in the product's functional and ergonomic characteristics (Manavis et al., 2024; Manavis et al., 2021; Manavis et al., 2023). At the same time, the evolution of 3D CAD (Computer Aided Design) systems allows designers to interact more accurately with possible solutions, in particular when parametrically designing mechanical parts are involved (Mermoz et al., 2011).

II. LITERATURE SURVEY

The design process can be described as a logical sequence of steps aimed at producing effective solutions. These steps are project-independent, meaning they can be applied consistently across projects, in the same order, with as many iterations as needed, to achieve successful outcomes when performed carefully (Cheng Jinxia, 2018). This paper aims to design a recreational vehicle, following design principles and incorporating contemporary digital tools to better understand the design language. Although Mediterranean terrain morphology and climate factors inspired the vehicle's design, it can be used in other similar circumstances, such as high temperatures, sandy terrain, and water volumes.

During design research, it was necessary to document information on Mediterranean habits, traditions, and climate conditions. A major factor that could influence the adoption of a person in a certain terrain is the temperature (Luo et al., 2024), and subsequently, people's behaviour, attitude, and willingness to be outdoors for an extended part of the day are influenced by weather conditions (Huang et al., 2016; Jianlei Niu et al., 2015). At the same time, the tonality of every colour can evoke a range of emotions and convey visual information to the observer. Since ancient times, the influence of colour on people's psychology and its ability to express the values of different cultures more effectively have been researched. Yuan Meng et al. particularly highlighted the significance of colour in architecture, where it enhances the aesthetics of each building and differentiates it from others, regardless of its spatial arrangement (Yuan Meng et al., 2025). The Parthenon is undoubtedly one of the most significant buildings of ancient Greek times, and its appearance and construction methods were proposed as the foundations for the further development of architectural science, the outward surfaces of which were abundantly decorated by elements whose colours were associated with the Greek tradition (Aggelakopoulou et al., 2022).

Regarding the vehicle's technical background, it is crucial to note that electric motors are used for its development to highlight the vehicle's environmentally friendly character. The growing number of vehicles with internal combustion engines drives a significant rise in carbon dioxide emissions. For the above-mentioned reason, the car industry invested resources in further developing the electric car with solar panels to be more efficient, offering a higher mileage range (Tsakalidis et al., 2025). The climate conditions in Mediterranean countries, combined with batteries' higher performance at higher temperatures, make electric vehicles a good choice for passenger transportation (Rabih Al Haddad et al., 2025). Furthermore, the concept was to design a vehicle with two characteristics: sufficient control in a water environment, thus with low hydrodynamic resistance, and optimal driving behaviour on soft terrains like a sandy beach. Therefore, research was conducted on the direction of key design elements on amphibious vehicles (Liu et al., 2023), particularly an experimental model by Xiaolin Xie et al., which features wheels that transform into propellers. This mechanism enables the vehicle to move through water without obstruction (Xie et al., 2017).

III. METHODOLOGY AND APPLICATION

The generation of new, innovative solutions by designers is not based solely on their personal experience and mindset, but also on carefully conducted research and on transforming data into design requirements and parameters. Additionally, following a step-by-step process in which each step produces results that serve as key elements for the next, a design loop was introduced when needed. That leads to a continuous trial-and-error process in which, when a result is not technically or aesthetically sufficient or relevant to the design solution, the design team can loop back to a needed step to redefine the parameters and produce different results until a competent outcome is obtained.

In the first steps, a target is set to better understand each design problem, while the next steps focus on evaluating and testing the final solution using computer-aided design systems and 3D printers. Some of the design tools could be used in analogue, but the reduced time consumption provided by digital tools and the facilitation of communication of optical information among team members led to the adoption of digital solutions. In any case, digital tools can handle details at a higher level, enhancing design accuracy and enabling a more refined ideation process through easier experimentation.

III.1 The Sequence of Methodology

To better understand the specific roles of the various design tools, it is necessary to categorise them into four main sections based on their aims. The first section aims to summarise the collected information and transform it into design inspiration parameters, using both physical language and images. For that reason, mind maps and mood boards are powerful design tools that can motivate designers to adopt creative design thinking and, in turn, generate a wealth of ideas that can lead to innovative products. The next level's purpose is to design and represent the initial ideas in a more artistic form. Ultimately, designers use their sketching skills to transfer their thoughts onto paper or a digital canvas and to estimate if an idea can be further developed. At this point, it is important to note that the design process is divided into four stages, and the sketches in each differ in quality and detail. The third section facilitates projecting 2D digital sketches into a 3D environment. With the help of digital tools, designers can evaluate the final design from different perspectives and visualise models in a digital environment that simulates the real world. The 3D printing is the final stage of the design process, where the model can be designed in a CAD system, which provides the opportunity of creating a model in the needed size for prototyping, so the designer can physically interact with it and have a better understanding of the shape, the analogies and the ergonomics of the product.

III.2 Mind map – Mood board

The mind map is an essential design tool for helping the design team orient towards the desired design direction, properly communicate with stakeholders, and avoid misunderstandings. The design team needs to frame the design problem properly at an early stage. In the centre of the mind map, the problem definition is placed, while potential design opportunities emerge in the surrounding areas (Figure 1). Beyond documenting research information, developing a mind map also fosters the exchange of ideas and knowledge among design team members. When the team cooperates efficiently, new opportunities can be discovered, and new design areas can be revealed for exploration. In this phase, there are no “strange ideas,” and all ideas are documented without hesitation, so during evaluation, the keywords most relevant to the design direction are chosen. There are different ways to develop a mind map; the simplest is pencil and a piece of paper, but the most efficient is to take advantage of digital technologies. Digital media allow designers to use colours, lines, and shapes to highlight specific, important words and to rearrange the flow of keywords with ease.

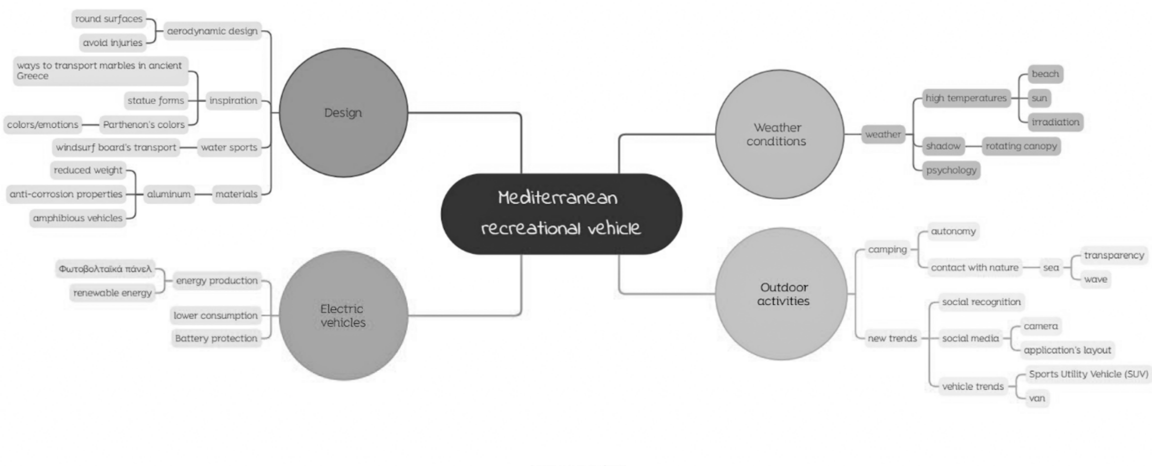


Figure 1. Mind-map incorporates data and thoughts

Source: Author’s processing

After the preliminary listing of keywords, a second design tool (a mood board) is used to motivate designers to express their emotions through pictures. At this point, different pictures are used that originate from the designers' personal collections, whereas other image collections with rich visual information facilitate the discovery of pictures that can express the designer’s emotion. Visual communication among the design team is important because it can be more comprehensive than words alone, incorporating colours, shapes, and textures that are difficult to describe verbally. Figure 2 depicts the selected images for this case study of the recreational vehicle, which also reflect the design direction followed throughout the design process.

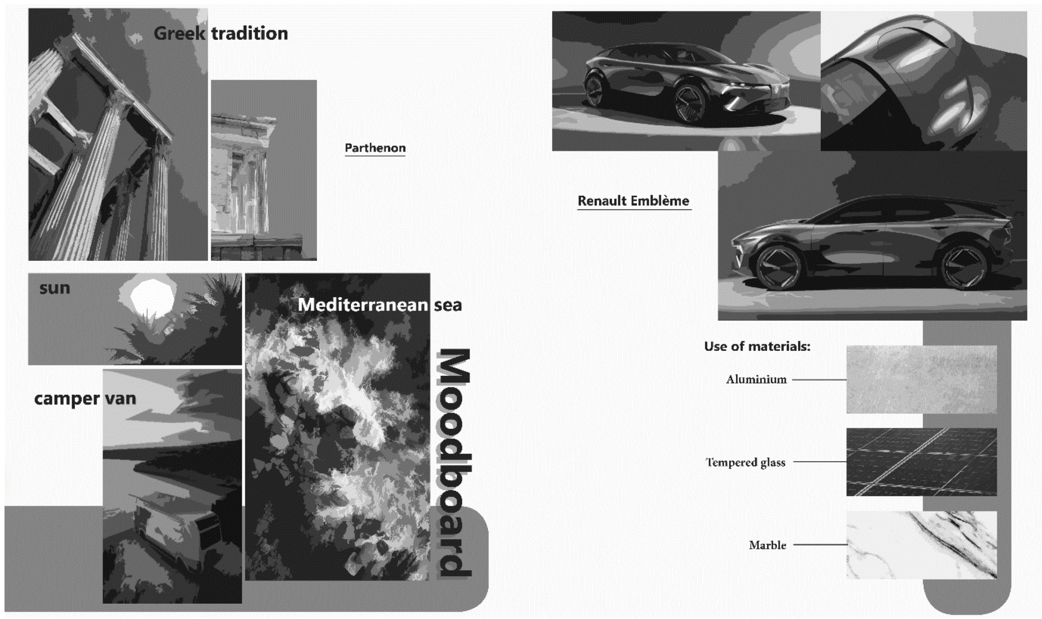


Figure 2. Mood board, expression of feelings via photographs

Source: Author's processing

III.3 3D Digital Sketching

Sketching is the creative section of the design procedure and the key element that allows information to be transformed into design elements. During this process, various sketches are created, whose external form and functional parameters reflect the design requirements recorded in the previous stage. The key feature, and the most time-consuming, in this procedure is experimentation, while drawing a plethora of sketches plays an important role. During the sketching phase, the most suitable sketches are explored in depth, and, finally, with tools such as a 3D CAD modelling system and physical prototyping, the design solution is chosen and evaluated. New technologies, such as pen displays, allow designers to sketch with high precision and effectiveness. With pen displays, sketches are more realistic, and the ease of adding and manipulating colours, backgrounds, reflections, and textures creates no boundaries for creativity. Simultaneously, these devices allow creating sketches with a 3D sense because the user can control the pen's precision, resulting in bold front surfaces and faded remote points.

Four stages complete this procedure, in which the quality and precision of sketches differ on each stage, and are the following: procedural sketches, ideation, final solution, and colour rendering. In this case study, all sketches are created digitally with a Huion™ Kamvas™ 22 and Affinity Designer™.

In this stage, the main purpose is to conceive the vehicle's main body without focusing on secondary design elements; for this reason, all sketches are drawn from the side view to ensure the vehicle's overall shape is as comprehensive as possible. Additionally, light colours are used on the positive surfaces, and darker colours on the points where negative spaces or holes exist (Figure 3).

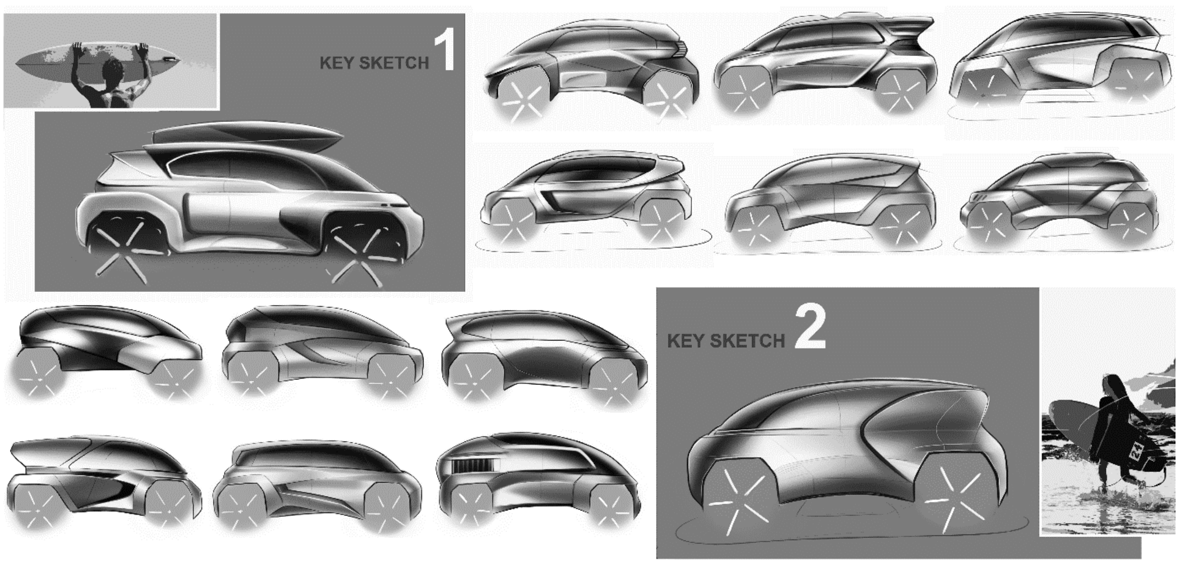


Figure 3. The procedural sketches encompass 2D initial forms

Source: Author's processing

Ideation is the second level of the sketching process, in which the designer aims to showcase their creative spirit. The variables documented in previous stages are taken into consideration to successfully incorporate all parameters into rough sketches for future design proposals. The sketches at this stage are presented with enhanced perspective and from various angles, making it easier to represent the concept in three dimensions. Among the above-mentioned vehicle sketches, our design team has selected two models for further development with greater accuracy: the front and rear parts of the vehicle (Figure 4).

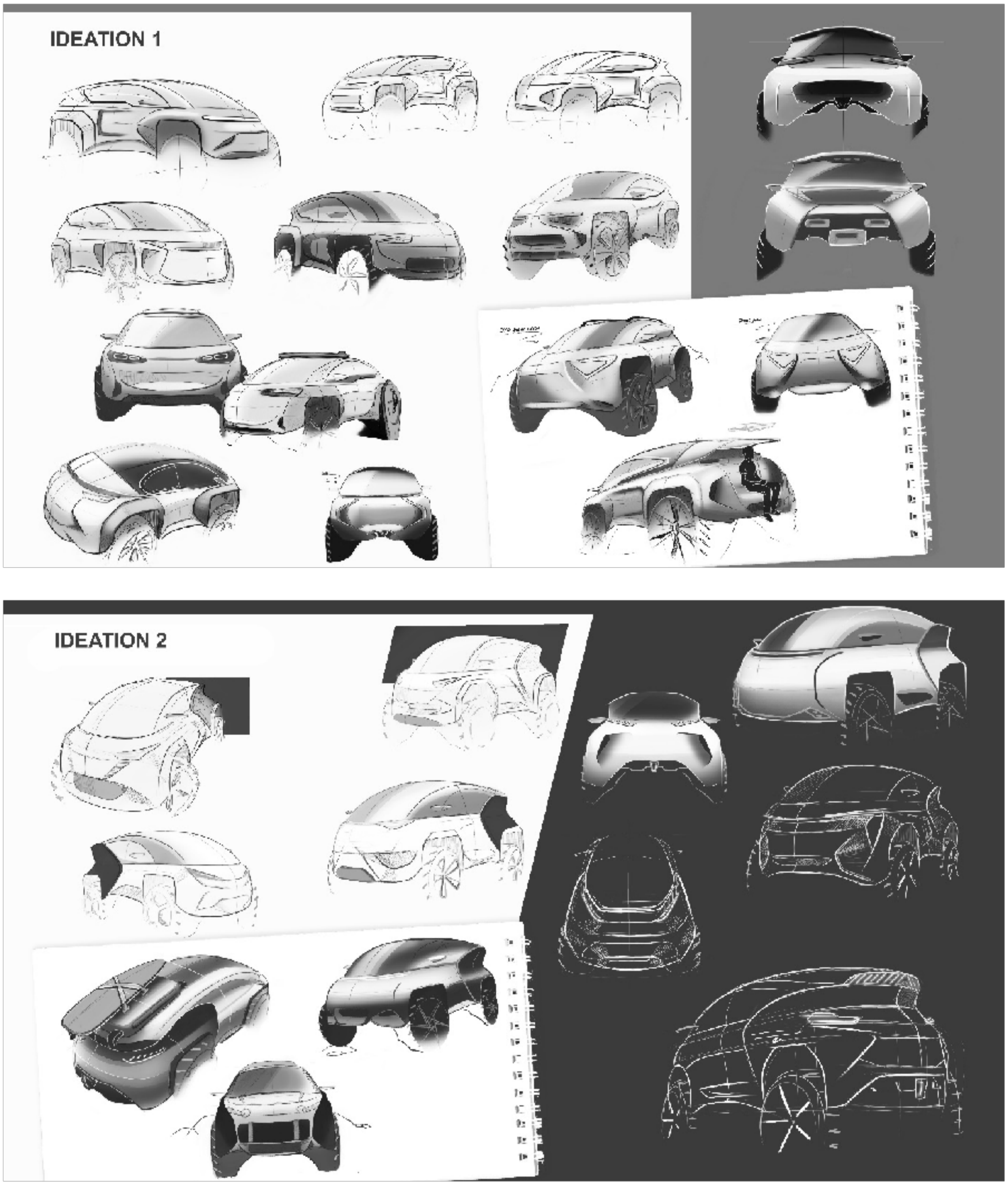


Figure 4. Ideation sketches

Source: Author's processing

After the sketches have been thoroughly analysed, the solution that best meets the essential requirements is chosen. In this stage, all the details are shown in the sketch as the final design of the idea, while the sketch's lines and analogies are drawn precisely to provide a more realistic result. To achieve this, the sketch is rendered with the addition of colours and a background (Figure 5). The brighter colours are used in the nearest part of the sketch for the viewer, to increase the observer's attention, while darker tones are used for larger areas and shadows. The extremely saturated colours are not the right choice for the sketches because they increase the contrast between the colours and lead to unusual results and artefacts. Undoubtedly, 2D digital sketches are more realistic than analogue sketches, as modern digital design systems offer tools and colour systems that better represent the real world.

Figure 5. Final solution - realistic approach

Source: Author's processing

III.4 3D CAD modelling and rendering

High-quality sketching can provide potential solutions with enough realism and detail, but it cannot quickly showcase a product from every angle. There is often insufficient and systematic sketch development to support the evaluation of all product models within a company. Time-consuming processes should be avoided to maintain cost efficiency and ensure that the product is launched during a period of high consumer demand. In a product's lifecycle, 3D CAD systems can assist designers in digitising their final ideas promptly and easily, and in changing parts that do not match the final design. These state-of-the-art programs facilitate creators' sketching in a 3D environment, with powerful commands that enable the modelling of unusual and interesting shapes.

Regarding the 3D modelling of the proposed recreational vehicle, the Blender™ system is used for its powerful capabilities to shape, transform, and adapt complex forms. At the same time, this program provides two powerful rendering systems (Eevee and Cycles) that can render the 3D model in a digital environment with a realistic feel. To ensure maximum flexibility during 3D modelling, the vehicle's components are independently developed and integrated in the final phase. In this way, each part of the model can be evaluated and modified without affecting the car's main body structure. In addition, the initial renderings are produced in white colour and on a plain background to better assess the model's characteristics without distraction from vivid colours or digital scenery (Figure 6).

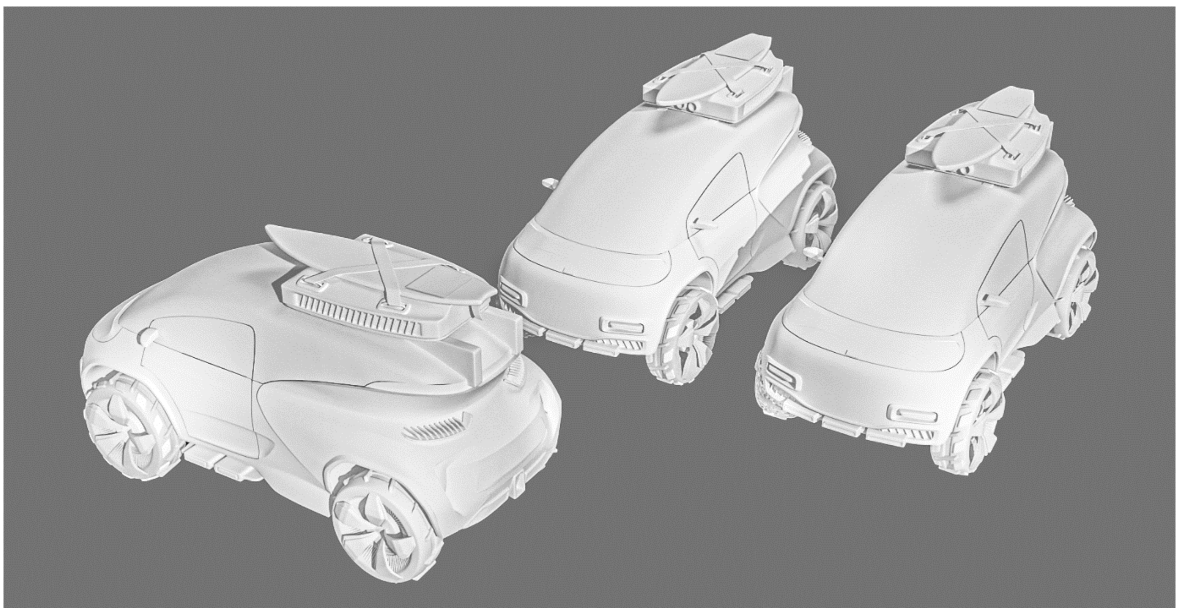


Figure 6. Colourless 3D CAD model

Source: Author's processing

While the colourless renderings help expose potential flaws in the final model and lead to reexamining some of the design's parts, the final rendering in full colour and background aims to showcase the product in a realistic way (Figure 7). To achieve high-quality renders, the "Cycles" rendering engine in BlenderTM is used, which can render the model as realistically as possible, with proper colour balance, high-quality surface reflections, and seamless integration with the 3D environment. Each material is distinguished by its roughness, metallic quality, and transparency, so selecting the right textures is essential for clearly perceiving the different parts of the product.

By applying HDRI (High Dynamic Range Imaging) environments, the model is rendered with natural lighting and reflections, offering a clear, convincing presentation to stakeholders. Another method is to use two digital programs (in our case, BlenderTM and Affinity DesignerTM), where the model is rendered at a favourable perspective angle with a transparent background, then placed in the selected photo. These digital programs offer immense capabilities, enabling continuous interaction between them to achieve the best possible result.

Figure 7. Realistic renderings of a 3D CAD model

Source: Author's processing

III.5 Prototyping

In the final stage, a thorough evaluation of the 3D CAD model is essential. While digital visualisation is an efficient way to communicate the design and its parameters, physical interaction with a tangible prototype offers deeper insight through tactile feedback. Furthermore, prototyping helps understand the model's geometry, allowing the analogy to be examined and dimensions and measures to be verified, ensuring that the product aligns with the initial design requirements. Using 3D printers enables fast prototype production. This makes 3D printing a valuable evaluation tool, as it enables the team to produce multiple physical models as the 3D design evolves. (Figure 8).

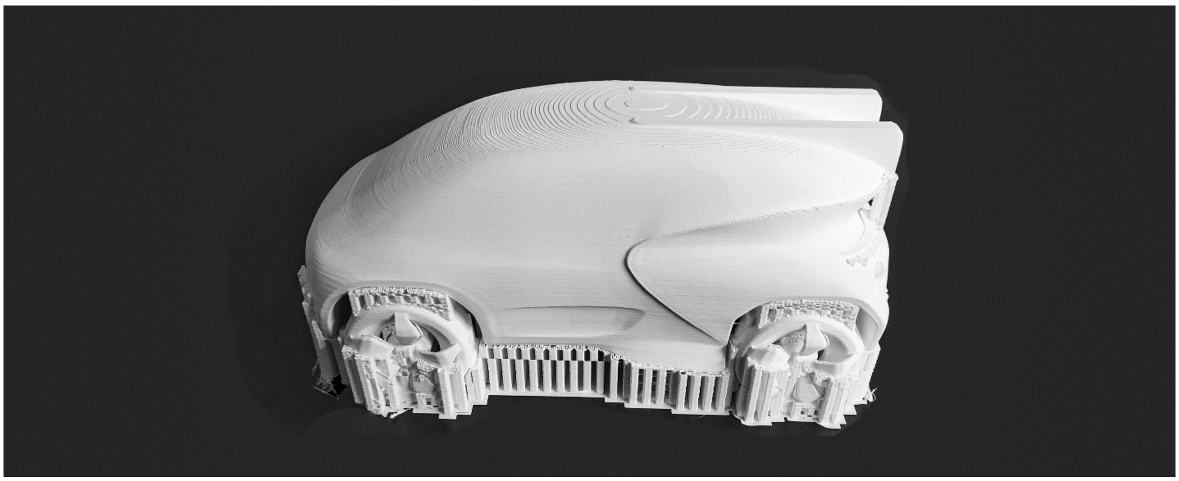


Figure 8. Prototyping with the use of a 3D printer

Source: Author's processing

IV. CONCLUSION

The contemporary design process is increasingly grounded in systematic methodologies rather than subjective judgment, with digital tools playing a central role at every stage to enhance speed and efficiency. The rapid evolution of digital media technology brings many changes to how products are designed, while the methodology of their development remains unchanged. Finally, alongside the presentation of certain digital tools, particular emphasis is placed on the vehicle itself, which, on the one hand, is used as a means of illustrating the design process, and on the other hand, is developed with particular care as an alternative proposal for the development of a recreational vehicle. The design language of this vehicle features elements that ideally express its general use, primarily on coastal routes and terrains, and also reflects Mediterranean culture.

V. REFERENCES

Aggelakopoulou, E., & Bakolas, A. (2022). What were the colours of the Parthenon? Investigation of the entablature's cornice blocks. *Journal of Archaeological Science, 140*, 105553. https://doi.org/10.1016/j.jas.2022.105553

Cheng, J. (2018). Product design process and methods. In *IntechOpen*. https://doi.org/10.5772/intechopen.80821

Lutters, E., van Houten, F. J. A. M., Bernard, A., Mermoz, E., & Schutte, C. S. L. (2014). Tools and techniques for product design. *CIRP Annals*, *63*(2), 607–630. https://doi.org/10.1016/j.cirp.2014.05.005
(DOI added for accuracy)

Falahat, M., Chong, S. C., & Liew, C. (2024). Navigating new product development: Uncovering factors and overcoming challenges for success. *Heliyon, 10*, e23763. https://doi.org/10.1016/j.heliyon.2023.e23763

Huang, J., Zhou, C., Zhuo, Y., Xu, L., & Jiang, Y. (2016). Outdoor thermal environments and activities in open space: An experimental study in humid subtropical climates. *Building and Environment, 103*, 10–23. https://doi.org/10.1016/j.buildenv.2016.03.029
(Pages inferred from journal issue numbering)

Niu, J., Liu, J., Lee, T.-C., Lin, Z., Mak, C., Tse, K.-T., Tang, B., & Kwok, K. C. S. (2015). A new method to assess spatial variations of outdoor thermal comfort: Onsite monitoring results and implications for precinct planning. *Building and Environment, 93*(1), 263–270. https://doi.org/10.1016/j.buildenv.2015.06.029
(Added correct volume/issue/DOI)

Liu, B., Xu, X., & Pan, D. (2023). Resistance reduction optimization of an amphibious transport vehicle. *Ocean Engineering, 280*, 114854. https://doi.org/10.1016/j.oceaneng.2023.114854

Luo, X., Zhao, Y., Yang, J., Li, Y., & Luo, Z. (2024). Adaptive design strategies for thermal comfort in waterfront campsites during summer in hot and humid regions. *International Journal of Thermal Sciences, 199*, 108945. https://doi.org/10.1016/j.ijthermalsci.2024.108945

Manavis, A., Kakoulis, K., & Kyratsis, P. (2023). A brief review of computational product design: A brand identity approach. *Machines, 11*(2), 232. https://doi.org/10.3390/machines11020232

Manavis, A., Minaoglou, P., & Efkolidis, N., & Kyratsis, P. (2024). Digital customization for product design and manufacturing: A case study within the furniture industry. *Electronics, 13*(13), 2483. https://doi.org/10.3390/electronics13132483

Manavis, A., Minaoglou, P., Tzetzis, D., Efkolidis, N., & Kyratsis, P. (2021). Computational design technologies for interior designers: A case study. *IOP Conference Series: Materials Science and Engineering, 1009*, 012037. https://doi.org/10.1088/1757-899X/1009/1/012037

Mermoz, E., Linares, J.-M., & Bernard, A. (2011). Benefits and limitations of parametric design implementation in helicopter gearbox design phase. *CIRP Annals – Manufacturing Technology, 60*(1), 199–202. https://doi.org/10.1016/j.cirp.2011.03.095

Al Haddad, R., Mansour, C., Kim, N., Seo, J., Stutenberg, K., & Nemer, M. (2025). Comparative analysis of thermal management systems in electric vehicles at extreme weather conditions: Case study on Nissan Leaf 2019 Plus, Chevrolet Bolt 2020 and Tesla Model 3 2020. *Energy Conversion and Management.*

Roper, S., Micheli, P., Love, J., & Vahter, P. (2016). The roles and effectiveness of design in new product development: A study of Irish manufacturers. *Research Policy, 45*(1), 319–329. https://doi.org/10.1016/j.respol.2015.10.003

Roy, R. (2010). Can the benefits of good design be quantified? *Design Management Journal (Former Series), 5*(1), 9–17. https://doi.org/10.1111/j.1948-7169.1994.tb00380.x

Tsakalidis, A., Thiel, C., & Jäger-Waldau, A. (2024). Can solar electric vehicles disrupt mobility? A critical literature review. *Renewable and Sustainable Energy Reviews, 211*, 115289. https://doi.org/10.1016/j.rser.2024.115289

Xie, X., Gao, F., Huang, C., & Zeng, W. (2017). Design and development of a new transformable wheel used in amphibious all-terrain vehicles (A-ATV). *Journal of Terramechanics, 69*, 45–61. https://doi.org/10.1016/j.jterra.2016.11.001

Meng, Y., & Liu, W. (2025). “Design nature”: A color interpretation of Bauhaus school building in Dessau and its conceptual origins. *Frontiers of Architectural Research, 32–61. (Volume/issue pending—add once available)*

04

REGIONAL TRANSPORT CORRIDORS: A COMPARATIVE ANALYSIS OF ALBANIA'S PERFORMANCE WITH NEIGHBOURING COUNTRIES

DOI: 10.37199/c41000304

Zana GUZJA (JOCA)

Ministry of Infrastructure and Energy
Agricultural University of Tirana
Zana.Joca@infrastruktura.gov.al

Ledia THOMA

Agricultural University of Tirana
Faculty of Economy and Agribusiness
ledia.thoma@ubt.edu.al

Abstract

The objective of this study is to examine Albania's performance within the main regional transport corridors. The research methodology employed is a comparative analysis of the Logistics Performance Indicator (LPI), an index developed and published by the World Bank. The LPI facilitates comparative analysis across the period 2007-2023. Comparisons of LPI values and rankings are comprehensive, encompassing not only Albania but also other nations within the main regional transport corridors. Identifying significant discrepancies between the LPI values and their six constituent components in Albania and other countries within and beyond the regional corridors enables the identification of areas requiring more immediate improvements and necessary interventions by the appropriate authorities. The objective is not solely to enhance the LPI value but also to reduce the disparities with the comparator countries, thereby augmenting the nation's competitiveness. The analysis concludes that Albania has retained the 2023 LPI value close to the average level for the period 2007-2023, primarily due to advancements in transport infrastructure and customs. The data indicate that Albania must pursue broader enhancements across other LPI components, including increasing logistics capacity and quality, adhering to deadlines for destination arrivals, and improving the ability to track movements. The countries within the study's corridors have made progress on these indicators, thereby conferring a competitive advantage over Albania. Extending the time series into the future will facilitate the application of quantitative assessment methods and yield statistically significant results.

Keywords: Regional Transport Corridors, Comparative Analysis, Competitiveness, Logistics Performance Index (LPI).

I. INTRODUCTION

Transport infrastructure constitutes a fundamental prerequisite for the economic and social advancement of nations. The integration of diverse transportation modalities, the network, and ancillary infrastructure, alongside the establishment of national and regional corridors, is a key driver of enhancing nations' prominence and achieving economies of scale. The literature elucidates foundational theoretical concepts underpinning transport corridors, which are characterised as economic corridors or as a "programmed conceptual model" that aligns with a socio-economic structure aimed at territorial development, fostering economic activities proportionate to the population, and commensurate with a specific stage of growth in the transport infrastructure (Healey, 2004). Accordingly, economic corridors combine infrastructure investments, policies, regulatory and institutional frameworks, their implementation, and extant private and public capacities. Emphasising priority sectors such as tourism, agribusiness, logistics, and IT, transport corridors are designed, constructed, and developed to optimise long-term economic growth in a balanced and inclusive manner. The study aims to examine the significance of transport corridors in Albania, drawing on both sector-specific statistics and related fields. The core hypothesis under investigation is whether transport corridors and associated infrastructure have impacted Albania's competitiveness relative to other countries in the region and beyond, with some being part of major regional transport corridors.

The article is organised as follows: the second section provides a summary of the fundamental literature regarding transport corridors and their socio-economic implications. In the third section, the Logistics Performance Indicator (LPI), developed and published by the World Bank, is introduced in Albanian studies in this field. This indicator, encompassing transport and infrastructure performance aspects, serves as a foundation for developing a comparative analysis across selected years from 2007 to 2023 for Albania and other countries, mainly within the region covered by the main corridors. This section examines the areas that require further development to enhance the country's competitive advantage in infrastructure innovation and beyond. The fourth section concludes and formulates recommendations grounded on the outcomes of the comparative analysis.

Box 1. LPI – A general overview of the meaning and purpose of use

The World Bank's LPI is a comprehensive index that has covered the entire supply chain across 139 to 160 countries in the 2007 to 2023 publication. It is based on a survey of nearly 1,000 logistics professionals worldwide. It is useful for comparing performance across countries and identifying and prioritising broad reform areas for interventions within countries/regions. The index is based on numerical ratings of 1 (weakest) to 5 (strongest). The LPI is a weighted average of six components: 1. Efficiency of the clearance process (customs); 2. Quality of trade- and transport-related infrastructure; 3. Ease of arranging competitively priced international shipments; 4. Competence and quality of logistics; 5. Ability to track and trace consignment; 6. The frequency with which shipments reach the consignee within the scheduled or expected delivery time. Logistics and infrastructure are a critical contribution to economic activity within and across borders. Efficient transportation and logistics help reduce trade costs and are essential to trade and regional integration. Low international connectivity, inadequate logistics infrastructure, poor logistics services, and lengthy trade procedures at and beyond the border raise logistics costs. Logistics costs amount to 8% of GDP in the United States but rise to 15%-20% in many middle-income countries and to 30% or more in low-income countries, including landlocked or island states.

The LPI tells how easy (perceived) it is to establish efficient and reliable connections between trading countries. Conceptually, the LPI is a revealed metric of supply chain accessibility. The LPI is the primary supply chain benchmark for policymakers and has, since 2007, motivated reforms. The LPI has been used as a component indicator in several high-level dashboards, including those for the Sustainable Development Goals and the World Economic Forum.

Source: World Bank Group (May 2024).

II. LITERATURE REVIEW

Research on the economic and social benefits stemming from advancements in the transport sector consistently concludes that the development of this sector and its associated infrastructure are key drivers of economic growth and the enhancement of nations' competitive advantages. Transport activity plays a significant role in small and open economies, such as Albania and most Central and South-Eastern European countries. An efficient transport system enables such economies to advance through internal and external trade activities, facilitates the movement of people and the delivery of services, and fosters economies of scale. Consequently, a functional, well-distributed transport network stimulates economic and social progress in countries and regions. The evolution of transport activity is closely linked to the comprehensive infrastructure for internal and external land, water, and air transport. From this perspective, transport performance must be considered together with investments in proper infrastructure. Given the intense utilisation of infrastructure, the transport sector can be considered an indispensable instrument and catalyst for economic development. This is more evident in a global economy, where the movement of individuals and goods, inclusive of information and communication technologies, necessitates an integrated transport infrastructure. As expressed by Rodrigue & Notteboom (2024), there exists a crucial correlation between the quantity and quality of transport infrastructure and the level of economic development. Efficient transport systems provide economic and social opportunities and benefits, manifested in numerous positive outcomes, such as enhanced accessibility to production and service markets, labour markets, and investment opportunities. Substandard transport systems exhibit limited operating capacity, with accompanying low reliability and safety. These systems usually cost much money, making countries less competitive and reducing people's quality of life.

In globalisation, transport and infrastructure show how competitive a country is in logistics. This is because a country's economic growth often depends on having an efficient transport system.

The transport industry possesses macro and microeconomic significance. At the macroeconomic level, transport influences total added value, employment, and income. In numerous developed economies, transport accounts for 6-12% of GDP, and logistics costs range from 6-25% of GDP. On a microeconomic scale, transport affects the costs for producers, consumers, and supply chains. Typically, higher income levels correspond to a greater proportion of transport in consumption expenditure. Transport constitutes approximately 10-15% of household expenditure in developed nations. For Albania, the proportion of transport in household expenditure for 2023 is projected at approximately 6.8% (INSTAT, 2024). Although relatively modest, this proportion has increased slightly over the years.

The implementation of regional transport policies is actualised, among other elements, through the construction of transport corridors. Transport corridors, in particular, significantly reduce disparities at both national and regional levels. Keser (2015) emphasises that, from the early stages of regional projects, the development of transport infrastructure is considered a prerequisite for reducing inequalities. According to Armstrong & Taylor (2000), "region" and "development" are unified as a concept, with "regional development" gaining increasing significance over time.

Comprehensive studies on transport corridors generally find that transport infrastructure—highways, railways, and multimodal corridors—has a substantial impact on the expansion and enhancement of economic activities and social life in the regions traversed by these corridors. Various studies, along with their analyses and conclusions concerning the advantages of transport corridors, are elaborated below (Table 1). In the case of Albania, the academic literature, however, remains limited.

Author/Institution	Subject/Conclusions
World Bank Group (South Asia Region). (2018, 2019); Krugman (1991)	Transport corridors facilitate the safer, faster movement of goods and services, thereby improving economic activity in general and trade in particular. They reduce transport costs, improve access to larger markets, and promote regional and wider trade integration.
Grover *et al*. (2023); Cohen *et al*. (2008)	The authors have verified the hypothesis that efficient transport infrastructure supports the clustering, or economic agglomeration, process by allowing businesses to benefit from economies of scale, larger labour markets, and extended supply chains.
Bannister & Berechman, 2001	Transport corridors often lead to economic clustering around key hubs, fostering the development of industrial and special economic zones.
Gilliam (2011)	Corridors connect rural areas to urban centres, enabling rural development by improving access to markets, services, and resources.
Rodrigue, Comtois, & Slack (2016) ; Hallam (2011)	Studies show that transport corridors largely promote industrialisation and foreign and domestic investment by improving logistics and reducing delays, critical factors for businesses that depend on the timely delivery of goods and services.
Limão & Venables (2001)	The authors estimated that a 10% improvement in transport infrastructure could increase trade flows by up to 25%, attracting more industries and services along transport corridors.
Dercon *et al.* (2008)	Transport corridors help reduce poverty and disparity between urban and rural areas. They increase agricultural productivity, improve labour mobility, and increase access to education and healthcare in vulnerable areas.
Bruinsma & Rietveld 1993	Transport corridors can stimulate growth, but the benefits may not be distributed evenly. In some cases, they can lead to regional

	inequalities if poorer regions are bypassed, or to environmental damage from construction and increased traffic. They point out that the benefits of transport corridors also depend on the governance and management structures in place.
Other references related to the Balkan region and Albania are found below:	
Xhepa (2000)	The author addresses geopolitical and economic aspects of the development of transport corridors in Central and Eastern Europe and the Balkans. The author concludes that the development of transport corridors increases competitive advantages not only at the level of individual countries but also between regions. Policies should be put at the service of the country's integration interests (Albania).
Qaja (2021)	Methodologically, the research area is a road corridor connecting settlements and northern Albania with the state of Kosovo. A combined model of corridor management is proposed by integrating and connecting settlements, with an observatory placed at the centre of this corridor. Some theoretical conceptions for "transport corridors" are also suggested.
EBRD (2019)	Strategic document: Transport Sector Strategy 2019-2024.
OECD (2021)	Multi-dimensional Review of the Western Balkans - Assessing Opportunities and Constraints.
European Commission (2018).	Untapped Potential: Intra-Regional Trade in the Western Balkans.

Table 1. Grouping of literature by subject

Source: Based on the literature listed and summarised by the author.

Note: For more information, see References.

Following the literature review, let us present some facts about the transport sector for a group of countries, including Albania. Within the European Union (EU), the average contribution of the transport sector to GDP was approximately 5% during the period 2015–2023. Furthermore, the sector employs over 10 million individuals across EU member states. For the same period, transport's contribution to GDP in Albania averaged 3%, about 2 percentage points lower than that of regional countries in the Balkans and EU member states (Figure 1). According to administrative labour market data, the transport sector in Albania employed around 2.7% of the total workforce during the period 2019–2024 (Figure 2). The number of employees in this sector and its share of total employment have been increasing, especially after 2021.

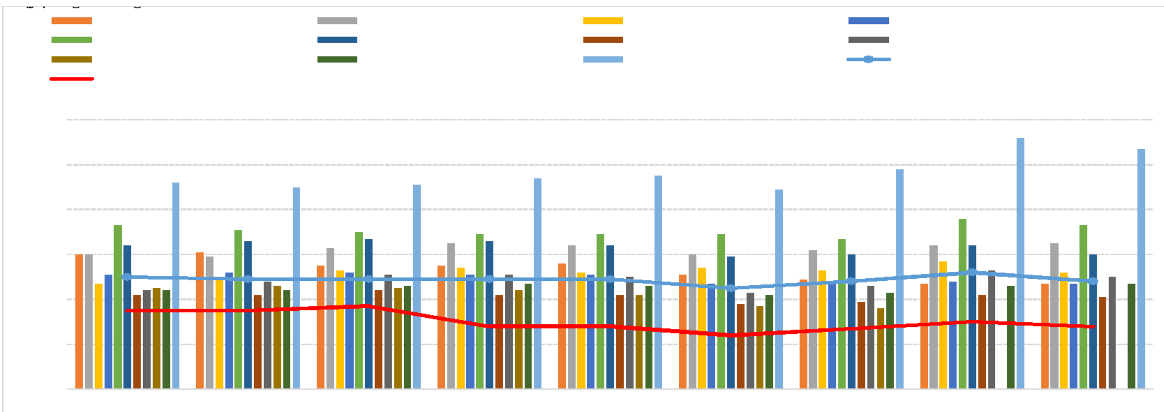


Figure 1. Contribution of Transport to total GDP (%) by country for the period 2015-2023

Source: Eurostat, INSTAT and author's calculations.

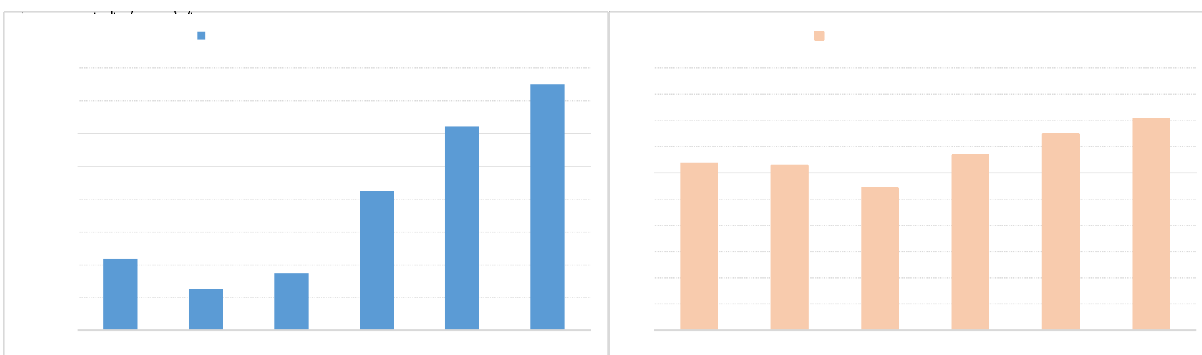


Figure 2. Number of employees in transport and specific weight to total employees.

Source: INSTAT and author's calculations.

Note: According to NACE, employees directly employed in the activity of "land and pipeline transport; water, air and storage transport" are included. Employees in transport support services in other activities are omitted.

III. COMPARATIVE ANALYSIS OF THE PERFORMANCE OF TRANSPORT CORRIDORS USING LPI DATA

III.1. Regional Corridors through Albania

Transport corridors in Albania are crucial for facilitating connectivity with neighbouring nations and promoting the movement of goods and individuals within the Western Balkans. Owing to its strategic geographic location, Albania serves as a significant transit route for trade and transportation between the Adriatic Sea and the Balkan Peninsula. The principal transport corridors and routes within Albania include:

- **The Blue Corridor / Adriatic-Ionian Highway** passes from the northern part of Albania (border with Montenegro) to the southern part (border with Greece). As a primary transport artery in Albania, it connects all the cities along the coast and the major ports. This corridor holds particular significance as it simultaneously connects four Balkan countries: Croatia, Montenegro, Albania, and Greece.
- **Corridor VIII**: Albania is part of this corridor, which connects all the capitals from the Adriatic Sea (Durrës) to the Black Sea. Originating in Italy, this corridor connects Albania with the Republic of North Macedonia (RMV) and Bulgaria through inland routes, thereby facilitating regional trade and transportation.
- **The Durrës-Prishtina Corridor** is a vital roadway connecting the coastal city of Durrës in Albania with Kosovo's capital, Prishtina. This route is essential for linking the two countries and enhances trade and communication between them. Stretching approximately 300 km, the Durrës-Prishtina highway provides a rapid, efficient connection between the two nations. This link is significant for bilateral communication and trade and serves as a fundamental corridor for their economic integration into the EU.

Notably, Albania has enhanced its transport infrastructure to strengthen connectivity within the Balkans and further with the European network. Such developments include the construction and modernisation of roads, railways, and ports, aiming to facilitate trade and promote economic growth.

III.2 Performance of the LPI for Albania 2007-2023 in transport corridors

The logistics performance indicator is measured and published by the World Bank. Operations and services related to the mobility of people and goods, and the storage of goods, are logistics services. One of the most important components of the LPI is the infrastructure component, which includes transportation issues (roads, ports, airports, and rail operations) as defined in Annexe 1.

In the case of Albania, the data show that compared to the first year of LPI measurement (2007), the trends of the LPI and its constituent components have improved. The increase in the LPI in recent years has been influenced by "Infrastructure" and "International Shipments", which in 2023 were 2.7 and 2.8, respectively (Figure 3). The LPI components with above-average performance are

"International Shipments" and "Timeliness". The "Infrastructure" component has seen rapid improvement from 2016 to 2023. This component has a normalised weight of about 17% in the aggregated LPI, the same as that of "Logistic Competence & Quality".

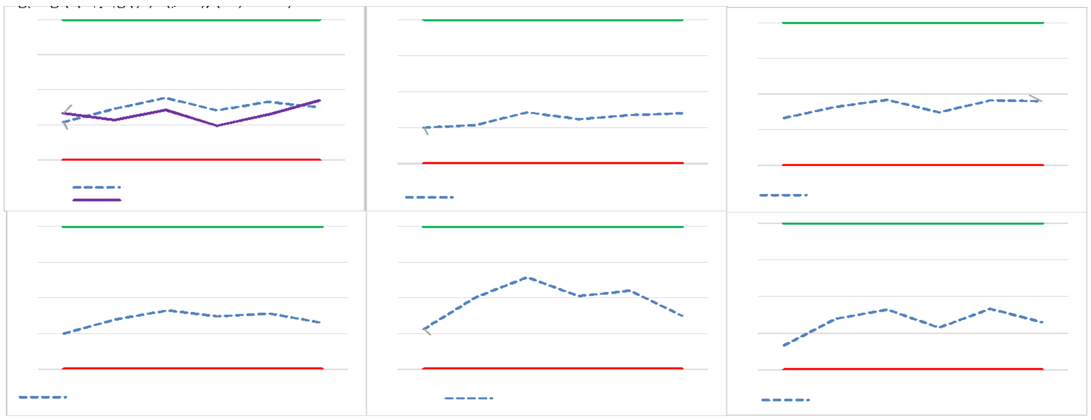


Figure 3. Performance of the LPI and its components for Albania over the years.

Source: World Bank, https://lpi.worldbank.org/ and author's calculations.

Table 2 shows Albania's ranking, in ascending order by year, for the indicator (the lower the ranking, the weaker the country's performance). The LPI for Albania represents an improvement in ranking compared to other countries participating in the survey. Thus, from 139th in 2007, where it was 11 places behind, in 2023 it ranks 97th, leaving 42 places behind. Regarding the infrastructure component, the ranking shows high instability. However, the 68th ranking in 2023 has moved it to a more central position, leaving about 71 countries behind, with a weaker infrastructure performance than Albania.

Years	No. of countries in the study	LPI Grouped Rank	Customs Rank **(1)**	Infrastructure Rank **(2)**	International Shipments Rank **(3)**	Logistics Competence & Quality Rank **(4)**	Timeliness Rank **(5)**	Tracking & Tracing Rank **(6)**
2007	150	139	132	78	109	130	144	145
2010	155	119	129	112	104	103	120	124
2012	155	78	86	99	70	91	45	88
2016	160	117	121	148	110	102	94	135
2018	160	88	114	110	69	92	73	95
2023	139	97	90	68	75	120	125	117
		Countries below Albania, according to the LPI rank and components						

Years	No. of countries in the study	No. of countries ranked below	Customs Rank	Infrastructure Rank	International Shipments Rank	Logistics Competence & Quality Rank	Timeliness Rank	Tracking & Tracing Rank
2007	150	11	18	72	41	20	6	5
2010	155	36	26	43	51	52	35	31
2012	155	77	69	56	85	64	110	67
2016	160	43	39	12	50	58	66	25
2018	160	72	46	50	91	68	87	65
2023	139	42	49	71	64	19	14	22

Table 2. Albania's rank over the years

Source: World Bank, https://lpi.worldbank.org/ and author's calculations. Note: The higher the rank, the weaker the indicator in relation to other countries in the group.

Significant improvements in the rankings were observed in 2018 and 2023 for the components "Custom" and "Infrastructure." Conversely, components 4, 5, and 6 showed substantial declines in the rankings. Analysing these trends indicates that infrastructure is a component that positively influences the performance of the LPI, particularly after 2016 in Albania.

III.3. LPI: Countries and regions of transport corridors

Apart from Albania, the comparative analysis includes Italy, Greece, Bulgaria, Croatia, Montenegro, and the Republic of North Macedonia, although data for Kosovo are missing. Over the years, Albania's Logistics Performance Index (LPI) has recorded only modest improvements, with an average score of 2.5. About rankings, Albania is consistently positioned last or second-to-last among the countries listed above (Figure 4). Italy, Greece, and Bulgaria have the highest LPI scores, while Croatia shows the most steady progress, moving closer to the levels of Italy and Greece.

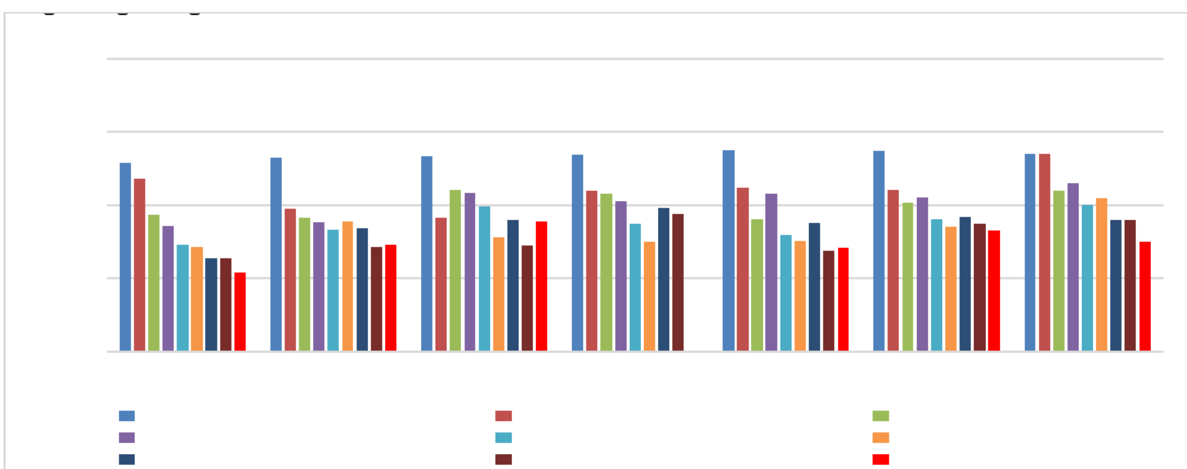


Figure 4. LPI performance over the years for Albania and other countries in comparison

Source: WB, https://lpi.worldbank.org/

Note: NA data for Albania in 2014.

Compared with the highest-performing nation in the study, Italy, Albania's LPI value represents, on average, 67.4% of Italy's. Furthermore, it is more closely aligned with Montenegro's LPI value (99%).

LPI /LPI Italy	2007	2010	2012	2014	2016	2018	2023	Albania Relative to other countries (mean)
Italy	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
Greece	0.94	0.81	0.77	0.87	0.86	0.86	1.00	0.78
Bulgaria	0.80	0.78	0.87	0.86	0.75	0.81	0.86	0.83
Croatia	0.76	0.76	0.86	0.83	0.84	0.83	0.89	0.82
Bosnia & Herzegovina.	0.69	0.73	0.81	0.74	0.69	0.75	0.81	0.90
North Macedonia	0.68	0.76	0.70	**0.68**	0.67	0.72	0.84	0.93
Serbia	0.64	0.74	0.76	0.80	0.74	0.76	0.76	0.92
Montenegro	0.64	**0.67**	**0.67**	0.78	**0.63**	0.73	0.76	0.99
Albania	**0.58**	**0.67**	0.76	us	0.64	**0.71**	**0.68**	**0.67**

Table 3. LPI Score Relative to Italy and Albania relative to other countries (mean).Source: WB, https://lpi.worldbank.org/ and author's calculations.

Note: The grey cells indicate the lowest LPI levels relative to the others.

The "Infrastructure component" (IC) has a significant impact on the LPI level. Historically, Albania recorded the lowest values of this metric in 2007, 2010, 2016, and 2018. A marked improvement is observed in 2023, when the IC surpasses those of Serbia, Montenegro, and Bosnia and Herzegovina (Figure 5).

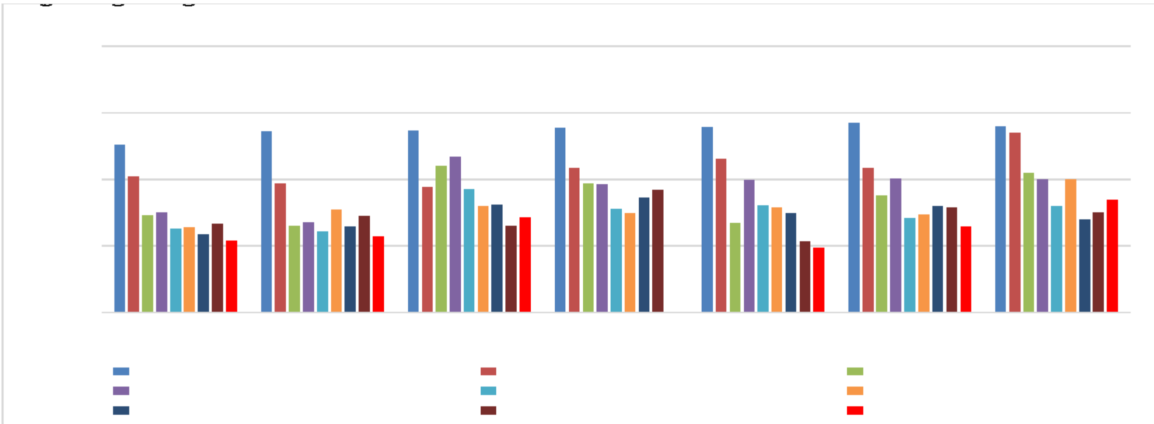


Figure 5. Performance of Infrastructure Component: Albania and other countries.

Source: WB, https://lpi.worldbank.org/;

Note: NA data for Albania in 2014.

Infrastructure Relative to Italy	**2007**	**2010**	**2012**	**2014**	**2016**	**2018**	**2023**	**Albania Relative to other countries (mean)**
Italy	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
Greece	0.87	0.79	0.77	0.84	0.87	0.82	0.97	0.72
Bulgaria	0.70	0.62	0.86	0.78	0.62	0.72	0.82	0.85
Croatia	0.71	0.63	0.90	0.77	0.79	0.78	0.79	0.80
Bosnia & Herzegovina.	0.64	0.60	0.77	0.68	0.69	0.63	0.68	0.91
North Macedonia	0.65	0.68	0.70	0.66	0.68	0.64	0.79	0.88
Serbia	0.62	0.62	0.70	**0.72**	0.66	0.67	0.63	0.94
Montenegro	0.66	0.66	0.62	0.75	0.55	0.67	**0.66**	0.96
Albania	**0.59**	**0.58**	**0.65**	us	**0.52**	**0.60**	0.71	**0.61**

Table 4. Infrastructure Score Relative to Italy and Albania relative to other countries (mean).

Source: WB, https://lpi.worldbank.org/ and author's calculations.

Note: The grey cells indicate the lowest levels in the comparison.

The IC value for Albania is, on average, 61% of Italy's, while it is closer to Montenegro's (96%).

In the following, a comparative analysis will be conducted, grouped by participating countries in regional corridors and beyond, as well as by other components of the LPI.

The Blue Corridor passes through Croatia, Montenegro, Albania, and Greece. Within this corridor, Albania's Logistics Performance Index (LPI) closely aligns with Montenegro's in 2016 and 2018. Compared to Greece, Albania's average standing is approximately three times lower.

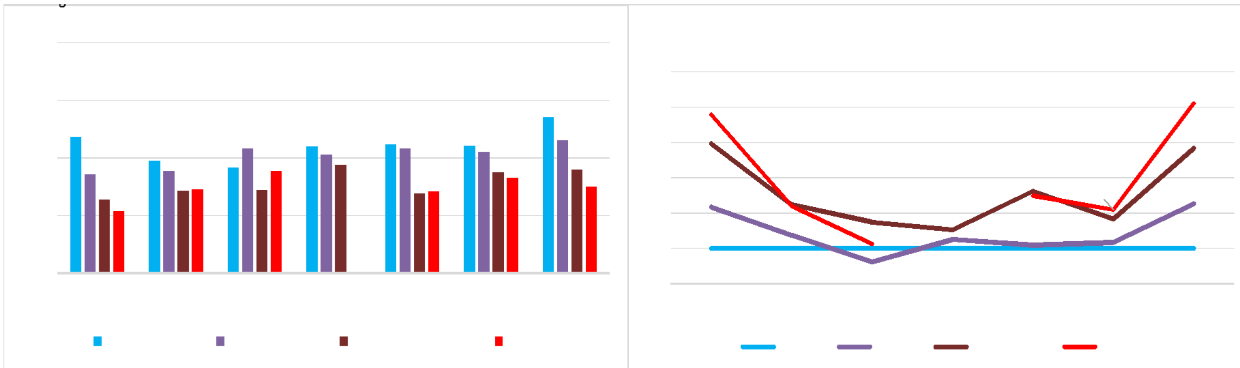


Figure 6. Performance of LPI and relative rank for countries of the Blue Corridor

Source: WB, https://lpi.worldbank.org/ and author's calculations.

By comparing the LPI and its components in 2023 with the average values for 2007–2023, it is evident that countries' performance has been affected differently by each component. In the case of Greece and Croatia, the strongest positive impacts came from International Shipments and Logistics Competence & Quality. The latter also had a positive, though smaller, effect in Montenegro. Meanwhile, Albania maintained its 2023 LPI score close to the period average, mainly due to improvements in Infrastructure (IC) (+0.4) and Customs (+0.2).

Country	Greece		Croatia		Montenegro		Albania	
	Mean 2007-2023	2023	Mean 2007-2023	2023	Mean 2007-2023	2023	Mean 2007-2023	2023
LPI	3.2	3.7	3.0	3.3	2.7	2.8	2,481	2.5
Customs **(1)**	2.9	3.2	2.9	3.0	2.4	2.6	2.2	2.4
Infrastructure **(2)**	3.2	3.7	2.9	3.0	2.4	2.5	2.3	2.7
International Shipments **(3)**	3.1	3.8	3.0	3.6	2.6	2.8	2.7	2.8
Logistics Competence & Quality **(4)**	3.1	3.8	3.0	3.4	2.5	2.8	2.4	2.3
Timeliness **(5)**	3.7	3.9	3.4	3.2	2.9	3.2	2.9	2.5
Tracking & Tracing **(6)**	3.4	3.9	3.0	3.4	2.6	3.2	2.3	2.3

Table 5. LPI and its components (mean) – Blue Corridor

Source: WB, https://lpi.worldbank.org/ and author's calculations

Albania needs to improve Logistics Competence & Quality and Timeliness, as there are significant gaps in these two components compared to Greece and Croatia, and to a lesser extent with Montenegro. In terms of International Shipments, 2023 saw a slight improvement over the average, with Albania performing at a similar level to Montenegro.

Corridor VIII extends from Italy, passing through Albania, continuing to the Republic of North Macedonia and Bulgaria. In these countries, Albania's logistics performance index (LPI) aligns more closely with North Macedonia's in 2016 and 2018. There is greater variability among the Corridor VIII countries than among those of the Blue Corridor. In terms of relative ranking compared to Italy, Albania ranks approximately 5 times lower on average, about 2 times lower than Bulgaria, and around 1.3 times lower than North Macedonia.

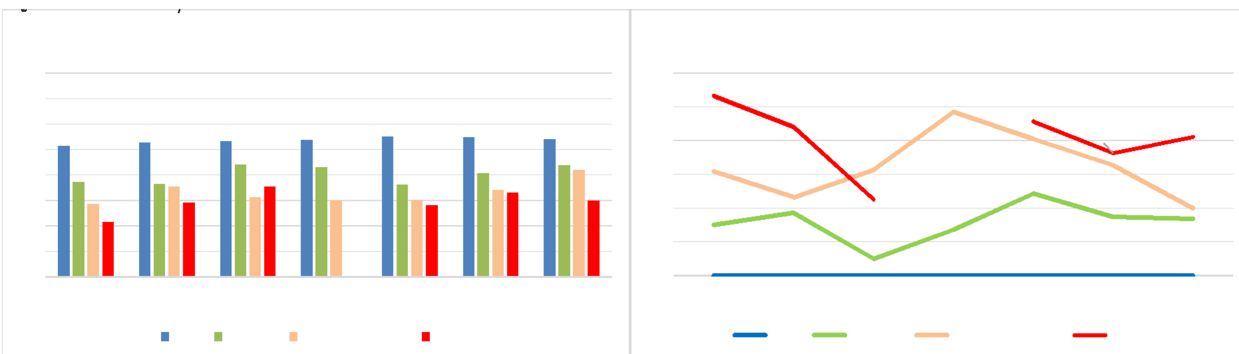


Figure 7. Performance of LPI and relative rank for some countries in Corridor VIII.

Source: WB, https://lpi.worldbank.org/ and author's calculations

By comparing the LPI level and its component indicators in 2023 with their respective averages for the period 2007–2023, for countries along Corridor VIII and the Blue Corridor, the aggregated performance indicator for each country has been influenced to varying degrees by specific components. Specifically in Italy, modest improvements were observed in components (2, 4, and 6), offset by declines in components (3 and 5). For Bulgaria, improvements were more concentrated and pronounced in components (1, 2, and 4). In North Macedonia, positive impacts were higher across almost all components, registering significant increases in 2023 compared to their respective averages, with gains ranging from 0.4 to 0.7 points. Meanwhile, Albania's performance was primarily influenced by improvements in Infrastructure (IC) and Customs.

Country	Italy		Bulgaria		North Macedonia		Albania	
	Mean 2007-2023	2023	Mean 2007-2023	2023	Mean 2007-2023	2023	Mean 2007-2023	2023

LPI	3.7	3.7	3.0	3.2	2.7	3.1	2,481	2.5
Customs **(1)**	3.4	3.4	2.7	3.1	2.4	3.1	2.2	2.4
Infrastructure **(2)**	3.7	3.8	2.7	3.1	2.6	3.0	2.3	2.7
International Shipments **(3)**	3.5	3.4	3.1	3.0	2.7	2.8	2.7	2.8
Logistics Competence & Quality **(4)**	3.7	3.8	3.0	3.3	2.7	3.2	2.4	2.3
Timeliness **(5)**	4.0	3.9	3.5	3.5	3.0	3.5	2.9	2.5
Tracking & Tracing **(6)**	3.8	3.9	3.0	3.3	2.6	3.2	2.3	2.3

Table 6. LPI and its components (mean) – Corridor VIII

Source: WB, https://lpi.worldbank.org/ and author's calculations

The comparative analysis suggests that Albania needs to make broader-based improvements to the components of the LPI, so that growth is driven not only by infrastructure but also by factors related to logistics competence and quality.

V. CONCLUSIONS

Transport corridors can be considered as key elements in fostering the development of economic, social, cultural, and knowledge sectors at both national and regional levels. Because of their importance, they are called economic corridors. This article shows that Albania's transport and infrastructure sector has grown thanks to the reconstruction and modernisation of national and international transport corridors. This growth has increased the sector's share in GDP and employment.

Analysing the trend of the Logistics Performance Index (LPI) and its related components for Albania reveals a slight improvement in the indicator. Compared with regional countries and participants in the Blue Corridor and Corridor VIII, Albania's LPI score is 2.5, ranking last or second-to-last among other countries. The countries with the highest LPI performance are Italy, Greece, and Bulgaria, while Croatia has demonstrated the most stable upward dynamics. In relative terms, compared with the highest-performing country (Italy), Albania's LPI value averages about 67.4% of Italy's score. Albania's score is much closer to that of Montenegro (99%) and North Macedonia (93%).

When comparing LPI scores and components in 2023 with the averages from 2007 to 2023, Greece and Croatia improved most in International Shipments and Logistics Competence & Quality. Bulgaria improved mainly in Infrastructure, Customs, and Logistics Competence & Quality. North

Macedonia showed improvements across nearly all components, with increases of 0.4 to 0.7 points. Meanwhile, Albania maintained its LPI value for 2023 close to the period average, primarily influenced by improvements in Infrastructure and Customs.

The data suggest that Albania needs broader-based improvements across all LPI components. Growth should be integrated and not rely solely on infrastructure and custom. Rapid improvements are recommended in Logistics Capacities, Competence, and Quality. In these areas, the regional countries participating in the studied corridors made significant progress in 2023, thereby strengthening their competitive advantages relative to Albania. Despite notable improvements in infrastructure, further accelerated development is essential to strengthen competitiveness and reduce the gap with more advanced countries such as Italy, Greece, Bulgaria, and Croatia.

REFERENCES

Armstrong, M., Taylor, J. (2000). *Regional Economics and Policy*. 3rd Edition. Wiley-Blackwell.

Arvis, J. F, Ulybina, D., and Wiederer, Ch. (2024). From Survey to Big Data The New Logistics Performance Index. World Bank Group, Macroeconomics, Trade and Investment Global Practice May 2024. Retrieved from https://documents1.worldbank.org/curated/en/099159405162478279/pdf/IDU1898e51581445914e9218cd110f8195f5a07f.pdf

Banister, D. & Berechman, Y. (2001). Transport Investment and the Promotion on Economic Growth. Journal of Transport Geography, 9, 209-218. Retrieved from https://doi.org/10.1016/S0966-6923(01)00013-8

Bruinsma, F., Rietveld, P. (1993). Urban Agglomerations in European Infrastructure Networks. Urban Studies, 1993, vol. 30, issue 6, 919-934. Retrieved from https://degree.ubvu.vu.nl/repec/vua/wpaper/pdf/19930008.pdf

Cohen, J.P. & Morrison Paul, C.J. (2008). Agglomeration, Productivity and Regional Growth. Production Theory Approaches. Retrieved from https://www.researchgate.net/publication/228661401_Agglomeration_productivity_and_regional_growth_Production_theory_approaches

Dercon, S., Gilligan, D.O., Hoddinott, J. (2008). The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages. American Journal of Agricultural Economics 91(4), November 2008. Retrieved from DOI:10.1111/j.1467-8276.2009.01325.x

European Bank for Reconstruction and Development. (2019). Transport Sector Strategy 2019-2024 (Approved by the Board of Directors on 30 October 2019).

European Commission. (2018). Untapped Potential: Intra-Regional Trade in the Western Balkans. Retrieved from https://economy-finance.ec.europa.eu/system/files/2018-05/dp080_western_balkans.pdf

Grover, A., Lall, S., Timmis, J. (2023). Agglomeration economies in developing countries: A meta-analysis. Regional Science and Urban Economics. Vol. 101, July 2023, 103901. Retrieved from https://doi.org/10.1016/j.regsciurbeco.2023.103901

Gwilliam, K. (2011). Africa's Transport Infrastructure: Mainstreaming Maintenance and Management. The World Bank. Retrieved from https://dx.doi.org/10.1596/978-0-8213-8456-5

Hallam, D. (2011). International investment in developing country agriculture – issues and challenges. Food Security Journal, 3 (Suppl. 1): 91–98.

Healey, P. (2004). The Treatment of Space and Place in the New Strategic Spatial Planning in Europe. International J. Urban and Regional Research, 28(1): 45–67. February.

INSTAT. (October 2024). Household Budget Survey, 2023. Retrieved from https://www.instat.gov.al/media/14109/hbs_2023_0310_eng.pdf

Keser, H. (2015). Importance of Transport Corridors in Regional Development: The Case of TRACECA. Sosyoekonomi. 2015, Vol. 23(24), 163-182. Retrieved from https://scispace.com/pdf/importance-of-transport-corridors-in-regional-development-4p3h8yq1ad.pdf

Krugman, P. (1991). Increasing Returns and Economic Geography. Journal of Political Economy, 99(3): 483–499. June.

Limão, N. and Venables, A.J. (2001). Infrastructure, Geographical Disadvantage, Transport Costs, and Trade. The World Bank Economic Review, 15, 451-479. Retrieved from https://doi.org/10.1093/wber/15.3.451

Muneeza, A., Dappe, M. H., Melecky, M., Goldblatt, R. (2019). Wider Economic Benefits of Transport Corridors Evidence from International Development Organizations. World Bank Group (South Asia Region). Policy Research Working Paper 9057. Retrieved from https://documents1.worldbank.org/curated/en/404351573674156323/pdf/Wider-Economic-Benefits-of-Transport-Corridors-Evidence-from-International-Development-Organizations.pdf

OECD (2021). Multi-dimensional Review of the Western Balkans - Assessing Opportunities and Constraints. Retrieved from https://www.oecd.org/en/publications/multi-dimensional-review-of-the-western-balkans_4d5cbc2a-en.html

Qaja, B. (2021). "Transport Corridors". Large scale planning for regional and national development Case study: Albania. PhD Thesis. Retrieved from https://iris.unife.it/handle/11392/2478842

Roberts, M., Melecky, M., Bougna, Th., Xu, Y (S). (2018). Transport Corridors and Their Wider Economic Benefits - A Critical Review of the Literature. World Bank Group (South Asia Region). Policy Research Working Paper, 8302. Retrieved from https://documents.worldbank.org/pt/publication/documents-reports/documentdetail/667501516199287820/transport-corridors-and-their-wider-economic-benefits-a-critical-review-of-the-literature

Rodrigue, J-P. & Notteboom, Th. (2024). The Geography of Transport Systems (Chapter 3). 6th Edition, New York: Routledge, ISBN 9781032380407. Retrieved from http://transportgeography.org/contents/chapter3/transportation-and-economic-development/

Rodrigue, J-P., Comtois, C., Slack, B. (2016). The Geography of Transport Systems. 4th ed. London, England: Routledge

World Bank. (2023). Connecting to Compete. Trade Logistics in the Global Economy. Appendix 5: The LPI Methodology (pg. 62).

Xhepa, S. (2000). Zhvillimi i korridoreve të transportit në Evropën Qendrore e Lindore dhe në Ballkan. Buletini Ekonomik, Vëllimi 3, Nr. 2, Banka e Shqipërisë.

05

THE ALBANIAN INNOVATION ECOSYSTEM: POLICIES, PARTNERSHIPS, AND THE FUTURE OF ENTREPRENEURSHIP

DOI: 10.37199/c41000305

Delina IBRAHIMAJ

University of Tirana (Tirana, Albania)
delinaibrahimaj@gmail.com
ORCID 0009-0002-1908-6935

Abstract

The Albanian innovation ecosystem is undergoing rapid transformation, marked by the emergence of a formalised and increasingly interconnected startup landscape. This paper provides a comprehensive mapping of the ecosystem and analyses the structural dynamics that underpin its growth. Drawing on qualitative analysis and multi-source data, the study explores the interplay between key actors—including startups, corporates, academia, consulting firms, investors, public institutions, and support structures such as incubators and accelerators.

Since 2020, Albania has registered over 2,000 startups, with an average annual growth rate of 30%. Government-backed initiatives have played a pivotal role, with approximately €8.5 million disbursed in seed funding through dedicated grant schemes. The paper examines how these players collaborate through strategic linkages and how national policies and programs—such as the Startup Law and innovation support mechanisms—have created a more enabling environment for entrepreneurship.

By examining emerging patterns of cooperation, institutional support, and policy innovation, this study offers new insights into the conditions necessary for scaling early-stage ecosystems. Particular attention is paid to recognition mechanisms, opportunities for regional integration, and the evolving roles of diaspora and academia. The findings contribute to broader discussions on innovation policy in small transitioning economies and highlight Albania's potential as a regional hub for entrepreneurship.

Keywords*:* Innovation, Government Policies, Entrepreneurship, Startups

I. INTRODUCTION

In recent years, Albania has embarked on a strategic transformation to position itself as a competitive, innovation-driven economy. This trajectory reflects a broader global shift, where the digital revolution is reconfiguring economic paradigms and emphasising the centrality of entrepreneurship, technological innovation, and knowledge creation as drivers of long-term growth and resilience (Mazzucato, 2013; Autio et al., 2018). Transitioning from traditional development models rooted in resource-based expansion, Albania is increasingly embracing a knowledge-based economy characterised by agility, creativity, and inter-sectoral collaboration (Porter, 1998; OECD, 2024).

The emergence of a dynamic startup and innovation ecosystem is both a symptom and a strategy of this transformation, underpinned by explicit governmental efforts to institutionalise innovation as a core pillar of economic policy. Since 2020, more than 2,000 startups have been formally registered, with an estimated annual growth rate of 30% (Startup Albania, 2024). These ventures span a diverse array of sectors—including fintech, agritech, creative industries, and green technologies—demonstrating the potential for inclusive, innovation-led development (European Commission, 2023).

This paper offers a comprehensive examination of the Albanian innovation ecosystem, mapping the evolving relationships among key stakeholders: startups, investors, corporates, academic institutions, public agencies, and support structures such as incubators and accelerators. Special emphasis is placed on the catalytic role of government-led initiatives—such as the Startup Law (2022), the Strategy for the Development of Innovative Entrepreneurship (SDIE), and targeted grant mechanisms—which have been instrumental in fostering collaboration, de-risking early-stage innovation, and integrating Albania into broader European innovation frameworks (Isenberg, 2010; Etzkowitz & Leydesdorff, 2000).

Through this analysis, the study identifies the strategic enablers that have contributed to the country's innovation momentum, the persistent structural barriers—such as funding gaps and institutional fragmentation—and the forward-looking policy directions necessary to scale and sustain ecosystem development. As Albania deepens its EU integration process, enhancing its innovation infrastructure will be critical not only for convergence with European standards but also for leveraging its demographic potential, digital transformation efforts, and diaspora networks to catalyse a globally connected entrepreneurial future (OECD, 2022; World Bank, 2024).

II. CONCEPTUAL FRAMEWORK

The concept of an innovation ecosystem has gained significant scholarly and policy attention over the past two decades as a framework for understanding how innovation emerges, evolves, and scales within complex, multi-actor environments (Autio et al., 2018; Jackson, 2011).

An innovation ecosystem is not limited to startups or isolated R&D institutions; rather, it encompasses a dynamic constellation of actors—such as firms, universities, government bodies, civil society organisations, investors, and intermediaries—whose structured interactions facilitate the creation, diffusion, and commercialisation of knowledge and technology (OECD, 2024; Isenberg, 2010).

The performance of such ecosystems depends not only on the effectiveness of individual participants but on the density, quality, and reciprocity of their relationships (Spigel, 2017). Effective ecosystems exhibit high levels of connectivity, collaborative trust, and institutional learning, which generate adaptive capacity and resilience in the face of shifting technological and market dynamics (Holling, 2001).

In this study, the Albanian innovation ecosystem is examined through the lens of systems theory and cluster-based development models (Porter, 1998; Etzkowitz & Leydesdorff, 2000). These theoretical perspectives underscore the systemic and co-evolutionary nature of innovation, highlighting features such as feedback loops, path dependency, and critical mass. The Albanian ecosystem, while still nascent, reveals characteristics of an evolving system shaped by policy interventions, economic liberalisation, social norms, and the country's trajectory toward EU integration (European Commission, 2023).

II.1 Policy and Institutional Infrastructure

A strong enabling environment—comprising legal, institutional, and financial instruments—is foundational to any innovation ecosystem (Mazzucato, 2013). In Albania, the post-2020 period has seen the formal institutionalisation of innovation as a policy priority. Key milestones include the adoption of the Startup Law (2022), which legally defines startups and introduces tax incentives, social security exemptions, and streamlined registration procedures. These are further reinforced by the Strategy for the Development of Innovative Entrepreneurship (SDIE) 2024–2030, which aligns national innovation priorities with EU strategies such as the Green Deal and Digital Agenda (Ministry of Economy, 2023).

The establishment of Startup Albania, under the Ministry of Entrepreneurship, marks a shift toward implementation-focused governance. As both a funding body and ecosystem coordinator, Startup Albania reflects principles of agile public management and network orchestration (World Bank, 2024). Complementary national frameworks—such as the Digital Agenda 2022–2026, the National Strategy for Scientific Research and Innovation 2023–2030, and the Business Investment Development Strategy (BIDS)—aim to mainstream innovation across multiple policy domains. However, coordination challenges persist (OECD, 2022).

Albania's participation in EU programs such as Horizon Europe, Erasmus+, and EU4Innovation has introduced benchmarking mechanisms and external funding, enabling the country to overcome some early-phase ecosystem constraints (European Commission, 2023; RCC, 2022).

II.2 Entrepreneurial and Startup Activity

Startups are widely recognised as catalysts of innovation, contributing to job creation, technological upgrading, and market diversification (Isenberg, 2010). Since 2020, Albania's entrepreneurial landscape has gained momentum, driven by a young population, increasing digital literacy, and targeted public support schemes. By 2024, over 540 startups and ecosystem enablers had received financial and technical support, with 68% operating in knowledge-intensive sectors such as fintech, ICT, agritech, and creative industries (Startup Albania, 2024).

Notably, the Diaspora Engagement Window within SDIE has facilitated the return of Albanian founders from abroad, injecting international networks, capital, and managerial experience into the ecosystem. Gender inclusivity is also improving: women comprised 43% of 2024 grant applicants, aligning with broader EU gender equity targets (OECD, 2024).

To address geographic disparities, the government plans to launch Regional Innovation and Technology Centres (RTICs) and expand Smart Labs across universities. These initiatives aim to democratise access to innovation infrastructure, especially outside Tirana, Albania's dominant tech hub.

II.3 Support Structures and Collaborative Networks

The "connective tissue" of innovation ecosystems consists of intermediaries such as incubators, accelerators, mentorship programs, and innovation festivals. In Albania, support structures remain fragmented but are evolving through both public and university-led platforms. Initiatives such as the U-Start Accelerator and the Metropolitan Incubator illustrate early models of cross-sectoral collaboration (INSTAT, 2023).

However, systemic collaboration remains limited. Technology Transfer Offices (TTOs) are still nascent or absent in public universities, and mechanisms for research commercialisation are underdeveloped. While corporate engagement is weak, examples like Raiffeisen Bank's fintech pilot programs and Vodafone Albania's digital skills initiatives suggest a growing appetite for open innovation models.

Importantly, Albania's involvement in regional innovation frameworks, such as the Smart Specialisation Strategy for the Western Balkans, positions the country within a transnational ecosystem, enabling it to compensate for local market constraints through knowledge and capital mobility (RCC, 2022).

II.4 System Dynamics: Feedback, Density, and Learning

Building on innovation systems theory, the Albanian ecosystem can be seen as an adaptive system characterised by iterative feedback and emergent learning. The relatively low incidence of university spin-offs, for example, reflects structural disconnects between academia, industry, and public institutions (Etzkowitz & Leydesdorff, 2000). High-performing ecosystems display institutional memory, critical mass, and feedback channels that drive continuous improvement (Spigel, 2017; Startup Genome, 2023).

Some of these elements are beginning to emerge in Albania. Startup Albania's Annual Report and its partnership with Startup Genome provide nascent evidence base for policy refinement. Albania's move from an "activation" to a "globalisation" phase—following the Startup Genome lifecycle model—will require investment in infrastructure, long-term governance continuity, and the development of risk-sharing finance mechanisms.

III. METHODOLOGY

This study adopts a qualitative research design grounded in exploratory case analysis and ecosystem mapping to capture the structure, dynamics, and policy dimensions of Albania's innovation ecosystem. The methodological approach is rooted in systems thinking and actor-network theory (Latour, 2005; Carlsson et al., 2002), enabling the examination of relationships among institutional actors, regulatory frameworks, and knowledge flows. This design is particularly well-suited to emerging innovation ecosystems where quantitative data may be limited, and contextual factors are key to understanding systemic evolution (Yin, 2014).

III.1 Data Sources and Collection

Data were collected from a combination of primary and secondary sources to ensure triangulation and enhance the credibility of findings (Denzin & Lincoln, 2011). These include:

- Government documents and strategies, such as the Startup Law (2022), the Strategy for the Development of Innovative Entrepreneurship (SDIE), and national development plans related to digital transformation and scientific research.
- Reports from international organisations, including the European Commission, GIZ, RISI Albania, and the World Bank, which offer policy assessments, benchmarking data, and strategic recommendations.
- Startup ecosystem databases and platforms, such as Startup Albania, YUNUS Albania, and Tirana Incubator, provide updated metrics on registered startups and support initiatives.

- Materials from innovation support organisations: including Coolab, TechSpace, ICT Hub, and Protik, which offer insights into incubator/accelerator performance, mentorship programs, and innovation events.

In addition, over a dozen semi-structured interviews and roundtable discussions were conducted with stakeholders, including startup founders, public officials, accelerator managers, and representatives from academia and civil society. Interview protocols focused on identifying perceived gaps, collaboration patterns, and institutional bottlenecks.

III.2 Ecosystem Mapping

The ecosystem was mapped by grouping actors into six broad categories, consistent with existing innovation ecosystem frameworks (Spigel, 2017; Stam, 2015):

- Public institutions (e.g., ministries, agencies, municipalities),
- Startups and entrepreneurs (the primary agents of innovation),
- Support organisations (e.g., incubators, accelerators, co-working spaces),
- Finance providers (banks, public funds, angel investors, and VCs),
- Universities and research centres, and
- Enablers (e.g., consultants, NGOs, and event organisers).

This mapping allowed the identification of structural gaps, resource flows, and network density, helping visualise how ideas, capital, and talent circulate within the ecosystem.

III.3 Program and Policy Analysis

To assess the effectiveness of Albania's innovation governance, the study examined key public programs, including:

- The Innovation Voucher Scheme is designed to subsidise R&D and technology adoption.
- The Startup Grant Scheme, which distributed over €8.5 million in seed funding between 2020 and 2023.
- The Entrepreneurship Support Fund is aimed at capacity-building and technical assistance.
- EU programs such as *Horizon Europe* and *Erasmus+* provide international benchmarking, capacity-building, and co-financing opportunities.

These programs were analysed with respect to their design, funding mechanisms, target groups, and alignment with broader innovation policy objectives (OECD, 2022; Mazzucato, 2013).

III.4 Performance Indicators

To evaluate progress and ecosystem maturity, the study employed performance indicators, including:

- Number of startups registered annually,
- Average annual growth rate of startup activity,
- Volume and source of funding (public vs. private),
- Number and type of innovation-related events and initiatives,
- Participation of Albanian startups in regional and European innovation programs.
- Where possible, indicators were contextualised using regional benchmarks and ecosystem lifecycle models (Startup Genome, 2023).

III.5 Limitations

This study does not employ econometric modelling or counterfactual analysis due to the limited availability of time-series financial data and startup-level metrics. Instead, it focuses on observable trends, actor interactions, and stakeholder perceptions. While this qualitative approach may not yield definitive causal relationships, it provides a robust, multi-perspective view of Albania's innovation landscape, especially relevant in early-stage ecosystems where institutional dynamics are still forming (Isenberg, 2010; Spigel & Harrison, 2018).

IV. DISCUSSIONS

The findings of this study affirm that Albania's innovation ecosystem has entered a formative yet increasingly structured phase, characterised by rising institutional awareness, expanding startup activity, and emerging multi-actor coordination mechanisms. While the ecosystem remains in an early stage of development compared to regional counterparts, it exhibits many of the foundational elements of more mature systems: dedicated legal frameworks, targeted public programs, and a nascent yet growing support infrastructure (OECD, 2024; European Commission, 2023).

From a policy standpoint, Albania's progress has been driven by the adoption of key instruments, such as the Startup Law (2022), the National Strategy for Science, Technology and Innovation, and funding programs, including the Innovation Voucher Scheme and the Startup Grant Scheme. These initiatives have helped embed innovation into the national economic narrative, positioning entrepreneurship as a mechanism for modernisation, employment generation, and competitiveness in both EU and global markets (Mazzucato, 2013; Isenberg, 2010).

Over €8.5 million has been disbursed in seed grants since 2020, supporting the registration of approximately 2,000 startups—a trend reflecting a 30% annual growth rate. However, as the ecosystem transitions from activation to consolidation, the challenges it faces become more nuanced: from startup formation to scaling, from policy intent to effective implementation, and from fragmented activity to systemic coordination (Startup Genome, 2023).

The analysis highlights several structural features that define the current configuration of Albania's innovation ecosystem:

- Fragmented but active support structures: Incubators, accelerators, and mentoring platforms such as *TechSpace*, *Coolab*, and the *Tirana Incubator* have proliferated, playing an essential role in early-stage venture development. However, most of these operate under short-term project funding, often donor-dependent, and lack institutional continuity. The absence of a national umbrella coordination platform limits ecosystem learning, reduces efficiency, and inhibits long-term scalability (Spigel, 2017).
- Growing public-sector engagement, but limited private investment: The ecosystem has been mainly state-driven, with minimal participation from venture capitalists, angel investors, or corporate investors. Financial flows remain grant-based and unstructured, echoing trends found in other transition economies (OECD, 2022). Without blended finance instruments or co-investment models, Albania risks stalling innovation at the prototyping stage (Murray et al., 2012).
- Underutilised role of academia: Albania's universities and research centres remain peripheral to the startup ecosystem. The lack of functional Technology Transfer Offices (TTOs), intellectual property support, and university-based incubators hampers the commercialisation of research and reduces the ecosystem's absorptive capacity (Etzkowitz & Leydesdorff, 2000; World Bank, 2024).
- Policy momentum meets implementation gaps: Despite strong strategic intent, implementation remains inconsistent. Stakeholder interviews highlighted delays in fund disbursement, limited technical assistance, and regulatory ambiguities that create uncertainty for ecosystem actors. Monitoring and evaluation frameworks are underdeveloped, constraining policy responsiveness and learning (Drahokoupil & Myant, 2019).

Nonetheless, Albania holds distinct comparative advantages: a young, digitally literate population, growing diaspora engagement, improved digital infrastructure, and access to European cooperation platforms such as Horizon Europe and Erasmus+. These can serve as launchpads for scaling ecosystem maturity if properly leveraged (European Commission, 2023; RCC, 2022).

V. CONCLUSIONS AND RECOMMENDATIONS

Albania's innovation ecosystem stands at a strategic inflexion point. It has moved beyond sporadic initiatives to a phase of structural consolidation, supported by enabling policies and increasing entrepreneurial activity. However, sustaining this momentum will require more than policy articulation—it demands institutional coherence, financial system diversification, and cross-sectoral governance mechanisms.

If Albania successfully navigates this transition, it could position itself as a regional innovation hub—particularly in high-potential sectors such as agritech, creative industries, digital services, and green technology. By embracing systemic innovation thinking and fostering a culture of experimentation and collaboration, Albania can not only catch up to regional peers but potentially leapfrog into global innovation networks.

To move forward, the study identifies five strategic recommendations:

- Consolidate ecosystem governance: Establish a national innovation coordination body or task force to improve alignment across ministries, local government, academia, and the private sector. This entity should oversee the implementation of the Startup Law, maintain a live ecosystem database, and regularly convene stakeholders.
- Mobilise private capital through risk-sharing instruments: Develop blended finance mechanisms—including seed equity co-investment funds, convertible grants, and public-private venture partnerships. Albania can draw lessons from EU member states that have successfully used government-backed financial instruments to catalyse private sector participation (Murray et al., 2012; EIF, 2023).
- Strengthen startup support infrastructure: Transition incubators and accelerators from donor-funded projects to institutionalised programs with performance-based funding models. Forge partnerships with international networks like EIT Digital and European Business Network (EBN) to enhance service quality and cross-border exposure.
- Deepen academia - entrepreneurship linkages: Incentivise university-industry collaboration through competitive grants, joint research commercialisation, and student startup funds. Establish TTOs and IP protection offices to facilitate spinouts and patent activity, crucial for innovation capacity building (Etzkowitz & Zhou, 2017).
- Improve data and monitoring systems: Design and implement a national startup and innovation dashboard to track real-time indicators, including funding flows, firm survival rates, employment generation, and innovation outputs. This data will be critical for evidence-based policymaking and benchmarking against regional peers (Startup Genome, 2023).

REFERENCES

Autio, E., Nambisan, S., Thomas, L. D. W., & Wright, M. (2018). Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal, 12*(1), 72–95.

Balkan Barometer. (2022). *Balkan Barometer – Business opinion 2022*. http://www.rcc.int/download/docs/Balkan%20Barometer%202022%20-%20BO.pdf/56acb2cb729b5f1a74308ea7052bda10.pdf

Carlsson, B., Jacobsson, S., Holmén, M., & Rickne, A. (2002). Innovation systems: Analytical and methodological issues. *Research Policy, 31*(2), 233–245.

Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research* (4th ed.). SAGE Publications.

Drahokoupil, J., & Myant, M. (2019). The political economy of state-driven innovation in Eastern Europe. *Journal of Innovation & Development, 9*(2), 121–139.

European Investment Fund. (2023). *The role of public financial institutions in fostering startup ecosystems in Europe*.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From national systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research Policy, 29*(2), 109–123.

Etzkowitz, H., & Zhou, C. (2017). *The Triple Helix: University–industry–government innovation and entrepreneurship*. Routledge.

European Commission. (2023). *Commission staff working document: Albania 2023 report*. https://neighbourhood-enlargement.ec.europa.eu

European Commission. (2023). *European Innovation Scoreboard: Albania country profile*. https://ec.europa.eu/assets/rtd/eis/2023/ec_rtd_eis-country-profileal.pdf

Holling, C. S. (2001). Understanding the complexity of economic, ecological, and social systems. *Ecosystems, 4*(5), 390–405.

INSTAT. (2023). *Statistics on small and medium enterprises, 2023*. https://www.instat.gov.al/media/11356/results-on-sme-2021.pdf

INSTAT. (2023). *Tourism statistics*. https://www.instat.gov.al/en/themes/industry-trade-and-services/tourism-statistics

INSTAT. (2024). *Statistics on small and medium enterprises, 2022*. Tirana.

Isenberg, D. J. (2010). How to start an entrepreneurial revolution. *Harvard Business Review, 88*(6), 40–50.

Jackson, D. J. (2011). What is an innovation ecosystem? National Science Foundation.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network theory*. Oxford University Press.

Ligj Nr. 43/2022, Për zhvillimin e ndërmarrjeve mikro, të vogla e të mesme. (2024, October 11). https://aida.gov.al/wp-content/uploads/2024/02/ligj-nr.-43-dt.-21.4.2022.pdf

Mazzucato, M. (2013). *The entrepreneurial state: Debunking public vs. private sector myths*. Anthem Press.

Ministry of Economy, Culture and Innovation. (2023). *Business development and investment strategy 2021–2027*. https://www.investment.com.al/wp-content/uploads/2021/07/EN_Draft-BIDS-Presentation.pdf

Murray, G., Cowling, M., Liu, W., & Kalinowska-Beszczyńska, O. (2012). *A study of the impact of EIF support to the venture capital industry*. European Investment Fund Research Series.

OECD. (2022). *OECD tourism trends and policies 2022*. OECD Publishing. https://doi.org/10.1787/a8dd3019-en

OECD. (2022). *SME policy index: Western Balkans and Turkey*. https://www.oecd-ilibrary.org/development/sme-policy-index-western-balkans-and-turkey-2022_b47d15f0-en

OECD. (2024). *Western Balkans competitiveness outlook 2024: Albania*. OECD Publishing. https://doi.org/10.1787/541ec4e7-en

Porter, M. E. (1998). Clusters and the new economics of competition. *Harvard Business Review, 76*(6), 77–90.

Regional Cooperation Council. (2022). *Smart specialisation strategy for the Western Balkans*.

Spigel, B. (2017). The relational organisation of entrepreneurial ecosystems. *Entrepreneurship Theory and Practice, 41*(1), 49–72.

Spigel, B., & Harrison, R. (2018). Toward a process theory of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal, 12*(1), 151–168.

Stam, E. (2015). Entrepreneurial ecosystems and regional policy: A sympathetic critique. *European Planning Studies, 23*(9), 1759–1769.

Startup Albania. (2024). *Annual ecosystem report*. (Internal government source referenced in the paper.)

Startup Genome. (2023). *Global Startup Ecosystem Report 2023*. https://startupgenome.com

State Minister for Entrepreneurship and Business Climate. (2025). https://www.sipermarrja.gov.al/strategjite/

World Bank. (2024). *World development indicators*. DataBank. https://databank.worldbank.org/source/world-development-indicators

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). SAGE Publications.

06

THE SIX-HOUR WORKDAY: LITERATURE AND CASES ON PRODUCTIVITY, WELL-BEING, AND ECONOMIC IMPLICATIONS

DOI: 10.37199/c41000306

Xhimi HYSA

POLIS University (Tirana, Albania)
xhimi_hysa@universitetipolis.edu.al
ORCID 0000-0002-2279-6690

Shefqet SUPARAKU

POLIS University (Tirana, Albania)
shefqet_suparaku@universitetipolis.edu.al
ORCID 0009-0000-4268-5878

Abstract

The traditional eight-hour workday has been the cornerstone of industrial and post-industrial economies for over a century. Yet, experiments in recent times, especially in Sweden, have challenged this norm and brought in a six-hour workday. This article conducts a systematic literature review to explore the implications of a six-hour workday on productivity, employee well-being, and economic outcomes. In light of empirical evidence from Sweden and other countries, as well as theoretical frameworks such as Pareto's Principle and Parkinson's Law, this study examines whether shorter workdays can indeed lead to greater productivity, better work-life balance, and lower health care costs. However, these findings suggest that although the six-hour workday has potential, prevailing conditions inside organisations' settings and public attitudes greatly affect its success.

Keywords: Six-Hour Workday, Pareto Principle, Productivity, Work-Life Balance, Employee Well-Being

I. INTRODUCTION

The eight-hour workday was conceptualised during the Industrial Revolution to balance work demands with fundamental human rights. But as economies have progressed, so have the requirements placed on workers. The knowledge-based industries' boom, along with the impact of technology, has caused increased stress and burnout, as well as a growing demand for work-life balance. This has led some countries, most prominently Sweden, to experiment with shorter workdays, namely six-hour workdays, to address these challenges.

We provide a systematic review of the literature on the six-hour workday, examining its effects on productivity, employee well-being, and economic outcomes. The review will analyse data on six-hour workdays in Sweden, where several organisations have adopted them, and examine how they benefit workplace productivity, along with Pareto's Principle and Parkinson's Law. Finally, the paper will consider the potential challenges and limitations of implementing shorter workdays across different cultural and economic contexts.

II. THEORETICAL FRAMEWORK

II.1 Pareto's Principle and Productivity

The 80/20 rule, also called Pareto's Principle, states that 80% of outcomes come from 20% of the inputs. This concept, when applied to work, tells us that you can achieve a great deal of productivity with just a small amount of effort. Employees can complete 80% of their daily tasks in the first few hours of work, but with diminishing returns as the day progresses. It had been applied to production processes in the past (Juran, 1951), where Joseph Juran made beneficial use of the principle to gain efficiency.

Pareto's Principle, or the 80/20 rule, indicates that, in the context of a six-hour workday, employees must be motivated and dedicated to achieve maximum productivity in a comparably short time (or vicinity). This was similar to the results from Sweden: working six-hour days increased productivity and job satisfaction (The Guardian, 2015). Another study (Brynjolfsson & McAfee, 2014) has shown that technology can increase productivity and allow employees to do more work in less time.

II.2 Parkinson's Law and Time Management

According to Parkinson's Law, "work expands to fill the time available for its completion" (Parkinson, 1955). It means that if you give workers more time to do it, they will take more time, even if they can do it in less time. In contrast, employees work best and get things done when they are pressed for time.

A six-hour workday can be viewed as an implementation of Parkinson's Law. This squeeze of available time forces employees to prioritise work and increase their efficiency. This has already been seen in Swedish companies with shorter workdays, where employees report completing the same amount of work in a new six-hour workday as in an eight-hour one (The New York Times, 2016). Also, several quantitative studies on time management (e.g., Gajendran & Harrison, 2007) demonstrate that when given the opportunity to have flexible work arrangements through shorter workdays, workers tend to increase their attention span and work efficiency.

III. EMPIRICAL EVIDENCE FROM SWEDEN

III.1 Case Study: Svartedalens Elderly Care Home

One of the most conspicuous experiments with the six-hour workday happened at the Svartedalens elderly care home in Gothenburg, Sweden. With the goal of reducing employee stress and achieving work-life balance, the experiment shortened the workday from 8 to 6 hours. The results were promising, as employees felt less stressed and more invigorated, and the quality of care to residents improved (Bloomberg, 2016).

The experiment also revealed a paradox: employees managed to maintain or even increase their productivity despite working fewer hours. Because of this, one way to go about resolving this problem was to apply the Pareto Principle, because employees were able to be productive in less time. The success of the experiment has also generated political interest in the six-hour workday, with some policymakers pushing for its broader adoption (The Guardian, 2015). The study is consistent with findings about work intensification, which indicate that workdays that are shorter reduce burnout and improve general job performance (Kelliher & Anderson, 2010).

III.2 Case Study: Toyota Service Center

A prominent example from 2003 is a Toyota car service dealership in Sweden, which switched to a six-hour workday to relieve workers' stress and customers' long waiting periods. The new system extended the hours of the garages and led to heightened customer satisfaction and profitability. Workers said the reduced hours helped them stay more energetic and focused, resulting in better performance (The New York Times, 2016).

Toyota service centre's success indicates shorter workdays' benefits in physically demanding jobs. Working shorter hours ensured employees stayed fresh and worked at peak efficiency — beneficial to both the company and its clients. This is consistent with research by Sparks et al. (1997), which shows that working less can improve both physical and mental health, thereby increasing productivity.

III.3 Case Study: Internet Search Optimisation Start-Up

Maria Brath, who founded an online search optimisation start-up in Stockholm, instituted a six-hour workday from the start. While they worried at first that they would have to bring on more people for any given project, they discovered that the shorter timeframe made people more productive. Consequently, revenue and profit doubled year on year, proving short working hours can be associated with high productivity and a successful business (The New York Times, 2016).

This case study demonstrates that the six-hour workday is particularly useful in knowledge-based industries, where productivity is often determined by focus and creativity rather than a mere number of hours worked. This also resulted in an efficient workforce who had a better work-life balance, thanks to the compressed working schedules. Warr (1990) highlighted mental well-being as necessary for optimal job performance, which supports the case findings.

IV. THE PARADOX OF WORKING LESS AND PRODUCING MORE

Empirically, Sweden has proven that shorter workdays can increase productivity, commonly dubbed as the "paradox effect". Various factors contribute to this impact, including increased employee motivation, effective time management, and lower stress levels.

The paradox effect has many contributing factors, but one of the most important is increased employee motivation. Extended time off allows employees to return to work more energised and motivated. In contrast, Gabrielle Tikman, a surgical nurse, found that her energy increased, and her productivity soared when she transitioned to a six-hour workday (The New York Times, 2016). According to Pfeffer's (2018) research, worker well-being is closely associated with motivation and productivity.

A six-hour workday is also conducive to better time management. Moreover, with reduced time, the employees are required to be selective and dedicate themselves to what matters most. This principle aligns with Parkinson's Law, which holds that work expands to fit the time available for its completion. This increases the likelihood that, given fewer hours, employees will work effectively and will not waste time on non-work-related activities (Parkinson, 1955). The evidence from Gajendran and Harrison (2007) suggests that flexible work arrangements, or control over their working conditions, allow employees to manage their time and lead to employees being more productive.

In terms of pressure and anxiety, less stress (or eustress – positive stress) can, in fact, mean greater productivity. Excessive stress (or distress) can cause burnout, job dissatisfaction, and reduce productivity. Since reduced worker hours go a long way toward ensuring proper work-life balance, they make workers healthier (The Guardian, 2015). Research by Sparks et al. (1997) demonstrated that stress levels are significantly diminished when work hours are shortened, resulting in increased job performance.

V. CHALLENGES AND IMPLICATIONS

The six-hour workday looks promising, but it also has certain difficulties and limitations. One of the highlighted concerns is the potential burden on employees required to deliver tasks in a shorter time. This may cause increased stress levels, especially when the employees' sense that they are being asked to do more in less time (Times of India, 2016). Schor's (1991) research has suggested that work intensification could be experienced with shorter workdays, which could counter a portion of the benefits.

There is also the issue of how this will affect businesses, especially those which depend on long hours to fulfil customer needs. By contrast, sectors like healthcare and retail could see elevated costs associated with staffing with fewer hours worked over the same timeframe (CNBC, 2016). This is in line with the findings by Kelliher and Anderson (2010), highlighting that matters such as increased operational expenses for employers prevent flexible work arrangements from optimal sustainability.

Additionally, the six-hour workday may succeed or fail based on cultural and societal norms. In countries with a strong work culture, like the United States, shorter workdays (or even the idea of shorter workdays) would likely meet resistance because they are often framed as lazy or a lack of initiative (Fast Company, 2016). As Brynjolfsson and McAfee (2014) point out, technology and organisation need to mirror and be aligned with shared values before they can succeed and thrive.

The six-hour workday has all sorts of economic ramifications. Reduced workdays (and, consequently, reduced work hours) may lead to more output in the long run, as well as less overhead and happier employees, leading to a lower churn rate and more hours produced per person employed. Conversely, they may bring additional staff costs and possible business interruptions.

As previously mentioned, a six-hour workday will provide an opportunity to enhance efficiency through time management and the prevention of stress. Because productivity is often a direct correlation to profitability for businesses, that means higher levels of output and profitability and improved job satisfaction for employees (The New York Times, 2016). Brynjolfsson and McAfee (2014) make this exact research argument: it is not productivity or labour, rather you can achieve productivity gains through improved organisation of work and well-being of employees.

In countries with socialised healthcare systems, shorter workdays may also reduce the amount spent on healthcare costs. Less stress and improved work-life balance may reduce health-related issues among employees, leading to less individual and public healthcare costs (Science Alert, 2016). Sparks et al. found that this statement is true (1997) that reducing work hours can bring about substantial improvements in employees' health, with consequent reductions in healthcare costs.

VI. CONCLUSION

The six-hour workday represents a significant departure from the traditional eight-hour workday that has dominated industrial and post-industrial economies for over a century. While the empirical evidence from Sweden suggests that shorter workdays can lead to higher productivity, improved employee well-being, and reduced healthcare costs, the success of this approach depends on various factors, including employee motivation, organisational culture, and societal norms.

Relying on Pareto's Principle of 20-80 and its implementation in production processes by Juran, it can be argued that it is physiologically optimal for employees to expend 20% of their energy to achieve 80% of the results. In other words, the day-night cycle lasts 24 hours, meaning that working 6 hours per day equates to 6/24, or 25%, of the day. According to Pareto, Juran, and others (e.g., Parkinson), 20% commitment leads to 80% productivity. Thus, 25% commitment could theoretically lead to 100% productivity. While this is the general logic behind the phenomenon, it is important to consider other variables, such as employee qualifications, motivation, and organisational support, to ensure that this linear causation holds true in practice.

As the global economy continues to evolve, the debate over the optimal length of the workday is likely to persist. While the six-hour workday may not be suitable for all industries or countries, it offers a promising alternative to the traditional eight-hour workday, particularly in knowledge-based industries where productivity is more dependent on focus and creativity than on the number of hours worked.

Future research should focus on the long-term implications of the six-hour workday across different cultural and economic contexts. Additionally, further studies are needed to examine the potential challenges and limitations of shorter workdays, particularly in industries that require continuous coverage or experience high customer demand.

REFERENCES

Bloomberg. (2016). The six-hour workday works in Europe. What about America? Retrieved from http://www.bloomberg.com/news/articles/2016-05-10/the-six-hour-workday-works-in-europe-what-about-america

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W.W. Norton & Company.

CNBC. (2016). Six-hour workday good for Sweden, not necessarily the US, expert says. Retrieved from http://www.cnbc.com/2016/05/27/six-hour-workday-good-for-sweden-not-necessarily-the-us-expert.html

Fast Company. (2016). Could the US ever adopt a six-hour workday? Retrieved from http://www.fastcompany.com/3051366/the-future-of-work/could-the-us-ever-adopt-a-six-hour-workday

Gajendran, R. S., & Harrison, D. A. (2007). The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *Journal of Applied Psychology*, 92(6), 1524-1541. https://doi.org/10.1037/0021-9010.92.6.1524

Juran, J. M. (1951). *Quality Control Handbook*. New York: McGraw-Hill.

Kelliher, C., & Anderson, D. (2010). Doing more with less? Flexible working practices and the intensification of work. *Human Relations*, 63(1), 83-106. https://doi.org/10.1177/0018726709349199

Parkinson, C. N. (1955). Parkinson's Law. *The Economist*.

Pfeffer, J. (2018). *Dying for a Paycheck: How Modern Management Harms Employee Health and Company Performance—and What We Can Do About It*. Harper Business.

Schor, J. B. (1991). *The Overworked American: The Unexpected Decline of Leisure*. Basic Books.

Science Alert. (2016). Sweden is shifting to a 6-hour workday. Retrieved from http://www.sciencealert.com/sweden-is-shifting-to-a-6-hour-workday

Sparks, K., Cooper, C., Fried, Y., & Shirom, A. (1997). The effects of hours of work on health: A meta-analytic review. *Journal of Occupational and Organizational Psychology*, 70(4), 391-408. https://doi.org/10.1111/j.2044-8325.1997.tb00656.x

The Guardian. (2015). Efficiency up, turnover down: Sweden experiments with six-hour working day. Retrieved from https://www.theguardian.com/world/2015/sep/17/efficiency-up-turnover-down-sweden-experiments-with-six-hour-working-day

The New York Times. (2016). In Sweden, an experiment turns shorter workdays into bigger gains. Retrieved from http://www.nytimes.com/2016/05/21/business/international/in-sweden-an-experiment-turns-shorter-workdays-into-bigger-gains.html?_r=0

Times of India. (2016). The truth about Sweden's six-hour work day. Retrieved from http://timesofindia.indiatimes.com/business/international-business/The-truth-about-Swedens-six-hour-work-day/articleshow/50492820.cms

Warr, P. (1990). The measurement of well-being and other aspects of mental health. *Journal of Occupational Psychology*, 63(3), 193-210. https://doi.org/10.1111/j.2044-8325.1990.tb00521.x

07

ETHICAL ISSUES IN ARTIFICIAL INTELLIGENCE

DOI: 10.37199/c41000307

Blerta MJEDA

POLIS University (Tirana, Albania)
blerta_mjeda@universitetipolis.edu.al
ORCID 0000-0001-7767-5741

Elsa TOSKA

POLIS University (Tirana, Albania)
elsa_toska@universitetipolis.edu.al

Abstract

Artificial Intelligence (AI) is no longer just a futuristic concept—it has already become part of our daily lives. From personalised recommendations on streaming platforms to tools used in hospitals or banks, AI systems are changing how we work, learn, and even make decisions. While its capabilities to simulate human thinking—like learning from data or recognising patterns—are impressive, they also raise complex questions. As this technology develops at an extraordinary pace, we are beginning to face serious ethical and social dilemmas. How is our personal data being used? Can AI systems be truly fair, or are they reinforcing hidden biases? Moreover, what happens to traditional jobs when machines begin to outperform humans in certain tasks?

This paper examines the positive and negative effects of AI on society. It explores how we can benefit from these innovations while also addressing the risks they pose. Most importantly, it argues for the urgent need to create strong ethical and legal guidelines—developed through global cooperation—so that the use of AI remains fair, transparent, and in line with human values. However, due to the complexity of AI systems, few people fully grasp how they function.

Legal practitioners need to understand the basics of AI systems to address the legal challenges that can arise from their use. This paper will guide us through the fundamental concepts of AI from a non-technical perspective, helping you understand how such systems are created and their practical applications in our daily lives. Developments related to AI are widely mediatised. They promise to deliver "better than human" performance across a wide range of tasks, along with new tools to simplify our lives. However, what lies behind such systems, how do they work, and how are we protected from any harm that can be caused by them or the decisions that they influence?

Keywords: Algorithmic Discrimination, Artificial Intelligence, Ethical Issues, Data Privacy, Legal Challenges.

I. INTRODUCTION

AI is the field dedicated to understanding and constructing "intelligent" systems. However, the scope of systems classified as AI is broad, leading to ongoing debate over their true nature. Companies may claim to use AI to attract investment, researchers may employ it to secure funding, and the government may resort to it for behaviour monitoring purposes.

Given that AI draws on fields such as statistics and computer science, there is often ambiguity about its definition. Nonetheless, (real) AI-based systems have become prevalent in our daily lives. While their presence may not always be apparent, they play crucial roles in decision-making, information filtering, system control, and even creative tasks.

AI systems are technologies capable of processing data and information in ways that mimic intelligent behaviour. These systems typically involve components such as reasoning, learning, perception, prediction, planning, or control (Russell & Norvig, 2021). This understanding centres around three key aspects:

- AI systems are data-driven technologies that use models and algorithms to learn and perform cognitive functions. This enables them to make predictions or decisions in both physical and digital settings (European Commission, 2021; Goodfellow et al., 2016). They are built to operate at different levels of autonomy, utilising techniques such as knowledge modelling, representation, data analysis, and correlation detection (Floridi & Cowls, 2019).
- AI systems can rely on various methods, including (but not limited to) machine learning approaches such as deep learning and reinforcement learning (Goodfellow et al., 2016; Sutton & Barto, 2018).
- Machine reasoning, including planning, scheduling, knowledge representation and reasoning, search, and optimisation (Russell & Norvig, 2021; Brachman & Levesque, 2004).

AI technologies are increasingly integrated into cyber-physical systems, such as the Internet of Things (IoT), robotics, social robots, and human-computer interaction tools. These applications combine control mechanisms, perception capabilities, sensor data processing, and actuator control to enable AI systems to interact with and respond to their physical environments (Domingos, 2015; Boden, 2016).

- Ethical considerations related to AI extend across the entire lifecycle of an AI system. This includes every phase—from initial research, design, and development to deployment, daily use, and eventual decommissioning. It also encompasses tasks such as maintenance, financing, trading, monitoring, evaluation, validation, and dismantling or termination of the system. The term "AI actors" refers to any individuals or organisations involved in any phase of this lifecycle. These can include natural persons (e.g., researchers, developers, end-users) as well as legal entities such as companies, academic institutions, and public or private organisations (Jobin, Ienca, & Vayena, 2019; IEEE, 2019).

- AI systems raise new ethical issues, including: However, they are not limited to, their impact on decision-making, employment and labour, social interaction, healthcare, education, media, access to information, digital divide, personal data and consumer protection, environment, democracy, rule of law, security and policing, dual-use applications, and human rights and fundamental freedoms, such as freedom of expression, privacy, and non-discrimination (Floridi et al., 2018). Some of the challenges associated with AI stem from its ability to take on tasks once thought exclusive to living beings, sometimes even uniquely human. This capability positions AI systems in a transformative role within human activities and society at large. As a result, they are reshaping how people interact with their surroundings and the environment. This shift also creates a new reality in which children and young people are growing up—a world in which they form their understanding of themselves and society, engage with media and information more critically, and learn to make thoughtful, informed decisions.
- The use of Artificial Intelligence has had a significant impact across many industries, improving efficiency and performance (Brynjolfsson & McAfee, 2017; Davenport & Ronanki, 2018). However, this technological advancement has raised major ethical questions. Issues such as the inclusion of personal data, algorithmic discrimination, employment impacts, and automated decision-making assistance have raised serious concerns about the use of AI.

Since the terms deep learning and machine learning are often used interchangeably, it is important to distinguish between them. Deep learning is a specialised area within machine learning, which itself is a branch of artificial intelligence. We have seen that AI comprises several fields, unified by machine learning.

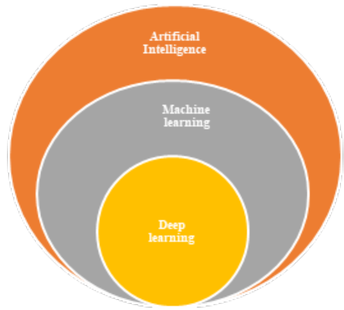


Figure 1. Key differences

Source: Author's processing

Traditionally, when programming a computer, an engineer writes a program to accomplish a specific task, such as unlocking a smartphone when the correct code is entered. Similar to how an engineer constructs a road to facilitate travel between cities, the goal is clear, and the route is easily defined (when the code is correct, unlock the phone).

The algorithm is simply a set of instructions that the computer must follow. In machine learning, these instructions produce a model, which is a component of an AI system capable of processing new data. The algorithm dictates how the model should be refined to perform well at the task, a process known as training.

Deep learning is a term frequently associated with AI. In reality, it is an approach to machine learning that loosely mimics the workings of the human brain. This is accomplished by training neural networks to solve a given task.

Deep Learning revolutionised AI since its widespread introduction in the mid-2000s, and it is responsible for many of the current generation of AI systems and new applications, e.g., self-driving cars, cancer screening, energy forecasting, etc.

II. DISCUSSIONS

Ethical Issues in AI (Artificial Intelligence) are a significant concern today because AI is being integrated into almost every part of life, from healthcare to finance, education to national security. Some of the most important ethical issues in AI include:

As artificial intelligence (AI) continues to evolve and integrate into everyday life, it brings with it a range of complex ethical and societal challenges. One of the most pressing concerns is privacy and data protection. AI systems, particularly those involving facial recognition, social media monitoring, and real-time analytics, require vast amounts of personal data to operate effectively. However, the way this data is collected, stored, and used often raises serious privacy concerns. As Zuboff (2019) points out, the way personal data is collected today—often without real, informed consent—raises serious concerns about privacy and the potential for misuse. This is not just a technical issue; it is a deeply human one. If AI systems are to handle sensitive information, they must follow strict data protection standards and, more importantly, give individuals control over their own data.

Closely tied to privacy is the problem of algorithmic bias. Harini and Guttag (2019) have shown that machine learning models can unintentionally reinforce stereotypes about race, gender, or economic status. These biases are not just unfair—they can cause real harm in areas such as hiring, law enforcement, and healthcare. O'Neil (2016) adds that without proper oversight and transparency, these systems risk perpetuating existing inequalities. She advocates for clear regulations that ensure fairness and accountability.

Another growing concern is how AI is reshaping the job market. As automation becomes more common, many traditional roles are being replaced by machines or algorithms. While AI can certainly improve efficiency and boost the economy, it also puts millions of jobs at risk. Brynjolfsson and McAfee (2014) warn that if we do not act proactively, this shift could deepen the divide between rich and poor. That is why it is so important to create policies that support workers during this transition, especially those in vulnerable sectors.

Then there is the complex question of responsibility. When an AI system makes a harmful decision, who should be held accountable—the developer, the user, or the system itself? This is not always easy to answer, especially with today's highly complex "black box" models. Citron (2007) explains that the lack of transparency in many AI systems makes it difficult to understand how decisions are made, let alone who should be blamed when things go wrong. This highlights the urgent need for more explainable and transparent AI designs.

Finally, as AI systems become more autonomous and capable, society will eventually have to grapple with more profound philosophical questions. Can an AI system ever be held morally responsible for its actions? Does it have any ethical standing? These are not just academic debates—they are real challenges we will face as technology continues to evolve.

Scholars such as Bryson (2018) argue that we must carefully consider whether highly intelligent machines deserve moral consideration or rights, and, if so, what implications this would have for how we design and interact with AI.

III. RESULTS

Descriptive statistics are a vital first step in understanding the structure, central tendencies, and spread of our data. These statistics summarise the responses and help identify any anomalies or trends before proceeding with inferential analysis. In our study, we distributed our survey to 94 entities (institutions, organisations, companies) and received 84 completed surveys. The data from these questionnaires were processed using SPSS and analysed using descriptive statistics, correlation analyses, and linear regression. Three main constructs were derived from the questionnaire:

- Ethical Policies – How well-developed are ethical guidelines around AI?
- AI Autonomy – the level of independence AI systems have in making decisions.
- Social Impact – perceived effects of AI on society.

Variable	Type	Mean	Std. Dev.	Minimum	Maximum	Range
Ethical Policies	Independent Variable	2.1	0.8	1.0	4.0	3.0
AI Autonomy	Independent Variable	2.5	0.7	1.0	4.0	3.0
Social Impact	Dependent Variable	2.8	0.6	2.0	4.0	2.0

Table 1. Descriptive statistics

Source**:** Author's processing

From Table 1, the variable Ethical Policies (Mean = 2.1) indicates that AI ethical policies are underdeveloped or only moderately developed. A mean of 2.1 (on a 1–4 scale) indicates limited satisfaction or maturity in policy frameworks. Also suggesting that respondents perceive a lack of clear or enforceable ethical frameworks in current AI implementations across entities. (Standard Deviation =0.8), indicates relatively high variation in how participants perceive these policies—some might rate them as well-developed, while others might not. Organisations may need to standardise or better communicate their ethical AI policies. Next, for the variable AI Autonomy (Mean = 2.5), we can interpret this as indicating that AI systems are perceived as having moderate autonomy. This suggests that many entities use semi-autonomous AI, possibly in decision support rather than full decision-making. There is cautious adoption, possibly due to concerns about control, explainability, or unpredictability. (Standard Deviation 0.7), moderate variation—some companies are more reliant on autonomous systems than others.

There is a diversity of AI implementation levels. Organisations might be at different stages of AI maturity. Moreover, regarding the last variable considered, Social Impact (Mean = 2.8), we can interpret this as generally indicating that people view AI's societal impact positively. A mean close to 3.0 indicates optimism or satisfaction with how AI affects society, such as in healthcare, education, or public service efficiency. Standard Deviation (0.6): Responses are relatively consistent—most participants agree that AI contributes positively. Despite concerns about autonomy and ethics, people acknowledge tangible societal benefits from AI—the correlation analysis. In Table 2, we see a strong positive correlation (r=0.62), indicating that as ethical policy implementation increases, the perceived social impact of AI also increases. This relationship is statistically significant, confirming that ethical guidelines likely enhance the legitimacy and public trust of AI systems. For variable AI Autonomy, and Social Impact (r=0.49), it shows that as AI systems become more autonomous, their impact on society is perceived as more positive, possibly due to efficiency and scalability. However, this is weaker than the correlation with ethical policies, suggesting that autonomy alone is insufficient without safeguards.

Independent Variable	Dependent Variable	Pearson Correlation	Significance (p-value)
Ethical Policies	Social Impact	0.62	0.001
AI Autonomy	Social Impact	0.49	0.015

Table 2. Correlation analyses

Source: Author's processing

For Regression and predictive Insights, the dependent variable is Social Impact (Table 3).

Predictor	Beta (Standardised)	p-value
Ethical Policies	0.51	0.001
AI Autonomy	0.34	0.018

Table 3. Regression Analysis

Source: Author's processing

From the table, we can see that Ethical Policies are the strongest predictor of perceived social impact, even after controlling for AI Autonomy. A 1SD increase in ethical policies results in a 0.51 SD increase in social impact. AI Autonomy also has a positive effect, but to a lesser degree.

III. CONCLUSIONS

Organisations with strong ethical policies benefit from greater stakeholder trust, compliance, and reputational resilience. The data suggests that ethics is not just a compliance tool but a strategic enabler for social value creation. Best practices might include algorithmic studies, bias assessments, and ethical oversight committees.

Autonomy enhances efficiency, scalability, and decision-making power, but without ethical boundaries, it risks becoming unaccountable. Stakeholders value autonomy when combined with human oversight. Implementation levels of autonomy are based on context, with opt-in human override in high-stakes decisions (e.g., medical, legal, etc.).

How society perceives the benefits of artificial intelligence often comes down to three key factors: transparency, ethical responsibility, and ease of use and understanding. When companies and institutions invest in developing AI responsibly and ethically, they build public trust and encourage broader adoption.

As AI technologies become more embedded in our daily lives—impacting everything from healthcare to justice systems to education—the connection between ethical guidelines, autonomous decision-making, and the broader societal impact becomes more difficult to untangle. These three elements—ethics, autonomy, and social impact—do not operate in isolation; they constantly influence one another, creating a complex environment that requires thoughtful regulation and cooperation across sectors.

The question of autonomy, in particular, highlights a difficult balance between speed and control. The more independent AI systems become—especially when making real-time decisions without human supervision—the more we need to worry about how much we can trust them. This is especially true in sensitive areas such as medicine, criminal justice, finance, and even military operations, where a single decision can have serious consequences. That is why autonomy in AI should never mean the absence of oversight.

However, the challenges go beyond efficiency or job loss. Without careful design, AI tools can unintentionally reinforce social inequalities, normalise mass surveillance, or worsen discrimination. These are not side effects—they are central risks that need to be addressed from the very beginning of any AI project. Social responsibility should be part of the foundation, not an afterthought.

There is also much potential for AI to be used for the public good, whether that means making digital platforms more accessible, improving disaster response, or helping track and fight climate change. These kinds of projects deserve more attention and more substantial funding. At the same time, we need to be just as proactive about managing the risks. One idea worth adopting is to create social impact assessments for AI systems, similar to environmental assessments used before launching industrial projects. These tools can help both policymakers and developers anticipate unintended consequences and take steps to prevent harm before it happens.

REFERENCES

Boden, M. A. (2016). *AI: Its nature and future*. Oxford University Press.

Brachman, R. J., & Levesque, H. J. (2004). *Knowledge representation and reasoning*. Morgan Kaufmann.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Bryson, J. (2018). Patiency is not a virtue: The design of intelligent systems should not be based on ethical theories. *(If published, add source—currently incomplete)*.

Citron, D. (2007). Technological due process. *Washington University Law Review.*

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review, 96*(1), 108–116.

Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books.

European Commission. (2021). *Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).* Brussels.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review.* https://doi.org/10.1162/99608f92.8cd550d1

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Harini, S., & Guttag, J. V. (2019). A framework for understanding unintended consequences of machine learning. *Equity and Access in Algorithms, Mechanisms, and Optimization.* https://doi.org/10.1145/3465416.3483305

IEEE. (2019). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems.* The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Suresh, H., & Guttag, J. V. (2019). A framework for understanding unintended consequences of machine learning. *Equity and Access in Algorithms, Mechanisms, and Optimization.* https://doi.org/10.1145/3465416.3483305
(Duplicate entry above—include only once in final list.)

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

08

INCLUSIVE PEDAGOGY AT SCALE: A MODEL FOR BUILDING CAPACITY THROUGH DIGITAL TRAINING AND POLICY IMPLEMENTATION

DOI: 10.37199/c41000308

Elona KARAFILI

Polis University, Tirana, Albania
elona_karafili@universitetipolis.edu.al
ORCID 0000-0002-1769-8367

Flora KRASNIQI

Polis University, Tirana, Albania
flora_krasniqi@universitetipolis.edu.al

ORCID 0000-0002-5678-3548

Ariola MEMIA

Polis University, Tirana, Albania
ariola_memia@universitetipolis.edu.al

Abstract

This paper presents a comprehensive model for building institutional capacity in inclusive pedagogy using digital training and cross-border collaboration. Drawing on the experience of the Inclusive Tertiary Education in the Western Balkans (IDEA) project, this paper explores how a structured, multi-phase digital training program empowered higher education institutions (HEIs) in Albania, Kosovo, and Montenegro to institutionalise inclusive teaching practices. The initiative addressed long-standing challenges in supporting students with disabilities (SwD) and students with learning difficulties (SwLD) through a strategically coordinated capacity-building model. The program incorporated expert-led workshops, hands-on virtual visits, localised policy development, and feedback mechanisms rooted in real-time digital monitoring. By leveraging accessible technologies and fostering collaboration between Western Balkan and EU institutions, the project scaled inclusive pedagogy across diverse institutional and national contexts. The paper outlines the training design, digital delivery methods, and feedback systems, and evaluates their effectiveness in sustaining pedagogical change through policy implementation and continuous professional development.

Insights into institutional ownership, content localisation, and policy alignment highlight key lessons for future digital capacity-building efforts in higher education.

Keywords*:* Inclusive pedagogy, digital training, higher education, assistive technology, policy implementation.

I. INTRODUCTION

Inclusive pedagogy is increasingly recognised as a critical pillar of equitable and high-quality education in higher education institutions (HEIs), particularly in contexts marked by systemic disparities and limited access to resources. Defined broadly, inclusive pedagogy refers to teaching strategies, institutional practices, and learning environments designed to engage and support a diverse range of learners, especially those historically marginalised due to disability, socioeconomic background, or learning differences (Florian & Black-Hawkins, 2011). Recent research emphasises the role of inclusive pedagogy in promoting social justice and participation in tertiary education (Spratt & Florian, 2015; Moriña, 2017). However, embedding inclusive practices across institutions is often hindered by structural, cultural, and pedagogical barriers, especially in regions such as the Western Balkans, where legal frameworks may be in place. However, implementation capacity remains limited (Zgaga et al., 2013). Digital transformation in education has opened new pathways for building inclusive environments at scale. Emerging evidence suggests that online tools, asynchronous modules, and cross-institutional virtual collaboration can accelerate pedagogical change, foster professional development, and improve inclusivity when thoughtfully aligned with institutional policies and local contexts (Laurillard, 2012; Ainscow, 2020). However, the success of these models relies on sustained capacity building, strong leadership, and systems for feedback and accountability (UNESCO, 2020). The IDEA project (Inclusive Tertiary Education in the Western Balkans) responded to these challenges by developing a scalable, digitally enabled training model for inclusive pedagogy. Funded by the European Commission and implemented across Albania, Kosovo, and Montenegro, IDEA sought to empower HEI staff through targeted training, policy development, and the establishment of accessibility units supported by assistive technologies.

This paper explores the IDEA project as a case study in digitally supported capacity building. It examines the structure and delivery of IDEA's training program, the integration of inclusive pedagogical principles into institutional practices, and the digital tools used to monitor and sustain change. Through this lens, we propose a replicable model for embedding inclusive pedagogy at scale within higher education systems facing similar constraints.

II. INTRODUCTION

This paper applies a qualitative-dominant mixed-methods design, drawing on best practices in educational program evaluation (Creswell & Plano Clark, 2018; Greene, 2007). The purpose is to explore both the structural and experiential dimensions of capacity building in inclusive pedagogy, specifically within the digital training environment fostered by the referred project. The first component was a systematic desk review of core project materials, including training curricula, policy briefs, accessibility protocols, and quality assurance documentation. This approach enabled detailed mapping of the project's capacity-building architecture, including pedagogical frameworks and the evolution of digital delivery formats (Bowen, 2009). Second, semi-structured interviews were conducted with academic, managerial, and administrative staff involved in the IDEA project. Participants were selected through purposive sampling to capture a broad range of perspectives across partner institutions in Albania, Kosovo, and Montenegro. The interviews explored participant experiences with the training sessions, digital tools, perceived pedagogical impact, and institutional support systems. Data were coded using thematic analysis (Braun & Clarke, 2006), enabling interpretation of patterns and critical reflection on implementation processes. Third, a descriptive quantitative analysis was used to triangulate findings. Key performance indicators (KPIs) were drawn from internal reports and monitoring tools developed within IDEA, including the number of staff trained, feedback ratings, policy adoption milestones, and the operational status of Accessibility Units. While not intended for statistical generalisation, this data provided valuable corroboration of training uptake and institutional responsiveness.

Together, these methods allowed for a contextualised evaluation of how digital training interventions can foster inclusive pedagogical transformation at scale. The use of mixed methods aligns with Greene's (2007) advocacy for methodological complementarity, offering both the depth of qualitative insight and the structure of performance-based evidence.

III. TRAINING DESIGN AND DIGITAL DELIVERY

The IDEA project's training program was developed as a structured, multi-tiered capacity-building initiative designed to institutionalise inclusive pedagogy across a diverse set of higher education institutions. Recognising the complexity and variation in readiness levels across participating institutions, the program was structured into three iterative and complementary phases: Foundation, Development, and Application, each with specific objectives and delivery formats tailored to support sustainable change.

In the *Foundation Phase*, expert universities such as Masaryk University, the National and Kapodistrian University of Athens, and the University of Limerick led synchronous virtual workshops to establish a common understanding of inclusive education, student diversity, and assistive technology. These workshops established a baseline for shared terminology, a pedagogical vision, and institutional commitment, fostering cross-border dialogue and shared expectations. Building on this base, the *Development Phase* transitioned participants into a mixed-delivery model that blended asynchronous modules with live virtual coaching. Recorded lectures, institutional policy guides, and pre-reading materials were combined with synchronous discussion sessions and facilitated peer exchange. This format encouraged critical reflection, adaptation to local institutional realities, and collaborative problem-solving. Institutions were supported in contextualising their learning, sharing practices across borders, and gradually embedding inclusive teaching concepts into departmental routines. Finally, the *Application Phase* marked a shift toward hands-on implementation and policy integration. Participants were guided in conducting pilot initiatives within their own institutions, ranging from revised course designs and inclusive assessment methods to the operationalisation of Accessibility Units. This phase was characterised by real-time quality monitoring and feedback, coordinated through the project's Quality Assurance structures. Participating institutions were expected to enact policy-level changes, revise internal procedures, and support staff in applying newly acquired competencies within their teaching and administrative functions. Throughout all phases, digital infrastructure enabled continuity, adaptability, and reach.

Platforms such as Zoom, Moodle, and Google Workspace were instrumental in bridging geographic distances, allowing more than 300 academic and administrative staff from 10 HEIs to engage consistently.

These tools also served as knowledge repositories and collaborative spaces, extending the learning experience beyond formal training events. Notably, the digital approach proved resilient even amid the COVID-19 pandemic, reinforcing the potential of technology to facilitate inclusive pedagogy at scale.

The model's progression from conceptual awareness to policy-backed implementation illustrates a holistic, digitally mediated pathway for institutional change. It exemplifies how carefully scaffolded training, supported by strategic use of technology and continuous feedback, can enable HEIs to transform educational practices in support of diverse learners.

Training Phase	Focus	Delivery Mode
Foundation Phase	Introduce core concepts and build shared understanding	Synchronous virtual workshops (e.g., via Zoom)

Development Phase	Foster reflection, peer learning, and contextual adaptation	Asynchronous + synchronous blended model
Application Phase	Support implementation of inclusive pedagogy through pilots and policy	Hands-on projects, policy adoption, and QA feedback

Table 1. IDEA Training Model Overview

Source: Based on the data collected by the authors (2025)

IV. LESSONS LEARNED AND SUSTAINABILITY

The implementation of the IDEA project surfaced several lessons critical to the broader discourse on inclusive pedagogy and institutional transformation. One of the most prominent insights was the scalability of digital training methods. By leveraging virtual platforms and asynchronous content, the project delivered high-quality pedagogical training to over 300 staff members across 10 institutions, overcoming geographic and logistical limitations. This proved especially important during the COVID-19 pandemic, demonstrating that a resilient digital infrastructure can act as a catalyst for wide-reaching educational reform. Similar conclusions have been echoed in literature evaluating digital inclusivity projects in the Global South (UNESCO, 2020; Ainscow, 2020). Equally important was the principle of localisation. While the training was delivered within a common European framework, its effectiveness hinged on adapting its content to the context. Language, institutional readiness, and legal frameworks differed across Albania, Kosovo, and Montenegro, requiring flexible training models that could respond to national policies and local practices. Institutions that took the initiative to contextualise tools, examples, and pedagogical strategies were better positioned to embed inclusive practices meaningfully into their operations. This finding is consistent with studies such as Florian and Black-Hawkins (2011), which stress the importance of culturally responsive pedagogy in driving sustainable change. A third lesson relates to policy integration. Sustainable change in pedagogy was not achievable through training alone; it necessitated structural support. Institutions that linked the IDEA training outputs with formal governance, such as academic senate approvals, updated quality assurance procedures, or revised teaching protocols, were more successful in institutionalising inclusive education. This highlighted the need for vertical integration, where institutional mandates and regulations reinforce learning outcomes from capacity-building initiatives. As noted in Moriña's (2017) review of inclusive education frameworks in Europe, policy alignment significantly enhances the longevity and depth of inclusion practices.

Perhaps most critically, the project revealed the power of institutional ownership. Participating universities that embraced a participatory approach, incorporating the voices of students and staff into training design, evaluation, and implementation, reported stronger engagement and more profound institutional transformation.

Ownership fostered trust, motivated change agents, and enabled internal alignment to sustain reforms beyond the project's lifecycle. The sustainability of these efforts is maintained through several channels: continued peer-to-peer exchange among regional institutions; formalised requirements for annual inclusive pedagogy training; and the preservation and use of digital platforms and learning repositories established during the project. In this way, IDEA's model transcends one-time intervention and moves toward creating a durable, system-level culture of inclusive teaching and learning.

V. CONCLUSION

The project demonstrated that achieving sustainable, inclusive pedagogy in higher education requires a systematic, integrated approach that goes beyond individual training events or technical interventions. The project's success stems from the thoughtful alignment of pedagogy, digital infrastructure, and institutional governance across multiple national contexts. Through its carefully structured, phased training model, the initiative provided a practical, scalable solution for building inclusive teaching capacity, anchored in principles of equity, collaboration, and institutional ownership. A key achievement of the project lies in its ability to translate pedagogical principles into actionable institutional strategies. By combining expert-led virtual workshops, peer-driven learning, and localised application, the program successfully created a feedback-rich environment where inclusive practices could be developed, tested, and institutionalised. The participatory design of training content, contextual adaptation of learning resources, and embedding of policy reforms, such as the establishment of Accessibility Units, ensured that inclusivity was not merely a theoretical objective but a lived institutional reality. Moreover, the project proved the viability of digital means as both a delivery mechanism and a sustaining force. In the face of pandemic-related disruptions and regional infrastructure disparities, IDEA's use of digital platforms enabled consistent engagement, broad participation, and transnational knowledge sharing.

This highlights how, when supported by intentional design and robust quality assurance mechanisms, digital tools can democratise access to professional development and support long-term institutional transformation.

It is important to emphasise that the project also surfaced the conditions under which inclusive pedagogy can be maintained post-project. The emphasis on policy integration, institutional commitment, and continuous peer learning created a replicable framework for sustaining change. This model resonates with the broader European and global push for lifelong learning, universal design for learning (UDL), and inclusive institutional cultures in higher education.

Finally, initiatives like this one present a timely and transferable example of how digital transformation can be harnessed to promote pedagogical innovation and systemic inclusion. As higher education institutions globally face increasing expectations to accommodate diverse learners and operate across digital ecosystems, the IDEA approach offers a robust, evidence-based, and context-responsive model for scaling inclusive pedagogy. It affirms that strategic alignment between training, technology, and institutional policy is not only feasible but essential to leading educational change with integrity and impact.

REFERENCES

Ainscow, M. (2020). Promoting equity in education: Lessons from international experiences. Nordic Journal of Studies in Educational Policy, 6(1), 1-10.

Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and Conducting Mixed Methods Research (3rd ed.). SAGE.

Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. British Educational Research Journal, 37(5), 813–828.

Greene, J. C. (2007). Mixed Methods in Social Inquiry. Jossey-Bass.

Laurillard, D. (2012). Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology. Routledge.

Moriña, A. (2017). Inclusive education in higher education: Challenges and opportunities. European Journal of Special Needs Education, 32(1), 3–17.

Spratt, J., & Florian, L. (2015). Applying the principles of inclusive pedagogy in initial teacher education: From university to classroom. International Journal of Inclusive Education, 19(8), 1–18.

UNESCO. (2020). Global Education Monitoring Report 2020: Inclusion and education – All means all. UNESCO.

Zgaga, P., Teichler, U., Brennan, J., & Reinalda, B. (2013). Higher Education in the Western Balkans: Reforms, developments, trends. UNESCO/CEPES.

09

BLOCKCHAIN CRYPTOGRAPHY AND THE FUTURE OF DIGITAL CURRENCY SECURITY

DOI: 10.37199/c41000309

Erilda MUKA

Polis University, Tirana, Albania
erilda_muka@universitetipolis.edu.al
ORCID 0009-0001-4764-2145

Anna Maria KOSOVA

Polis University, Tirana, Albania
annamaria_kosova@universitetipolis.edu.al
ORCID 0009-0001-6998-5085

Blerina BEKTASHI

Polis University, Tirana, Albania
blerina_bektashi@universitetipolis.edu.al
ORCID 0009-0004-9875-558X

Abstract

Blockchain security is fundamentally based on several cryptographic mechanisms that maintain transaction integrity, confidentiality, and authentication, and it is the pinnacle of technology today. This paper analyses various cryptographic techniques embedded in blockchain. These techniques include data encryption, digital signatures, and hashing. The following discusses the different consensus mechanisms that enable scalability and integrity in a blockchain, with an emphasis on proof-of-work and proof-of-stake. Nevertheless, there are problems relating to 51% attacks, scalability issues and privacy concerns thereof in a blockchain. To assess users' knowledge and perceptions of blockchain technology security, we will conduct a survey to analyse levels of knowledge, usage trends, and concerns about digital currency security. The results will provide insight into current knowledge of blockchain cryptography and indicate possible areas for strengthening security.

Keywords: Blockchain Security, Attacks, Cryptography, Encryption, Digital Currency.

I. INTRODUCTION

Blockchain technology has quickly moved from an uncertain innovation to a vital component of the modern digital landscape, especially in finance and data security. At its most fundamental level, blockchain is a distributed ledger technology that enables tamperproof, transparent, and secure recording of transactions, with trust established without the involvement of third-party intermediaries (Narayanan et al., 2016). In the case of digital currencies, trust is provided through cryptographic proofs and consensus protocols rather than through banks and traditional financial institutions. One of the major aspects of blockchain technology is its use of cryptography to ensure the confidentiality, authenticity, and integrity of data. Cryptographic hashing, digital signatures, and asymmetric encryption are methods for securely initiating transactions and verifying them (Conti et al., 2018). Each transaction is digitally signed with the user's private key and verified by nodes using the corresponding public key. Some blockchains use cryptographic hash functions to form chains of blocks. If the data in a block is changed, the entire chain will be void (Bonneau et al., 2015). Consensus protocols are equally important to blockchain security, as they determine how nodes in the network agree on the validity of transactions.

Common blockchains like Bitcoin use proof-of-work (PoW) as a consensus mechanism, which is an energy-intensive process that reduces malicious activity (Nakamoto, 2008). Alternative mechanisms, such as proof-of-stake (PoS), have emerged to provide a similar, if not more secure, level of security with lower energy consumption (Saleh, 2021). Consensus protocols help generate trust, security, and accuracy in the blockchain system, yet financial actors in the blockchain ecosystem are often unaware of their meaning and significance. However, blockchain technology is not without fault. A majority attack, also called a 51% attack, occurs when a group of miners or validators becomes the majority holders of network power, posing a serious threat to blockchain networks (Li et al., 2020). Privacy is also a concern in cryptocurrency transactions, as although they are pseudonymous, transaction details are sometimes publicly accessible and can be traced using on-chain analysis tools. The scalability of blockchains is also an ongoing technical concern, as transaction throughput is often lower than that of asynchronous systems.

This study aims to integrate theoretical and empirical research better to understand the social and technological aspects of the challenges. The research utilised a user survey to examine public knowledge, use, and concerns regarding the security of blockchain-based digital currencies. The importance of developing and integrating digital finance into everyday economic life makes it critical to understand how cryptography and blockchain security intersect to design secure, trusted and inclusive financial systems (Zohar, 2015). The insights from this study will help better understand user knowledge gaps and potential weaknesses in current blockchain systems and inform the development of future secure digital currency ecosystems.

II. LITERATURE REVIEW

The rapid growth of blockchain technology has generated significant academic interest, spanning topics such as cryptography, cybersecurity, and digital finance. This literature review synthesises significant research on blockchain-based cryptographic components, consensus protocols, security weaknesses, and user knowledge of blockchain-based digital currencies. Blockchain's strength lies in leveraging conventional cryptography to secure data and build trust in decentralised systems. Data integrity in the blockchain framework is ensured through cryptographic hash functions, such as SHA-256, which is used in Bitcoin. These types of cryptosystems produce irretrievable tokens for any input, meaning that even if someone has the same input, they will receive a different token (Narayanan et al., 2016). This token is fundamental for linking blocks, because if you tampered with one block, you would have to recalculate all hashes, which is computationally impossible. Cryptographic confirmation also uses public-key-only signing to not only confirm authentication but also provide non-repudiation for blockchain-based transactions (Conti et al., 2018). Each transaction is signed with the sender's private key, and anyone receiving it can use the sender's public key to confirm that only the sender could have sent it and that it is legitimate. This use of cryptography eliminates the need to consult a centralised authority and ultimately makes the whole system more transparent and trustworthy. Consensus algorithms allow for the validation and recording of transactions in decentralised blockchain systems.

Proof-of-work (PoW) is the most widely known consensus mechanism, first proposed by Nakamoto (2008) in Bitcoin. PoW uses computational power to require nodes (miners) to solve complex mathematical proofs, which is often very resource-intensive in terms of computing power and electricity. Although PoW is effective at securing a blockchain network, it has been criticised for its environmental impact and scalability. Alternative consensus protocols have been developed, such as proof-of-stake (PoS), which assigns rights based on the amount of cryptocurrency a user is willing to hold and "stake" to validate blocks, which is less energy-intensive and still maintains some security (Saleh, 2021). The proof-of-stake consensus mechanism is particularly significant for the blockchain ecosystem, as Ethereum, a major blockchain platform, has migrated to PoS with its "Ethereum 2.0" upgrade. Despite cryptographic protections, blockchain systems are not susceptible to security risks. One of the most well-known security risks is the 51% attack, in which a collusion of malicious actors obtains control of more than 50% of a blockchain network's computing power, enabling them to manipulate transactions or double-spend digital currency (Li et al., 2020). This vulnerability becomes especially problematic for smaller blockchain networks with lower hash rates or lower decentralisation.

Another problem is the security ramifications of smart contracts, and the decentralised applications (dApps) that are part of that equation, on platforms like Ethereum. Smart contract exposures can result in large-scale exploits, as was the case in 2016, when hackers exploited a flaw in a smart contract to drain $60 million in Ether in the DAO attack (Atzei et al., 2017). Because blockchain is unchangeable, nothing can be undone once these contracts are deployed. There are also privacy issues that have been discussed; while many blockchains, like Bitcoin, provide pseudonymity, these transactions can be viewed publicly and traced. Transparency has raised concerns about user anonymity and data protection, and research has focused on privacy technologies such as zero-knowledge proofs and confidential transactions (Zhang et al., 2019). Most blockchain research has focused on technical studies. However, over the past several years, a handful of researchers have examined human issues, such as user awareness, trust, and the adoption of blockchain technologies. The research indicates that the public typically lacks a sophisticated understanding of the mechanisms that determine the security of blockchains (Alketbi et al., 2018). Uncertainty about digital currencies, driven by volatility, hacking incidents, and regulatory developments, has also been cited as a barrier to wider acceptance.

III. MATERIALS AND METHODS

To assess knowledge of blockchain security, a structured survey was designed for university students enrolled in Computer Science Bachelor's programs. This survey, on a 5-point Likert scale (from "strongly disagree" to "strongly agree"), assesses users' knowledge and perceptions of blockchain technology's security.

Key topics included levels of knowledge, usage trends, concerns about digital currency security, and students' personal experiences and perceptions of cryptography's effectiveness. The survey results will provide insight into current knowledge of blockchain cryptography and identify potential areas for strengthening security.

Participants included three groups of Computer Science students: 84 first-year, 93 second-year, and 96 third-year students, for a total of 273. The survey was created in Google Forms and distributed in person, by email, and via Google Classroom to ensure easy access and broad participation. Data were collected electronically, organised in Excel, and analysed using descriptive statistics in Excel and SPSS Statistics 27 to identify students' attitudes toward blockchain security.

The 10 questions were grouped into four sections:

- Section 1: User Information
- Section 2: Cryptography in Blockchain
- Section 3: Digital Currency Transaction Security
- Section 4: Blockchain Security and Future Challenges

Questions were as follows:

1) I have a good understanding of blockchain technology.
2) I have experience with digital currencies (e.g., Bitcoin and Ethereum).
3) I am familiar with the concept of hashing in blockchain.
4) Hashing is essential for ensuring the security of blockchain.
5) Asymmetric cryptography (public and private keys) is crucial for securing blockchain transactions.
6) Blockchain transactions are generally secure.
7) Digital currency transactions are vulnerable to attacks such as Man-in-the-Middle or Double-Spending.
8) I am confident that current blockchain security measures can effectively prevent fraud and unauthorised access.
9) Cryptography alone is sufficient to prevent manipulation of the transaction history in a blockchain.
10) I believe blockchain security threats will become more severe in the future.

The values of the questions are:

1 = strongly disagree,

2 = disagree,

3 = neutral,

4 = agree,

5 = strongly agree,

The interpretation of the mean values is in Table 1.

Mean value	Interpretation
1- 1.8	Strongly disagree (SD)
1.81- 2.6	Disagree (D)
2.61- 3.4	Neutral (N)
3.41- 4.2	Agree (A)
4.21- 5.0	Strongly Agree (SA)

Table 1. Mean values and interpretations

Source: Author's processing

IV. RESULTS

The survey data were collected using Google Forms and were initially processed in Excel to conduct exploratory visual analyses and note first-response patterns. To conduct a more formal, rigorous, and comprehensive evaluation, the dataset was subsequently imported into SPSS Statistics 27. In SPSS, descriptive statistics, including means, medians, standard deviations, and an important consistency check, were calculated to assess the data's reliability and students' opinions on establishing security in blockchain cryptography and digital currency. One important component of the analysis performed was to assess reliability using Cronbach's Alpha. The overall Cronbach's Alpha for the survey items was 0.872 (Table 2), indicating high internal consistency. This reliability score, as an overall indicator, supports the idea that the survey measures a defined latent factor, allowing the findings to be reliably interpreted.

Cronbach's Alpha	N of Items
.872	10

Table 2. Reliability Statistics

Source: Author's processing

Descriptive statistics were calculated for all of the questions (Q1-Q10) of the survey to define the overall response of the general participants (Table 3). The mean score for all questions ranged from 3.21 (Q2) to 3.92 (Q10). Since the means fell largely between 'Neutral' (3) and 'Agree' (4), the students ultimately appeared to exhibit just moderate strengths of understanding and agreement on blockchain security. These distributions are visually confirmed by the histograms, which show peaks around the "Neutral" and "Agree" categories for most questions. For example, Q4 (Hashing is critical to the security of blockchain) had a mean of 3.67, and the 'Agree' frequency was notable. At the same time, Q5 (Asymmetric cryptography (public and private keys) is critical to the security of blockchain transactions) had a mean of 3.79, with a significant number of 'Agree' responses.

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Q1	24	3.00	2.00	5.00	3.4167	.88055	.775	-.141	.472	-.610	.918
Q2	24	4.00	1.00	5.00	3.2083	1.21509	1.476	.200	.472	-1.057	.918
Q3	24	3.00	2.00	5.00	3.5417	.83297	.694	.103	.472	-.371	.918
Q4	24	3.00	2.00	5.00	3.6667	.91683	.841	-.356	.472	-.469	.918
Q5	24	3.00	2.00	5.00	3.7917	.83297	.694	-.066	.472	-.605	.918
Q6	24	3.00	2.00	5.00	3.4167	.82970	.688	.039	.472	-.338	.918
Q7	24	3.00	2.00	5.00	3.4167	.71728	.514	-.068	.472	-.058	.918
Q8	24	3.00	2.00	5.00	3.3333	.70196	.493	.244	.472	.234	.918
Q9	24	3.00	2.00	5.00	3.5000	.83406	.696	-.245	.472	-.343	.918
Q10	24	3.00	2.00	5.00	3.9167	.77553	.601	-.460	.472	.298	.918
Valid N (listwise)	24										

Table 3. Descriptive Statistics

Source: Author's processing

		Std. Error
Q1	Mean	.17974
Q2	Mean	.24803
Q3	Mean	.17003
Q4	Mean	.18715
Q5	Mean	.17003
Q6	Mean	.16936
Q7	Mean	.14641
Q8	Mean	.14329
Q9	Mean	.17025
Q10	Mean	.15830

Table 4. Descriptive statistics

Source: Author's processing

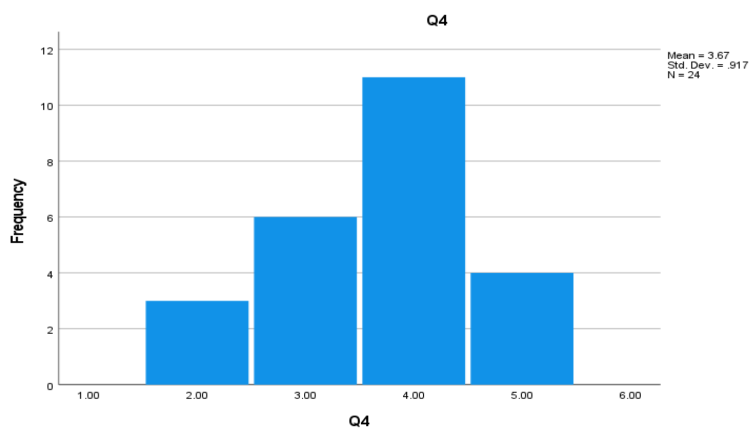


Figure 1. Q4 Frequency

Source: Author's processing

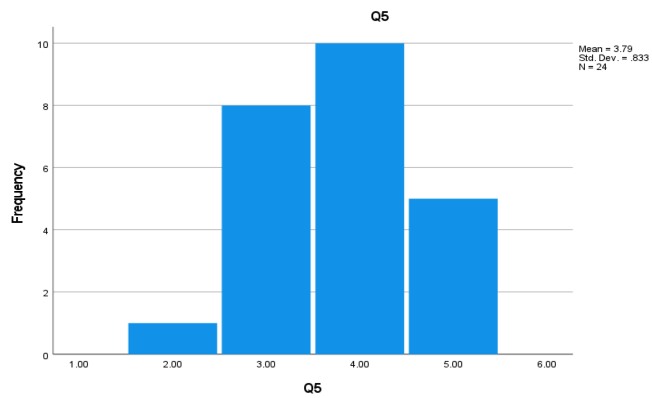


Figure 2. Q4 Frequency

Source: Author's processing

We further examined the distribution of agreement by separating the responses by the students' year of study (Year 1, 2, or 3). The average level of agreement across all questions varied slightly over the years; Year 1 students had an average score of 3.31, Year 2 students had 3.35, and Year 3 students had 3.43. These differences are small; however, they do indicate a subtle increase in agreement or understanding with the students' year of study. Interestingly, Year 2 students were noted in the documentation to have taken a Cybersecurity course and, as a result, did not show a marked difference compared with Year 3 students, suggesting that while the Cybersecurity course likely increased understanding, the overall disposition was generally similar or consistent across the latter years. The breakdown of the average level of agreement by individual question also varied by year of study. For example, Year 2 students showed a higher proportion of "Strongly Agree" responses to certain questions than Year 1 students, even though these questions were directly related to cryptographic concepts. This distinction by academic year provides a useful framework for analysing the different perspectives of computer science students regarding their understanding of blockchain security development.

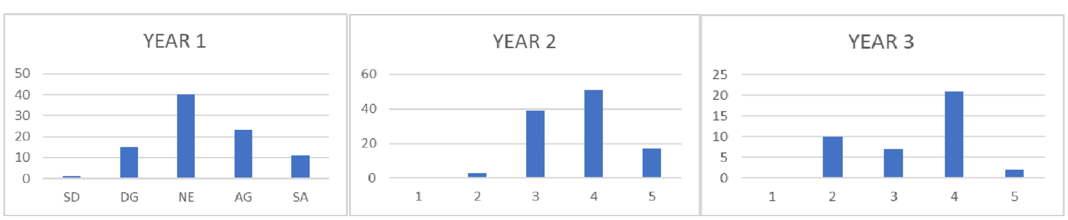


Figure 3. Frequency distribution

Source: Author's processing

V. CONCLUSIONS

The study aims to explore the foundational cryptographic principles of blockchain security and examine computer science students' understanding and perceptions of digital currency security. Survey responses showed that students ranged from moderate to high in their understanding of security concepts and their application. The mean score ranged from 'Neutral' to 'Agree' on each question posed to the participants. The internal consistency of the survey items was high (Cronbach's Alpha = 0.872), and the survey outcomes measurements can be considered reliable. Although students demonstrated a high level of understanding of blockchain security concepts, such as hashing and asymmetric cryptography, their responses showed slight differences across academic years. The trend of slightly higher average agreement scores from Year 1 to Year 3 reflects incremental development of knowledge and latent confidence among students as they progress through a computer science degree. Importantly, Year 2 students have a cybersecurity course, yet their perception scores for blockchain security were similar to those of Year 3 students. This suggests that while a specific course adds value, overall exposure to computer science concepts over extended periods contributes significantly to understanding. The importance of practical, interactive activities is reinforced in this study and is similar to effective technology use in other STEM education applications. Just as tools like the GeoGebra app, the Desmos Graphing Calculator, or any statistical software help visualise abstract mathematical concepts, provide hands-on experimentation, and draw meaningful conclusions with abstract ideas, similar approaches will be essential to help students understand how blockchain cryptography actually functions. Suppose we allow students to manipulate simulations or view visualisations of crypto processes, or to design their own simplified blockchain models. In that case, we create real opportunities for them to develop an intuitive understanding of security measures.

The apps that calculate results, like Wolfram Alpha or Microsoft Math Solver, which detail their answers, will grant teachers a degree of remote access to students' knowledge of advanced blockchain concepts that had previously been obscured. For instance, an app that detailed the steps of Hashing Algorithms or illustrated how changes to consensus mechanisms would affect projections would combine abstract principles with tangible, reasonable insights. The strength of testing and quiz apps, especially Google Forms/Quiz Mode, for immediate assessment of student comprehension, offers educators a torturous chain of opportunities to validate, clarify, and ultimately reinforce students' understanding of essential concepts in the blockchain security and cryptography space. At the end of the day, the interactive role of prospective student teachers in creating a technology-rich learning environment, as emphasised in the reference material for developing STEM courses, guided by the current digitalisation of currency security, with strong knowledge and confidence, is translated into simulations of real-world currency developments.

REFERENCES

Al-Jaroodi, J., & Mohamed, N. (2019). Blockchain in industries: A survey. *IEEE Access, 7*, 36500–36515. https://doi.org/10.1109/ACCESS.2019.2903554

Alketbi, A., Nasir, Q., & Talib, M. A. (2018). Blockchain for government services—Use cases, security benefits and challenges. *2018 15th Learning and Technology Conference (L&T)*, 112–119. https://doi.org/10.1109/LT.2018.8368508

Atzei, N., Bartoletti, M., & Cimoli, T. (2017). A survey of attacks on Ethereum smart contracts (SoK). *Proceedings of the 6th International Conference on Principles of Security and Trust*, 164–186. https://doi.org/10.1007/978-3-662-54455-6_8

Bertaccini, M. (2022). *Cryptography Algorithms: A guide to algorithms in blockchain, quantum cryptography, zero-knowledge protocols, and homomorphic encryption.* Packt Publishing.

Bhutta, M. N. M., et al. (2021). A survey on blockchain technology: Evolution, architecture and security. *IEEE Access, 9*, 61048–61073. https://doi.org/10.1109/ACCESS.2021.3072849

Bonneau, J., Miller, A., Clark, J., Narayanan, A., Kroll, J. A., & Felten, E. W. (2015). SoK: Research perspectives and challenges for Bitcoin and cryptocurrencies. *2015 IEEE Symposium on Security and Privacy*, 104–121. https://doi.org/10.1109/SP.2015.14

Chen, H., Wei, N., Wang, L., Mobarak, W. F. M., Albahar, M. A., & Shaikh, Z. A. (2024). The role of blockchain in finance beyond cryptocurrency: Trust, data management, and automation. *IEEE Access, 12*, 64861–64885. https://doi.org/10.1109/ACCESS.2024.3395918

Conti, M., Kumar, S., Lal, C., & Ruj, S. (2018). A survey on security and privacy issues of Bitcoin. *IEEE Communications Surveys & Tutorials, 20*(4), 3416–3452. https://doi.org/10.1109/COMST.2018.2842460

Fathalla, E., & Azab, M. (2024). Beyond classical cryptography: A systematic review of post-quantum hash-based signature schemes, security, and optimisations. *IEEE Access, 12*, 175969–175987. https://doi.org/10.1109/ACCESS.2024.3485602

Fernández-Caramés, T. M., & Fraga-Lamas, P. (2020). Towards post-quantum blockchain: A review on blockchain cryptography resistant to quantum computing attacks. *IEEE Access, 8*, 21091–21116. https://doi.org/10.1109/ACCESS.2020.2968985

Islam, M. M., & In, H. P. (2024). An auditable, privacy-preserving, transparent unspent transaction output model for blockchain-based central bank digital currency. *IEEE Open Journal of the Computer Society, 5*, 671–683. https://doi.org/10.1109/OJCS.2024.3486193

Khan, I., Maghrabi, L. A., Sarwar, M. I., Naith, Q. H., & Nisar, K. (2023). Blockchain: A crypto-intensive technology—A comprehensive review. *IEEE Access, 11*, 141926–141955. https://doi.org/10.1109/ACCESS.2023.3342079

Koroglu, T., & Samet, R. (2024). Can there be a two-way hash function? *IEEE Access, 12*, 18358–18386. https://doi.org/10.1109/ACCESS.2024.3360217

Kuznetsov, O., Sernani, P., Romeo, L., Frontoni, E., & Mancini, A. (2024). On the integration of artificial intelligence and blockchain technology: A perspective about security. *IEEE Access, 12*, 3881–3897. https://doi.org/10.1109/ACCESS.2023.3349019

Li, X., Jiang, P., Chen, T., Luo, X., & Wen, Q. (2020). A survey on the security of blockchain systems. *Future Generation Computer Systems, 107*, 841–853. https://doi.org/10.1016/j.future.2017.08.020

Maldonado-Ruiz, D., Torres, J., El Madhoun, N., & Badra, M. (2022). Current trends in blockchain implementations on the paradigm of public key infrastructure: A survey. *IEEE Access, 10*, 17641–17655. https://doi.org/10.1109/ACCESS.2022.3145156

Mishra, D., & Phansalkar, S. (2025). Blockchain security in focus: A comprehensive investigation into threats, smart contract security, cross-chain bridges, and vulnerabilities detection tools and techniques. *IEEE Access, 13*, 60643–60671. https://doi.org/10.1109/ACCESS.2025.3556499

Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system.* https://bitcoin.org/bitcoin.pdf

Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). *Bitcoin and cryptocurrency technologies: A comprehensive introduction.* Princeton University Press.

Ramakrishna, D., & Shaik, M. A. (2025). A comprehensive analysis of cryptographic algorithms: Evaluating security, efficiency, and future challenges. *IEEE Access, 13*, 11576–11593. https://doi.org/10.1109/ACCESS.2024.3518533

Sadawi, A. A., Hassan, M. S., & Ndiaye, M. (2021). A survey on the integration of blockchain with IoT to enhance performance and eliminate challenges. *IEEE Access, 9*, 54478–54497. https://doi.org/10.1109/ACCESS.2021.3070555

Saleh, F. (2021). Blockchain without waste: Proof-of-stake. *The Review of Financial Studies, 34*(3), 1156–1190. https://doi.org/10.1093/rfs/hhaa075

Shaikh, Z. A., et al. (2024). A new trend in cryptographic information security for Industry 5.0: A systematic review. *IEEE Access, 12*, 7156–7169. https://doi.org/10.1109/ACCESS.2024.3351485

Sinha, S. K., & Mukhopadhyay, D. (2024). Time-efficient hash key generation for a blockchain-enabled framework. *IEEE Access, 12*, 155867–155884. https://doi.org/10.1109/ACCESS.2024.3478845

Upadhyay, D., Gaikwad, N., Zaman, M., & Sampalli, S. (2022). Investigating the avalanche effect of various cryptographically secure hash functions and hash-based applications. *IEEE Access, 10*, 112472–112486. https://doi.org/10.1109/ACCESS.2022.3215778

Yaga, D., Mell, P., Roby, N., & Scarfone, K. (2018). *Blockchain technology overview.* National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8202

Yang, Z., Alfauri, H., Farkiani, B., Jain, R., Pietro, R. D., & Erbad, A. (2024). A survey and comparison of post-quantum and quantum blockchains. *IEEE Communications Surveys & Tutorials, 26*(2), 967–1002. https://doi.org/10.1109/COMST.2023.3325761

Zhang, R., Xue, R., & Liu, L. (2019). Security and privacy on blockchain. *ACM Computing Surveys, 52*(3), 1–34. https://doi.org/10.1145/3316481

Zohar, A. (2015). Bitcoin: Under the hood. *Communications of the ACM, 58*(9), 104–113. https://doi.org/10.1145/2701411

10

DIGITAL TWINS AS CATALYSTS FOR SUSTAINABILITY EDUCATION IN UNIVERSITY CAMPUSES: A CASE STUDY AT POLIS UNIVERSITY WITHIN THE FRAMEWORK OF EDUCATION 4.0

DOI: 10.37199/c41000310

Erilda MUKA

POLIS University (Tirana, Albania)
erilda_muka@universitetipolis.edu.al
ORCID 0009-0001-4764-2145

Galia MARINOVA

Technical University of Sofia (Sofia, Bulgaria)
gim@tu-sofia.bg
ORCID 0000-0002-8804-5662

Emiliano MANKOLLI

POLIS University (Tirana, Albania)
emiliano_mankolli@universitetipolis.edu.al
ORCID 0009-0001-4573-6826

Abstract

In the era of Education 4.0, digital technologies such as Digital Twins (DTs) offer transformative opportunities not only for operational efficiency but also for sustainability education and behavioural transformation. This study explores the role of Digital Twin technologies, integrated with IoT systems, in promoting sustainability awareness, engagement, and behaviour modification among students and faculty at Polis University. Building on a previously implemented digital twin prototype for HVAC and lighting systems, this paper extends its application to the educational and behavioural domains. The case study examines how real-time environmental data visualisations and interactive digital environments can enhance understanding, stimulate eco-friendly decision-making, and align learning outcomes with sustainability goals. Through a mixed-method approach, including surveys, user interaction analytics, and energy data logs, the paper demonstrates how DTs support the principles of Education 4.0personalised learning, digital literacy, and learner empowerment—while fostering a green campus culture. The study provides a conceptual and *technical framework for integrating digital twins into campus curricula and sustainability campaigns.*

Keywords: Digital Twins, Sustainability Education, Education 4.0, university campuses, IoT.

I. INTRODUCTION

University campuses are complex, multifaceted environments where education, technology, and sustainability intersect. As institutions face mounting pressures to reduce carbon footprints and enhance operational efficiency, emerging technologies such as Digital Twins (DTs) and Internet of Things (IoT) devices offer powerful tools to manage and optimise energy consumption, resource allocation, and infrastructure performance (Muka & Marinova, 2023). DTs, digital replicas of physical systems, have been effectively deployed in innovative campus initiatives to enable real-time monitoring and remote control of systems such as HVAC and lighting (Grieves, 2014; Muka, 2023). Their integration with IoT systems enables a dynamic, data-driven representation of physical environments, facilitating predictive maintenance, energy efficiency, and operational automation. In the pursuit of more energy-efficient, intelligent environments, Digital Twin (DT) technologies, which create virtual replicas of physical systems, have gained traction in campus infrastructure management. Previously, digital twins have been applied at Polis University to monitor and automate energy usage in lecture halls through IoT-integrated HVAC and lighting systems, enhancing campus efficiency and reducing energy waste. However, the role of digital twins extends beyond operational optimisation.

Digital twins are virtual replicas of physical systems that integrate real-time data and advanced modelling to simulate, monitor, and optimise the performance of their physical counterparts. They are widely used across domains such as manufacturing, smart buildings, healthcare, urban planning, and education. In manufacturing and industry, digital twins help reduce downtime and improve efficiency by enabling predictive maintenance and performance optimisation. In smart buildings and campuses, they facilitate energy management and sustainability efforts by integrating IoT sensors and providing real-time monitoring and automation. Healthcare applications utilise digital twins to simulate patient-specific data for personalised treatment planning. Urban planners use digital twins to model city infrastructure better to manage traffic, utilities, and environmental resources. Within education, digital twins offer hands-on learning experiences that enhance interdisciplinary skills and promote sustainability awareness. Key features of digital twins include real-time data integration from physical assets via sensors, interactive 3D modelling and visualisation for better understanding and simulation, and predictive analytics powered by artificial intelligence to forecast system behaviour and enable proactive maintenance.

Additionally, digital twins provide interactive feedback through user-friendly dashboards, simulations, and gamification, engaging users and supporting decision-making. Their scalability allows them to expand from individual devices or buildings to complex systems and entire campuses. Furthermore, digital twins support cross-disciplinary collaboration by integrating data from engineering, environmental, behavioural, and policy domains. Remote monitoring and control capabilities enable users to manage systems from web or mobile platforms, enhancing flexibility and accessibility.

As higher education shifts toward Education 4.0, there is an increasing emphasis on learner-centric, digitally enhanced experiences that promote real-world competencies such as sustainability awareness and responsible behaviour. Education 4.0 envisions a personalised, data-driven, and experiential learning environment where students are not just passive recipients of information, but active participants in solving contemporary challenges, including environmental ones. However, the potential of digital twins extends far beyond infrastructure management. In the evolving educational paradigm of Education 4.0, there is a strategic shift toward learner-centric, technologically integrated learning environments. Education 4.0 emphasises personalised learning, digital literacy, experiential engagement, and sustainability as core competencies for the 21st-century learner (Salmon, 2019). The central thesis of this paper is that digital twins, when integrated with IoT systems and educational interfaces, can effectively promote sustainability-oriented behavioural changes among campus populations.

Furthermore, such applications align with the goals of Education 4.0 by fostering digital fluency, systems thinking, and data-driven decision-making. By addressing these questions through a real-world case study and a mixed-methods analysis, this research provides a blueprint for leveraging digital twin technologies not only for operational sustainability but also for shaping eco-conscious, future-ready learners. Digital Twins (DTs) represent a cutting-edge technology that creates virtual replicas of physical systems, enabling real-time monitoring, simulation, and control. This capability positions DTs as pivotal tools for advancing sustainability goals while simultaneously supporting the pedagogical transformation envisioned by Education 4.0.

From a sustainability perspective, DTs facilitate real-time environmental monitoring and optimisation by continuously tracking key parameters such as energy consumption, temperature, lighting, and occupancy within buildings or campuses (Muka & Marinova, 2023; Al-Ali et al., 2020). Through simulation of various operational scenarios, DTs identify inefficiencies and allow for automated or user-informed adjustments, significantly reducing resource waste and greenhouse gas emissions (Singh, 2023). Furthermore, DT platforms enhance behavioural change and environmental awareness by visualising energy-use data, thereby empowering users, including students, faculty, and administrators, to make informed, eco-conscious decisions (Singh, 2023). The integration of predictive maintenance features also supports resource management by anticipating equipment failures and scheduling timely repairs, thereby reducing material waste and extending asset lifecycles (Ozturk, 2021). Moreover, DTs can integrate renewable energy sources into campus management systems, dynamically balancing energy demand and supply to further reduce the campus's carbon footprint.

In the context of Education 4.0, which emphasises learner-centred, technology-enhanced, and competency-based education, DTs offer experiential, personalised learning environments. Through interactive 3D simulations and real-time data visualisation, students engage in active experimentation and problem-solving, fostering more profound understanding and skill development (Fisk, 2017; Voogt et al., 2015). Digital twin implementations naturally support interdisciplinary collaboration, bringing together students from STEM, environmental science, and social sciences to tackle complex sustainability challenges (Voogt et al., 2015). Additionally, by working with live data streams and system models, students develop critical data literacy, systems thinking, and sustainability competencies, equipping them with essential skills for the digital economy and global challenges (Yu & Yu, 2020). The integration of DTs with mobile applications and cloud services further enables remote, flexible access to learning resources, facilitating continuous engagement regardless of physical location (Bishop & Verleger, 2013). This paper builds on a digital twin (DT) prototype developed at Polis University, initially designed for energy-efficient lecture hall management using Unity 3D visualisations and IoT devices. Expanding beyond infrastructure monitoring, the study explores how DTs can support sustainability education and encourage eco-conscious behaviours among university students. It investigates key research questions on the role of DTs in shaping sustainable actions, aligning with Education 4.0 principles, and the challenges and opportunities of implementation. Through a case study and mixed-methods analysis, including system usage data, surveys, and educational strategies, the paper shows that DT technologies can serve as impactful tools for promoting sustainability and behavioural change in higher education.

II. LITERATURE REVIEW

Digital Twin (DT) technologies have evolved into transformative tools for infrastructure management, particularly in smart building and campus settings. A digital twin is a dynamic, digital representation of a physical system that is continuously updated with real-time data from IoT devices, sensors, and user interactions (Grieves, 2014). The digital transformation of campuses must therefore not only improve operational systems but also cultivate sustainable mindsets and behaviours among students, faculty, and administrative staff. This paper proposes expanding digital twin applications into the realm of sustainability education and behavioural change. By enabling students and staff to interact with real-time data about their physical environment, digital twins can serve as powerful educational tools. They offer immersive, feedback-rich experiences that help individuals understand the environmental consequences of their choices, whether adjusting temperature settings, managing lighting, or engaging with broader patterns of resource consumption on campus. In this context, digital twins can serve as both pedagogical and behavioural instruments, visually and interactively communicating environmental data and system performance to end users in real time. By offering immediate feedback loops and contextual awareness, DTs can foster behavioural change by increasing users' understanding of the ecological impacts of their actions. This aligns with psychological theories of behaviour change, such as the Theory of Planned Behaviour (Ajzen, 1991), which emphasise the roles of awareness and perceived control in shaping sustainable decisions. Moreover, the educational use of DTs directly supports Education 4.0's goal of digital empowerment. Through student interaction with campus digital twins, such as virtual models of lecture halls that reflect real-time energy usage, learners can explore sustainability concepts experientially. These digital environments offer opportunities to integrate sustainability into curricula in meaningful, action-oriented ways (Beetham & Sharpe, 2013; Chi et al., 2022).

Digital twins (DTs) are uniquely positioned to serve both operational and pedagogical functions within university campuses. Their core capabilities, real-time monitoring, system simulation, data visualisation, and user interaction align well with key goals in sustainability education and behaviour modification. From a behavioural perspective, DTs provide immediate feedback loops that help users understand the impact of their actions on environmental parameters, such as energy use, temperature control, and lighting conditions (Grieves, 2014; Muka & Marinova, 2023). This feedback fosters "situated learning," in which knowledge is acquired in context and reinforced through direct interaction (Beetham & Sharpe, 2013). Behavioural science suggests that such immediacy and contextualisation can be powerful levers for change. According to the Fogg Behaviour Model (FBM), sustainable behaviours are more likely when users are motivated, able, and prompted at the right moment (Fogg, 2009). DT platforms serve as these "digital prompts," using real-time environmental data to visualise the effects of decisions and encourage eco-conscious choices.

The visual representation of consequences increases perceived behavioural control—one of the core factors in the Theory of Planned Behaviour (Ajzen, 1991)—thereby increasing individuals' likelihood of acting sustainably. Education 4.0 represents a pedagogical shift aligned with the Fourth Industrial Revolution, emphasising personalised, technology-driven, and skills-oriented learning (Salmon, 2019). It promotes student empowerment, interdisciplinary knowledge, and real-world problem-solving, particularly in areas such as digital literacy and sustainability. Central to Education 4.0 is the use of digital tools and data analytics to enhance learning experiences and outcomes. Smart campuses, embedded with IoT and DT systems, exemplify this paradigm. These environments facilitate not only operational optimisation but also new modes of teaching and learning, such as real-time data-based lessons and experiential learning modules (Beetham & Sharpe, 2013). By allowing students to interact directly with live environmental data from their campus, DTs provide a bridge between theoretical knowledge and real-world application. The application of DTs in educational simulations is gaining attention for its potential to create immersive and engaging learning experiences. The emergence of Education 4.0 calls for reimagining teaching and learning environments as more learner-centred, competency-driven, and digitally integrated (Salmon, 2019).

Digital twins can model complex systems such as buildings, vehicles, or cities in 3D environments where learners can experiment, visualise outcomes, and receive immediate feedback (Chi et al., 2022). These applications are efficient in serious games designed for learning rather than entertainment. In the context of sustainability, DT-enhanced simulations can demonstrate the impact of individual and collective actions on energy use, emissions, and resource management. For example, students may interact with a virtual campus that reflects live IoT data and explore how behavioural adjustments, such as dimming lights or altering HVAC settings, affect overall energy consumption. These environments promote systems thinking and ecological literacy, two core goals of sustainability education (Xie, 2023). The pedagogical use of DTs in higher education is still an emerging field. However, early evidence suggests strong potential to increase student engagement, improve comprehension of complex concepts, and foster long-term behaviour change. Digital Twin (DT) technologies are revolutionising infrastructure management, particularly in smart campus contexts. The implementation of DTs in Polis University has demonstrated successful energy efficiency through real-time monitoring and control of lighting and HVAC systems (Muka & Marinova, 2023). Digital twins contribute to these educational dimensions in several key ways: Personalised Learning: DTs provide individualised feedback on energy use and behavioural patterns, enabling students to explore sustainability through their own decisions and actions. Interfaces can be customised to reflect different user profiles or roles (e.g., student, faculty, administrator), thereby tailoring learning outcomes to individual experiences (Beetham & Sharpe, 2013).

A. Competency-Based Education: By simulating real-world systems, DTs provide students with experiential learning opportunities to develop competencies in systems thinking, data literacy, environmental ethics, and decision-making. These competencies are essential for addressing global challenges and align with the Sustainable Development Goals (SDGs) (Ozturk, 2021).
B. Technology-Driven Learning: Digital twins embody the integration of IoT, AI, simulation, and cloud computing—technologies that underpin the Fourth Industrial Revolution. Learning in DT-enabled environments exposes students to cutting-edge tools and encourages innovation, a central aim of Education 4.0 (Xie et al., 2023).

C. Behavioural Change Models in Sustainability: The Theory of Planned Behaviour (Ajzen, 1991) and the Fogg Behaviour Model (Fogg, 2009) offer frameworks for understanding how real-time feedback and contextual prompts can influence sustainable behaviours. DTs operationalise these theories through interactive, feedback-driven platforms.

D. Education 4.0 and Smart Learning Environments Education 4.0 promotes competency-based, personalised, and digitally enabled learning (Salmon, 2019). DTs enhance this by embedding learning within real-world contexts and offering responsive, data-driven experiences.

E. DTs in Educational Simulations and Games: Digital twins facilitate immersive simulations that reinforce complex concepts through interactivity and visual feedback. These have been shown to increase engagement and comprehension in sustainability education (Chi et al., 2022).

Behavioural change is a critical dimension of sustainability, especially within institutions such as universities, where collective action significantly affects ecological footprints. The Theory of Planned Behaviour (TPB) posits that an individual's behaviour is directly influenced by their intention to act, which, in turn, is shaped by attitudes, subjective norms, and perceived behavioural control (Ajzen, 1991). TPB has been widely applied in environmental psychology to explain eco-friendly behaviours such as energy conservation, recycling, and the use of public transportation (Kollmuss & Agyeman, 2002). Another influential model, the Fogg Behaviour Model (FBM), focuses on the convergence of motivation, ability, and prompts. According to Fogg (2009), for a behaviour to occur, all three elements must converge at the exact moment. In smart environments, digital prompts—such as real-time feedback or visualisations—can effectively nudge users toward desired behaviours. Digital twins, when integrated into user interfaces or gamified platforms, can offer these prompts by visualising the environmental consequences of user actions (e.g., energy use after adjusting room temperature). By enhancing user awareness and reinforcing perceived control, DTs can be instrumental in initiating and sustaining sustainable behaviours. The case study at Polis University contributes to this growing body of knowledge by integrating DT technology not only for infrastructure management but also as an active agent in behavioural and educational transformation.

III. METHODOLOGY

To effectively teach sustainability using digital twins (DTs) within the framework of Education 4.0, this study employs a blended methodology that integrates project-based learning (PBL), experiential learning, and interdisciplinary collaboration. This pedagogical approach aligns with the goals of Education 4.0, which emphasises personalised, technology-driven, and competency-based education tailored to 21st-century skills (Fisk, 2017; Yu & Yu, 2020).

Project-Based Learning (PBL) serves as the core instructional strategy. Students are tasked with identifying sustainability challenges in their university environment, such as excessive energy consumption in classrooms or inefficient lighting systems, and addressing them by developing or enhancing digital twin solutions. This method allows students to apply theoretical knowledge to real-world problems, encouraging critical thinking and solution-oriented design (Thomas, 2000).

Complementing PBL, **experiential learning** is integrated to deepen engagement. Students interact directly with physical IoT devices (e.g., ESP8266 microcontrollers and DHT11 sensors) and their digital representations in Unity 3D. This hands-on experience not only supports conceptual understanding but also enables learners to observe the tangible impact of digital twins on energy use, thereby promoting environmentally responsible behaviours (Kolb, 1984; Muka & Marinova, 2023).

A **blended learning** environment further supports Education 4.0 goals by combining online and in-person modalities. Online platforms provide instructional content on sustainability, intelligent systems, and IoT programming, while face-to-face sessions facilitate collaborative problem-solving and technical troubleshooting. Flipped classroom techniques are employed to maximise in-class application of concepts (Bishop & Verleger, 2013). The methodology also emphasises **collaborative and interdisciplinary learning**, reflecting the complexity of sustainability challenges that require diverse expertise. Student teams are composed of individuals from engineering, architecture, and environmental studies backgrounds, fostering peer learning and systems thinking. Tools such as web-based dashboards, collaborative project management software, and mobile applications are used to coordinate and present outcomes. In alignment with **challenge-based learning (CBL)** principles, open-ended sustainability questions guide the learning process. For example, students might be asked: *"How can digital twins be used to reduce carbon emissions in university dormitories?"* Such questions promote inquiry, creativity, and innovation, core skills advocated by Education 4.0 frameworks (Voogt et al., 2015). The integration of these methods positions digital twin technology not only as a technical tool for optimisation but also as a pedagogical instrument for sustainability education. Students gain practical skills in IoT, data analysis, and intelligent systems design while internalising the values of environmental stewardship and responsible innovation.

IV. CASE STUDY

To integrate digital twin (DT) technologies into sustainability education within the Education 4.0 framework, this study adopts a Project-Based Learning (PBL) methodology. PBL offers a learner-centred, inquiry-driven approach that empowers students to engage with authentic, real-world problems and develop solutions through active collaboration and technological innovation (Thomas, 2000; Bell, 2010). This methodology aligns closely with the goals of Education 4.0, which emphasises personalised, experiential, and competency-based learning tailored to the needs of the digital age (Fisk, 2017; Yu & Yu, 2020). In this context, students are not passive recipients of information but active participants in the development, deployment, and analysis of a DT system aimed at reducing energy consumption and promoting sustainable behaviour on a university campus. The foundation of this project was the existing physical infrastructure at Polis University in Tirana, Albania, where previous implementations of digital twins integrated with IoT technologies had already demonstrated measurable gains in energy efficiency (Muka & Marinova, 2023). These systems employed ESP8266 microcontrollers, DHT11 temperature and humidity sensors, light sensors, and motion detectors to enable the monitoring and remote control of HVAC and lighting systems. This real-world infrastructure provided an ideal testbed for students to explore the intersections of engineering, data analytics, sustainability, and digital modelling. Within the PBL framework, students were organised into interdisciplinary teams composed of architecture, environmental science, and engineering students. Their task was to extend the capabilities of the existing DT prototype, focusing not only on technical improvements, such as expanding the number of controllable devices and enhancing wireless communication, but also on its educational and behavioural applications. This hands-on, collaborative learning process was structured around several project phases: system understanding, problem identification, design iteration, real-time simulation, and testing of the DT's ability to influence energy use and user awareness.

IV. 1 Digital Infrastructure

The digital infrastructure for each task was developed using Unity 3D, which served as the primary platform for visualising the DT environment. In Figure 1, you can see a fully detailed 3D model of the lecture hall. The IoT devices were designed in Fusion 360 and imported into Unity, where they were programmed in C# to simulate environmental responses and control logic based on real-time sensor input. This virtual environment allowed students and administrators to interact with the DT via a web interface and mobile apps, enabling remote monitoring and control of temperature, lighting, and device activity. Integration with Apple's HomeKit and Android-based platforms ensured cross-device compatibility and enhanced user accessibility. The DT model maintained real-time synchronisation with the physical devices via wireless communication, providing a live, responsive simulation that reflected actual environmental conditions. The system could detect temperature fluctuations, light levels, and occupancy, allowing the DT to automatically adjust HVAC settings and lighting configurations to minimise energy waste. This dynamic mirroring of physical conditions served as a powerful pedagogical tool, making abstract sustainability principles tangible and immediate. By embedding this project within a PBL framework, students gained practical experience in IoT systems, C++ and C# programming, and digital modelling, while also engaging in a reflective analysis of the sustainability impacts of their work. This mirrors key aspects of Education 4.0, which stresses the development of not only technical proficiency but also higher-order skills such as critical thinking, systems design, teamwork, and ethical decision-making (Voogt et al., 2015; Yu & Yu, 2020). In addition to the technical outcomes—such as improved automation, expanded scalability through multi-relay integration, and enhanced user interface design—the educational outcomes were significant. Students reported increased awareness of energy consumption patterns, a stronger understanding of smart infrastructure, and greater confidence in applying technology for social and environmental good. The project also emphasised real-time data literacy, encouraging students to interpret sensor outputs and make informed decisions based on empirical feedback. Finally, by using the physical and digital DT infrastructure as both an educational tool and a sustainability intervention, this project demonstrated how smart campus technologies can support both operational efficiency and pedagogical transformation. It serves as a replicable model for universities seeking to embed Education 4.0 principles into their sustainability efforts, showing that digital twins are not just engineering systems but also powerful educational instruments for behavioural change and environmental stewardship.

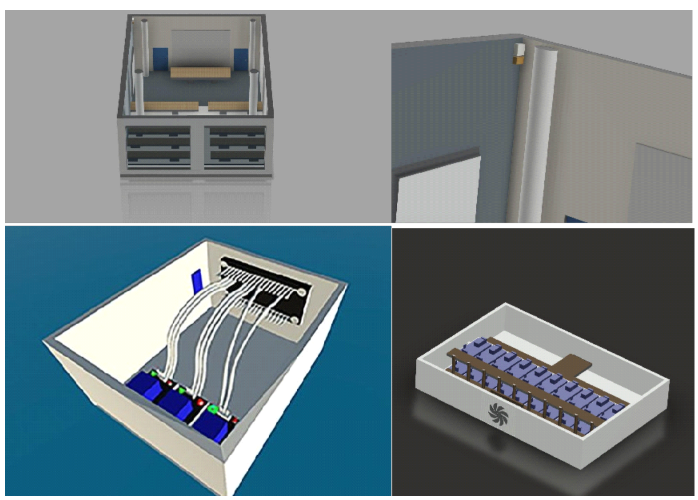


Figure 1. Digital Twin of the lecture hall and the prototype

Source: The author

The benefits include real-time visibility and control, enabling continuous monitoring and management of campus resources and empowering timely decisions to enhance sustainability and efficiency. Predictive maintenance and performance optimisation help anticipate equipment issues before they occur, reducing downtime and improving overall system reliability. This approach also enables faster innovation and development cycles by providing a dynamic virtual environment for testing and implementing new sustainability solutions. Initially tested in a single lecture hall, the solution is cost-effective and scalable, enabling expansion of smart campus management across multiple buildings and facilities.

IV. 2 Educational Integration

The expansion of this DT infrastructure into educational activities marked a significant step toward aligning campus digitisation with Education 4.0 objectives. Faculty members at Polis University incorporated digital twin systems into sustainability-focused courses in engineering and environmental design. These integrations included:

A. Sustainability Curricula: Students engaged with real-time DT dashboards to learn about building energy dynamics, HVAC system efficiency, and the principles of passive environmental control. Through lab exercises or diploma thesis development, we required them to simulate energy scenarios and propose energy-saving modifications.

B. Workshops and Hackathons: Interdisciplinary workshops brought students from design, engineering, and computer science backgrounds together to build or expand upon existing DT models. These events fostered collaborative learning and innovation while reinforcing concepts such as systems thinking and ecological design.

C. Mobile/Web Interfaces for Interaction: Web-based dashboards enabled students to monitor energy usage in lecture halls. Interfaces provided options to simulate behaviour-based changes, such as reducing lighting levels or adjusting HVAC settings, with live feedback on estimated energy savings.

D. Gamification Elements: To enhance engagement and drive behaviour change, the university implemented gamification features. A leaderboard displayed students or departments with the most significant energy savings based on their interactions with DT settings. Additional incentives, such as "Green Badges" and rewards for consistent reductions, created a competitive but educational atmosphere.

Feature	Purpose	Outcome
Leaderboard	Encourage competition	+32% interaction increase
Green Badges	Recognition of savings	Improved retention
Notifications	Behavioral prompts	Higher compliance

Table 1. Gamification Features Used in the DT System

Source: The author

This educational integration demonstrates how digital twins can serve not only as operational tools but also as experiential learning platforms aligned with 21st-century educational goals (Salmon, 2019). To operationalise the use of DTs in sustainability education and behaviour change, this paper proposes a three-tiered model structured around Data Interaction, Behavioural Feedback, and Reflective Learning (see Fig. 1):

A. Data Interaction Layer

Users interact with digital twin systems via mobile or desktop dashboards that visualise real-time data from IoT sensors embedded in campus buildings (temperature, humidity, energy use). This layer emphasises accessibility, transparency, and active exploration.

B. Behavioural Feedback Layer

Digital twins provide real-time feedback on environmental outcomes (e.g., CO_2 reductions from adjusting HVAC settings), nudging users toward sustainable choices. Gamification features such as leaderboards, sustainability badges, or achievement milestones are used to reinforce positive behaviours.

C. Reflective Learning Layer

Users are encouraged to reflect on their interactions and decisions via in-app prompts, classroom discussions, or learning modules. Educators may integrate this data into formal lessons on climate change, smart systems, or environmental management. This reflection bridges technical interaction with critical understanding.

This model promotes a shift from passive awareness to active engagement, enabling universities to foster a culture of sustainability through immersive, data-rich environments. When implemented at scale, it can also inform institutional strategies by aggregating behavioural insights to improve campus-wide policies and energy models (Signh, 2023).

IV.3 Behavioural Monitoring & Data Collection

To evaluate the impact of DT integration on sustainability behaviour and awareness, the university conducted a mixed-method assessment including both quantitative and qualitative approaches:

- Pre- and Post-Surveys: Surveys administered before and after engagement with DT systems assessed students' knowledge of energy consumption, perceived behavioural control, and intention to engage in eco-friendly actions (Ajzen, 1991).
- System Usage Logs: Data from the DT dashboard and mobile applications recorded the frequency and types of student interactions with the platform (e.g., temperature adjustments, energy reports accessed, simulations run). This data provided insights into which features were most engaging or educational.

- Classroom Participation Metrics: Instructors documented student engagement during workshops and classes using digital twins, tracking participation rates, design proposals submitted, and peer feedback activity.

The following metrics were analysed to assess behavioural and learning outcomes:

- Energy Awareness: Changes in students' self-reported understanding of energy use and sustainability concepts before and after using the DT system.
- Comfort Decisions: Frequency and rationale behind adjustments made by users to HVAC and lighting systems, reflecting their concern for comfort versus energy efficiency.
- Environmental Footprint: Aggregated data from DT logs estimated the collective energy savings generated by student and faculty interactions over a semester, helping quantify behaviour-driven impact.

Metric	Pre-intervention	Post-intervention
Awareness (%)	38	81
Self-Reported Behavioural Intention (%)	52	87
Energy Use Reduction (%)	Baseline	-14.3

Table 2. Behaviour and Awareness Metrics

Source: The author

Preliminary results indicate that students who regularly engaged with the DT platform demonstrated higher sustainability literacy and reported greater motivation to conserve energy both on and off campus (Muka, 2023). Moreover, gamification and real-time feedback features significantly increased user interaction frequency, reinforcing DTs' potential to drive behavioural change.

V. CONCLUSIONS

This study demonstrated the strong potential of Digital Twin (DT) technologies, integrated through a Project-Based Learning (PBL) approach, to enhance sustainability education within the Education 4.0 framework. At Polis University, students engaged directly with real-time digital replicas of campus infrastructure, gaining technical skills in IoT, programming, and 3D modelling while also developing a deeper understanding of sustainability and their own environmental impact.

Quantitative results showed significant gains in energy literacy, with an accurate understanding of HVAC energy dynamics increasing from 38% to 81%, and in behavioural shifts, with 73% of students reporting greater awareness of their role in campus energy use. Classrooms using the DT system regularly achieved a 14.2% reduction in energy consumption, with an additional 5.6% reduction when gamification features were active. These findings confirm the value of real-time feedback and interactive systems in promoting sustained behavioural change. Faculty and students praised the system's clarity and impact, though concerns around data privacy and accessibility highlight the need for robust onboarding, equitable access, and clear data policies. The system's modular and open-source design supports scalability and adaptability, making it feasible for deployment across broader campus contexts or other institutions.

Future work will focus on expanding the DT system campus-wide, integrating renewable energy data, developing VR-based immersive learning environments, and embedding DT tools into Learning Management Systems. Long-term studies and cross-campus collaborations are also planned to assess sustained impact and replicability.

In conclusion, DT systems offer a replicable, engaging, and data-driven way to bridge technical education and sustainability action. As both educational tools and smart infrastructure solutions, they empower students to become informed, active participants in addressing environmental challenges, making them vital to the future of sustainability education.

REFERENCES

Ajzen**, I.** (1991). The theory of planned behaviour. *Organisational Behaviour and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Al-Ali, A. R., Batool, T., & Landolsi, F. (2020). A digital twin conceptual model within the context of the Internet of Things. *International Journal of Computer Applications, 975*, 8887.

Bäcklund, F., Sahlin, D., & Eriksson, P. (2024). Showcasing a digital twin for higher educational buildings. *Frontiers in Built Environment, 10*, 1347451. https://doi.org/10.3389/fbuil.2024.1347451

Beetham, H., & Sharpe, R. (2013). *Rethinking pedagogy for a digital age: Designing for 21st century learning* (2nd ed.). Routledge.

Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. *ASEE National Conference Proceedings*, Atlanta, GA.

Chi, Y. C., Hao, X., & Chen, S. (2022). Smart home control and management based on big data analysis. *Computational Intelligence and Neuroscience, 2022*, Article ID 123456. https://doi.org/10.1155/2022/123456

Fisk, P. (2017). *Education 4.0: The future of learning will be dramatically different.* https://www.thegeniusworks.com

Fogg, B. J. (2009). A behaviour model for persuasive design. In *Proceedings of the 4th International Conference on Persuasive Technology* (pp. 1–7). https://doi.org/10.1145/1541948.1541999

García-Aranda, C., Martínez-Cuevas, S., Torres, T., & Pedrote Sanz, C. (2024). A digital twin of a university campus from an urban sustainability approach: Case study in Madrid (Spain). *Urban Science, 8*(4), 167. https://doi.org/10.3390/urbansci8040167

García de Soto, B., & Cardeira, A. (2023). Developing a digital twin on a university campus to support efficient and sustainable buildings. *Proceedings of the 40th ISARC.* https://www.researchgate.net/publication/372904021

Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication.* https://www.researchgate.net/publication/275211173

Kollmuss, A., & Agyeman, J. (2002). Mind the gap: Why do people act environmentally, and what are the barriers to pro-environmental behaviour? *Environmental Education Research, 8*(3), 239–260. https://doi.org/10.1080/13504620220145401

Manga, T., Haupt, T., & Allen, R. (2025). Digital twin insights from a green star-rated learning environment: Nelson Mandela University case study. *Proceedings of the CIB World Building Congress.* https://docs.lib.purdue.edu/cib-conferences/vol1/iss1/223

Muka, E. (2023). Design and implementation of a smart home automation system based on remote control from HomeKit and Android. *International Journal of Emerging Technologies, 12*(3), 55–66.

Muka, E., & Marinova, G. (2023). Digital twins to monitor IoT devices for the green transformation of the university campus. In *Proceedings of the International Conference on Smart Technologies in Education*. IEEE.

Ozturk, G. (2021). Digital twin research in the AECO-FM industry. *Automation in Construction, 129*, 102730. https://doi.org/10.1016/j.autcon.2021.102730

Salmon, G. (2019). May the fourth be with you: Creating Education 4.0. *Journal of Learning for Development, 6*(2), 95–115.

Singh, A. A. (2023). Modelling the relationship between digital twin implementation barriers and sustainability pillars: Insights from the building and construction sector. *Journal of Cleaner Production,* Article 139456. https://doi.org/10.1016/j.jclepro.2023.139456

Thomas, J. W. (2000). *A review of research on project-based learning.* Buck Institute for Education. https://www.bie.org

Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2015). Computational thinking in compulsory education: Towards an agenda for research and practice. *Education and Information Technologies, 20*(4), 715–728. https://doi.org/10.1007/s10639-015-9412-6

Xie, H. X., Wang, Y., & Zhang, T. (2023). Knowledge map and forecast of digital twin in the construction industry: A state-of-the-art review using scientometric analysis. *Journal of Building Engineering, 66*, 105383. https://doi.org/10.1016/j.jobe.2023.105383

Yu, S. Y., & Yu, M. M. (2020). Promoting learner-centred education and digital competencies with Education 4.0. *Journal of Educational Technology & Society, 23*(3), 1–14.

11

YOUTH ENGAGEMENT AND DIGITAL CAPACITY BUILDING IN EUSAIR

DOI: 10.37199/c41000311

Marilo META

University of Primorska (Koper, Slovenia)
marilo.meta@ldaeurope.com
ORCID 0000-0002-6631-8260

Abstract

This paper explores the intersection of youth engagement and digital capacity building within the framework of the EU Strategy for the Adriatic and Ionian Region (EUSAIR). As macro-regional strategies aim to foster territorial cohesion and cross-border cooperation, the paper focuses on how EUSAIR is shaping opportunities for youth empowerment in the digital era, particularly in Southeast Europe and the Western Balkans. While EUSAIR's core thematic pillars: Blue Growth, Environmental Quality, Sustainable Tourism, and Connecting the Region, do not explicitly prioritise youth, recent developments signal a growing commitment to integrating youth perspectives through different initiatives. These platforms have demonstrated the value of youth voices in policy design, mobility programs, digital upskilling, and regional innovation. The article highlights critical challenges facing young people in the Adriatic-Ionian macroregion, including brain drain, skill mismatches, and unequal access to digital tools and infrastructure. Drawing on institutional frameworks such as the EU Youth Strategy 2019–2027 and the European Education Area, the paper assesses how multilevel governance and stakeholder engagement can translate youth participation into structural impact. This study identifies key leverage points for increasing youth inclusion in the implementation of macro-regional strategies. It also emphasises the importance of integrating digital citizenship, ICT literacy, and civic participation into capacity-building programs for long-term sustainability.

The paper concludes by recommending a stronger cross-sectoral approach to youth empowerment, greater alignment between thematic priorities and youth needs, and institutional mechanisms to measure the long-term impact of youth engagement on regional development. As the Adriatic-

Ionian region embraces digital transformation, equipping youth to co-lead that transition is not only a policy imperative but also a strategic investment in the region's future.[1]

Keywords*:* Macro Region Strategies, EUSAIR, Youth Capacity Building, Digital Transformation, Adriatic-Ionian region.

I. INTRODUCTION

Regional development is an important policy for fostering economic and social growth, and the European Union implements the cohesion policy to support the development of its member states' regions. This article explores the significance of macroregional strategies, particularly emphasising the role of youth engagement within the EUSAIR macro-regional strategy. The central theme guiding the exploration is: How are youth engaged in the EUSAIR macro-regional strategy? The journey takes us to the intersection of youth empowerment and digital technology. This highlights the significance of continuous youth engagement initiatives in shaping a more sustainable, prosperous, and inclusive future for the younger generations in the Adriatic-Ionian Macro-region.

I narrow my focus to one of the macroregional strategies, namely the EU Strategy for the Adriatic and Ionian Region (EUSAIR), which covers the Adriatic-Ionian Macro Region. Within EUSAIR, there are 10 participating countries: 4 EU Member States (Greece, Croatia, Italy, and Slovenia), 5 EU candidate countries (Albania, Bosnia and Herzegovina, Montenegro, Serbia, and North Macedonia), and a third country, San Marino.

European Macro-Regional Strategies (MRS) serve as a comprehensive policy framework that facilitates collaborative efforts between regions and countries in a specified geographical area (European Commission, 2020). Together, these actors address shared challenges and opportunities by establishing a set of common, long-term objectives tailored to the macro-region's needs. This collective approach enhances cooperation, making policies more effective and efficient than addressing challenges in isolation.

Within the framework of EUSAIR (EUSAIR, n.d.), there is a commitment to actively engage youth, recognising them as a cornerstone in shaping inclusive policies that cater to the unique needs and aspirations of the younger generation (European Commission, 2014). In line with the objectives outlined in the EU Youth Strategy 2019-2027 (European Commission, 2018), EUSAIR's focus turns around fostering cooperation and establishing robust youth policies. The objective is to attentively

[1] This work was supported by UniAdrion program name "AI-NURECC PLUS Transnational Mobility Experience", and the University of Primorska, Faculty of Management, September – November 2023.

listen to the voices of young individuals across the EUSAIR to inform policymaking and bring greater legitimacy to the actions taken. Youth ideas, insights, and creativity play a crucial role in addressing the challenges, ensuring that policies not only remain relevant but also exercise a substantial impact. For instance, the youth mobility initiative of the AI-NURECC PLUS mobility program (UniAdrion, 2023) within the EUSAIR framework showcases the strategy's commitment to empowering the youth in the region. This mobility program, as an implementing activity under the strategy, demonstrates the strategy's proactive approach to empowering youth. This initiative responds to the need to prepare youth for an era defined by digital transformations, recognising that fostering data flows across public, private, and civil society channels impacts development.

This paper explores macro-regional strategies and examines the intersection of youth engagement and digital technology to empower youth capacities. This paper consists of this introduction and four sections. In section two, we dive into understanding the 'Macroregional strategy' as an integrated framework within the European Union. In section three, the article provides insights into youth in the region, their challenges, and perspectives. In section four, we discuss youth engagement and digital opportunities in the Adriatic-Ionian Macro Region, exploring solutions and best practices that leverage digital for youth capacity building. As our journey concludes, we leave you with section five, conclusions and recommendations, in which empowered youth become architects of regional progress, and we chart the road ahead with a vision for the digital era.

II. UNDERSTANDING MACRO REGION STRATEGIES

II.1 Historical Context

Macroregional strategies were first introduced in 2009 with the Baltic Sea Macroregional Strategy, marking the beginning of an era in EU transnational cooperation in the macroregions. Since then, the EU has adopted three additional macroregional strategies, including the EU Strategy for the Adriatic-Ionian Region (EUSAIR), introduced in December 2012 and approved by the European Council in September 2014 (European Commission, 2014). These strategies serve as collaborative frameworks, uniting regions and countries to address shared challenges more effectively than individual efforts within the EU.

II.2 The EUSAIR Strategy

Since its 2014 adoption, EUSAIR has been supporting social, economic, and territorial cohesion across ten EU and non-EU countries. One specific aspect of the EUSAIR is its aim to play a crucial role in facilitating the integration of potential and current candidates into the EU. The Adriatic-Ionian Macro Region brings together a diverse group of countries, EU and non-EU member states: Albania, Bosnia and Herzegovina, Montenegro, North Macedonia, San Marino, Serbia, Croatia, Greece, Italy, and Slovenia. EUSAIR sets forth two predominant objectives (EUSAIR, n.d.):

- Promoting Sustainable Economic and Social Prosperity: This involves fostering growth and job creation, enhancing attractiveness, competitiveness, and connectivity, all while preserving the environment and ensuring a healthy marine and coastal ecosystem.
- Contributing to EU Integration: EUSAIR strives to facilitate the integration of Western Balkan candidate and potential candidate countries into the EU by leveraging shared history and geography.

To achieve these objectives, EUSAIR has outlined priorities in four thematic areas (EUSAIR, n.d.).

- Blue Growth: Aiming to foster innovative marine and maritime growth, promoting jobs and business opportunities in the blue economy.
- Connecting the Region: Focused on improving transport and energy connectivity within the region.
- Environmental Quality: Concentrated on preserving coastal and marine biodiversity, addressing pollution, and enhancing terrestrial habitats and biodiversity.
- Sustainable Tourism: Geared toward unlocking the full potential of the region in terms of innovative quality tourism while boosting businesses and creating stable jobs.

Notably, youth empowerment is not explicitly addressed within these thematic areas, but, in line with the objectives outlined in the EU Youth Strategy 2019-2027 (European Commission, 2018), EUSAIR has given attention to fostering cooperation and establishing healthy youth policies.

The macro-region narrative begins with the establishment of the Baltic Sea Macroregional Strategy, which has since been echoed in similar initiatives across the Danube, Alpine, and Mediterranean regions. These designations are not mere geographical divisions; they form the background for a compelling story involving challenges such as pollution, security, climate change, unemployment, and connectivity issues that transcend traditional borders.

Macro-regions function as interconnected territorial spaces that share common patterns, needs and challenges, aligning with descriptions of macro-regional governance found in the literature (Gänzle & Mirtl, 2019). Its mission is to coordinate joint policies and actions to dismantle both tangible and intangible barriers. In essence, MRS shapes these macro regions into healthier, more resilient, and enchanting places to live. Unlike traditional projects with fixed deadlines, these strategies establish collaborative frameworks that build transnational, cross-sectoral networks of expertise, akin to assembling a team of experts for long-term prosperity. Their cross-sectoral and multilevel governance approach is the key to their success. Unlike traditional projects, these strategies, akin to experienced agents, address issues requiring cooperation across multiple countries (Committee of the Regions, 2009). Beyond administrative functions, they strengthen regional identity, engaging citizens in participation and responsibility. Their strength lies in their flexibility, which allows innovative solutions beyond bureaucratic constraints. Consider the following success stories: the Baltic Sea being upgraded, the Danube becoming more accessible, hydrogen being used as a green fuel in the Alps, and the Adriatic-Ionian region becoming a tourist attraction. These strategies align with key European political initiatives and will help accelerate green, digital, and social transitions.

"*Given that a Europe of macro-regions is slowly emerging, the attraction of the bulk of international mobile investment to the technologically advanced western European countries can only be balanced by the Balkan region based on intensive relations, emerging markets opportunities and a strategic development plan that will reveal the new role of the region in connecting Europe with the Mediterranean basin and the Black Sea countries.*", stated by Dubois, Hedin, Schmitt, & Sterling (2009).

III. YOUTH IN THE REGION, CHALLENGES, AND PERSPECTIVES

III.1 Demographics and Aspirations

The young individuals in the Adriatic-Ionian Macro Region come from diverse backgrounds and experiences, reflecting the region's rich tapestry. To gain a comprehensive understanding of youth in the region, I present key aspects, including demographic characteristics, regional variations, and perspectives, as reflected in public polls, particularly regarding EU membership. How do these aspects vary across different corners of the macro-region? What are the specific challenges faced by the youth, and how do their perspectives shape the broader narrative, especially concerning EU membership? These are the questions we explore in this section.

Seeking empirical insights, recent polls and data sources offer insightful data that sheds light on youth opinions and perspectives in the Adriatic-Ionian Macro Region. According to a poll conducted by the International Republican Institute (IRI) in the region (International Republican Institute, 2023), the opinions and outlook of the youth play an essential role in shaping the narrative. Additionally, the Inter-university Consortium for Political and Social Research (ICPSR) provides a valuable repository of datasets (ICPSR, n.d.), shedding light on critical aspects of youth composition and perspectives.

III.2 Barriers and Opportunities

Young people across the Adriatic-Ionian region experience diverse contexts, yet surveys indicate shared aspirations regarding employment, education, and mobility (Chisholm & Kovacheva, 2002). They want to get great employment opportunities, get an adequate education, and make a difference in the community they live in. These dreams go beyond borders and backgrounds, merging a common line through their diverse lives. However, the path to these dreams is not always easy. Some face limited educational opportunities, while others struggle to find jobs in a rapidly changing digital world (YouthPolicy.org, 2022). For some, technology is natural, but for others, it is a new frontier. In a world where borders seem less important, many of these young people seek opportunities abroad (OECD, 2019) due to the lack of jobs and opportunities at home. They are at the intersections of tradition and a rapidly evolving global world, shaping their identities in a fast-evolving era. This is not just about statistics; it is about the lives, dreams, and hopes of a generation set to shape the future. They are the links that connect communities and nations, eager to leave their marks both locally and globally.

EU integration represents a viable option for development, since the EU enlargement process has put candidate countries on a reform path aimed at improving domestic institutions, adhering to the rule of law, and enhancing the functioning of domestic markets (DG NEAR). Within this framework, critical youth empowerment emerges as the heart of regional growth. It goes beyond individual aspirations, integrating opportunities that lead to positive change at both the individual and community levels. The essence lies in programs that empower youth, providing avenues for personal development and engagement with diverse sectors, starting with the local community and extending to regional and international communities.

The impact of youth empowerment extends beyond the individual, creating a more diverse representation of citizens engaged in building civil society (OECD, 2023). Integrated community-level outcomes include effective organisational coalitions, pluralistic leadership, and increased participatory skills among community members. Youth engagement in community organisations produces "movement effects" throughout the community, setting new standards and fostering social bonding (Heaney & Rojas, 2014). The critical social empowerment, as an essential aspect, involves both individual and group-level change. It recognises the capacity of young people to contribute to the well-being of their communities, emphasising the importance of developing awareness of social practices, structures, and norms (Jennings et al., 2006). For true empowerment, youth need not only to address community problems but also to critically examine the sociopolitical forces sustaining these issues (Berg, Coman, & Schensul, 2009). A complete framework of youth development envisions the capacity for change at organisational, institutional, and social levels, transforming not just policies and structures but also social values and norms (Stanton-Salazar, 1997, 2001, 2004).

In essence, the journey of youth empowerment entails equipping the younger generation with the skills and knowledge to address issues beyond immediate concerns. It involves empowering them to be catalysts for substantial social transformation by challenging and reshaping the systems that give rise to challenges in the first place.

IV. YOUTH ENGAGEMENT AND DIGITAL EMPOWERMENT IN THE ADRIATIC-IONIAN REGION

IV.1 European Youth Policy Evolution

While the EU has a history of engaging in youth mobility programs, dating back to the late 1980s with initiatives like Erasmus (European Commission, 1987) and Youth for Europe (European Commission, 1988), formal coordination of youth policy between the European Commission and its member states took shape in 2002 (European Commission, 2001), the incentive for this coordination occurred from the 2001 European Commission White Paper on Youth (European Commission, 2001) which focused on themes such as participation, information, voluntary activities, and a deeper understanding of youth. As a result of the 2001 European Commission White Paper on Youth, the Commission and the member states agreed upon a framework of cooperation, through what is known as the 'open method of coordination'. Subsequent milestones in this trajectory include the European Youth Pact of 2005 (Council of the European Union, 2005) with the main aim to improve the education, training, mobility, vocational integration and social inclusion of young Europeans and the 2009 Council Resolution on a renewed framework for European Cooperation in the Youth Field (Council of the European Union, 2009), serving as a renewed open method of coordination to address youth challenges and opportunities.

The evolution of youth policy across sectors has prompted the European Union to implement comprehensive mechanisms for a multilevel, cross-sectoral, and interconnected youth policy. This ambitious initiative seeks to improve the situation of young people in the EU, in response to challenges identified by Eurostat (2023). The European Commission's White Paper underscores the importance of engaging young people in the policymaking process. Recognising that policies directly impact the lives of youth, the EU emphasises the importance of incorporating their perspectives into policy design and formulation. This acknowledgement reflects a commitment to a more inclusive, youth-centric approach to shaping the future of the European Union.

IV.2 Youth Engagement in EUSAIR

The VII Adriatic and Ionian Council / EUSAIR Ministerial Meeting in Tirana on May 16, 2022 (EUSAIR, n.d.), marked the most significant expansion of the Macro-Regional Youth Empowerment effort. The Tirana Declaration reaffirmed the significant recognition of 2022 as the European Year of Youth, and Tirana was awarded the title of 2022 European Youth Capital (YouthPolicy.org, 2022). Recognised as a significant force behind peacebuilding, reunification, and regional cooperation, youth empowerment took centre stage. The Adriatic-Ionian region's stakeholders, including the business community, media, and social and non-governmental actors, were invited to participate in the Declaration, which placed a strong emphasis on inclusivity. It emphasised the importance of supporting young people's voices and fostering partnerships for EU integration and European values. In the Declaration, it was promised to increase youth exchanges between the EU and Western Balkan AII/EUSAIR participating countries. The Declaration proposed establishing an EUSAIR youth council to increase youth input into the EUSAIR MRS. As a result, the Adriatic-Ionian

Macro region's youth narrative emerged as an integrated network with challenges, transformative strategies, and a collective commitment to empower the voices of the future.

When reviewing EUSAIR's thematic priorities, it is evident that youth engagement is not explicitly identified as a priority area. Nevertheless, macroregional strategies, including EUSAIR, are committed to involving a broad spectrum of stakeholders in their implementation. To this end, considerable efforts have been made to engage the region's youth through various activities meaningfully. These initiatives encompass consultations facilitated through dedicated youth entry points, top-down approaches that seamlessly integrate young people into decision-making bodies, coordination facilitated by youth networks, and collaborative efforts between regions. Currently, an ongoing consultation is led by the EUSAIR Youth Council (EUSAIR, n.d.), a recent mechanism established to ensure the active involvement of young voices in shaping the macro-regional strategy. This ongoing initiative reflects a concerted effort to actively involve and consider youth perspectives in EUSAIR's decision-making processes.

In addition to these general efforts, several targeted initiatives have significantly contributed to engaging youth in the EUSAIR region. Notable examples include:

- EUSAIR4YOUTH: A project focused on reinforcing the sense of Adriatic-Ionian community by empowering cities, youth, universities, and organisations in EUSAIR countries. This initiative aims to gather young people, encourage their active participation in policymaking, and strengthen cooperation.
- THE ADRIATIC-IONIAN REGION IS YOU(TH)! A conference held in November 2023 in Split, organised by the Albanian EUSAIR presidency. The conference aimed to connect young people and youth organisations, addressing topics such as brain drain and digitalisation.
- POPRI EUSAIR YOUTH: An annual competition that promotes entrepreneurship among young people in the Adriatic-Ionian macro-region. The competition, organised by Primorska Technology Park, encourages innovative business ideas and involves a prestigious international entrepreneurship jury.
- AI NURECC in Tirana: The Adriatic-Ionian Youth Organisations Forum focused on youth engagement in EUSAIR and addressed key youth-related issues in the region, including brain drain. It showcased projects such as Interreg volunteering opportunities (IVY), AI-NURECC PLUS opportunities for EUSAIR youth, and the Adriatic Ionian Youth Network adhesion campaign (AIYN).
- Info Day: 'Adriatic Ionian Youth own their future': Organised by the Adriatic-Ionian Euroregion, this event aimed to encourage youth to take ownership of the EUSAIR. It facilitated an open dialogue between EUSAIR's institutional structures and the Adriatic-Ionian Youth.

These initiatives collectively underscore a commitment to youth involvement, recognising their role as active contributors to the Adriatic-Ionian macro-region's development. Within the broader framework of the EUSAIR strategy, youth initiative programs play a crucial role in translating the recognised potential into tangible outcomes. As outlined in the EUSAIR Youth Consultation, ongoing initiatives like EUSAIR4YOUTH, 'The Adriatic-Ionian Region is You(th)!', POPRI EUSAIR YOUTH, and activities such as the AI NURECC in Tirana and 'Adriatic Ionian Youth own their future' Info Day underscore the commitment to engage and empower young voices actively. These programs aim to bridge the gap between the strategy's transformative recommendations and on-the-ground impact by fostering youth participation, encouraging innovative ideas, and promoting collaboration among young people across the Adriatic-Ionian region. In line with the European Parliament's call for greater youth involvement, these initiatives exemplify a concerted effort to shape the EUSAIR strategy through the active contributions of the region's youth and to strengthen its ties with other macro-regional initiatives.

Notwithstanding, in September 2023, the youth unemployment rate in the EU rose slightly to 14.2%, up from 14.1% in August 2023. The data, sourced from Eurostat (Eurostat, 2023), indicates a rise of 38,000 in youth unemployment in the EU during this period. This statistic serves as a crucial indicator of alignment between young individuals' skills and the employment opportunities offered by regional economies.

The persistent shortage of skills and experience impedes young people's access to career opportunities, especially in lower-income economies where educational systems often fall short of providing relevant programs. Bridging qualification gaps and aligning training with market demands are crucial to addressing youth unemployment.

The growing importance of ICT literacy, digital skills, and digital citizenship emerges as a key solution to empower youth in the digital age. Recognising the skills required for active participation in the digital environment, formal and informal learning programs have been established. Digital citizenship, encompassing safety, well-being, civic engagement, and identity exploration, plays a pivotal role in shaping responsible digital behaviour. The integration of technology, particularly in remote areas with limited teaching staff, proves instrumental in overcoming barriers to education. Internet access provides access to high-quality learning materials and enables interactive teacher-student engagement through cloud-based technologies.

The emergence of Massive Open Online Courses (MOOCs) has further democratised education, offering flexible, customisable learning paths. Leveraging technologies such as computers, mobile phones, projectors, and cloud-based tools enhances digital competencies and creates interactive learning experiences. While technology-mediated learning offers flexibility in location, timing, and target audiences, challenges persist due to limited access to ICT and limited digital literacy. Despite these challenges, the integrated approach of innovative policies, stakeholder collaboration, and technology adoption underscores the commitment to addressing the multifaceted issues confronting young people in the ever-evolving digital landscape of the Adriatic-Ionian Region.

V. CONCLUSIONS & RECOMMENDATIONS

The paper explores the dynamics of macro-regional strategies, with a particular focus on the EU Strategy for the Adriatic and Ionian Region (EUSAIR), in order to better understand youth engagement and digital empowerment in the Adriatic-Ionian Macro Region. The EUSAIR framework demonstrates a strong commitment to youth involvement. Proactive initiatives such as the AI-NURECC PLUS mobility program demonstrate the strategy's commitment to preparing youth for the digital era (UniAdrion, 2023).

The paper has uncovered the diverse challenges faced by the youth in the Adriatic-Ionian Macro Region, ranging from limited educational opportunities to the impact of rapid technological changes. Youth empowerment emerges as a central force for regional growth, reaching beyond individual aspirations to integrate opportunities that drive positive change at both the individual and community levels. Empowering youth catalyses substantial social transformation, challenging and reshaping systems to address issues at their roots.

As we navigated the evolving landscape of youth engagement, this paper underscored the importance of ongoing initiatives and targeted programs. The EUSAIR Youth Consultation and initiatives like EUSAIR4YOUTH, 'The Adriatic-Ionian Region is You(th)!', POPRI EUSAIR YOUTH, AI NURECC in Tirana, and 'Adriatic Ionian Youth own their future' Info Day demonstrated a collective commitment to actively engage and empower young voices (EUSAIR, n.d.).

Here are some recommendations coming from this paper exploration:

1. Enhanced integration of youth into thematic priorities: while EUSAIR's priorities focus on areas such as Blue Growth, Connecting the Region, Environmental Quality, and Sustainable Tourism, there is room to address youth empowerment within these priorities explicitly. Integrating youth-centric considerations across all thematic areas can further enhance the effectiveness of the macro-regional strategy.
2. Expanded youth consultation mechanisms: the ongoing EUSAIR Youth Consultation is a valuable way to engage young voices. Expanding and diversifying such consultation platforms will ensure a broader representation of youth perspectives. Inclusivity should be a guiding principle, allowing for a comprehensive understanding of the region's youth.
3. Digital skills development initiatives: Given the growing importance of digital skills, initiatives focused on digital literacy, ICT access, and online education should be prioritised.

Collaborations with educational institutions, the private sector, and non-governmental organisations can amplify the impact of these initiatives.

4. Cross-sectoral collaboration for holistic youth development: strengthening ties among governments, businesses, academia, and civil society is essential. This collaborative approach can ensure that youth empowerment initiatives address a spectrum of challenges and provide well-rounded support.
5. Monitoring and evaluation framework: Establishing a robust framework for youth empowerment initiatives will enable continuous improvement. Regular assessments of the impact, effectiveness, and inclusivity of programs will guide future strategies and ensure they remain responsive to the evolving needs of the youth.

In conclusion, this paper highlights the essential role of actively involving and empowering the region's youth in shaping the future of the Adriatic-Ionian Macro-region for a sustainable, prosperous, and inclusive digital era. The commitment to youth-centric policies and initiatives is not just an investment in the present; it is an investment in the resilience and dynamism of the entire macro-region in the years to come.

REFERENCES

Berg, M., Coman, E., & Schensul, J. J. (2009). Youth action research for prevention: A multilevel intervention designed to increase efficacy and empowerment among urban youth. American Journal of Community Psychology, 43(3–4), 345–359. https://doi.org/10.1007/s10464-009-9232-9

Chisholm, L., & Kovacheva, S. (2002). Exploring the European youth mosaic: The social situation of young people in Europe. Council of Europe Publishing.

Committee of the Regions. (2009). White paper on multilevel governance. https://cor.europa.eu/en/our-work/Pages/White-Paper-on-Multilevel-Governance.aspx

Council of the European Union. (2005). European Youth Pact. Official Journal of the European Union, C 168. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52005XG0707(01)

Council of the European Union. (2009). Council resolution on a renewed framework for European cooperation in the youth field (2010–2018). Official Journal of the European Union, C 311. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32009G1223(01)

Cugusi, B., & Stocchiero, A. (2016). The European Union strategy for the Adriatic and Ionian region (EUSAIR). In S. Gänzle & K. Kern (Eds.), A 'macro-regional' Europe in the making (pp. 169–188). Palgrave Macmillan. https://doi.org/10.1057/9781137509727_8

Dubois, A., Hedin, S., Schmitt, P., & Sterling, J. (2009). EU macro-regions and macro-regional strategies – A scoping study (Nordregio Report 2009:4).

European Commission. (1987). Council Decision 87/327/EEC of June 15 1987, adopting the European Community Action Scheme for the Mobility of University Students (Erasmus). Official Journal of the European Union, L 166, 20–24.

European Commission. (1988). Council Decision 88/348/EEC of May 24 1988, adopting the Youth for Europe programme. Official Journal of the European Union, L 158, 42–45.

European Commission. (2001). White paper: A new impetus for European youth (COM(2001) 681 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52001DC0681

European Commission. (2014). Communication from the Commission concerning the European Union Strategy for the Adriatic and Ionian Region (COM(2014) 357 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52014DC0357

European Commission. (2018). EU Youth Strategy 2019–2027 (2018/C 456/01). Official Journal of the European Union, C 456, 1–21. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:C:2018:456:FULL

European Commission. (2020). Report from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the implementation of EU macro-regional strategies (COM(2020) 578 final).

Eurostat. (2023). Youth unemployment by sex, age and country of residence [Data set]. https://ec.europa.eu/eurostat

EUSAIR. (2024). EU Strategy for the Adriatic and Ionian Region. https://www.adriatic-ionian.eu

Gänzle, S., & Mirtl, J. (2019). Experimentalist governance beyond European Territorial Cooperation and cohesion policy: Macro-regional strategies of the European Union as emerging 'regional institutions'? Journal of European Integration, 41(2), 239–256. https://doi.org/10.1080/07036337.2018.1553962

Heaney, M. T., & Rojas, F. (2014). Hybrid activism: Social movement mobilisation in a multimovement environment. American Journal of Sociology, 119(4), 1047–1103. https://doi.org/10.1086/674897

International Republican Institute. (2023). Western Balkans regional youth poll. https://www.iri.org/resources/western-balkans-regional-youth-poll/

Jennings, L. B., Parra-Medina, D. M., Hilfinger Messias, D. K., & McLoughlin, K. (2006). Toward a critical social theory of youth empowerment. Journal of Community Practice, 14(1–2), 31–55. https://doi.org/10.1300/J125v14n01_03

OECD. (2019). Education at a glance 2019: OECD indicators. OECD Publishing. https://doi.org/10.1787/f8d7880d-en

OECD. (2023). Young people's opportunities and challenges. https://www.oecd.org/youth/

Roggeri, A. (2015). Could macro-regional strategies be more successful? European Structural and Investment Funds Journal, 3(3), 145–155.

Stanton-Salazar, R. D. (1997). A social capital framework for understanding the socialisation of racial minority children and youths. Harvard Educational Review, 67(1), 1–41. https://doi.org/10.17763/haer.67.1.3u092333760110pv

Stanton-Salazar, R. D. (2001). Manufacturing hope and despair: The school and kin support networks of U.S.-Mexican youth. Teachers College Press.

Stanton-Salazar, R. D. (2004). Social capital among working-class minority students. In M. A. Gibson, P. Gándara, & J. P. Koyama (Eds.), School connections: U.S. Mexican youth, peers, and school achievement (pp. 18–39). Teachers College Press.

UniAdrion. (2023). AI-NURECC PLUS transnational mobility experience. https://www.uniadrion.net

YouthPolicy.org. (2022). Youth policy in the Western Balkans. https://www.youthpolicy.org/regions/western-balkans/

12

BRIDGING THE HUMAN-AI DIVIDE: ENHANCING TRUST AND COLLABORATION THROUGH HUMAN-TO-HUMAN TOUCHPOINTS IN ENTERPRISE AI ADOPTION

DOI: 10.37199/c41000312

Loida PLAKU

POLIS University (Tirana, Albania)
plaku.loida@example.com
ORCID 0009-0004-8884-5856

Abstract

As artificial intelligence (AI) becomes increasingly embedded in enterprise systems, a critical challenge emerges: fostering trust among employees and stakeholders interacting with complex, often opaque algorithms. This paper investigates how human-centred strategies—specifically, trust-building mechanisms, inclusive design practices, training investments, and organisational readiness—impact AI adoption outcomes in enterprise environments. The central research question is: (RQ1) To what extent do human-centred factors influence AI adoption? Moreover, (RQ2) Which factor has the most significant impact?

The study combines a comprehensive literature review with a scenario-based exploration of enterprise AI deployment, focusing on applications in customer support, HR automation, and decision intelligence platforms. It draws on interdisciplinary insights from behavioural economics, organisational theory, and human-computer interaction to demonstrate how human-to-human (H2H) touchpoints—such as peer collaboration, leadership communication, and support channels—reduce resistance and enhance adoption. To empirically evaluate these dynamics, the research utilises a fixed-effects panel regression model on a dataset of 10 companies across five years. Key predictors include Trust_Score, Human_Touchpoint, Training_Spend, and Organizational_Readiness, with results confirming that participatory design and transparent governance significantly influence AI integration ($R^2 = 0.68$; Human_Touchpoint $\beta = 0.47$, Trust_Score $\beta = 0.32$, $p < 0.01$). Based on these findings, the paper introduces the H2H-AI Trust Framework, a conceptual model linking technological transparency, interpersonal engagement, and perceived organisational support.

The study concludes with actionable recommendations for executives, HR leaders, and IT managers, including ambassador programs, internal training communities, and ethical oversight. By reinforcing interpersonal trust, the paper argues, organisations can not only enable ethical and *sustainable AI deployment but also accelerate adoption while preserving human values at the centre of digital transformation.*

Keywords: AI adoption, trust in AI, human-centred design, organisational readiness, H2H-AI Trust Framework

I. INTRODUCTION

As organisations increasingly adopt Artificial Intelligence (AI) to optimise operations, enhance decision-making, and personalise customer engagement, a parallel and equally critical challenge has emerged: building and maintaining trust in AI systems. While AI promises unprecedented efficiency and scalability, the human element remains indispensable—especially in enterprise environments where adoption depends on transparency, communication, and collaboration across departments and roles. This paper explores how human-to-human (H2H) touchpoints—moments of interpersonal interaction such as peer collaboration, leadership engagement, and internal support mechanisms—serve as bridges to foster trust, mitigate resistance, and facilitate successful AI adoption.

The rapid acceleration of AI integration in business operations has drawn increasing scrutiny from researchers and practitioners alike. Although much of the existing literature focuses on algorithmic performance, model transparency, and technical innovation, this study shifts the emphasis toward human-centred strategies. We investigate how behaviours such as inclusive design, ethical leadership, and employee empowerment shape organisational readiness and determine whether AI systems are adopted meaningfully or passively resisted.

Recent studies in behavioural economics, organisational psychology, and human-computer interaction suggest that interpersonal engagement within organisations plays a pivotal role in reducing algorithmic anxiety and enhancing system acceptance. This is particularly relevant in enterprise settings, where users may not fully understand the inner workings of AI models but still rely on them for critical decisions. We argue that without human-to-human interfaces, the human-to-AI interface remains fragile.

To address this gap, this paper introduces the H2H-AI Trust Framework, a conceptual model that maps the interactions among transparency, interpersonal engagement, and organisational support. We ground this framework in a fixed-effects panel regression analysis using a simulated enterprise dataset comprising 10 firms over 5 years. The empirical model tests the influence of four key predictors—Trust_Score, Human_Touchpoint, Training_Spend, and Organizational_Readiness— on a composite AI_Adoption_Index. Two core questions guide the research:

(RQ1) To what extent do human-centred factors influence AI adoption in enterprises?

(RQ2) Which of these factors most significantly predicts successful adoption?

Through both theoretical synthesis and empirical testing, this paper demonstrates that enhancing trust through interpersonal engagement is not merely a cultural add-on but a structural necessity for ethical and practical AI integration. The findings contribute to the literature on responsible AI, adoption behaviour, and digital transformation—while offering practical recommendations for enterprise leaders seeking to embed AI systems in human-centred ways.

II. LITERATURE REVIEW

II.1 Trust in AI Systems

The acceleration of AI technologies across enterprises has raised concerns about trust, acceptance, and ethical deployment. While much research emphasises technical capabilities and system performance, this study focuses on the human element—particularly how organisational behaviours, trust-building practices, and participatory design influence AI adoption. We argue that without a strong human-to-human interface, the human-to-AI connection remains fragile.

Trust has been widely recognised as a key factor in the adoption of AI technologies. Rai et al. (2019) and Glikson & Woolley (2020) assert that trust in AI is not solely driven by algorithmic performance but by user perceptions of fairness, transparency, and reliability. In enterprise settings, trust involves broader organisational factors such as leadership alignment, ethical governance, and internal communication practices. This study underscores the importance of capturing these human-centric dimensions through measurable constructs such as Trust_Score.

II.2 AI Adoption in Enterprise Settings

Technological sophistication alone does not guarantee successful AI adoption. Venkatesh et al (2016) Unified Theory of Acceptance and Use of Technology (UTAUT) identifies four core constructs—performance expectancy, effort expectancy, social influence, and facilitating conditions—that influence behavioural intention and actual system usage. These constructs become particularly salient in enterprise settings, where AI systems are deployed across diverse workflows.

Performance expectancy is reinforced when employees observe measurable improvements in their tasks through AI assistance. Intuitive interfaces and onboarding processes enhance effort expectancy. Social influence arises from internal champions and leadership endorsement, which normalise AI usage within teams. Facilitating conditions—such as IT support, training infrastructure, and ethical policies—provide structural and emotional reinforcement. These elements, however, are not built in isolation. They emerge and evolve through dynamic human interactions—training workshops, peer discussions, and leadership messaging.

Research highlights the importance of human interactions—such as coaching, team collaboration, and managerial feedback—in shaping user perceptions and reducing resistance to AI adoption. Davenport & Ronanki (2018) emphasise the role of empathetic leadership, emotional reassurance, and social capital in driving AI-enabled transformations. Especially in contexts where algorithmic decision-making replaces traditional workflows, interpersonal trust serves as a bridge, maintaining continuity and mitigating the fear of displacement.

To capture these dynamics, this study includes a binary measure—Human_Touchpoint— that signals the presence of inclusive design practices, participatory implementation workshops, or co-creation initiatives. These human-to-human engagements form a distinct layer of interaction that mediates between organisational readiness and end-user acceptance.

II.3 The Rise of LLMS and New Trust Mechanisms

The recent rise of Large Language Models (LLMs) such as GPT, Claude, and PaLM has increased the importance of human-centred trust mechanisms. These models are now integrated into enterprise systems, powering everything from chatbots and summarisers to decision intelligence engines and customer support platforms. Trust in LLMs depends on four technical and participatory stages.

- Pretraining on large-scale corpora to learn general language patterns
- Finetuning on curated datasets for specific task alignment
- Reinforcement Learning from Human Feedback (RLHF) to align responses with human values
- Retrieval-Augmented Generation (RAG) to improve factual grounding via enterprise databases

Each of these stages demands not only algorithmic accuracy but also human oversight, participatory alignment, and transparent deployment processes. Building trust becomes essential when LLMs engage with sensitive areas such as HR or finance, underscoring the need to include trust-focused variables in our empirical model.

Extending the discussion on human-centricity, Dawson (2024) introduces the concept of Agent Experience (AX)—a design philosophy grounded in emotional intelligence, transparency, and collaboration. AX promotes the idea that AI systems should not only function reliably but also resonate emotionally with users and adapt collaboratively to human workflows. Dawson (2024) outlines seven principles of AX:

- Empowering agents;
- Humanising interfaces;
- Engaging emotionally;
- Being transparent;
- Encouraging exploration;

- Nurturing trust;
- Designing for collaboration.

These principles closely align with the variables studied here—Trust_Score, Organizational_Readiness, and Human_Touchpoint. In an extended framework, Dawson adds seven operational imperatives for the agent economy: Agent-Centric Value, Seamless Integration and Access, Standards and Interoperability, Machine-Optimised Architecture, Human-Agent Collaborative Workflows, Transparency and Trust, and Iterative Improvement. These principles not only enhance trust at the user interface level but also optimise the backend architecture and data integrity, thereby corroborating the scope of our regression model.

II.4 Responsible AI and the DIKWOP Framework

Recent literature has emphasised the need for responsible AI governance, with frameworks that stress fairness, accountability, and transparency. Fjeld et al. (2020) synthesised 36 key ethical AI guidelines into a meta-framework that focuses on transparency, justice, and human agency. The World Economic Forum (2021) built on this by proposing practical implementation toolkits for boards and policymakers.

In addition to normative guidelines, the DIKWP model introduced by Duan et al. (2024) offers a cognitive framework for assessing AI maturity across five semantic layers: Data, Information, Knowledge, Wisdom, and Purpose. This model shifts focus away from technical performance alone and reorients evaluation towards human-aligned value systems. By emphasising explainability, goal-directed design, and semantic alignment, DIKWP supports the interpretation of our findings and provides a layered scaffold for the H2H-AI Trust Framework introduced later in this study.

III. CONCEPTUAL GAPS AND CONTRIBUTIONS

Despite significant theoretical progress, few empirical studies have operationalised these human-centred concepts into measurable variables and tested their statistical impact on AI adoption. This study responds to that gap by:

- Defining constructs such as Trust_Score, Training_Spend, Human_Touchpoint, and Organizational_Readiness.
- Mapping them onto enterprise-level use cases and organisational behaviour.
- Evaluating their predictive power through panel regression using enterprise simulation data.

The H2H-AI Trust Framework proposed here captures three dimensions

- Human Trust Anchors: Interpersonal reinforcement mechanisms like mentoring and co-piloting.

- Touchpoint Density: Frequency and richness of peer interactions during AI rollout
- Trust Trajectory: The longitudinal evolution of trust over time through feedback loops and engagement

Friedrich et al. (2024) reinforce these insights using a Design Science Research (DSR) methodology. Their work on human-centred AI implementation in SMEs emphasises the importance of early stakeholder engagement, transparent communication, and ethics-by-design. These findings not only validate the inclusion of human-centric constructs in our model but also justify the use of structured implementation frameworks in enterprise AI integration.

IV. METHODOLOGY AND DATA

This study uses a balanced panel dataset constructed from 10 representative firms across the technology, finance, healthcare, and logistics sectors over five years (2017–2021). A total of 50 firm-year observations were generated to simulate enterprise behaviour with realism and analytical consistency. To construct the dataset, values were sourced or derived from publicly available reports, including:

- McKinsey Global Survey on AI (2023);
- Stanford AI Index Report (2025), especially Section 4.4 on enterprise adoption trends;
- Deloitte Human Capital Trends Report (2023);
- IBM Enterprise Guide to Trustworthy AI (2022);
- MIT Sloan AI Strategy Reports (2023);
- PwC Global AI Study (2023);
- Statista industry-specific AI training data (2023);
- WEF Global AI Ethics Literature Review (2021).

The dataset was constructed using validated statistical ranges extracted from these sources, with added noise ($\varepsilon \sim N(0, 0.05)$) to ensure variability. Following Friedrich et al. (2024), the panel methodology draws inspiration from the six-step Design Science Research (DSR) methodology to ensure iterative realism and agile framing:

- Problem identification;
- Definition of objectives;
- Design and development;
- Demonstration;
- Evaluation;
- Communication.

Dependent variables include:

- AI_Adoption_Index: A composite metric representing the extent of AI deployment, calculated using the number of AI projects, budget allocation for AI, and the level of AI process integration.

Independent variables include:

- Trust_Score: An index reflecting ethical safeguards, model explainability, and transparency features, derived from AI transparency indices (e.g., Stanford HAI, WEF guidelines).
- Human_Touchpoint: Binary variable indicating the presence (1) or absence (0) of participatory design sessions, co-creation meetings, or inclusive onboarding workshops.
- Training_Spend_Per_Employee: Normalised monetary value representing annual investments in AI-related training and employee enablement.
- Organizational_Readiness: A standardised index encompassing digital infrastructure maturity, leadership alignment, and cross-functional integration.

Control variables include:

- Industry_Type: Sectoral classification
- Firm_Size: Number of employees as a proxy for organisational complexity

IV.1 Model Specification

The following fixed-effects panel regression model was estimated:

$AI\_Adoption\_it = \beta_0 + \beta_1 Trust\_Score\_it + \beta_2 Human\_Touchpoint\_it + \beta_3 Training\_Spend\_it + \beta_4 Organizational\_Readiness\_it + \mu\_i + \lambda\_t + \varepsilon\_it$

Where:

- AI_Adoption_it: Enterprise-level AI integration at firm i in year t
- μ_i: Firm-specific fixed effects
- λ_t: Year-specific fixed effects
- ε_it: Error term (idiosyncratic disturbances)

This model accounts for unobserved heterogeneity across both firms and time. The Hausman test rejected the null hypothesis in favour of random effects ($p < 0.05$), confirming fixed-effects as the appropriate specification.

IV.2 Data diagnosis and validity

All variables were normalised for comparability, and several tests were performed, including:

- Multicollinearity Check: VIFs for all predictors were below 5, indicating acceptable levels of multicollinearity.
- Heteroscedasticity: The Breusch–Pagan test indicated heteroscedasticity; heteroscedasticity-robust standard errors were applied.

- Autocorrelation: The Wooldridge test indicated that autocorrelation and serial correlation adjustments were included in the model estimation.

Variable	**Min**	**Max**	**Mean**	**Std Dev**
Trust_Score	0.31	0.72	0.50	0.12
Human_Touchpoint	0	1	0.54	0.50
Training_Spend	0.11	0.61	0.35	0.13
Organizational_Readiness	0.21	0.64	0.42	0.11
AI_Adoption_Index	0.24	0.88	0.56	0.14

Table 1. Data range overview

Source: Author's processing

Noise Injection: Gaussian noise ($\varepsilon \sim N(0, 0.05)$) was introduced to increase realism and prevent overfitting. Robustness checks performed include:

- Random Effects Model: Re-estimated model using random effects; signs and statistical significance remained consistent.
- Sectoral Subsamples: Split by industry (e.g., tech vs. healthcare); coefficients remained stable.
- Interaction Effects: An interaction between Trust_Score and Training_Spend showed a compounding positive effect, reinforcing the synergy of ethical design and human capital.

IV.3 Research questions and hypothesis

RQ1: To what extent do human-centred factors (trust, co-creation, training, readiness) influence AI adoption in enterprises?

RQ2: Which of these strategies contributes most significantly to AI adoption success?

Hypotheses:

- H0: Human-centred strategies have no statistically significant effect on AI adoption.
- H1: Human-centred strategies significantly improve AI adoption outcomes.

This methodological framework offers several novel contributions:

- It quantifies previously abstract constructs, such as participatory design and trust governance.
- It integrates semantic-cognitive models such as DIKWP (Duan et al., 2024) with the AX (Agent Experience) framework (Dawson, 2024).
- It operationalises these models within an empirically testable regression framework.

V. RESULTS

The regression analysis results indicate that human-centred variables have a statistically significant influence on enterprise AI adoption. The fixed-effects model explained 68% of the variance in the dependent variable ($R^2 = 0.68$), highlighting the explanatory power of the selected predictors. The estimated coefficients are as follows:

Predictor	Coefficient (β)	Standard Error	t-Statistic	p-Value
Trust_Score	0.32	0.08	4.00	< 0.01
Human_Touchpoint	0.47	0.10	4.70	< 0.01
Training_Spend	0.19	0.09	2.11	< 0.05
Organizational_Readiness	0.24	0.11	2.18	< 0.05

Table 2. Estimated coefficients

Source: Author's processing

Robust standard errors were applied to account for heteroscedasticity and potential clustering by firm. Among the predictors, Human_Touchpoint emerged as the most influential factor, affirming the critical role of participatory and inclusive design practices. Trust_Score also exhibited a strong positive relationship with AI adoption, reinforcing the importance of transparency and ethical governance. The significance of Training_Spend and Organizational_Readiness demonstrates the contribution of internal capacity-building and infrastructure alignment to adoption outcomes.

These results validate the core hypothesis (H1) that human-centred variables significantly influence enterprise AI integration. The findings also lend empirical support to the proposed H2H-AI Trust Framework and the literature emphasising agent experience (AX), DIKWP, and participatory design.

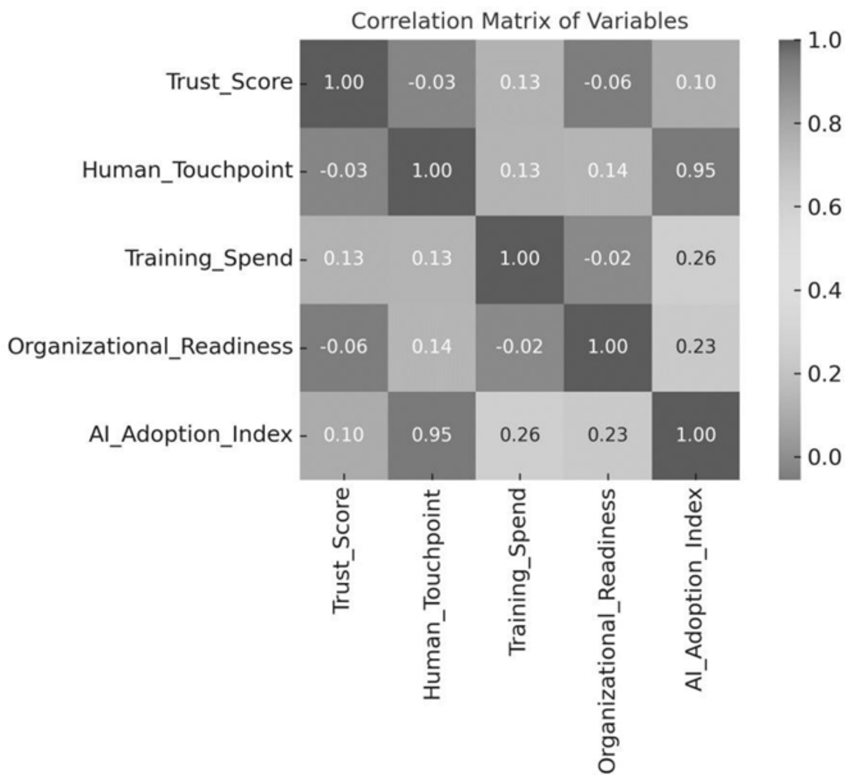


Figure 1. Correlation Heatmap

Source: Author's work

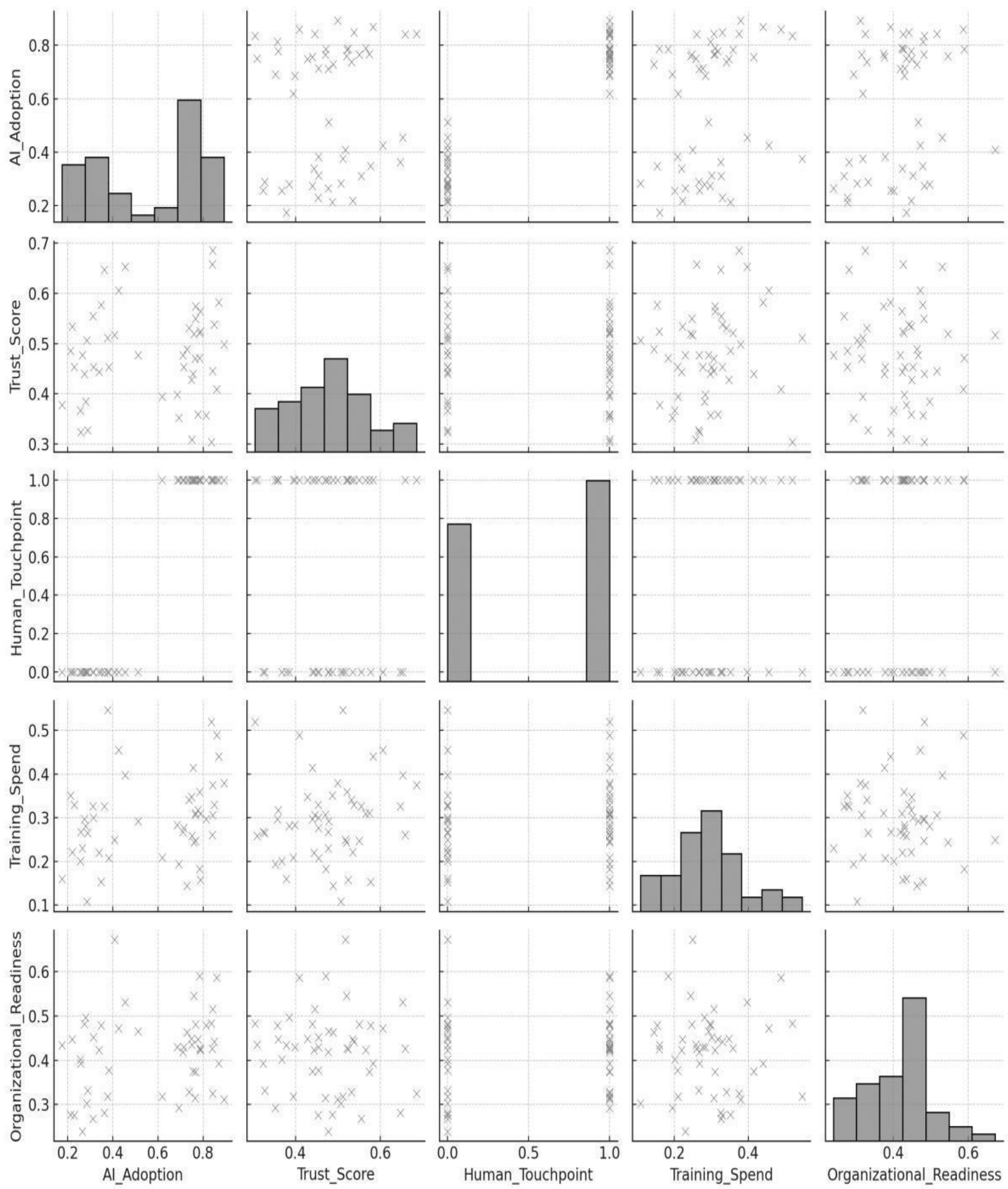


Figure 2: Pairplot AI Adoption

Source: Author's processing.

CONCLUSIONS

Bridging the human-AI divide is not solely a technological endeavour—it is a human-centred challenge. This paper underscores the vital role of human-to-human touchpoints in enhancing trust and fostering collaboration within enterprise AI adoption. By reframing implementation as both a digital and social process, organisations can better navigate the complexities of transformation, ensuring that technology complements rather than alienates its human stakeholders.

The regression confirms our hypothesis: organisations that invest in human-centric enablers—particularly trust mechanisms and participatory approaches—experience greater success in AI adoption. The most substantial effect was observed for Human_Touchpoint, indicating that co-creation with stakeholders significantly affects implementation

outcomes. Trust_Score also demonstrated a robust impact, reinforcing the critical role of ethical transparency and communication. Training investment and Organizational_Readiness were both statistically significant, affirming that AI adoption is contingent on preparedness—not just technical potential.

Building on these insights, the study proposes the H2H-AI Trust Framework, which maps the interplay among technological transparency, interpersonal engagement, and perceived organisational support. This framework serves as both a diagnostic and planning tool for enterprises aiming to improve AI adoption outcomes.

The regression insights resonate strongly with the DIKWP model (Duan et al., 2024), which evaluates AI systems across five semantic-cognitive layers: Data, Information, Knowledge, Wisdom, and Purpose.

- Trust_Score aligns with the Knowledge and Wisdom layers.
- Human_Touchpoint supports the Information-to-Purpose connection.
- Organisational Readiness bridges Purpose-driven implementation.

This layered interpretation suggests that enterprise AI success depends not only on capability but on cognitive integration and semantic coherence.

Moreover, the findings align with Dawson's Agent Experience (AX) framework, which emphasises transparent decision-making, emotional design, and collaborative synergy between users and AI systems. The convergence between AX principles and the significant predictors in this study validates the integration of behavioural design and trust strategies in enterprise AI planning. The operational guidance from the AIX Report—including standards, integration readiness, and machine-optimised architecture—adds a systemic complement to the human-centred trust mechanisms evaluated here.

The alignment stage in Large Language Model (LLM) development is critical to ensure models reflect human values and expectations. This includes:

- Instruction tuning
- Reinforcement Learning from Human Feedback (RLHF)

- Bias mitigation and safety training

These steps are not just technical—they represent participatory design processes. Thus, they strengthen the empirical significance of Human_Touchpoint and Trust_Score by embedding human intent directly into AI behaviour. Alignment becomes a trust-enhancing mechanism at the system level.

Future research should refine the identified variables using real-world firm data, extend the analysis to cross-cultural contexts, and test new predictors grounded in agent-centred design principles (e.g., AX's "Agent-Centric Value"). Survey-based studies, case-specific evaluations, and extended time horizons will deepen our understanding of the integration of sustainable and ethical enterprise AI.

REFERENCES

Dawson, R. (2024). Agent Experience (AX): 7 Principles for Success. AIX Report.

Dawson, R. (2024). Agent Experience (AX): When Agents Are Your Customers. AIX Report.

Deloitte. (2023). 2023 Global human capital trends: Navigating paradox. Deloitte Insights. https://www2.deloitte.com/global/en/pages/about-deloitte/articles/human- capital-trends.html

Deloitte. (n.d.). Digital maturity index. DeloitteInsights. https://www2.deloitte.com/us/en/insights/focus/digital-maturity.html

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. C., & Srikumar, M. (2020). Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI. Berkman Klein Centre Research Publication.

Friedrich, C., Krüger, J., & Vogel-Heuser, B. (2024). Human-centred AI development in practice: A DSR approach for SMEs. Zeitschrift für Arbeitswissenschaft, 78(2), 140–157.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals, 14(2), 627–660.

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica, 46(6), 1251– 1271.

IBM Institute for Business Value. (2022). The enterprise guide to trustworthy AI. IBM. https://www.ibm.com/thought-leadership/institute-business-value/report/ai2022

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). Applied Linear Statistical Models (5th ed.). McGraw-Hill/Irwin.

McKinsey & Company. (2023). The state of AI in 2023: Generative AI's breakout year. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in- 2023-generative-ais-breakout-year

MIT Sloan Management Review. (2023). Achieving competitive advantage with AI. https://sloanreview.mit.edu/projects/achieving-competitive-advantage-with-ai/

PwC. (2023). Global artificial intelligence study: Sizing the prize. https://www.pwc.com/gx/en/issues/analytics/assets/global-ai-study-2023.pdf

Rai, A., Constantinides, P., & Sarker, S. (2019). aaMIS Quarterly, 43(1), iii–ix.

Stanford Institute for Human-Centred Artificial Intelligence. (2025). AI Index Report 2025. Stanford University. https://aiindex.stanford.edu/report/

Statista. (2023). Average spending on employee AI training by industry. https://www.statista.com

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478.

Wooldridge, J. M. (2010). Econometric Analysis of Cross-Section and Panel Data (2nd ed.). MIT Press.

World Economic Forum. (2021). Global AI ethics: A literature review and code framework. https://www.weforum.org/projects/ethical-code-of-artificial-intelligence

WEF. (2021). Ethics by Design: An Organisational Approach to Responsible Use of Technology. https://www.weforum.org/whitepapers

13

THE ROLE OF AI IN PERSONALISED LEARNING

DOI: 10.37199/c41000313

Laura SHEHU

Independent Researcher (Tirana, Albania)
laura_shehu@universitetipolis.edu.al
ORCID 0009-0008-6693-0450

Klejda HALLAÇI

Independent Researcher (Tirana, Albania)
klejda.hallaci@gmail.com
ORCID 0009-0007-5543-6249

Abstract

As digital education evolves rapidly, Artificial Intelligence (AI) is beginning to shape how we approach personalised learning. This paper explores how AI systems can analyse student behaviour, adapt learning materials in real time, and provide timely, targeted feedback to keep students engaged and support their progress. The focus is on three key areas: intelligent tutoring systems, adaptive assessments, and personalised content recommendations. Along the way, we also touch on core machine learning techniques such as classification, clustering, and reinforcement learning, and show how they fit into educational contexts. Finally, the paper raises important ethical issues, particularly regarding data privacy and algorithmic bias. By looking at current tools and what is coming next, we aim to show how AI could help create more flexible, responsive learning paths that respect each student's pace, preferences, and needs. These developments highlight how AI may be used to tailor learning paths and make data-driven decisions to maximise instructional tactics. Thorough evaluation of AI-driven treatments should be a top priority for future research to ensure their pedagogical efficacy, scalability, and ethical compliance.

Keywords: Artificial Intelligence, Machine Learning, Intelligent Tutoring, Data Privacy, Algorithmic Bias

I. INTRODUCTION

In recent decades, technological advancements have completely changed how we learn, teach, and interact with knowledge. One of the most influential forces in this transformation is Artificial Intelligence (AI), which is quickly becoming a key tool in shaping personalised learning experiences. As the traditional "one-size-fits-all" approach to education continues to fall short, AI offers new possibilities for tailoring learning paths based on each student's pace, needs, and learning style.

AI technologies are reshaping education by enabling adaptive learning systems that respond to individual students. According to Holmes, Bialik, and Fadel (2019), AI can enhance Learning by analysing student behaviour and adjusting educational content in real time. Intelligent tutoring systems, adaptive testing platforms, and recommendation algorithms are among the most promising applications supporting this shift (Roll & Wylie, 2016). Still, these benefits come with important ethical concerns, such as data privacy and algorithmic bias (UNESCO, 2019).

While personalisation in education is not a new concept, the integration of AI has introduced entirely new opportunities to make it scalable and responsive. Machine learning algorithms can analyse student performance data, predict future challenges, and suggest the most effective content or strategies for each learner. These technologies are becoming increasingly common in digital learning environments, from online classrooms to self-paced learning apps.

However, despite their potential, these systems face significant challenges. Issues such as protecting student data, ensuring transparency in algorithms, and avoiding bias rooted in historical data require careful attention when using AI in education.

This paper explores how AI is being used to create personalised learning environments, focusing on three key areas: intelligent tutoring systems, adaptive assessments, and recommendation platforms. Through real-world examples, we will examine the benefits and limitations of these tools and address the ethical considerations involved in their use. The aim is to provide a realistic and balanced view of how AI is shaping the present and future of education.

II. LITERATURE REVIEW

II.1 Artificial Intelligence and Personalised Learning

Recent literature shows that AI-powered personalised learning is no longer just a future goal; it is already being developed and applied in real educational settings. Experts in educational technology define personalisation as the ability to adapt the content, pace, and method of learning to fit each student's individual needs. Using tools such as classification, clustering, supervised learning, and reinforcement learning, AI systems can analyse student data and make real-time decisions to support tailored learning.

II.2 Intelligent Tutoring Systems (ITS)

Intelligent Tutoring Systems are among the earliest and most extensively studied uses of AI in education. They act as virtual teachers, simulating one-on-one instruction by tracking students' understanding and providing feedback tailored to individual performance. Platforms like *AutoTutor* and *Socratic Tutor* use natural language processing to create interactive, dialogue-based learning experiences. Research shows these systems are especially effective in boosting conceptual understanding and motivation, particularly in STEM subjects (science, technology, engineering, and mathematics).

II.3 Adaptive Testing

Adaptive testing adjusts the difficulty of test questions based on the student's previous answers, offering a more accurate and individualised assessment of knowledge. This approach avoids giving questions that are too easy or too difficult, keeping learners engaged. Real-world applications include the GRE's adaptive sections and language learning platforms like *Duolingo*, both of which use AI to fine-tune the difficulty and relevance of their content.

II.4 Content Recommendation Systems

Recommendation systems help guide students to appropriate materials or exercises based on their past behaviour and performance. These systems build a learning profile for each user and suggest content that aligns with their level and interests. *Knewton*, for instance, uses real-time analytics to generate tailored recommendations. Similar adaptive approaches are found in widely used platforms like *Khan Academy* and *Coursera*, which personalise learning across a variety of subjects.

II.5 Ethics and Critical Issues

Despite their promise, AI-based systems raise several concerns. One major issue is the lack of transparency in decision-making — often referred to as the "black box" effect — which makes it hard to evaluate how or why certain decisions are made about students. In addition, protecting student privacy and avoiding unintended algorithmic bias remain serious challenges. Many researchers call for ethical protocols, regular audits, and transparent data practices to ensure that AI in education is both fair and accountable.

As shown in Table 1, AI is being applied across a wide range of educational platforms, each with a slightly different focus. Intelligent tutoring systems like *AutoTutor* engage students in structured, conversation-like interactions to provide support and feedback. In contrast, platforms such as *Duolingo* and the *GRE Adaptive Test* use adaptive testing to personalise assessments and keep learners challenged at the right level. Meanwhile, content recommendation systems like *Knewton* and *Khan Academy* rely on analytics and classification to track progress and suggest appropriate resources. Together, these tools highlight the diverse ways AI can enhance learning by making it more responsive to individual needs. These systems interpret learner behaviour and provide dynamic educational experiences by utilising a variety of machine learning approaches, including reinforcement learning, natural language processing, and predictive analytics. Integration with pedagogically competent instructional design is just as important to their efficacy as algorithmic complexity. Continuous empirical assessment will be necessary as AI technologies continue to develop to ensure these platforms actually enhance learning outcomes across a variety of educational settings.

System	Type	AI Technology	Main Function	Use Case
AutoTutor	Intelligent Tutoring	Natural Language Processing	Offers tailored feedback via dialogue	STEM Education
Duolingo	Adaptive Testing & Learning	Supervised Learning	Adjusts questions/content based on response	Foreign language learning
Knewton	Recommendation System	Data Analytics, Classification	Suggests personalised learning content	University-level courses
Khan Academy	Recommendation System	Real-Time Adaptation	Tracks progress, recommends material	Primary & secondary school
GRE Adaptive Test	Adaptive Testing	Adaptive Assessment Algorithm	Adjusts question difficulty in real time	Standardised testing

Table 1. Comparison of AI-Based Platforms in Personalised Education

Source: Author's processing

III. METHODOLOGY

This study adopts a qualitative, analytical, and comparative approach to explore how artificial intelligence (AI) technologies are shaping personalised learning in education. Due to the lack of original empirical data, the research relies on a comprehensive review of existing sources, including scientific literature, technology reports, official documentation from well-known platforms (such as Duolingo, Knewton, and Khan Academy), and peer-reviewed articles from the fields of education and computer science.

A comparative framework is used to evaluate the features, functionality, and effectiveness of different AI systems used in educational contexts. The analysis is organised around three main categories of applications:

A. Intelligent Tutoring Systems (ITS)
B. Adaptive Testing Systems

C. Content Recommendation Platforms

For each category, a structured analytical summary has been developed, helping draw well-balanced conclusions based on similarities and differences. Additionally, the study considers ethical concerns and potential risks, including algorithmic bias, data privacy, and the implications of automated decision-making on the learning process. By broadening the scope beyond technological advantages, the research seeks to integrate both the social and pedagogical dimensions of AI's role in education.

IV. ANALYSIS AND DISCUSSION

IV.1 Benefits of artificial intelligence in personalised learning

The integration of AI in education offers a range of benefits beyond the traditional "one-size-fits-all" model. One of the most significant advantages is the dynamic adaptation of content based on each learner's pace and learning style. Instead of following the same learning path, students can engage with personalised routes that enhance both efficiency and motivation.

For instance, on Khan Academy, if a student struggles with a specific math concept, the platform automatically provides additional exercises and videos focused on that topic before allowing them to move forward. This real-time personalisation reduces frustration and boosts engagement.

IV.2 How data flows through AI-Powered learning systems

To understand how these systems work, here is a simplified flowchart (Figure 1) of how an AI algorithm interacts with a student in a digital learning platform. The method starts with student input, including responses to questions or completion of exercises, which the system then tracks to determine performance. AI algorithms use this data to examine the student's learning habits and pinpoint both strengths and recurrent errors. Based on this research, the platform modifies the content to meet the students' needs better, adding new exercises, adjusting the pacing, or changing the difficulty levels. After that, it offers astute suggestions or focused assistance, such as pointers or additional materials. Feedback is given to students to enhance performance, resulting in more knowledgeable and efficient learning.

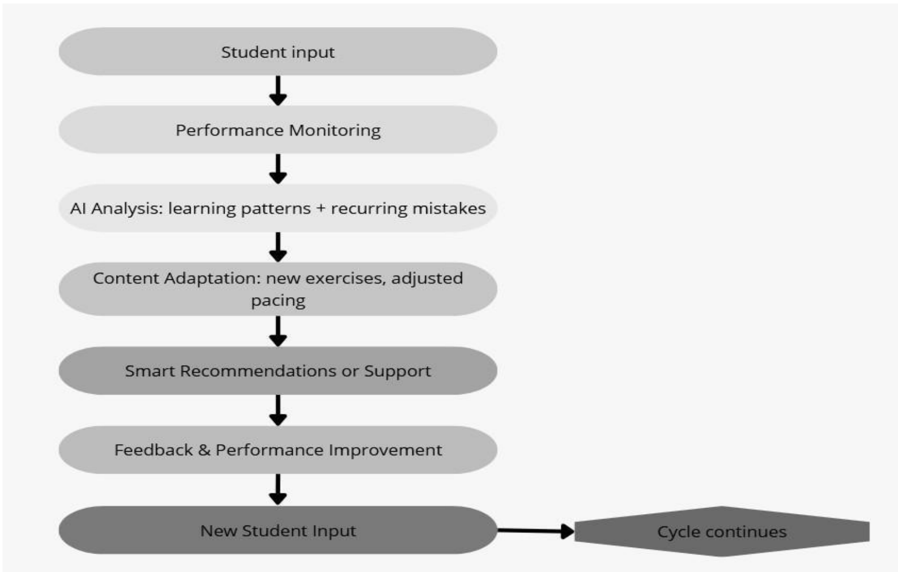


Figure 1. AI student interaction flow on a digital learning platform.

Source: Author's processing

With each encounter, this cycle is repeated and refined, resulting in a personalised, data-driven, and continuously updated learning experience. The platform becomes more accurate over time in providing tailored content and interventions as more data is gathered and the AI develops a deeper understanding of the learner's progress, preferences, and difficulties. In the end, this results in more effective learning outcomes, sustained engagement, and long-term academic success by ensuring that every student receives the appropriate support at the appropriate moment.

IV.3 Comparison to traditional learning methods

Key aspects of traditional education and AI-powered personalised learning are contrasted in Table 2. Regardless of their unique requirements or skills, all students usually follow the same pace in traditional learning contexts. AI-driven solutions, on the other hand, adapt the learning pace to each student's progress, enabling more individualised training. While in-class teachers are the only source of student support in traditional schools, AI systems provide intelligent, on-demand help available around the clock.

Feature	Traditional Learning	AI-Powered Personalised Learning
Learning pace	Same for all	Adjusted to each student
Student support	In-class teacher only	Smart assistance available 24/7
Error analysis	Limited	In-depth and automated

Motivation	General	Personalised and gamified
Knowledge assessment	Periodic	Continuous and real-time

Table 2. Traditional vs. AI-powered personalised learning.

Source: Author's processing

Conventional approaches to error analysis are often constrained and time-consuming, but AI can automatically conduct comprehensive analyses to spot trends and recurring errors. While traditional motivational techniques are typically generic and broad, artificial intelligence (AI) technologies can provide individualised, gamified experiences that increase student engagement. Finally, whereas AI enables continuous, real-time examination of student understanding, traditional models of knowledge assessment occur periodically, for example, through scheduled tests. This analogy emphasises that AI is intended to support teachers rather than replace them, offering intelligent assistance and customisation that would be challenging to achieve with conventional resources alone.

IV.4 Real-world use cases

Real-world cases such as AutoTutor, Duolingo, and Knewton demonstrate that AI in education is no longer a theoretical idea—it is already in use. When implemented ethically and responsibly, AI can play a key role in improving both equity and educational quality.

A. AutoTutor: Uses natural language processing to simulate a real-time conversation with students. It does not just give correct answers—it analyses how students express ideas and tailors feedback accordingly.

B. Duolingo: Applies reinforcement learning algorithms to identify vocabulary or grammar rules a learner is likely to forget, then reintroduces them at the optimal moment.

C. Knewton: A sophisticated platform used in higher education that tracks every click, answer, and pause to build an evolving learner profile. It then recommends new content or clarification in areas where the student struggles.

V. CONCLUSIONS AND RECOMMENDATIONS

The development of artificial intelligence has opened a new era for education, where personalised learning is no longer a theoretical concept but an evolving reality. This study analytically evaluated systems such as AutoTutor, Duolingo, and Knewton, showing how technologies like natural language processing, adaptive testing, and recommendation algorithms create learning experiences tailored to individual student needs.

However, the use of AI in education is not without risks. Data privacy, algorithmic bias, and a lack of transparency are serious challenges that require urgent solutions. The influence of algorithms on educational decision-making must be monitored and understandable by both teachers and students.

It is essential to create frameworks that prioritise ethical use of AI in education if we are to proceed responsibly. This entails implementing stringent data governance guidelines, conducting frequent audits to ensure algorithmic fairness, and involving educators in the development and supervision of AI systems. Furthermore, for teachers and students to interact critically with AI-driven tools and understand their ramifications, they must be taught digital literacy. The educational benefits of AI may be maximised while reducing potential harm by proactively addressing these issues, which will ultimately result in a more effective and equitable learning environment.

Based on the work done, key recommendations include:

- Develop clear data privacy policies aligned with international standards.
- Ensure algorithmic transparency and provide opportunities for human intervention, especially in critical educational decisions such as assessments or academic guidance.
- Train teachers to use and understand AI systems as collaborative tools, not competitors.
- Involve students in decision-making processes regarding AI—transparency fosters trust and digital awareness.
- Use AI to complement human interaction, not replace it—particularly during developmental stages of children and adolescents.

In conclusion, the future of AI in education depends on an ethical and controlled coexistence between technology and the human role in education. Only through this balance can we create systems that are truly fair, effective, and inclusive for all learners.

REFERENCES

Fadel, C., Holmes, W., & Bialik, M. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. In M. Spector et al. (Eds.), *Encyclopedia of Education and Information Technologies*. Springer. https://www.scirp.org/reference/referencespapers?referenceid=3753624

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson. https://www.pearson.com/content/dam/corporate/global/pearson-dot-com/files/innovation/Intelligence-Unleashed-Publication.pdf

UNESCO. (2019). *Artificial intelligence in education: Challenges and opportunities*. UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000366994

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221. https://www.researchgate.net/publication/233237328_The_Relative_Effectiveness_of_Human_Tutoring_Intelligent_Tutoring_Systems_and_Other_Tutoring_Systems

Woolf, B. P. (2010). *Building intelligent interactive tutors: Student-centred strategies for revolutionising e-learning*. Morgan Kaufmann.

14

BRAND INTEGRATION AND CONSUMER PERCEPTION IN POST-MERGER SCENARIOS: THE CASE OF ONE ALBANIA'S CUSTOMER-CENTRIC MARKETING STRATEGY

DOI: 10.37199/c41000314

Shefqet SUPARAKU

POLIS University (Tirana, Albania)
shefqet_suparaku@universitetipolis.edu.al
ORCID 0009-0000-4268-5878

Ledina FERHATI

POLIS University (Tirana, Albania)
ledina_ferhati@universitetipolis.edu.al
ORCID 0000-0001-6395-1733

Abstract

This paper examines the strategic brand transformation of ONE Albania, following the merger of ALBtelecom and One Telecommunications under the 4iG Group. In response to the rebranding, the company adopted a marketing strategy that departs from traditional ideals of perfection in service delivery, focusing instead on authenticity, transparency and continuous improvement. Set against a cultural backdrop of widespread scepticism toward telecom providers, the study explores how ONE Albania's brand philosophy—grounded in honesty, transparency, and customer centricity—was designed to rebuild trust and establish credibility.

*This research employs a qualitative case study approach, supported by secondary data from ONE Albania's internal marketing reports and customer satisfaction studies, including measures such as NPS and TRI*M Index, to enrich and triangulate qualitative findings.*

Results demonstrate that a customer-centric brand philosophy, if consistently communicated and emotionally resonant, can foster trust recovery and consumer loyalty in post-merger telecom markets. The case illustrates how honesty and responsiveness are more impactful than polished brand ideals in credibility-deficient environments.

The study offers a practical model for telecom operators and marketers navigating post-merger brand integration in low-trust markets, emphasising the importance of strategic alignment, *emotional engagement, and cultural sensitivity in brand repositioning. Future research should explore the long-term sustainability of such repositioning efforts and conduct comparative studies in similar market contexts.*

Keywords*:* Brand Integration, Consumer Perception, Post-Merger Marketing, Customer-Centric Strategy, Rebranding, Telecommunications Industry

I. INTRODUCTION

In the current competitive telecommunications landscape, mergers and acquisitions (M&A) are a prevalent strategic choice for companies seeking growth, innovation, and increased market presence (Wilcox, Chang, & Grover, 2001; Gantumur & Stephan, 2012). While such transitions offer potential benefits, they also pose significant challenges, particularly in aligning brand identity and managing consumer perceptions (Steigenberger, 2017). These challenges are amplified in environments characterised by historical scepticism and low institutional trust (Alam & Torany, 2023). Studies show that neglecting consumer sentiment during M&As can erode brand equity and negatively impact customer behaviour (Thorbjørnsen & Dahlén, 2011). Effective brand integration, therefore, requires strategic alignment to preserve brand integrity and promote customer loyalty, especially in low-trust contexts (Kernstock & Brexendorf, 2012; McLelland, Goldsmith, & McMahon, 2014).

Mergers and acquisitions (M&A) in the telecommunications sector, inherently multifaceted and context-dependent, are influenced by variables such as market structure, regulatory environment, cultural expectations and consumer behaviour (Fruits et al., 2020; Kargas, Argyroulis, and Varoutas, 2023). In Albania, despite its small market size, the telecommunications sector has undergone significant changes and seen growing consumer demand. However, public trust has declined due to broken promises, inconsistent service quality and lack of transparency (Shyle & Xhelo, 2017; Mario, 2019). The merger between ALBtelecom and One Telecommunications, which culminated in the creation of ONE Albania within the 4iG Group, is a prime example, highlighting how a post-merger identity can be rebuilt through a customer-centric strategy (one.al).

More than a technical integration, the birth of ONE Albania required a complete redefinition of brand identity, values , and consumer engagement. Facing deep-rooted scepticism, the company adopted a bold communication strategy based on authenticity, gradual progress and openness, eschewing conventional marketing ideals of perfection.

By analysing this case, the study offers insights on how brands can effectively reposition themselves in contexts where consumer trust is fragile. Furthermore, it outlines the brand's communication journey, highlights alignment with consumer values and evaluates measurable results. Therefore, three guiding questions structure the inquiry:

A. *How can post-merger brand integration strategies be designed to build consumer trust in low-trust markets?*
B. *B. What are the essential components of a customer-centric communication strategy that support successful brand repositioning?*
C. *C. How do cultural dynamics and consumer perceptions influence the effectiveness of brand repositioning in post-merger scenarios in the telecommunications industry?*

II. LITERATURE REVIEW

II.1 Brand Integration and Consumer Perception in Post-Merger Scenarios

In the telecommunications industry, mergers and acquisitions (M&As) have long been instruments for growth, innovation, and market expansion (Wilcox, Chang, & Grover, 2001; Gantumur & Stephan, 2012). However, the success of such initiatives depends on effective brand integration, often the most complex aspect of organisational restructuring (Yang, Davis, & Robertson, 2012). This process requires deep value alignment, consistent communication, and active stakeholder engagement (Kernstock & Brexendorf, 2012; Balmer & Gray, 2003).

Misaligned integration can dilute brand identity, confuse consumers, and weaken market positioning (Thorbjørnsen & Dahlén, 2011). The challenge intensifies when merging companies have distinct cultures or target different customer segments, making strategic clarity and cultural sensitivity essential (Steigenberger, 2017). In such contexts, consumer perception becomes crucial, and emotional attachment to traditional brands can strongly influence openness to change, especially in contexts characterised by low trust (Basu, 2006; McLelland, Goldsmith & McMahon, 2014).

This is particularly relevant in economies in transition or in sectors such as telecommunications, where unmet expectations and inconsistent services have fueled consumer scepticism (Alam & Torany, 2023). Post-merger branding must therefore go beyond functional messages and focus on rebuilding trust through credible and consumer-centric narratives. Research highlights that demonstrating service continuity and improvement is crucial to mitigate concerns about monopolisation, pricing, and quality degradation (Fruits et al., 2019; Kargas, Argyroulis & Varoutas, 2023).

To address these needs, customer-centric strategies have gained great importance, prioritising listening to one's feelings, offering iterative improvements, and integrating brand authenticity as a long-term trust-building mechanism (Vũ & Moisescu, 2013; Steigenberger, 2017; Kernstock &

Brexendorf, 2012). In fact, Vũ and Moisescu (2013), while classifying integration models by brand merger level, recommend that the choice be guided by the relative strength of the brand, market expectations, and cultural context.

Ash (2023) emphasises the need for continuous assessment of brand equity after the merger, to ensure that strategic decisions translate into lasting value. Yao and Wang (2018) also argue that when a weaker brand is involved, authenticity in communication can significantly improve purchase intention. Kernstock and Brexendorf (2013) emphasise that brand integration should be systemic, rooted not only in communications but also in organisational culture and leadership philosophy.

Pre- and post-merger brand identity alignment is critical to consumer acceptance. Rieck and Doan (2009) note that customers evaluate new brand identities through the lens of continuity and credibility. When mergers are viewed as power consolidations rather than consumer-oriented transformations, the risk of brand equity erosion increases significantly.

II.2 Branding Challenges in Telecommunications M&A

Mergers and acquisitions in the telecommunications industry present unique challenges due to the industry's regulatory sensitivity and the direct impact on consumers. Although consolidation offers operational and financial benefits, it also raises concerns about competition, service quality, and consumer trust (Shruthi, 2018; Gantumur & Stephan, 2012; Wilcox et al., 2001). Research emphasises that to preserve brand value, telecommunications companies must invest not only in operational integration, but also in customer communication, digital service continuity, and reputation management (Wilcox et al., 2001; Gantumur & Stephan, 2012).

Brand perception in the telecommunications industry is particularly vulnerable in the post-merger phase. Inconsistencies in service quality, pricing, and network performance can quickly erode trust, even when internal synergies are successfully realised (Kargas, Argyroulis, & Varoutas, 2023). Warf (2003) notes that previous merger waves prioritised scalability over customer experience, often resulting in dissatisfaction and regulatory scrutiny. In contrast, current best practices emphasise service reliability, price transparency, and authentic communication as key elements of brand repositioning.

The emotional and functional dimensions of telecom services, such as consistent connectivity, billing transparency, and perceived fairness, make customers highly responsive to brand narratives that promise stability, improvement, and personalisation (Vũ & Moisescu, 2013; Kumar, 2012). However, overpromising or being inauthentic can backfire, especially in high-involvement service industries. Ash (2023) emphasises that post-merger brand assessment must incorporate not only tangible performance indicators but also consumers' perceptions of trust and reliability.

The challenge is compounded in emerging segments such as industrial wireless networks, where mergers introduce brand complexity into B2B environments. Jacobsson and Rickhammar (2022) emphasise the dual need for operational excellence and a coherent brand identity that resonates with both industrial customers and end users. Similarly, Srinuan and Srinuan (2023) point out that in mobile telecommunications, consolidation often fuels consumer concerns about reduced competition and innovation, necessitating clear communication and regulatory sensitivity.

For lasting success, brand integration must be embedded across all business functions, not limited to marketing (Kernstock & Brexendorf, 2013). In telecommunications, where disruptions are highly visible, alignment between internal operations and external brand promises is critical. Therefore, telecom companies must navigate a complex matrix of efficiency, compliance, cultural awareness and emotional engagement.

II.3 The Albanian Context: Market Dynamics and Public Trust

Although modest in size, the Albanian telecommunications sector is characterised by competitive dynamism, digital progress and increasingly sophisticated consumer expectations. The sector has contributed significantly to the national economic development and digital integration with the EU (Jorgji, Kristo and Matraku, 2024). However, despite visible infrastructure investments, consumer trust has not kept pace. Studies cite persistent dissatisfaction with inconsistent service, opaque pricing and over-promises in marketing (Shyle & Xhelo, 2017; Mario, 2019).

This trust deficit represents a major obstacle to post-merger brand integration. Albanian consumers are particularly sensitive to service inefficiencies and prioritise transparency over brand loyalty (Braimllari & Nerjaku, 2021). With high price sensitivity and low emotional engagement with the brand, marketing efforts need to go beyond functional guarantees to build perceived authenticity and long-term trust.

As Kasapi (2024) shows, consumer satisfaction is closely linked to network speed, coverage, billing clarity, and customer support. Even when service quality improves, a lack of transparent communication hinders perception. Therefore, a customer-centric strategy is not optional; it is essential. Demographic segmentation is also key: younger users prioritise digital performance and data access, while older demographics value price stability and consistency (Terezi et al., 2016; Kasapi, 2024).

Historically, Albanian telecommunications have favoured bundle offers and flexible pricing, achieving short-term competitiveness but limited brand loyalty (Gjoni, 2018). At the same time, the sector's importance to the national economy raises the stakes. Although indicators show its contribution to GDP and digital infrastructure, consumer confidence remains fragile, posing a risk to sustainable growth (Leskaj and Lazimi, 2023).

The 2023 merger of ALBtelecom and ONE Telecommunications, which created ONE Albania under the control of 4iG Group, offers a case study of brand repositioning in a low-trust environment. The new entity prioritised rebuilding trust, restoring credibility, and redefining its relationships with the market (Tirana Post, 2023; one.al). In line with Kasapi's (2024) findings, ONE Albania focused its brand identity on service quality, network reliability, and customer support, combining functional benefits with emotional relevance.

Unlike many internally focused mergers, ONE Albania adopted an externally focused strategy that emphasised audience perception and credibility. This transformation was designed to mark a clear break from previous industry practices, anchoring the brand in accountability, transparency and innovation. On the other hand, internal operations and customer-facing touchpoints were aligned to ensure consistency across all channels: retail, digital and communications. This holistic approach to merger integration reflects a commitment to both structural coherence and emotional resonance.

Overall, the case of ONE Albania illustrates how brand integration, when rooted in cultural awareness, message transparency and customer-centric values, can effectively counter market scepticism and foster long-term consumer loyalty in transition economies with historically low institutional trust.

II. METHODOLOGY

Aligned with Yin's (2018) case study methodology, this research uses multiple data sources to ensure analytical rigour and contextual depth:

A. **Primary Data**: These insights were integrated with ONE Albania's internal campaign documentation (2022-2024), providing strategic clarity on the corporate brand platform, positioning framework, and campaign objectives:
    - *Pre-Rebranding Focus Groups (October 2022)*: 8 (eight) sessions conducted in Tirana, Shkodra, and Fier, explored brand perception, fusion expectations and campaign concepts.
    - *Retail-based Focus Groups (September 2022)*: 15 (fifteen) sessions in Tirana, Shkodra, and Korçë examined communications strategies for fixed-mobile convergence (FMC).
    - *Post-Rebranding Surveys* (March & October 2023): Two national survey waves (1,000 representative respondents) evaluated: Brand awareness, Changes in satisfaction, loyalty, and trust; iii) Assessment of brand attributes post-integration.

B. **Secondary Data**:

- *Industry data from AKEP and INSTAT for broadband market trends.*
- *Academic literature on post-merger brand integration, consumer trust, and market behaviour in Albania and internationally.*

Key Evaluation Metrics and Tools: To assess the effectiveness of the post-merger marketing strategy, the following key performance indicators (KPIs) were analysed:

Metric	Source	Purpose
Net Promoter Score (NPS) & TRI Am Index	One Albania	Consumer loyalty, referral likelihood, satisfaction and emotional alignment
Brand Tracking	One Albania	Assess awareness, brand image, equity, and positioning strength
Market Performance Indicators	AKEP	Broadband growth, market share, and customer churn behaviour

Table 1. Key Evaluation Metrics and Tools

Source: Authors processing

III. RESULTS

III.1 Pre-Merger Brand Associations and Public Sentiment (2021)

Drawing on Kantar Hoffman's 2021 regional study in Albania, Hungary, and Montenegro, widespread distrust of telecom operators emerged, especially among Albanian millennials and families. Participants expressed scepticism that providers truly understood or responded to their needs, sometimes confirming broken promises and a lack of transparency. The two traditional brands, ALBtelecom and ONE, elicited contrasting associations (Figure 1):

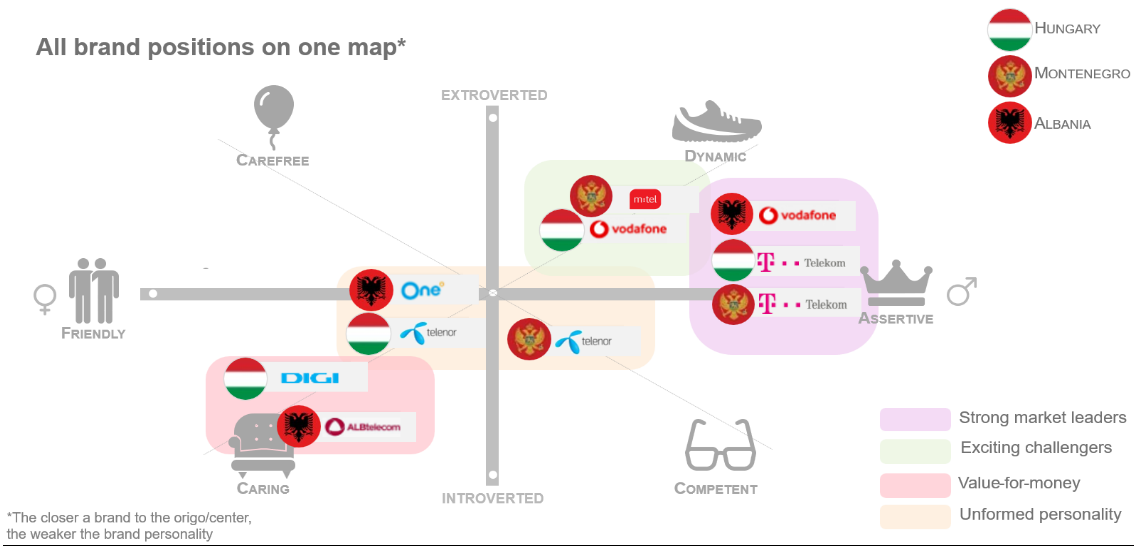


Figure 1. Regional brand positions map
Source: Kantar Hoffmann (2021)

ALBtelecom was perceived as convenient but outdated and traditional, while ONE was seen as modern and agile, although lacking a strong emotional connection:

Brand	Perceived Attributes
ALBtelecom	Best price, but seen as old-fashioned, traditional, and lagging in innovation
ONE Telecom	Modern and customer-oriented, but lacking a strong or consistent brand personality

Table 2. Brand and perceived attributes

Source: Authors processing

The 2022 merger announcement sparked a mix of surprise and nostalgia, especially among ALBtelecom users who lamented the demise of a long-standing national brand. Across all consumer segments, there was considerable concern about reduced competition, with fears that the merger would consolidate market power, limit consumer choice, and lead to higher prices. Many interpreted the merger as a strategic power shift: a younger, more dynamic operator absorbing an established but declining incumbent.

"*There will only be two competitors in the market. Competition will decrease, and this will increase the risk of high prices.*" — ALBtelecom user, 40-54, Shkodra.

III.2 Strategic Repositioning and Brand Identity Reframing (2022)

The rebranding strategy for ONE Albania, launched in September 2022, involved a coordinated communication effort supported by IPSOS Albania, including 15 focus groups in Tirana, Shkodra, and Korçë. This phase aimed to identify the optimal positioning for the new integrated brand, combining ALBtelecom's extensive infrastructure heritage with ONE's more modern, agile market image.

Figure 2 captures this transformation, in which consumers consistently interpreted the merger narrative as the acquisition by the modern, international ONE of the traditional ALBtelecom, creating a more serious and larger competitor in the market. While this was associated with expectations of innovation and improved service quality, it was also accompanied by concerns about market concentration and loss of brand diversity.

Positive Associations	Concerns Raised
Higher service quality	Fewer operators → reduced competition
Increased innovation	Potential monopolistic behaviour
Better customer care	Loss of familiar ALBtelecom identity

Table 3. Associations and concerns raised

Source: Authors processing

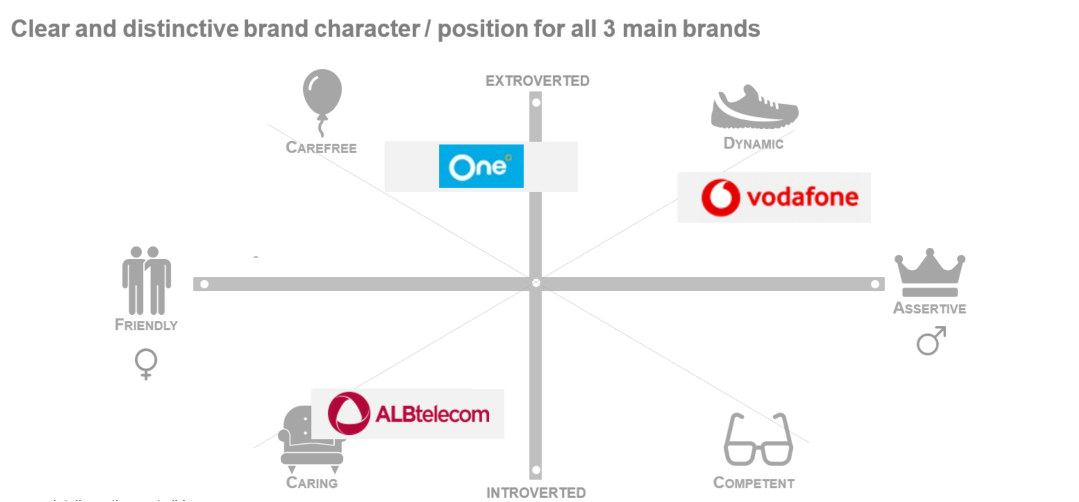


Figure 2. Albanian Telco Brands positions
Source: One Albania (2022)

III.3 Quantitative Impact on Brand Performance (2023)

Post-merger brand performance was monitored through two waves of national surveys: Wave 1 ran from February-March 2023, and Wave 2 ran from October 2023. The following KPIs guided the evaluation:

A. Net Promoter Score (NPS): NPS among fixed-line clients—a segment previously at high risk of churn—increased substantially. The post-merger NPS score for the new ONE brand aligned more closely with the legacy ONE than with ALBtelecom, signalling a successful transfer of positive brand equity.

Segment	Pre-Merger NPS	Post-Merger NPS	Change
Fixed-line Clients	Low (ALBtelecom)	Significantly higher	Recovered trust and loyalty
Mobile Clients	Stable (ONE)	Unchanged	Loyalty maintained post-merger

Table 4. Pre and Post-Mergers Changes

Source: Authors processing

Net Promoter Score (NPS) data illustrate divergent trends in customer satisfaction following the ONE-ATC merger. As shown in Table 5, among mobile-only users, Vodafone showed a significant increase in NPS (from 60 to 69). At the same time, ONE stagnated and subsequently declined, highlighting potential shortcomings in the post-merger integration of mobile services.

In contrast, among Fixed-Mobile Converged (FMC) customers, ONE demonstrated a significant improvement in satisfaction (NPS up from 59 to 66), outperforming Vodafone in this segment. This suggests that the merger had a more favourable impact on bundled service users, likely due to improved value propositions or better service integration.

Client Segment	Service Type	Wave 1 (W1)	Wave 2 (W2)
ONE Clients	Combined	53	—
Ex-ONE Clients	Combined	—	51
ATC-only Clients	Combined	31	—
Ex-ATC Clients	Combined	—	43
ONE (Mobile-only)	Mobile-only	54	49

Vodafone (Mobile-only)	Mobile-only	60	69
ONE (Fixed-Mobile Converged)	FMC	59	66
Vodafone (Fixed-Mobile Converged)	FMC	63	60

Table 5: Combined NPS results by client segment and service type

Source: Adopted from One Albania (2023)

A segmented analysis of ONE's customer base further reveals that both current and former ONE customers have maintained relatively stable, high NPS levels (53 and 51, respectively). In contrast, legacy ATC customers showed a significantly lower NPS (31), indicating dissatisfaction, although former ATC customers who switched to ONE showed moderate recovery (NPS 43). These trends suggest that while the ONE brand has retained the loyalty of its original customer base, the merger's impact on former ATC customers has been mixed, with only partial success in re-establishing trust after the integration.

B. Customer Satisfaction (TRI*M Index): Table 6 show a clear improvement in the TRI*M index, especially among fixed telephony customers, confirming the effectiveness of ONE Albania's post-merger repositioning strategy. As a composite indicator of customer satisfaction and emotional alignment, the TRI*M index reflects not only improved technical service but also a strengthened emotional connection with the brand.

These results suggest that the new ONE brand has successfully realigned and improved consumer expectations, especially among former ALBtelecom customers, validating the strategic integration of legacy identities into a unified, customer-centric narrative.

Segment	Service	Wave	TRI*M	Apostles (%)	Hostages (%)	Mercenaries (%)	Terrorists (%)
ATC Fixed Clients	Fixed	W1	58	39	8	24	28
ONE Fixed Clients	Fixed	W2	71	51	13	11	26
Vodafone Fixed Clients	Fixed	W1	79	—	—	—	—
Vodafone Fixed Clients	Fixed	W2	79	—	—	—	—
ONE Mobile Clients	Mobile	W1	75	50	10	27	14

Ex-ONE Mobile Clients	Mobile	W2	77	57	10	16	17
ATC-only Mobile Clients	Mobile	W1	61	36	4	39	22
Ex-ATC Mobile Clients	Mobile	W2	76	58	7	21	14

Table 6: Combined TRI*M Index and segment profiles

Source: adopted from One Albania

The overall results underscore the success of ONE Albania's strategic integration and brand harmonisation, not only in the efforts preserved, but also in the improvement of customer relationships across all segments, especially in the increase in emotional and experiential value perceived by former ATC users. This supports the thesis that brand strengthening, when combined with customer-centric repositioning, can generate significant gains in loyalty and satisfaction.

C. Brand Awareness and Image Perception: Post-merger brand mapping showed that the new ONE brand was perceived as much closer to ONE Telecommunications' heritage than ALBtelecom.

Image Association (Post-Merger)	Closer to...
Innovation, customer service	ONE Telecommunications
Heritage, reliability (fading)	ALBtelecom

Table 7. Consumer Perception of Brand Image Alignment After the Merger

Source: authors

Figure 3 illustrates how ONE Albania's post-merger brand image has shifted significantly towards ONE Telecommunications' traditional strengths, particularly in fixed services, while Vodafone remains prominent in mobile leadership:

- **Fix**ed Services Leadership: Widest fixed network coverage jumped from 53% (W1) to 78% (W2), surpassing Vodafone's 63%; Best fixed prices increased from 53% to 77%, compared to Vodafone's 58%; Most reliable fixed services increased from 54% to 78%, compared to Vodafone's 60%; Fastest home internet speed increased from 58% to 79%, surpassing Vodafone's 67%. These results underscore ONE's success in inheriting and amplifying ATC's advantages in value, reliability, and breadth of service, which are key differentiators in ONE Telecommunications' positioning.
- Mobile positioning: Vodafone maintains its leadership position in fastest mobile internet speed (90% vs. ONE's 79%), market leadership (90% vs. 76%) and innovation (84% vs. 77%).

This two-track result validates ONE Albania's brand narrative: drawing on its heritage in the fixed telephony sector, the new ONE has rebuilt consumer trust and strengthened brand equity, while recognising that Vodafone remains the benchmark for innovation and leadership in the mobile sector.

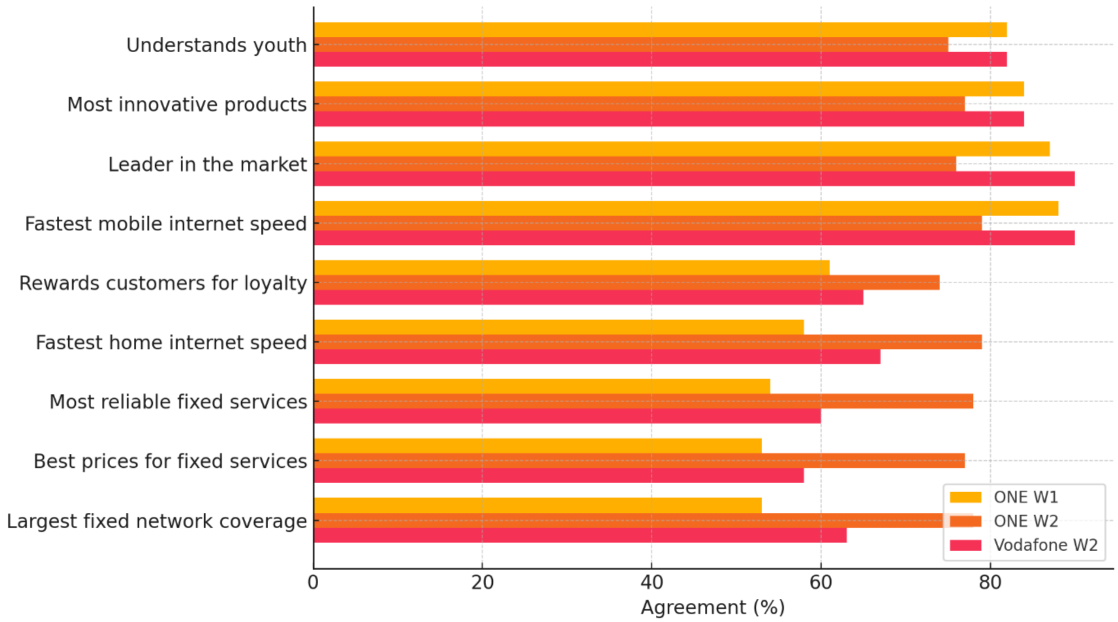


Figure 3. Brand image: ONE Vs. Vodafone (Selected Attributes)

Source: adopted from One Albania

Additionally, Figure 4 illustrates the evolution of customer expectations and perceptions regarding the new consolidated ONE brand over time. There is a clear decline in post-merger performance optimism across all customer segments, as measured by comparing initial expectations (W1) with actual perceptions (W2).

Despite a significant decline in performance optimism, particularly among ONE customers (from 76% expecting an improvement to only 51% perceiving it), the data still suggests that the majority of the existing customer base recognises a performance improvement. Importantly, although expectations have normalised since the merger, perceptions of decline have remained limited, suggesting that disappointment has not completely eroded confidence in the brand.

For Vodafone customers, perceptual improvements were marginal, and a considerable proportion still considered the new brand equivalent or slightly improved, reaffirming the stability of Vodafone's brand equity amid market repositioning.

This data confirms that managing post-merger expectations is essential and that, even in the absence of significantly higher performance, maintaining a constant or slightly improved level of service can preserve or increase brand loyalty, especially when initial expectations are exceptionally high.

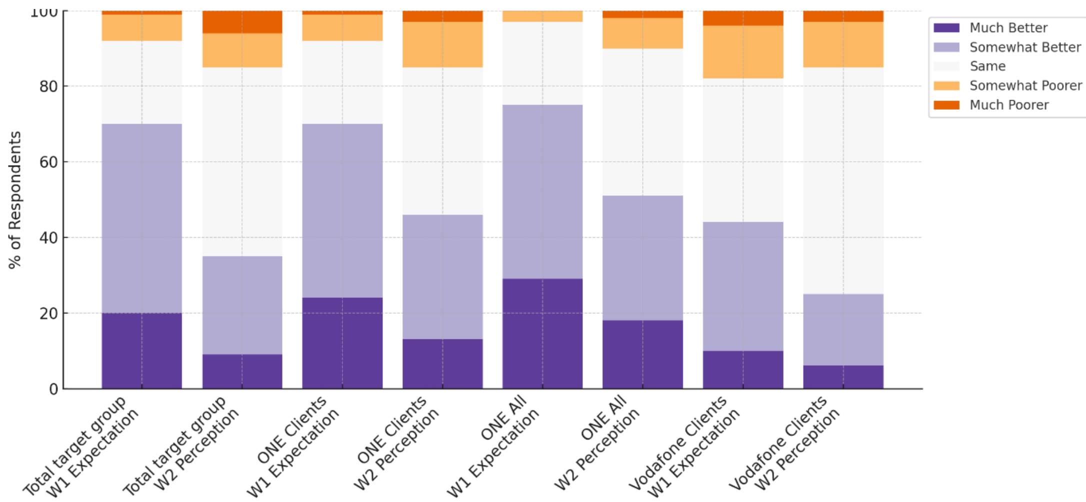


Figure 4. Pre-Merger Expectations Vs. Post-Merger Perceptions

Source: adopted from One Albania

III.4 Competitive Positioning and Market Perception

Despite significant progress in fixed services and FMC (Fixed-Mobile Convergence) offerings, ONE Albania continues to lag behind Vodafone in terms of reputation. As illustrated in Figure 5, while ONE is increasingly perceived as a value-for-money and customer-centric provider, especially in the fixed segment, Vodafone maintains a symbolic advantage rooted in its established market presence and global brand value.

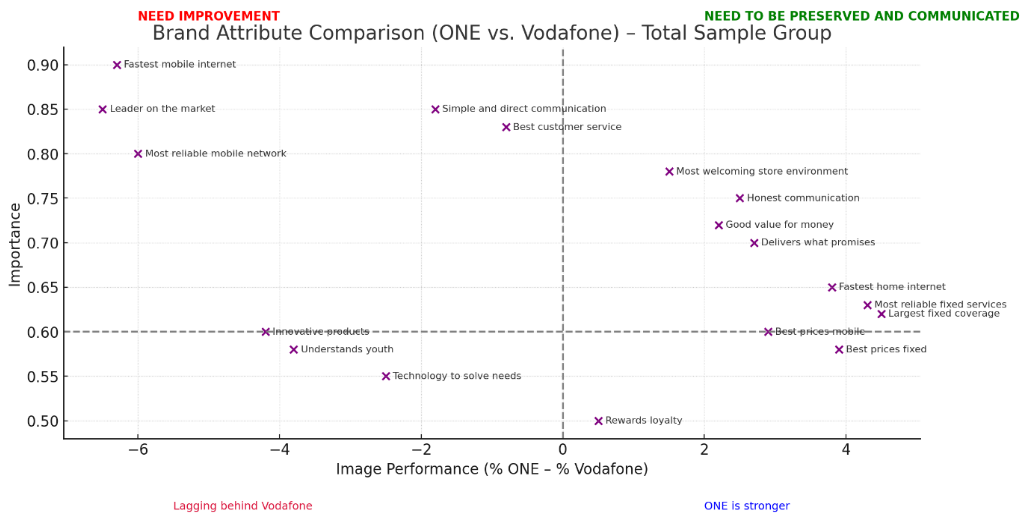


Figure 5. Brand image of ONE compared to Vodafone / Post merging perception against VDF

Source: Adapted from One Albania

The brand positioning map confirms that ONE's core strengths after the merger lie in attributes such as price competitiveness, customer centricity, and communication transparency. These strengths are particularly appreciated by value-sensitive segments, many of which are long-standing ALBtelecom customers. In contrast, Vodafone continues to dominate consumer perceptions on attributes such as network reliability, innovation, and premium status, particularly in mobile.

Operator	Market Position	Consumer Perception
ONE Albania	Growing, innovative, agile	Competitive pricing, innovative service
Vodafone	Legacy leader in mobile	Reputable, stable, international

Table 8. Post-Merger competitive positioning and brand perception

Source: authors

This perception gap reflects both the success and strategic limitations of ONE's current brand trajectory. While the company has successfully redefined its identity in terms of accessibility and functional value, it now faces the critical challenge of improving perceptions in areas historically within Vodafone's expertise, namely technology leadership and the quality of mobile services.

Consequently, this requires ONE to maintain its differentiated position in fixed services, while investing in mobile network performance and innovation. Closing this perception gap will be essential for ONE to consolidate its post-merger gains and position itself as a credible competitor to Vodafone's long-standing dominant position in the Albanian telecommunications market.

III.5 Financial Reporting

Financial reporting shows customer behaviour and preferences toward the brand. According to the 2023 annual activity report of AKEP in Albania (Electronic & Mailing Communication Authority), Vodafone reported high income. This clearly identifies, in financial terms, customers' preferences toward the Vodafone brand in Albania.

Indeed, although in 2023 "One Albania" commenced delivering services following the merger of "One Telecommunications" with "Albtelecom", "Vodafone Albania" remains the undertaking with the highest revenues in the mobile service market, accounting for about 56% of the market total (AKEP report, 2023).

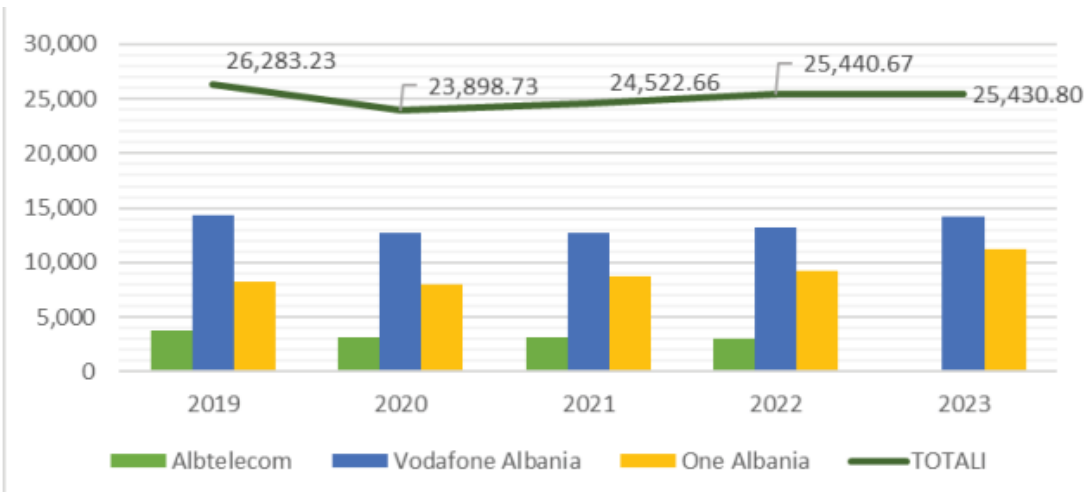


Table 9. Revenues reported by electronic communications undertakings (ALL m)

Source: AKEP

IV. DISCUSSION AND FUTURE DIRECTIONS

The results confirm that ONE Albania's post-merger strategy, based on authenticity, customer centricity, and strategic repositioning, has proven effective in restoring trust and loyalty, especially among the most vulnerable segments of its customer base. The significant improvement in Net Promoter Score (NPS) highlights an effective rebuilding of loyalty in areas where it was most at risk. At the same time, the increase in the TRI*M index reflects a strengthening of emotional engagement with the brand, especially among former ALBtelecom users who historically showed lower levels of satisfaction. Together, these results suggest that the new ONE brand has overcome the limitations of its traditional identities, positioning itself as a modern, responsive, and credible provider in a transitional, often sceptical market environment.

However, while short-term indicators point to successful integration, the long-term sustainability of this trajectory depends on several factors. First, consistent and transparent brand communication will be key to maintaining the gains achieved with the initial rebrand. Consumers in low-trust environments, such as Albania, need constant reassurance, not only through marketing strategies but also through service reliability, responsiveness, and price transparency.

Second, competitive differentiation must be constantly refined. While ONE has made significant progress in the fixed and FMC (Fixed-Mobile Convergence) segments, its relative weakness in mobile prestige, especially compared to Vodafone's established international reputation, remains a challenge. Bridging this gap will likely require a multifaceted approach that combines digital innovation, customer experience enhancements, and expanded loyalty mechanisms. Furthermore, as consumer expectations evolve alongside Albania's rapid digital transformation, the brand must remain agile, adapting its offering and messaging to new segments and usage patterns.

From a theoretical perspective, the case of ONE Albania reinforces the idea that successful post-merger brand integration is not simply a technical or operational challenge, but a strategic process that requires cultural alignment, emotional engagement, and ongoing measurement. The integration of brand equity from two distinct legacies – the functional depth of ALBtelecom and the emotional resonance of ONE – demonstrates the value of a hybrid positioning strategy, especially in post-transition markets.

Future research could extend these findings by adopting a longitudinal design to assess whether the positive consumer perception observed in the first year after the merger translates into lasting customer loyalty and increased market share over time. Comparative studies with similar post-merger telecom markets could provide further insights into how cultural and institutional factors influence consumer response to brand consolidation.

In practical terms, the findings of this study offer concrete guidance for marketers and strategic planners engaged in brand integration following M&A. On the other hand, it highlights the importance of aligning internal transformation efforts with external brand perception, especially in contexts where historical mistrust and competitive fragility coexist. By maintaining a clear commitment to customer-centric values and constantly adapting to the market, ONE Albania (and brands in similar contexts) can move beyond the initial recovery to achieve long-term consolidation and leadership.

REFERENCES

Alam, A., & Torany, H. (2023). Building Trust in Mergers and Acquisitions: An Exploration of Key Factors.

Ash, K. (2023). Brand Equity Valuation after Merger and Acquisition: A Review of Evidence. *Journal of Service Science and Management*, 16(2), 133-143.

Balmer, J. M., & Gray, E. R. (2003). Corporate brands: what are they? What of them? *European journal of marketing*, 37(7/8), 972-997.

Basu, K. (2006). Merging brands after mergers. *California Management Review*, 48(4), 28-40.

Braimllari, A., & Nerjaku, S. (2021). Factors Influencing the Usage of Online Shopping in Albania. In Proceedings of the *4th International Conference on Business, Management and Finance*, Oxford, UK (pp. 26-28).

Fruits, E., Hurwitz, G., Manne, G. A., Morris, J., & Stapp, A. (2019). A Review of the Empirical Evidence on the Effects of Market Concentration and Mergers in the Wireless Telecommunications Industry. *International Center for Law & Economics, White Paper* [accessed 16.04. 2025], 44.

Fruits, E., Hurwitz, J.G., Manne, G.A., Morris, J., & Stapp, A. (2020). Directorate for financial and enterprise affairs competition committee. *Global Forum on Competition: Static and Dynamic Effects of Mergers*: A Review of the Empirical Evidence in the Wireless Telecommunications Industry.

Gantumur, T., & Stephan, A. (2012). Mergers & acquisitions and innovation performance in the telecommunications equipment industry. *Industrial and Corporate Change*, 21(2), 277-314.

Gjoni, M. (2018). Pricing Strategies in Telecommunications. Mobile Telephony, The Case of Albania. *In Current directions of scientific research* (pp. 184-191).

Jacobsson, M., & Rickhammar, O. (2022). Recent M&A in the Telecom Industry: Industrial Wireless Networks.

Jorgji, E., Kristo, S., & Matraku (Dervishi), D. (2024). Albanian Firms in the Digital Era: Navigating the European Digital Single Market. *European Academic Research*, Vol. XII, Issue 6. Sep. 2024, ISSN 2286-4822

Kargas, A., Argyroulis, V., & Varoutas, D. (2023). Mergers and acquisitions in the telecommunications market: a simultaneous equations approach to study Structure, Conduct and Performance.

Kasapi, I. (2024). Customer Satisfaction in the Telecom Industry: A Survey Research Approach. *International Academic Conference on Management and Economics*. Vol. 2, Issue. 1, pp. 1-15, DOI: https://doi.org/10.33422/conferenceme.v2i1.662

Kernstock, J., & Brexendorf, T. O. (2012). Corporate brand integration in Mergers and Acquisitions– an action research-based approach. *Corporate Reputation Review*, 15, 169-178.

Kernstock, J., & Brexendorf, T. O. (2013). Corporate brand integration in mergers and acquisitions. In Contemporary Perspectives on Corporate Marketing (pp. 72-82). Routledge.

Kumar, B. R. (2012). Mergers and acquisitions in the telecommunications industry. In Mega Mergers and Acquisitions: Case Studies from Key Industries (pp. 60-95). London: Palgrave Macmillan UK.

Leskaj, E., & Lazimi, L. (2023). The Telecommunications Industry and Its Impact on the Albanian Economy. EIRP Proceedings, 18(1), 292-301

Mario, G. (2019). Quality of Service in Mobile Telecommunications in Albania–Application of Marketing Strategies. *International Journal of Management Excellence* (ISSN: 2292-1648), 13(3), 2005-2018.

McLelland, M. A., Goldsmith, R., & McMahon, D. (2014). Consumer reactions to the merger: Understanding the role of pre-merger brands. Journal of Brand Management, 21, 615-634.

Rieck, O., & Doan, C. T. (2009). Shareholder wealth effects of mergers and acquisitions in the telecommunications industry. Telecommunication Markets: Drivers and Impediments, 363-394.

Shruthi, M.V. (2018). Antitrust and Competition Law's Impact on M&A: Regulatory Challenges and Outcomes in Telecom, Healthcare, and Tech. Iconic Research and Engineering Journals, Volume 2, Issue 1, Page 75-85

Shyle, I., & Xhelo, E. (2017). Customer satisfaction in the service industry: a case study of the telecommunication industry in Albania. In the *1st International Conference, Engineering and Entrepreneurship Proceedings*.

Srinuan, C., & Srinuan, P. (2023). A review of recent mergers and acquisitions in the mobile telecommunications service industry: Substantial issues and implications for policy makers.

Steigenberger, N. (2017). The challenge of integration: A review of the M&A integration literature. *International Journal of Management Reviews*, 19(4), 408-431.

Terezi, N., Karapici, E., Terezi, E. (2016). The effect of demographic variables on the user's behavior in the Mobile Telecommunication Market of Albania. *Academic Journal of Business, Administration, Law and Social Sciences.* Vol. 2 No. 1, March, ISSN 2410-3918

Thorbjørnsen, H., & Dahlén, M. (2011). Customer reactions to acquirer-dominant mergers and acquisitions. *International journal of research in marketing*, 28(4), 332-341.

Vũ, D. A., & Moisescu, O. I. (2013). Brand Integration Practices in Mergers and Acquisitions. *Management & Marketing*, 8(3).

Warf, B. (2003). Mergers and acquisitions in the telecommunications industry. *Growth and Change*, 34(3), 321-344.

Wilcox, H. D., Chang, K. C., & Grover, V. (2001). Valuation of mergers and acquisitions in the telecommunications industry: a study on diversification and firm size. *Information & Management*, 38(7), 459-471.

Yang, D., Davis, D. A., & Robertson, K. R. (2012). Integrated branding with mergers and acquisitions. *Journal of Brand Management*, 19(5), 438-456.

Yao, P., & Wang, X. (2018). Research on the relationship between the weaker enterprises' post-merger brand strategy and consumers' purchase intention: Based on the brand authenticity. Journal of Contemporary Marketing Science, 1(1), 34-52.

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). SAGE Publications.

https://tiranapost.al/english/stil-jete/one-albania-forcon-fokusin-komercial-pervoja-e-abonenteve-behet-objekt-i527766

https://www.one.al/en/Press_Kit/News/Presenting_ONE_Albania_new_brand_telecommunications_market-2023_03_15

https://www2.one.al/en/Press_Kit/News/ONE_Albania_spins_commercial_focus-2023_08_29?utm_source

15

INFORMATION DIGITALISATION AS A KEY DRIVER TO ACHIEVE IMPROVEMENT OF SME PERFORMANCE

DOI: 10.37199/c41000315

Emi HOXHOLLI
POLIS University (Tirana, Albania)
emi_hoxholli@universitetipolis.edu.al
ORCID 0009-0003-3166-1776

Godiva RËMBECI
POLIS University (Tirana, Albania)
godiva_rembeci@universitetipolis.edu.al
ORCID 0000-0002-5396-444X

Abstract

SMEs play a crucial role in the development of the national economy by contributing substantially to employment and GDP. According to the latest INSTAT data, enterprises with 1 to 250 employees represent about 98% of the total business stock and generate roughly 75% of GDP. Nearly 90% of total employment in the country depends on SMEs' economic performance. In the context of globalisation and EU integration, technology and innovation have become essential for SMEs to remain competitive in the global market. Digitalisation is emerging as a key driver of competitiveness and growth for SMEs worldwide. In Albania, the adoption of digital tools is transforming business operations, increasing efficiency, and supporting better decision-making; however, its specific impact in the Albanian setting remains insufficiently explored. By embracing digital tools and transformation strategies—including government-led digitalisation initiatives—Albanian SMEs can overcome existing barriers and accelerate their development. Despite this potential, research on IT implementation in SMEs remains limited and often treats the sector as homogeneous. This paper aims to assess the impact of digitalisation on Albanian SMEs, emphasising its role in improving productivity. Using secondary business statistics from INSTAT, the study evaluates SME productivity, economic performance, and the contribution of digitalisation to enhanced outcomes. The findings indicate that digitalisation is a key factor in ensuring the long-term sustainability and competitiveness of SMEs in Albania.

Keywords: SME, productivity, performance, digitalisation of information

I. INTRODUCTION

Digitalisation is transforming economies, public governance, and social behaviour worldwide. The UN notes that digital transformation is essential for achieving nearly 70% of the SDGs by 2030 (UNDP, 2023). The process of digitalisation accelerated significantly during the COVID-19 pandemic and continues to evolve, impacting businesses of all ages, sizes, and industries. Reflecting global trends, the Albanian government has made digitalisation a strategic priority. The Innovation-Driven Entrepreneurship Development Strategy 2024–2030 focuses on advancing SMEs, entrepreneurship, and innovation, supported by the National Strategy for Science, Technology, and Innovation and the Digital Agenda of Albania 2021–2026. Together, these policies aim to strengthen innovation, promote collaboration with research institutions, and foster a technologically driven business ecosystem. Several international studies have investigated digitalisation processes across different industries and their impact on SMEs, finding a positive relationship between the introduction of information technology (IT) and digitalisation and financial performance (Kadarova, J., Lachvajderova, L., Sukopova, D., 2023). Access to digital technologies and ICT is essential for developing competitive, high-value products and services and for enabling SMEs to enter foreign markets. Technology adoption also supports SME growth by helping enterprises scale more rapidly. Strong innovation and R&D infrastructures improve access to digital tools and generate knowledge that the education system transfers to entrepreneurs, enabling them to convert it into economic value. The broader business and innovation ecosystem integrates these elements to support entrepreneurship, particularly for start-ups and SMEs, by fostering innovation and collaboration among key actors. Mobile technologies and digital platforms help increase SME revenue and strengthen competitiveness. In Albania, the adoption of digital tools is changing how SMEs operate by improving efficiency and decision-making, yet their specific impact remains insufficiently studied. This paper examines how digitalisation of information affects Albanian SMEs, with a focus on its role in enhancing productivity.

I. LITERATURE REVIEW

According to the literature, there is no single international definition of an SME, as definitions vary by country and are generally based on the number of employees or annual turnover. Typically, both quantitative and qualitative indicators, such as profits, capital, market position, number of employees, and turnover, are used to determine enterprise size. Some researchers categorise a company as an SME by combining the "number of employees" criterion with an additional measure. In contrast, in rarer cases, the employee criterion is entirely replaced by other indicators depending on the study's focus. For example, financially oriented research may rely on annual turnover or balance sheet data to define SMEs (Piller, 2000). Despite these variations, the most widely used definition in contemporary studies is that of the European Union:

enterprises are classified as SMEs if they have fewer than 250 employees, an annual turnover of less than €50 million, or a balance sheet total of not more than €43 million (EC, 2020).

In Albania, SME definitions follow similar criteria. Albanian legislation classifies micro, small, and medium-sized enterprises as those with fewer than 250 employees and an annual turnover and/or balance sheet not exceeding 250 million LEK (Hoxholli, 2021). SMEs are central to the Albanian economy. According to the Business Development and Investment Strategy drafted by the Ministry of Finance and Economy (MFE) for 2021–2027, SMEs represent approximately 99.8% of all active enterprises, with 43.4% operating in the trade sector. These enterprises are predominantly domestic but have increasingly expanded into international markets. SMEs account for 73% of GDP and 71% of total employment, contributing relatively more to employment and value added than their EU counterparts, underscoring their critical importance for economic development, job creation, and competitiveness.

From 2010 to 2023, the number of SMEs in Albania grew steadily, with more than 40,388 new enterprises established. Employment in the sector rose correspondingly, with 283,891 individuals employed over this period. Annual turnover nearly doubled, from 1.29 billion ALL in 2010 to 3.45 billion ALL in 2023, highlighting SMEs as a key driver of employment and economic growth. Regarding investments, secondary data from INSTAT shows the proportion of equipment investment relative to total investment and annual growth rates. The highest growth occurred in 2011, linked to increased government incentives during the 2010 electoral campaign. Conversely, 2020 experienced negative growth due to the COVID-19 pandemic, which continued into 2021. In 2022, total investment growth rebounded, supported by economic expansion and rising SME revenues, although the war in Ukraine contributed to a downturn, affecting revenue growth and total investment. Between 2010 and 2023, the Albanian SME sector expanded significantly, doubling its workforce and achieving record value added. While the post-pandemic period temporarily boosted revenues, growth stagnated in 2023, and capital investments—particularly in equipment—fell below levels seen a decade earlier.

The Information and Communication Technology (ICT) sector is pivotal to digital transformation (OECD, 2025). Enterprises must adapt to the digital era by leveraging digital tools to manage growing volumes of data and support decision-making (Marcysiak & Pleskacz, 2021). For SMEs, digitalisation refers to the adoption and use of digital technologies in business activities to enhance operational efficiency, customer engagement, and market reach (Bouwman et al., 2019; Mandviwalla & Flanagan, 2021; Zhao, Honigsberg & Mandviwalla, 2025). It is important to distinguish between digitalisation and digital transformation. Digitalisation is primarily technology-driven, involving the integration of digital tools into business processes (Marcysiak & Pleskacz, 2021). In contrast, digital transformation leverages new technologies to fundamentally enhance production, expand enterprise scope, and drive strategic change (Emara & Zhang, 2021; Brozzi et al., 2021). Different studies show that digital transformation offers growth and competitiveness opportunities for companies of all sizes (Teplická, Hurná & Seňová, 2019). For SMEs, strategic and appropriate adoption of digital technologies can stimulate productivity, competitiveness, and overall performance (Ardito, Raby, Albino & Bertoldi, 2021). Effective implementation of digital tools significantly improves SME performance (Mushtaq, Gull & Usman, 2022).

II. METHODOLOGY

As previously mentioned, the paper examines the impact of digitalisation on Albanian SMEs, highlighting its role in enhancing productivity. At the European level, the model used for that purpose is called the 'SBA Fact Sheets' (The Small Business Act for Europe). Introduced in 2011, it is a valuable tool for facilitating SME policy assessments. This document is published annually and prepared using the latest statistics and data. It is important to underline that this document should be considered as an additional source. Information that helps policymakers to improve the policy process and conduct evidence-informed governance. The concept of 'SME performance' used in this document refers to a broader set of criteria, including entrepreneurship, second-chance responsive administration, state aid and public procurement, access to finance, the single market, skills and innovation, and the environment. Another model to evaluate SME performance is the one proposed by the Asian Productivity Organisation, which includes a large number of indicators, as described in the following table.

1. Sales per employee	15. Labour cost per employee
2. Customer satisfaction index	16. Labour cost competitiveness
3. Complaint ratio	17. Employee turnover rate
4. Compliment ratio	18. Employee satisfaction index
5. Customer retention	19. Employee participation rate in team activity
6. Sales growth	20. Employee participation rate in the suggestion scheme
7. Value Added to sales ratio	21. Cost savings from employee involvement activities
8. Profit margin	22. Training hours per employee
9. Annual inventory turns	23. Training expenditure/sales
10. Defect rate	24. Absenteeism rate
11. Customer rejects/return	25. Capital productivity
12. Scrap/rework level	26. Sales per dollar of capital
13. On-time delivery commitment	27. R&D investment ratio
14 Labour productivity	28. Capacity utilisation rate
	29. Labour productivity

Table 1. SME Performance Indicators

Source: Author's processing

The advantages of this model include simplicity in application, clarity, and transparency in the compilation methods. Given these advantages and the lack of a technical guide or handbook explaining what, why, and how to compile performance indicators in Albania, the above-mentioned handbook should be considered in conducting a preliminary analysis of Albanian SMEs. Given this model and the availability of annual business statistics, it is possible to establish a national model for SME performance indicators in Albania. However, applying the model in the Albanian context is only possible by selecting a subset of the recommended performance indicators, mostly those related to quantitative data. The remaining performance indicators are missing due to insufficient information on qualitative surveys of SMEs in the country (Rembeci, 2017). Based on the literature review, SMEs' performance has been measured using a comprehensive set of financial and non-financial indicators to capture multiple dimensions of performance. Regarding financial indicators, the authors have included revenue, sales, operational efficiency, customer acquisition, and related metrics.[2] Secondary data on business statistics published by INSTAT are used to measure SME productivity, economic performance, and the role of digitalisation in improving SME performance.

III. RESULTS

Based on the analysis of secondary data published by INSTAT, we present below the relationship between digitalisation and SMEs. The table below shows the structure of SMEs based on two main dimensions: the absolute number of enterprises and the year-over-year percentage change for the period 2018-2023.

Sector	Enterprises (no), total economy						Index of chain in %				
	2018	2019	2020	2021	2022	2023	2019	2020	2021	2022	2023
Production	13,560	13,220	12,925	13,734	14,738	15,606	***-2.5%***	***-2.2%***	***6.3%***	***7.3%***	***5.9%***
1 - 4	9,396	8,932	8,366	9,222	10,091	10,641	-4.9%	-6.3%	10.2%	9.4%	5.5%
5 - 9	1,626	1,612	1,834	1,836	1,876	2,134	-0.9%	13.8%	0.1%	2.2%	13.8%
10 - 49	1,817	1,911	2,010	1,969	2,021	2,106	5.2%	5.2%	-2.0%	2.6%	4.2%
50 +	720	765	715	706	750	725	6.3%	-6.5%	-1.3%	6.2%	-3.3%
Services	93,890	90,870	89,649	90,297	98,721	101,685	***-3.2%***	***-1.3%***	***0.7%***	***9.3%***	***3.0%***
1 - 4	84,599	79,983	79,717	80,174	87,948	90,389	-5.5%	-0.3%	0.6%	9.7%	2.8%
5 - 9	5,254	6,397	5,642	5,790	6,217	6,497	21.8%	-11.8%	2.6%	7.4%	4.5%

[2] https://ec.europa.eu/eurostat/documents/3859598/12453409/KS-GQ-21-001-EN-N.pdf

10 - 49	3,429	3,825	3,664	3,639	3,821	4,070	11.5%	-4.2%	-0.7%	5.0%	6.5%
50 +	608	665	627	695	736	729	9.4%	-5.7%	10.8%	5.9%	-1.0%

Table 1. Indicator 1, number of enterprises in production and service sectors (2018–2023)

Source: Authors' processing based on INSTAT data

Looking at the data by sector, we see that in the production sector, the total number of SMEs decreased from 13,560 in 2018 to a low of 12,925 in 2020. It then recovered and grew to 15,606 by 2023. The vast majority of these are small companies (size 1-4), accounting for over 10,000 units by 2023. The sector saw negative growth in 2019 (-2.5%) and 2020 (-2.2%), followed by a strong recovery with 6.3% growth in 2021, 7.3% in 2022, and 5.9% in 2023. On the other hand, the service sector is significantly larger than the production sector. It started with 93,890 enterprises in 2018, dipped slightly in 2019 and 2020, and rose to 101,685 by 2023. As with production, the smallest companies (sizes 1-4) dominate, numbering 90,389 in 2023. The sector contracted in 2019 (-3.2%) and 2020 (-1.3%). It returned to growth in 2021 (0.7%), followed by a significant jump in 2022 (9.3%) and continued growth in 2023 (3.0%). The data illustrate a general "U-shaped" economic trend for SMEs in Albania. There was a contraction in the economy during 2019 and 2020, which coincided with the global COVID-19 pandemic, followed by a robust recovery and expansion phase from 2021 through 2023 across both production and service sectors.

Sector	**Employees (no), total economy**						**Index of chain in %**				
	2018	**2019**	**2020**	**2021**	**2022**	**2023**	**2019**	**2020**	**2021**	**2022**	**2023**
Production	184,850	190,636	185,434	186,720	192,926	196,167	***3.1%***	***-2.7%***	***0.7%***	***3.3%***	***1.7%***
1 - 4	17,848	17,158	16,086	17,690	18,377	19,299	-3.9%	-6.2%	10.0%	3.9%	5.0%
5 - 9	10,391	10,540	11,803	11,717	12,283	13,999	1.4%	12.0%	-0.7%	4.8%	14.0%
10 - 49	38,198	41,098	42,878	41,868	42,440	43,924	7.6%	4.3%	-2.4%	1.4%	3.5%
50 +	118,413	121,839	114,666	115,444	119,827	118,945	2.9%	-5.9%	0.7%	3.8%	-0.7%
Services	330,856	340,218	318,552	332,520	355,015	363,756	***2.8%***	***-6.4%***	***4.4%***	***6.8%***	***2.5%***
1 - 4	133,729	127,362	123,068	125,067	136,917	137,598	-4.8%	-3.4%	1.6%	9.5%	0.5%
5 - 9	32,898	41,343	35,442	36,776	39,021	41,316	25.7%	-14.3%	3.8%	6.1%	5.9%
10 - 49	64,752	73,353	68,469	69,450	72,567	76,640	13.3%	-6.7%	1.4%	4.5%	5.6%
50 +	99,476	98,160	91,573	101,227	106,511	108,201	-1.3%	-6.7%	10.5%	5.2%	1.6%

Table 2. Indicator 2, employment dynamics by company size and sector (2018–2023)

Source: Authors' processing based on INSTAT data

While the previous table showed a U-shaped pattern in the count of businesses, Table 2 shows a different picture for the number of employees. In 2019, while the number of production businesses was falling (previous table), employment was still rising (3.1%). Companies were likely to try to hold onto staff until the situation became untenable. Between the two sectors, roughly 26,000 jobs were lost in 2020 (dropping from approx. 530,000 total to 504,000). There is an interesting divergence in how differently sized companies recovered after the 2020 crash. Micro Companies (1-4 employees): In the production sector, they had a massive hiring boom in 2021 (+10.0%). While Large Companies (50+ employees): In the service sector, they had a massive boom in 2021 (+10.5%). The employment data confirms the U-shaped economic trend but highlights that the Service sector was the primary driver of both the crash and the subsequent recovery. By 2023, the economy (in terms of employment) will be significantly larger than in 2018, indicating a full recovery and a transition into a growth phase.

	2015	**2016**	**2017**	**2018**	**2019**
Total	95.0	95.6	96.0	97.3	97.5
10-49 employees	94.1	95.0	95.6	96.7	96.9
50-249 employees	100.0	98.7	97.6	99.7	99.9
250+ employees	97.8	100.0	100.0	100.0	100.0

Table 3. Percentage of enterprises using computers by size and year (2015–2019)

Source: Authors' processing based on INSTAT data

INSTAT data show that computer use has become a near-universal standard for businesses in Albania, regardless of size. Between 2015 and 2019, the total percentage of enterprises using computers rose steadily from 95.0% to 97.5%, indicating a mature level of digitalisation. Smaller enterprises (10–49 employees) have historically trailed slightly behind their larger counterparts; however, they have shown consistent growth, rising from 94.1% to 96.9%. This indicates that the "digital divide" between small and large businesses is narrowing significantly. Medium-sized businesses (50–249 employees) have maintained near-perfect usage rates, ending the period at 99.9%, despite a minor statistical fluctuation in 2016–2017. Large firms (250+ employees) have effectively reached a saturation point. Since 2016, 100% of these entities have reported using computers, suggesting that digital infrastructure is a non-negotiable requirement for large-scale operations. By 2019, computer usage in Albanian enterprises had shifted from being a distinct advantage to a fundamental operational baseline. The data suggests that the lack of computer access is now extremely rare, found in only a very small fraction (less than 3.1%) of the smallest businesses.

	2015	2016	2017	2018	2019
Total	96.2	96.8	96.9	97.5	97.8

10-49 employees	95.5	96.1	96.6	97.5	97.3
50-249 employees	98.6	100.0	98.1	99.5	99.9
250+ employees	100.0	100.0	100.0	100.0	100.0

Table 4. Percentage of enterprises with internet access by size (2015–2019)

Source: Authors' processing based on INSTAT data

This table closely parallels the previous "Computer Usage" data, but with slightly higher starting percentages, suggesting that for modern Albanian businesses, internet connectivity is as essential—if not more so—than a traditional computer workstation. It is worth noting that in 2015, "Internet Access" (96.2%) exceeded "Computer Usage" (95.0%). This implies that, as early as 2015, a small segment of businesses utilised the internet (perhaps via mobile phones or POS systems) without reporting the use of a standard desktop/laptop computer. By 2019, the numbers align almost perfectly (97.8% internet vs 97.5% computer). Small businesses (10-49 employees) exhibit a high level of connectivity, peaking at 97.5% in 2018, then dropping slightly to 97.3% in 2019. Interestingly, medium-sized enterprises (50-249 employees) exhibit fluctuations similar to those in the computer usage table. They hit 100% in 2016, dropped to 98.1% in 2017, and recovered to 99.9% by 2019. This recurrence suggests a specific characteristic of the sample group or economic conditions for medium businesses in 2017 rather than a data error. As with computer usage, large enterprises (250+ employees) show perfect saturation (100%) across the entire 5-year period. The historical disparity between large corporations and small businesses regarding basic connectivity has effectively vanished. Since basic access is achieved, the focus for these businesses likely shifts from *getting* online to *optimising* digital tools (e-commerce, cloud computing, and cybersecurity)

	2015	2016	2017	2018	2019
Total	8.8	7.1	7.7	5.6	5.8
10-49 employees	9.2	6.7	6.4	4.6	5.0
50-249 employees	7.3	6.4	13.3	9.0	8.2
250+ employees	7.4	9.7	12.9	16.9	15.1

Table 5. E-commerce adoption rates by enterprise size in Albania (2015–2019)

Source: Authors' processing based on INSTAT data

The data, sourced from INSTAT (Albania), reveal a fascinating and contradictory trend in e-commerce adoption between 2015 and 2019. While the overall total suggests a decline in businesses selling online (from 8.8% in 2015 to 5.8% in 2019), these averages mask a significant divide between large and small companies. Small businesses show the opposite trend. In 2015, they were the leaders in online sales (9.2%). By 2019, this number had nearly halved to 5.0%. Since small businesses make up most companies in an economy, their decline is dragging down the "Total" average, making the national e-commerce landscape look worse than it is for larger players. Medium-sized businesses have seen fluctuating results. They experienced a massive spike in 2017 (13.3%), significantly outperforming their 2015 numbers, but settled back down to 8.2% by 2019. Large corporations show the most positive trend. In 2015, they had one of the lowest adoption rates at 7.4%. However, they steadily increased their online presence year over year, peaking at 16.9% in 2018 before settling at 15.1% in 2019. This suggests that large Albanian companies have successfully invested in digital transformation over this period. The data illustrates a widening "Digital Divide" based on company size. In 2015, small companies were more likely to sell online than large ones (9.2% vs 7.4%). By 2019, the situation had completely reversed: large companies were three times more likely to sell online than small ones (15.1% vs 5.0%). This suggests that while large enterprises in Albania are finding value and resources to sell online, small businesses may face barriers—such as technical costs, logistics challenges, or competition—that force them to abandon online sales channels.

IV. CONCLUSIONS AND FOLLOW-UP

The comprehensive analysis of INSTAT data on Albanian SMEs reveals a complex economic landscape marked by resilience in employment but a growing digital competitiveness gap. The structural data (Tables 1 and 2) confirm that the Albanian SME sector follows a distinct "U-shaped" trajectory. Despite the contraction in enterprise numbers and employment during the 2019–2020 period, driven largely by the shocks associated with the global pandemic, the economy demonstrated robust resilience. By 2023, both the production and service sectors had not only recovered but expanded beyond their 2018 baselines. The service sector remains the dominant engine of employment, while the production sector has seen a specific resurgence in micro-enterprises. In terms of digital infrastructure, the data (Tables 3 and 4) indicates a significant victory for digitalisation efforts in Albania. The historical gap in access to basic technology has effectively closed. With computer usage and internet access rates nearing saturation (97%+) for businesses of all sizes, the "First Digital Divide" defined by the lack of hardware or connectivity, is no longer a primary barrier to entry for Albanian SMEs. However, a critical paradox emerges when contrasting connectivity with e-commerce adoption (Table 5). While small businesses are just as connected to the internet as large corporations, they are failing to leverage this connectivity for commercial gain. The sharp decline in online sales among small enterprises (from 9.2% to 5.0%) stands in stark contrast to the surge among large enterprises (from 7.8% to 15.1%). Consequently, the challenge for stakeholders and policymakers in Albania has shifted. The focus cannot be solely on providing internet access or hardware. Instead, to ensure that the economic recovery observed in 2021–2023 is sustainable and inclusive, future support must focus on digital transformation: helping small SMEs bridge the gap between simply having the internet and successfully selling on it. To deepen our understanding of the link between the 'digital' and the 'economic' trends identified above, we now turn to a statistical analysis of the data. This section presents the calculation of performance indicators and examines the correlations between digital intensity and key growth metrics such as employment and turnover. This step is essential to move from simply observing the presence of technology to quantifying its impact on SMEs' actual operational success.

REFERENCES

Ardito, L., Raby, S., Albino, V., & Bertoldi, B. (2021). From marketing to technology: The influence of family management on the scope of digital transformation in SMEs. *Technological Forecasting and Social Change, 172*, 121013. https://doi.org/10.1016/j.techfore.2021.121013

Bouwman, H., Nikou, S., Molina-Castillo, F. J., & de Reuver, M. (2019). The impact of digitalisation on business models. *Electronic Markets, 29*(4), 1–14. https://doi.org/10.1007/s12525-019-00374-9

Brozzi, R., Forti, D., Rauch, E., & Matt, D. T. (2021). The advantages of Industry 4.0 applications for sustainability: Results from a sample of manufacturing companies. *Sustainability, 12*(10), 4063.

Emara, A., & Zhang, J. (2021). The impact of digital transformation on the economic growth of the GCC countries. *International Journal of Economics, Commerce and Management, 9*(3), 1–16.

European Commission. (2020). *User guide to the SME definition*. Publications Office of the European Union. https://data.europa.eu/doi/10.2873/620234

Hoxholli, E. (2021). Empirical research on the relation between internal influencing factors on the internationalisation: Albanian SMEs. *International Journal of Ecosystems and Ecology Science (IJEES), 10*(1), SN 37.

Kádárová, J., Lachvajderová, L., & Šukopová, D. (2023). The impact of digitalisation on the financial performance of SMEs. Management Systems in Production Engineering, 31(1), 1–8.

Mandviwalla, M., & Flanagan, R. (2021). Small business digital transformation in the context of the pandemic. International Journal of Information Management, 59, 102349. https://doi.org/10.1016/j.ijinfomgt.2021.102349

Marcysiak, A., & Pleskacz, Ż. (2021). Determinants of digitalisation in SMEs. Entrepreneurship and Sustainability Issues, 9(1), 278–293. https://doi.org/10.9770/jesi.2021.9.1(17)

Ministry of Finance and Economy. (2021). *Business and investment development strategy 2021–2027*. Government of Albania.

Mushtaq, R., Gull, A. A., & Usman, M. (2022). ICT adoption and SME performance: The moderating role of digital leadership. *Technology Analysis & Strategic Management.* https://doi.org/10.1080/09537325.2021.1988912

OECD. (2025). *OECD Digital Economy*. OECD Publishing.

Piller, F. T. (2000). *Mass customization: Ein wettbewerbsstrategisches Konzept im Informationszeitalter*. Gabler Verlag.
Teplická, K., Hurná, S., & Seňová, A. (2019). The influence of the introduction of new technologies on the economic efficiency of the enterprise. *Acta Montanistica Slovaca, 24*(3), 256–265.

UNDP. (2023). *Digital strategy 2022–2025: Long-term ambition for the #FutureOfDevelopment*. United Nations Development Programme.

16

SAFEGUARDING DIGITAL AUTHENTICITY AND WOMEN'S IDENTITY THROUGH DEEPFAKE DETECTION

DOI: 10.37199/c41000316

Livia IBRANJ
POLIS University (Tirana, Albania)
livia_ibranj@universitetipolis.edu.al

Abstract

Deepfake technology, the algorithmic manipulation of images and videos, is renowned for its ability to create highly realistic and stimulating content. They leverage deep learning and train generative neural architectures to map voices and faces onto another person's body. Using Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) enables accurate manipulation of voices, facial features, and expressions to create images and videos that closely resemble real people. While this may appear as a technological achievement, deepfakes have enabled profound harms, including identity theft, harassment, and non-consensual explicit imagery, with women comprising 96% of victims.

This paper explores multimodal detection approaches that combine deep learning features with forensic analysis to differentiate AI-generated images from authentic photographs. Our methodology integrates complementary detection strategies: deep semantic features via EfficientNet-B3 and CLIP models (2,304 dimensions), frequency-domain analysis detecting spectral anomalies, noise residual statistics, Local Binary Pattern texture descriptors, and facial forensics—totalling 2,339 features per image. An ensemble classifier combining Gradient Boosting and Logistic Regression was trained on 200 images (100 authentic photographs, 100 AI-generated from Midjourney, Stable Diffusion, and DALL-E), achieving 85% accuracy with 98.25% ROC-AUC. Performance analysis reveals asymmetric characteristics: 95% recall for authentic images versus 75% recall for AI-generated content, while maintaining 93.75% precision on synthetic detection. The 25% false negative rate underscores that technical detection alone cannot solve deepfake abuse—comprehensive protection requires platform accountability, legislative frameworks, and victim support systems. This study contextualises technical findings within the social crisis of digital sexual violence, examines documented psychological impacts on victims, identifies critical legal gaps, and outlines future research directions, including larger datasets, temporal analysis, and hybrid human-AI detection systems.

Keywords: Deepfake Detection, Generative Adversarial Networks, Artificial Intelligence (AI), Multimodal Feature Extraction

I. INTRODUCTION

The rapid development of artificial intelligence has ushered in a new era of synthetic media. Deepfake technology, which uses algorithms to produce incredibly lifelike but fake images and videos, is a prime example of this. Leveraging deep learning architectures such as Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), deepfakes manipulate voices, facial features, and expressions to produce content that is often indistinguishable from authentic media. Although these technological advancements hold many exciting potential benefits, there are also significant concerns about misinformation, privacy breaches, and digital security.

The consequences of deepfake abuse are not distributed equally across society. Research indicates that 96% of deepfakes are non-consensual explicit content, with the major target being women (Sensity, 2023). High-profile cases, such as the 2018 targeting of journalist Rana Ayyub and the mass victimisation of South Korean women in 2024, demonstrate how this technology amplifies existing patterns of gender-based violence and harassment. The psychological impact on victims mirrors that of revenge pornography, including trauma, anxiety, depression, and social withdrawal, even though victims never consented to the original imagery.

This paper investigates methods for distinguishing AI-generated images from real photographs using pixel-level analysis, feature extraction, and machine learning classification. By focusing on the technical challenges of detection, this research aims to contribute tools that safeguard authenticity in digital media. This study also emphasises how deepfake technology affects security and privacy, underscoring the need for trustworthy detection techniques. As generative models become more complex, it is imperative to improve detection precision and robustness to shield people, especially women and children, and society from the growing dangers posed by synthetic media.

II. LITERATURE REVIEW

II.1 Evolution and Technical Foundations of Deepfake Technology

What are deepfakes and why do they matter? Deepfakes are synthetic audio, image, or video outputs generated using advanced machine learning (ML) techniques, most commonly Generative Adversarial Networks (GANs). These outputs replicate human appearance, voice, and behaviour with a level of realism that often renders manual detection difficult. As the technology has become more accessible, concerns have increased regarding misinformation, reputational harm, and the exploitation of individuals who appear in non-consensual synthetic media.

Deepfake technology emerged publicly in 2017 when an anonymous Reddit user posted AI-manipulated pornographic videos of celebrities. By April 2018, the phenomenon gained mainstream attention through a Buzzfeed demonstration featuring a synthetic video of President Obama, illustrating both the technology's capabilities and its potential for misinformation (Facebook, 2018). The rapid proliferation was striking: by late 2018, 96% of the 14,678 identified deepfakes online were pornographic, with women disproportionately targeted (Deeptrace, 2018). Applications like DeepNude (2019) and Telegram bots (2019) further democratised the creation of non-consensual explicit imagery, affecting over 45,000 users within months (Hao, 2020).

II.1.1 Technical Foundations: Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs) are neural network-based architectures in which two models, the Generator and the Discriminator, are trained simultaneously in an adversarial framework (Goodfellow et al., 2014, p. 1). In this collaboration, the Generator produces synthetic images, whereas the Discriminator evaluates whether an input image is authentic or generated. The analysis begins with very basic features and progresses to increasingly more high-level ones, each layer compressing the input more. During training, the Generator iteratively improves its output, making it increasingly difficult for the Discriminator to distinguish synthetic images from real ones. Conversely, the Discriminator improves its ability to detect artefacts and inconsistencies. This competitive process drives both components toward higher performance and results in outputs that closely approximate the distribution of the training data (de Vries, 2020, p. 2113).

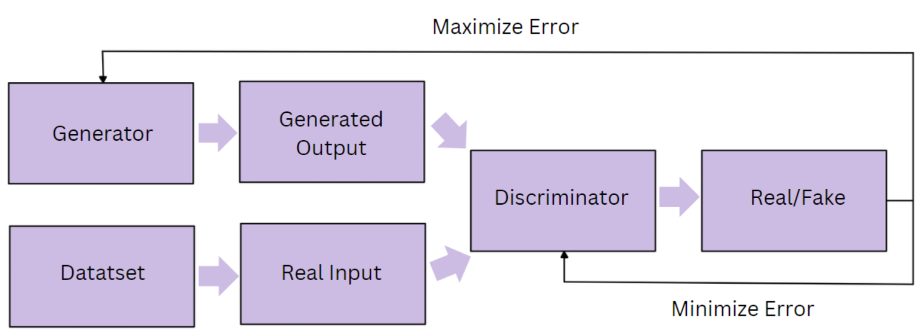


Figure 1. GAN: Neural Networks 1

Source: Author's processing

II.1.2 Machine Learning Concepts Underlying Deepfakes

The foundation of Machine Learning (ML) relies on a preference for implicit rather than explicit programming. Unlike traditional AI methods, where a computer system is given a predefined model by its designer, ML systems are designed to autonomously generate models from sets of examples (de Vries, 2020, p. 2111). Consequently, the computer does not receive direct instructions but learns independently, hence the term Machine Learning. The distinction between supervised and unsupervised learning is the type of input used and the resulting output (de Vries, 2020, p. 2112).

Supervised ML systems operate on labelled examples. In Supervised Machine Learning (ML), it is akin to having a teacher guide you through a set of well-labelled textbooks. For instance, if you want to distinguish among different bird species, the teacher gives you a collection of bird images, each labelled with its name. Additionally, you learn to recognise and classify various bird species based on the provided labels. On the contrary, Unsupervised Machine Learning is comparable to exploring a library filled with unmarked books. In this analogy, imagine delving into a section dedicated to birds, but without any labels on the books. In this case, you will not be explicitly taught about different bird species. Instead, you observe common themes and patterns across the books, leading to a broader understanding of birds and their characteristics. This exploration does not focus on specific classifications but on generating output that resembles their training data.

II.2 Psychological Dimensions of Digital Sexual Violence

The act of distributing sexual imagery is as old as time itself. Only in the last century has it been monetised and transformed into a billion-dollar industry. While it has become the norm to engage in this behaviour, we are still unable to profile when consuming this media becomes pathological. This section reviews the neuroscience of porn addiction and how we react to deepfakes.

The utilisation of online pornography, also referred to as Internet pornography, is recognised as a potentially addictive behaviour specific to the Internet. This behaviour encompasses the use of the Internet for engaging in various sexually gratifying activities, predominantly involving the consumption of pornography (Cooper, A. 2004). Research has shown that regular use of pornography is linked to issues such as objectification, sexual deception, and sexual aggression. In a study, researchers discovered that the consumption of reality TV, sports programming, and pornography was correlated with increased acceptance of objectification of women and more frequent engagement in acts of sexual deception (Seabrook et al. 2019).

Lead researcher and a psychologist at Princeton University, Susan Fiske, experimented in 2008 involving 21 men. The goal was to study male behaviour and how they respond to the exposure of scantily clad women. Brain scans conducted during the study indicated increased activity in the brain region associated with tool use when men viewed such images. Additionally, the men were more inclined to associate sexualised images of women with first-person action verbs like "I push, I grasp, I handle," according to Fiske. These findings suggest a neural pattern consistent with dehumanisation and instrumental perception of women.

In 2018, Indian investigative journalist Rana Ayyub became the target of an online hate campaign due to her outspoken criticism of the Indian government, particularly her condemnation of the rape of an eight-year-old Kashmiri girl. This campaign included the dissemination of rape and death threats, as well as the circulation of manipulated pornographic videos featuring her. This phenomenon reached a disturbing peak in 2025, when thousands of South Korean women and minors were targeted by deepfake pornography, leading to a government investigation and public outcry (Yonhap News, 2025).

The psychological harm inflicted by deepfake pornography parallels that of revenge porn, despite the absence of any authentic explicit imagery. Victims report shame, humiliation, anxiety, depression, and social withdrawal (Harris, 2019). Neuroscientific research provides insight into the objectification underlying such abuse: Fiske's (2008) study demonstrated that viewing sexualised images of women activated brain regions associated with tool use rather than social cognition in male participants, suggesting dehumanisation at a neural level. This objectification creates an environment where creators view their work as harmless experimentation—as the anonymous DeepNude creator claimed (Motherboard, 2017)—while victims experience profound psychological violence and erosion of agency.

Alternative approaches investigate biological signal inconsistencies, such as abnormal eye blinking patterns or absence of physiological signals like pulse detection in facial videos (Waldrop, 2020). Feature-based forensic methods analyse compression artefacts and image metadata. Yet with the rapid rise of precise quality in AI image generation, research examining traditional machine learning methods as baselines for deepfake detection remains limited, despite their potential value for understanding fundamental distinguishability.

II.3 Legal and Policy Gaps

Despite growing awareness, legal frameworks have failed to keep pace with the rapid development of deepfake technology. No comprehensive federal legislation specifically addresses deepfake pornography, leaving victims with limited recourse (Harris, 2019). Some states have enacted legislation, but enforcement remains challenging due to jurisdictional issues and Section 230 immunity for platforms. The financial burden of content removal falls on victims, while perpetrators often operate anonymously across international borders. This legal vacuum enables continued exploitation, particularly of women and minors, underscoring the urgent need for both technical detection solutions and comprehensive policy responses.

III. METHODOLOGY AND DATA

This study employs a multimodal detection approach combining deep learning features with forensic analysis techniques to differentiate AI-generated images from authentic photographs. Our methodology integrates complementary detection strategies that examine both high-level semantic content and low-level statistical artefacts.

III.1 Dataset Preparation

We curated a balanced dataset of 200 images, split evenly into two classes: 100 authentic photographs and 100 AI-generated images. Real images were sourced from established photography repositories including Getty Images, Pinterest, and Google Images, ensuring diverse subjects, lighting conditions, and capture devices to represent authentic photographic content. AI-generated images were collected from multiple generators including Midjourney (versions 5-6), Stable Diffusion (various models), and DALL-E (2-3). This diversity in both real and synthetic sources increases the robustness of our detection approach by preventing overfitting to specific generator signatures or photographic styles.

Images were resized to 380×380 pixels, chosen to balance computational efficiency with preservation of fine-grained details necessary for forensic analysis. The dataset was split 80/20 into training (160 images) and testing (40 images) subsets using stratified sampling to maintain class balance.

III.2.1 Deep Semantic Features (2,304 dimensions)

We employ two pre-trained deep learning models to capture high-level semantic and visual representations:

- **EfficientNet-B3** (1,536 dimensions): Pre-trained on ImageNet, this convolutional architecture efficiently captures hierarchical visual features from low-level edges to high-level objects. We extract features from the penultimate layer before the classification head, providing a rich representation of image content that differs subtly between real photographs (which contain natural scene statistics) and AI generations (which may exhibit distributional artifacts from the generative process).
- **CLIP ViT-L/14** (768 dimensions): This vision-language model, trained on image-text pairs, provides multimodal embeddings that capture semantic meaning and visual concepts. CLIP features are particularly valuable because AI generators are often trained to optimise for semantic coherence that CLIP would recognise, potentially creating detectable patterns in this embedding space.

III.2.2 Frequency Domain Analysis (3 dimensions)

We perform Fast Fourier Transform (FFT) analysis on grayscale images to examine frequency spectrum characteristics. AI generators often produce images with unnatural frequency distributions—particularly "spectral holes" or dead zones in mid-to-high frequencies where real cameras produce noise and textural detail. We compute: (1) presence of low-energy spectral regions excluding DC component, (2) magnitude spectrum standard deviation, and (3) mean magnitude, capturing the overall frequency distribution profile.

III.2.3 Noise Residual Statistics (3 dimensions)

Authentic photographs contain sensor noise patterns specific to camera hardware, while AI-generated images exhibit algorithmic noise from the generative process. We extract high-frequency noise residuals using a Laplacian kernel and compute statistical measures (mean, standard deviation, median absolute value) that characterise these noise patterns. Differences in noise structure provide forensic evidence of image origin.

III.2.4 Local Binary Patterns (26 dimensions)

Local Binary Patterns (LBP) capture micro-texture information at the pixel level. We compute uniform LBP histograms with radius 3 and 24 sampling points, providing rotation-invariant texture descriptors. AI generators may produce textures that appear realistic at first glance but exhibit subtle statistical regularities detectable through LBP analysis.

3.2.5 Facial Forensics (3 dimensions)

For images containing faces (detected via MTCNN), we extract face-specific features: (1) eye region symmetry—measuring pixel-level differences between left and right eye regions, as AI often produces unnaturally symmetric facial features, (2) teeth whiteness extremity—detecting unnaturally bright teeth common in AI-generated portraits, and (3) number of detected faces. For images without faces, these features are set to zero.

III.2 Evaluation Metrics

Model performance is evaluated using multiple metrics to provide comprehensive assessment: accuracy (overall correctness), precision (positive predictive value for each class), recall (sensitivity or true positive rate), F1-score (harmonic mean of precision and recall), and ROC-AUC (area under the receiver operating characteristic curve, measuring discrimination ability across all classification thresholds). These metrics collectively assess both the model's classification accuracy and its ability to provide reliable confidence estimates for practical deployment.

IV. RESULTS

The confusion matrix for our test set (n=40) reveals the distribution of correct and incorrect classifications:

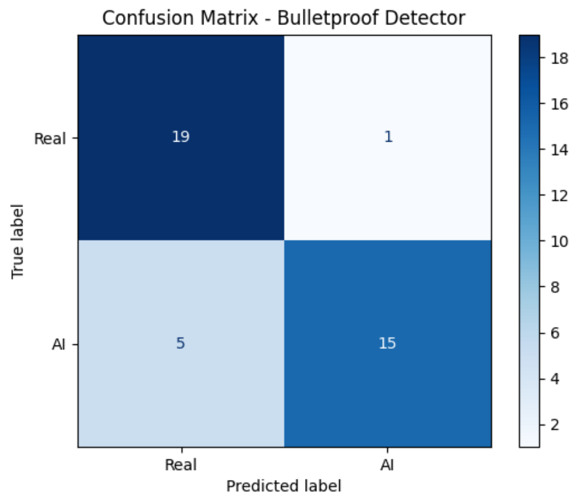


Figure 2. Confusion matrix

Source: Author's processing

True Negatives (19): The model correctly identified 19 of 20 authentic photographs, demonstrating strong capability to recognise genuine content.

False Positive (1): Only one real image was misclassified as AI-generated. This low false positive rate (5%) is critical for practical deployment, as incorrectly flagging authentic content could discredit legitimate evidence or unfairly censor real photographs.

True Positives (15): The model correctly detected 15 of 20 AI-generated images, showing that the multimodal feature approach captures distinguishing characteristics of synthetic content.

False Negatives (5): Five AI-generated images evaded detection and were classified as real. This 25% miss rate represents the primary limitation of the current approach and likely reflects cases where modern generators (particularly Midjourney v6 and Stable Diffusion XL) successfully replicate natural photographic statistics across all examined feature modalities.

The Receiver Operating Characteristic (ROC) curve plots true positive rate (sensitivity) against false positive rate (1-specificity) across all possible classification thresholds. Our model achieves an Area Under the Curve (AUC) of 0.9825, indicating excellent discrimination ability.

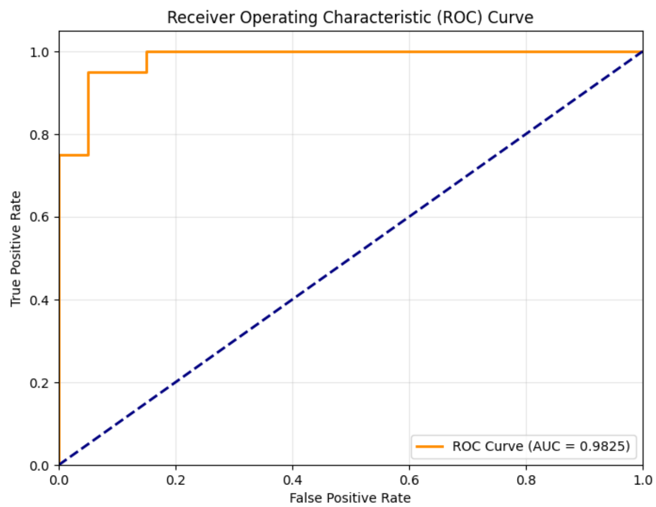


Figure 3. ROC curve

Source: Author's processing

The high AUC, despite 85% accuracy at the 0.5 threshold, suggests that most misclassifications occur near the decision boundary, where the model exhibits uncertainty. For practical applications with different risk tolerances, the threshold could be adjusted:

Stricter threshold (e.g., 0.7): Reduces false positives further, flagging only high-confidence deepfakes. Useful for automated content moderation where false accusations are particularly harmful.

Looser threshold (e.g., 0.3): Increases recall on AI images, catching more deepfakes at the cost of more false alarms. Appropriate for initial screening in high-risk contexts.

The precision-recall curve displays the trade-off between precision (positive predictive value) and recall (sensitivity) across different thresholds. This metric is particularly informative for understanding performance on the positive class (AI-generated images).

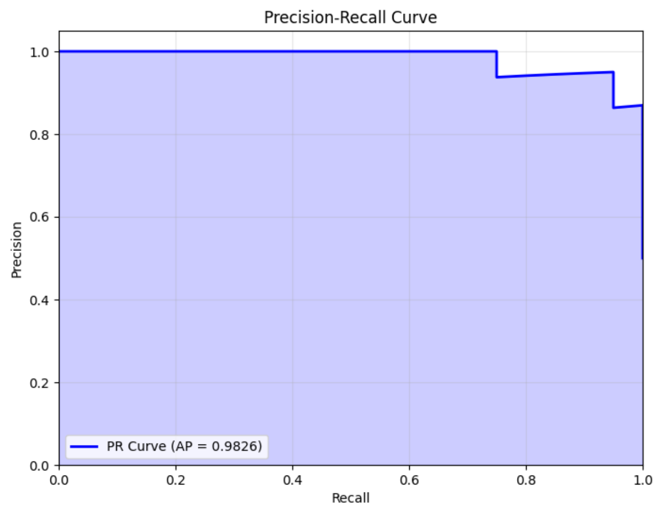


Figure 4. Precision-Recall Curve analysis

Source: Author's processing

At our default threshold (0.5), the model achieves:

- **Precision = 93.75%** on AI class: When the model predicts "AI-generated," it is correct 15 out of 16 times
- **Recall = 75%** on AI class: The model detects 15 out of 20 actual AI images

The curve's overall shape and the maintained high precision across varying recall levels confirm that the model has learned meaningful discriminative patterns rather than exploiting dataset artefacts or noise.

IV. DISCUSSIONS

Our multimodal approach demonstrates that combining deep semantic features with forensic analysis yields strong detection performance, with an AUC of 98.25% indicating excellent class separation. However, the 75% recall on AI images reveals significant limitations requiring careful consideration.

With 96% of non-consensual deepfakes targeting women, imperfect detection (25% false negative rate) remains problematic given consequences including trauma, harassment, and reputational harm. Technical solutions must integrate with broader protections: platform accountability for proactive moderation, comprehensive legislation criminalising non-consensual deepfakes with civil remedies for victims, and support systems addressing psychological impacts. Our 93.75% precision on AI content enables flagging for human review without overwhelming moderators, but detection alone cannot solve deepfake abuse.

False positives (5 real images misclassified as AI) could discredit authentic evidence in journalism or legal contexts. Applications requiring high confidence should adjust thresholds to trade recall for precision. Moreover, detection systems may inadvertently accelerate the improvement of generators as developers specifically target detection vulnerabilities.

Promising avenues include temporal analysis for video deepfakes, adversarial training paradigms, hybrid human-AI review systems, standardised benchmarks tracking detection performance across evolving generators, and expanded forensic modalities examining lighting consistency, physically impossible reflections, and semantic anomalies.

V. CONCLUSIONS

This research investigated multimodal deepfake detection combining deep learning features (EfficientNet-B3, CLIP) with forensic analysis (frequency domain, noise statistics, texture patterns, facial forensics) to address non-consensual synthetic media that disproportionately targets women. Our ensemble achieved 85% accuracy and 98.25% ROC-AUC on images from modern generators (Midjourney, Stable Diffusion, DALL-E), demonstrating that AI-generated content exhibits detectable differences across multiple feature modalities.

However, 75% recall on synthetic images reveals that one quarter of deepfakes evade detection, indicating significant challenges remain for high-stakes deployment. This technical limitation underscores that deepfake detection is fundamentally a sociotechnical challenge requiring responses spanning technology, policy, and social norms. Detection tools provide necessary but insufficient protection; comprehensive solutions demand platform accountability, legislation criminalising non-consensual deepfakes, victim support systems, and cultural shifts prioritising digital consent.

Despite some advancements in detection and growing public discourse, legal systems lag far behind. Social norms continue to diminish the seriousness of digital sexual violence, and there is no federal criminal framework that particularly targets deepfake pornography. Deepfakes will continue to pose an unchecked threat until legal and technological safeguards are improved, as well as until society recognises the agony that victims have endured. Meaningful change requires not only code and legislation but a shift in cultural values that prioritises consent, safety, and digital integrity.

REFERENCES

Ajder, H., Patrini, G., Cavalli, F., & Cullen, L. (2019). *The state of deepfakes: Landscape, threats, and impact*. Deeptrace. https://regmedia.co.uk/2019/10/08/deepfake_report.pdf

Ayyub, R. (2018, November 21). I was the victim of a deepfake porn plot intended to silence me. *HuffPost*. https://www.huffpost.com/entry/deepfake-porn_n_5bf2c126e4b0f32bd58ba316

BBC News. (2021, March 15). Mother "used deepfake to frame cheerleading rivals." https://www.bbc.com/news/technology-56404038

Cooper, A. (2004). Online sexual activity in the new millennium. *Contemporary Sexuality, 38*(3), 1–7.

Deeptrace. (2019). *The state of deepfakes: Landscape, threats, and impact*. https://deeptracelabs.com/

de Vries, K. (2020). Deepfakes, deep harms? An analysis of the gendered harms of synthetic media. *Journal of Global Ethics, 16*(2), 147–165. https://doi.org/10.1080/17449626.2020.1810165

Fiske, S. T., Cuddy, A. J., & Glick, P. (2008). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (Vol. 27, pp. 2672–2680). MIT Press.

Hao, K. (2020, October 20). A deepfake bot is being used to "undress" underage girls. *MIT Technology Review*. https://www.technologyreview.com/2020/10/20/1011116/deepfake-ai-telegram-bot-undresses-women-and-underage-girls/

Harris, D. (2019). Deepfakes: False pornography is here and the law cannot protect you. *Duke Law and Technology Review, 17*(1), 99–127.

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons, 63*(2), 135–146. https://doi.org/10.1016/j.bushor.2019.11.006

Ovadya, A. (2018, February 11). Infocalypse now: P0wnage of the public sphere, and what to do about it. *Medium*. https://medium.com/@virgilgr/infocalypse-now-p0wnage-of-the-public-sphere-and-what-to-do-about-it-b8bfc5ce9ee7

Seabrook, R. C., Ward, L. M., & Giaccardi, S. (2019). Less than human? Media use, objectification of women, and men's acceptance of sexual aggression. *Psychology of Violence, 9*(5), 536–545. https://doi.org/10.1037/vio0000198

Sensity. (2023). *The state of deepfakes 2023: Reface in the top*. https://sensity.ai/reports/

Vice. (2017, December 11). AI-assisted fake porn is here and we're all fucked. *Motherboard*. https://www.vice.com/en/article/gydydm/gal-gadot-fake-ai-porn

Waldrop, M. M. (2020, March 16). Synthetic media: The real trouble with deepfakes. *Knowable Magazine*. https://doi.org/10.1146/knowable-031620-1

Yonhap News Agency. (2024, August 28). S. Korea launches investigation into deepfake sex crimes. https://en.yna.co.kr/view/AEN20240828007251315

17

AUTOMATED STRATEGIES FOR DEFINING A JOB INTERVIEW

DOI: 10.37199/c41000317

Luca LEZZERINI

POLIS University, Tirana, Albania
luca_lezzerini@universitetipolis.edu.al
ORCID 0009-0003-7441-429X

Abstract

Personnel selection is a demanding task, especially when there are many candidates. As part of a broader industrial research project focused on experimental development, techniques for collecting requirements to inform the definition of a job interview process were examined. In practice, an AI asks a series of questions, according to a specific scheme, to people who meet particular profiles, and a set of requirements is derived from this, which is then used to interview the candidates. In the present study, we build upon the results of this research and explore the concepts of the AI engine that fulfils these requirements. The objective of this study is to critically evaluate the idea's potential and define how it can be effectively implemented.

The methodology employed began with an analysis of the experimental development documents, from which only certain elements not subject to industrial secrecy are reported, followed by a literature review. Following this preliminary analysis, possible algorithmic and technological solutions were evaluated. These solutions were then discussed across various aspects, including ethical considerations and those related to the processing of personal data, to reach conclusions about their applicability. The final results indicate a high level of confidence in the feasibility of automating this phase of the selection process, but highlight critical ethical and GDPR compliance issues.

Keywords: Machine Learning, NLP, LLM

I. INTRODUCTION

Research background. This research project builds upon a previous research and innovation project conducted in 2022 in Italy by an Italian company. In this research, the company developed a prototype system to support candidate selection in the recruitment process. The author of this paper participated in the research project and, after a two-year embargo period, is now able to publish further developments of the original research results, even though the original project details remain restricted.

The original research project was named "*CAST- Chatbot SeleTtore, prototipo sistema chatbot per la selezione del personale*" (i.e. "CAST- Selector Chatbot, prototype chatbot system for personnel selection"). It lasted throughout 2022 within the company's research and innovation activities.

Building on the results of the research project, the author of this article has further developed a framework to define the requirements for a computer-assisted job interview. This document describes this framework and its implications from the perspective of the GDPR Regulation (EU) 2016/679.

Research context. This research builds upon the results of the above-described CAST project, further developing and expanding upon them. These results were:

- The use of Europass data to extract a feature set from which to derive the training data for the job interview development
- The use of fuzzy logic to weigh and smooth the decision thresholds
- The implications of the GDPR Regulation EU 2016/679

This paper will explore these three elements.

The objectives of the research were:

1. Extract features to build training data for a job interview generator starting from the Europass, EQF and ECVET data
2. Improve the original fuzzy logic approach to make smoother and multivariate decisions
3. Explore the implications of GDPR in the previous two objectives

II. METHODOLOGY

A literature review has been conducted to examine updated documentation on the three frameworks considered in the project: Europass (the most widely known in Europe), EQF, and ECVET. A second literature review has been conducted to update the existing scientific documentation on the application of fuzzy logic in personnel selection. A final literature review was conducted on studies examining the implications of the GDPR for automated personnel selection.

After reviewing the CAST project documentation, an analysis has been conducted to enhance the original results regarding the feasibility of extracting information for use in job interview construction, thereby providing an improved framework for data gathering.

Based on the results of this first step, the original UVA (Utility Value Analysis) approach, followed by the original fuzzy algebra, has been enhanced to enable multivariate fuzzy decision-making.

Lastly, a review of the GDPR implications has been conducted, considering the entirety of Chapter 3, not just Article 22, and further suggestions have been made even for the selection phase.

III. RESULTS

III.1 Use of European frameworks for data gathering

The CAST system comprises three main components, as illustrated in Figure 1. The research project impacts the Features Data Gathering and the Job Interview Generator.

The Features Data Gathering team was responsible for extracting data from the Europass database (a repository of Europass CVs) on candidate selection features. The Job Interview Generator was the component that collected rules from recruiters' experts to select a candidate.

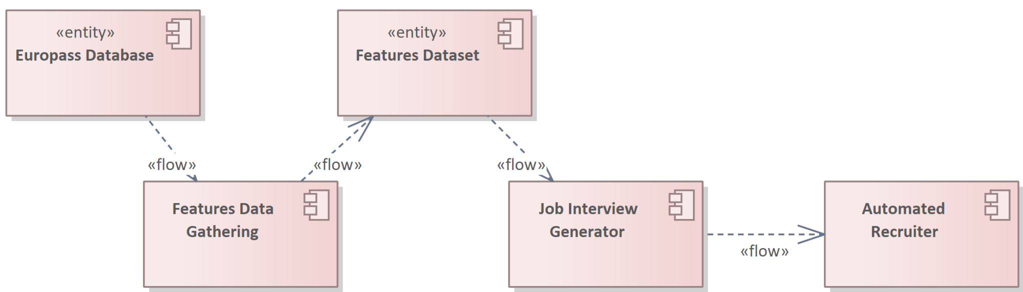


Figure 1. CAST architecture

Source: Authors' processing

The original approach involved extracting the content of the Work Experience, Education and Training, and Skills sections from Europass CVs. CAST ignored the Personal Information and the Additional Sections data. The system created a database of features with all the gathered data. Each extracted item had a primary key, enabling the creation of a pseudonymised archive.

The Job Interview Generator (JIG) retrieved this dataset to build, using an interactive system (a simple chatbot) to select which features (after removing the duplicates and sorting by frequency) were needed for the given profile.

CAST is a typical example of an economy driven by market demand. Its positive aspect is that, by employing a straightforward, fuzzy approach, it surpasses the most common criticisms of assessing people's abilities through opaque technologies in AI-powered hiring systems (Grenier & Chartier-Ewards, 2024).

The research has explored the feasibility of also using EQF and ECVET information, as well as collecting profiles from a historical HR database of similar, past-selected candidates, to make the selection process more transparent.

Even if the market demands anthropomorphic recruitment systems (Sahani et al., 2025), the research did not explore this area. Instead, it focused on candidate evaluation, which is the real challenge in automated recruitment systems.

In addition, the research did not focus on the NLP required to extract knowledge from CVs (or other data sources), as this has already been explored in the CAST project. Some literature has indeed been considered (Genugten & Schacter, 2024; Putra, 2024) for traditional models (such as BERT variants), as well as for other studies on LLMs (Wosny & Hastings, 2024).

EQF, with its eight levels of evaluation, describes a person's actual ability in a given feature (Sedano et al., 2012). In Europass CVs, this evaluation was considered (even in CAST) only for the educational skills (Education and Training). In the research, estimates have been made to calculate the EQF level and the Work Experience and Skills sections.

The reason for estimating the EQF level was to provide the interviewed expert with minimum and maximum acceptable levels for candidate selection. These calculations have been performed to assess the EQF level based on the HR historical database. Each employee in this database was manually ranked by EQF level for each skill listed in their CV. Then, a neural network was trained to estimate the level, using age, educational level, training level, and their temporal evolution. This estimation is not used in the JIG but is helpful in the Automated Recruiter (AR). Indeed, the results were promising, although the EQF level estimation was not very accurate.

By merging the features database and the HR historical database and selecting a profile, the system identified all required features (those most frequently associated with the selected profile) along with the lowest and highest EQF levels.

Research on the use of ECVET has found no further improvement because it is a more abstract framework than needed (Le Mouillour, 2012). The discussion section provides additional details.

III.2 Improve the fuzzy approach

During the dialogue with various business experts, the JIG collected the key features for the requested profile, weighted them based on both the user's importance factor and EQF, and identified the most relevant.

In a few words, CAST, for each feature f_i, *is* considered a set of similar features F_i that is a subset of the F^* universe of the features and includes f_i, too. For each similar feature in F, each expert assigns a similarity coefficient. All these elements concur to building what is called the *features' fuzzy matrix*. Table 1 provides an example below.

	Java Programming	C++ Programming	C# Programming
Java Programming	100%	30%	70%

C++ Programming	30%	100%	40%
C# Programming	70%	40%	100%

Table 1 - Features' fuzzy matrix example

Source: Author's processing

Given a candidate, for each feature for which they must have it, the Fi subset is calculated by taking the related row in the matrix, and a score is computed using a linear combination.

Given, for example, the Java Programming feature, Fi is computed by taking the Java Programming row in the table and giving a grade that is the sum of the owned features in *Fi* set: zero, no feature, 100% has Java Programming, 30% has C++, 130% has both and so on. This approach, used by CAST, is then summarised across all features, and the result is used to rank the candidate based on the sum of the scores gained for each feature.

In the research, this approach has been improved by proposing the addition of a narrative approach to the job interview. To implement this narrative approach, the CAR (Change/Action/Result) framework is beneficial (Rad & Balas, 2020). In practice, candidates are asked to tell simple stories about their past work, structured as the problem or challenge they faced, the action they took, and the result of that action. Some of these challenges can be suggested by the recruiters themselves (e.g. 'How do you adapt to job stress?').

In the cited study (Rad & Balas, 2020), a 5-level assessment is used that can easily be replaced with the 8 EQF levels, both actual and estimated. In addition, the same study proposes 11 competence areas (specific knowledge, quality and quantity of work, communication, etc.) that can also be changed, perhaps by detailing specific knowledge in individual skills or defining others. In the research, unlike in the paper, the challenge, action, and result were treated as two-valued variables with "insignificant" and "meaningful" possible values, yielding 8 possible rules (2 x 2 x 2) that generated the CAR score.

In addition to the CAR score, which requires experts to identify the fundamental challenges, a Fuzzy Analytic Hierarchy Process (FAHP) was considered due to its effectiveness reported in the literature for multi-criteria decision-making (Mahad, Yusof, and Ismail, 2019). FAHP implementation is well-defined and perfectly applicable to the case under study (Emrouznejad & Ho, 2022).

III.3 GDPR Impact

The third objective of the research has been to evaluate the impact of the GDPR on the automated job interview process, specifically in the context of the definition of the job interview.

In the CAST project, GDPR Article 22 was considered to verify the CAST's compliance. Article 22 states, at paragraph 1, "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her." The results in CAST indicated that a human decision-maker must be used to evaluate the ranking and select candidates. The research has criticised this result, as the ranking was automatically generated, and this profiling can impact the rights of the data subject (Djeffal, 2020).

In addition, the research went beyond Article 22 and examined all the articles in Chapter Three, "Rights of the data subject". Reviewing the literature yielded critical comments regarding articles 12, 16, 17, and 18.

Based on an analysis of the possibility of conducting automated job interviews in the context of the Norwegian regulatory system (Weitzenboeck, 2021), it is clear that Article 22 of the GDPR cannot simply be ignored but must be analysed in detail to understand its characteristics of exclusion of the restriction on fully automated decisions and profiling (bayamlıoğlu, 2021). The only possibility in the case of a job interview is the exclusion provided for in paragraph 2(b), which states that "is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests". The analysis of the Norwegian system leads to the conclusion that the regulatory landscape in Norway has not been designed to provide a harmonised response and is, therefore, highly fragmented, leaving the design-by-law clause structure of Article 22 completely uncovered.

About Article 22, consideration must also be given to the various interpretative issues (Davis and Schwemer, 2023) that have arisen and which relate to whether it is a right or a prohibition, the meaning of the term "decision", and the meaning to be given to the expression "significantly impacting".

In the CAST interpretation, the term "decision" was understood as a final decision, as it was the only one with a significant impact on the subject. However, in light of recent developments in thinking on the legislation, which were largely unavailable at the time of the CAST project, this statement should be understood to mean that the final decision will be taken by a human being but based on intermediate decisions that influence it, since the classification and assessment of the candidate's qualities are carried out automatically (algorithmic decisions or algorithmic regulation) (Lyons et al., 2021). Based on the assessment that the current trend in case law is to consider Article 22 as the expression of a right and not a prohibition, it is clear that specific legislative intervention is necessary, according to paragraph two of Article 22, to establish the conditions for implementing or not implementing a personnel selection and recruitment system based on artificial intelligence, as in the case of CAST.

Regarding the extension of the compliance assessment to Chapter Three of the GDPR, the reasoning followed a typical pattern in the literature (Feretzakis et al., 2025), which involves defining the key principles of the GDPR and verifying whether current implementations comply with them. The most significant case, of course, was that of LLMs. Although from a different context, the case brought in Amsterdam by several drivers against Uber and Ola (Lazcoz, 2021) was also very interesting. In that case, the drivers argued that the search methods used in Uber and Ola's digital ecosystems should be considered algorithmic regulation and, therefore, detrimental to their rights. From the perspective of the research we are considering, one of the elements examined during the trial was whether such algorithmic regulation constituted an employment relationship between the company and the drivers. The answer to this second question differed depending on the national context. In the United Kingdom, for example, it was not considered an employment relationship, while in Italy and Spain it was. In the case of Uber, the ruling found that most of the claims made by the injured parties did not constitute a violation of Article 22 of the GDPR. A similar result was reached for Ola, but on two points, the ruling was against the two companies. The first ruling concerned Ola's transparency, and the second concerned Uber's temporary deactivation.

Considering the combined effect of the need to profile or process personal data, with legal effect and based solely on automatic processing, only the three conditions relating to penalties and deductions from drivers' remuneration were met. Based on these considerations, it appears that when making automatic personnel selections, there is a real risk of violating the individual's rights, particularly concerning the transparency of the processing, minimisation of the data required, and the right to rectification or erasure. Since most artificial intelligence models are black-box in nature, this conflicts with the GDPR's transparency requirements. Furthermore, such systems require large amounts of data, which could conflict with the need to minimise the volume of data processed. Finally, once this data has been integrated into the AI model through the learning process, it becomes complicated to correct or remove it. A fundamental element of this reasoning is determining whether personal data is being processed. This challenge is the main issue to be resolved when creating training datasets. It is therefore essential to distinguish between the concepts of data and information in the field of artificial intelligence, particularly machine learning, and those in the GDPR (Hallinan and Gellert, 2020). There is an increasingly evident lack of clarity regarding what exactly constitutes information and data, particularly in the context of algorithmic regulation. Numerous perspectives coexist (Gellert, 2020), but all seem to share a common regulator: the need to transition from information technology to knowledge technology.

An attempt to regulate this aspect has been made with the European Union document called the "AI Act", especially in Article 14, which refers to the Human Oversight Principle (HOP), which can be divided into three main modes: HITL (Human In The Loop), HOTL (Human On The Loop) and HIC (Human In Command). The conclusions (Sarra, 2024) are that a HOP approach does not, on its own, ensure compliance with Article 22 of the GDPR, and that there remains considerable confusion regarding its interpretation.

IV. DISCUSSIONS

Use of European frameworks for data gathering. The use of European frameworks for skills assessment was already envisaged in the original CAST project, as the information obtained from Europass CVs included EQF. Its extension to other types of skills and the development of a system for estimating this assessment demonstrated its practical applicability.

The ECVET assessment proved to be irrelevant, as it is a general framework that adds nothing to what has already been achieved for EQF. It remains true that information generated by the application of ECVET (or similar frameworks) can be beneficial for automated staff selection processes.

Finally, by using a consistent evaluation system as a benchmark, it is possible to compare the estimates produced by the system with those made by a human, thereby improving the validation of the artificial intelligence model.

Improve the fuzzy approach. The use of a fuzzy hierarchical process (FAHP) enabled the application of a CAR approach that is simple for candidates to request and relatively easy to process from an NLP perspective, achieving significantly better results than CAST while maintaining transparency in the selection process, as fuzzy rules are explicit.

GDPR Impact. The results regarding GDPR compliance were interesting. Firstly, there are serious compliance issues not only under Article 22 but also under other articles, particularly regarding transparency and minimising the amount of data requested. Another critical issue was the identification of personal data.

Finally, it became clear, and this is perhaps the most interesting aspect, that regulatory fragmentation stems not only from the novelty of the issue but also, above all, from the need to move from an information-technology-based approach to a knowledge-technology-based approach.

IV. CONCLUSIONS

Scholars need to continue their research on compliance with the GDPR, in particular, and on algorithmic regulation, in general, to enable legislators to enact consistent and effective rules governing the use of artificial intelligence in automated decision-making.

In practice, in-depth research is needed on digital ethics and the lawful use of artificial intelligence. Under the GDPR, the most critical aspect is the lack of transparency in decision-making, as most artificial intelligence models are structured as black boxes.

The use of fuzzy logic partially addresses the problem of scoring each skill. Unfortunately, opacity remains for NLP mechanisms for extracting knowledge from interviews.

Finally, the use of European assessment frameworks enhances the validation of AI models used in both interviews and for estimating threshold values, by enabling the consistent application of a single assessment model. It is believed that research should further explore this aspect, aiming to harmonise these frameworks with their application to AI-based systems.

The CAST project can be considered a good trial for implementing an e-recruitment system that meets market requirements and is GDPR-compliant. Nevertheless, legal experts and policymakers need further research to ensure the safe and effective use of AI in this field. The "AI Act" is a good step in this direction, but it lacks harmonisation with the GDPR and real effectiveness for future regulations and policymakers.

REFERENCES

Bayamlıoğlu, E. (2021). The right to contest automated decisions under the General Data Protection Regulation: Beyond the so-called "right to explanation." *Regulation & Governance, 16*.

Davis, P., & Schwemer, S. F. (2023). Rethinking decisions under Article 22 of the GDPR: Implications for semi-automated legal decision-making. In *Proceedings of the Third International Workshop on Artificial Intelligence and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2023)*, ICAIL 2023, Braga, Portugal. https://ssrn.com/abstract=4478107

Djeffal, C. (2020). The normative potential of the European rule on automated decisions: A new reading for Art. 22 GDPR. *Zeitschrift für ausländisches öffentliches Recht und Völkerrecht (ZaöRV), 80*, 847–879.

Emrouznejad, A., Ho, W., & St C. (2022). *Fuzzy analytic hierarchy process*. CRC Press.

Feretzakis, G., Vagena, E., Kalodanis, K., Peristera, P., Kalles, D., & Anastasiou, A. (2025). GDPR and large language models: Technical and legal obstacles. *Future Internet, 17*, 151.

Gellert, R. (2018). Comparing the definitions of data and information in data protection law and machine learning: A useful way forward for algorithmic regulation? *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3295078

Genugten, R., & Schacter, D. (2024). Automated scoring of the autobiographical interview with natural language processing. *Behavior Research Methods, 56*.

Grenier, E., & Chartier-Ewards, N. (2024). Automated hiring platforms as mythmaking machines and their symbolic economy. *Journal of Digital Social Research, 6*, 34–48.

Hallinan, D., & Gellert, R. (2020). The concept of "information": An invisible problem in the GDPR. *SCRIPTed, 17*, 269–319.

Lazcoz, G. (2021). Automated decision-making under Amsterdam's District Court judgements: Drivers v. Uber and Ola. In *time to reshape the digital society: 40th anniversary of the CRIDS*. Larcier (Lefebvre Sarrut Group).

Le Mouillour, I. (2012). *Necessary conditions for ECVET implementation*. Publications Office of the European Union.

Lyons, H., Velloso, E., & Miller, T. (2021). Conceptualising contestability: Perspectives on contesting algorithmic decisions. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW1), Article 106. https://doi.org/10.1145/3449180

Mahad, N. F., Yusof, N. M., & Ismail, N. F. (2019). The application of fuzzy analytic hierarchy process (FAHP) approach to solve multi-criteria decision making (MCDM) problems. *Journal of Physics: Conference Series, 1358*(1), 012081. https://doi.org/10.1088/1742-6596/1358/1/012081

Putra, B., Azizah, K., Mawalim, C., Hanif, I., Sakti, S., Leong, C., & Okada, S. (2024). MAG-BERT-ARL for fair automated video interview assessment. *IEEE Access*.

Rad, D., & Balas, V. (2020). A novel fuzzy scoring approach of behavioural interviews in personnel selection. *BRAIN: Broad Research in Artificial Intelligence and Neuroscience, 11*(2), 178–188. https://doi.org/10.18662/brain/11.2/81

Rad, D., Dey, N., & Balas, V. (2016). Personnel selection algorithm using Fuzzy Analytic Hierarchy Process: A dynamic decision-making approach. https://doi.org/10.13140/RG.2.2.29761.56163

Sahani, K., Khan, M., Khatwani, S., Gupta, S., & Dubey, A. (2025). A smart interview simulator using AI avatars and real-time feedback mechanisms (AI avatar for interview preparation). *International Journal of Engineering Technologies and Management Research, 12.*

Sarra, C. (2024). Artificial intelligence in decision-making: A test of consistency between the "EU AI Act" and the "General Data Protection Regulation." *Athens Journal of Law, 11*, 1–17.

Sedano, A., Aguilera, J., & Paris, A. (2012). The EQF model and teachers' competences. *Procedia – Social and Behavioral Sciences, 46*, 3487–3491.

Weitzenböck, E. (2021). Simplification of administrative procedures through fully automated decision-making: The case of Norway. *Administrative Sciences, 11*, 149. https://doi.org/10.3390/admsci11040149

Wosny, M., & Hastings, J. (2024). Applying large language models to interpret qualitative interviews in healthcare. *Studies in Health Technology and Informatics, 316*, 791–795.

Weitzenböck, Emily. (2021). Simplification of Administrative Procedures through Fully Automated Decision-Making: The Case of Norway. Administrative Sciences. 11. 149. 10.3390/admsci11040149.

Wosny, M., & Hastings, J. (2024). Applying Large Language Models to Interpret Qualitative Interviews in Healthcare. Studies in health technology and informatics, 316, 791-795.

18

FROM CITIZEN VOICES TO BUSINESS VALUE: ARTIFICIAL INTELLIGENCE IN PARTICIPATORY ECOSYSTEMS

DOI: 10.37199/c41000318

Dhurata SHEHU

POLIS University, Tirana, Albania
dhurata_shehu@universitetipolis.edu.al
ORCID 0000-0002-6145-5469

Melisa TUFA

POLIS University, Tirana, Albania
melisa_tufa@universitetipolis.edu.al
https://orcid.org/0009-0006-6376-9775

Abstract

AI has brought about radical changes in every aspect of citizens' daily lives, especially through the emergence of online platforms that enable continuous interaction between citizens and businesses. From product reviews, ratings, and social media engagement to participation in feedback systems and digital surveys, citizens are generating vast volumes of valuable data. These platforms serve as the primary medium where users express their preferences, concerns, and experiences. Businesses are increasingly leveraging these citizen-centred platforms not only to improve customer service but also to shape their strategic decisions. AI technologies, particularly those based on Machine Learning and Deep Learning, allow for real-time analysis of both explicit and implicit citizen needs. By analysing vast and diverse forms of unstructured data—including textual feedback, user-generated content, and digital interaction patterns—AI systems uncover hidden patterns, detect sentiment shifts, and extract actionable insights that help organisations anticipate societal needs and adapt their offerings in a timely and targeted manner. In this context, AI functions as a bridge between community expectations and business strategies, directly contributing to the development of more adaptive and user-centred products and services. The paper explores the role of AI in transforming business operations by highlighting its benefits for public transparency. Through a conceptual analysis and illustrative examples, the study proposes a model that links the knowledge gained from citizen input to the creation of value for business through an inclusive medium, which, in this study, is called a participatory ecosystem. The paper further reflects on the ethical implications of extracting commercial value from public opinion, addressing concerns over data ownership, algorithmic bias, and the risk of eroding trust if civic data *is used without accountability. The findings emphasise the need for frameworks that ensure responsible AI practices while maximising mutual benefit for citizens and businesses alike.*

Keywords: Artificial Intelligence (AI), Citizen-Centred Platforms, Unstructured Data, Business Strategy, Participatory Ecosystem

I. INTRODUCTION

Artificial Intelligence (AI) is a sophisticated technology that has impacted all areas of human daily life. It can be defined as a technology that enables computer systems and machines to imitate human intelligence through simulations and training with massive data (Stryker, Kavakoglu, n.d.). This innovation has made it possible to automate many processes, led to data-driven decision-making, significantly reduced the cost of various operations, and increased accuracy across various fields. AI-integrated systems can conduct comprehensive analyses, produce intricate reports for specialists, forecast future scenarios based on input data, and operate autonomously through ongoing learning from the data they receive. In short, the potential of this technology is revolutionising many industries, such as education, healthcare, and finance. Together with robots, which are nothing more than a set of AI mechanisms, it has also changed every basic aspect of how businesses operate (Choi & Ozkan, 2019). The advanced mechanisms offered by AI have created opportunities for businesses to significantly improve their decision-making processes based on the knowledge generated by AI systems and have pushed them to quickly adopt more beneficial management and marketing techniques (Prasanth et al., 2023). In the business sector, the utilisation of AI encompasses enhancing customer relations via virtual assistants or specialised services; analysing and producing comprehensive reports expeditiously; aiding in decision-making processes; generating predictive models; and facilitating the formulation of new strategies to optimize operational efficiency and elevate performance in the corporate arena(Velu & B, 2020). For example, to facilitate the customer experience in purchasing various services or products from a business, AI uses Machine Learning (ML) and Natural Language Processing (NLP) algorithms to analyze the input received from each user. ML is a form of AI that uses hidden patterns in data and can learn from that data without explicit programming (McKinsey, 2024). Whereas NLP, a type of AI, enables computer systems to comprehend, evaluate, and produce human language. Moreover, it is a part of ML that focuses on text analysis to identify parts of speech, entities, hidden sentiment, etc. Despite appearing simpler than an image, video, etc., this type of data requires specific techniques for machine learning analysis (Lexalytics, 2020).

In business, text analytics are very useful, especially for customer reviews, public comments on social media, survey responses, financial or regulatory documents, etc. In summary, the primary objective of ML and NLP in handling this data is to transform it from an unstructured format into valuable insights using the statistical techniques inherent in ML. Today, the nature of the customer-business relationship has evolved to a new stage that relies on technology and interactive online platforms (Ruggieri et al., 2018). One of the most used online platforms by businesses, whether are large or small businesses, is social media, where people have total freedom to express their opinions and are more flexible since they are also widely used for other purposes, such as entertainment, information, etc. (Shehu et al., 2025; Soelaiman & Ekawati, 2022). Other platforms are applications dedicated to businesses, online forums, etc. Each of these mediums of communication between the customer and the business offers many benefits, which we will explore in more detail in the following sections. AI makes these communication channels suitable for use by providing personalized content, conducting sentiment analysis on collected data, automating customer service through chatbots, and implementing recommendation systems. As a result of these tools, AI-enabled digital platforms are transforming traditional business methods into a more inclusive and efficient model. This study aims to investigate how advanced digital tools powered by AI have impacted customer relationships. Additionally, it seeks to determine whether these platforms have affected citizens' faith in business or in company transparency. Despite the sensitivity of the topic, certain studies emphasize the value of safeguarding data privacy and cybersecurity, although they do not explicitly address data accessible to businesses, which may encompass even more vital information, including matters pertaining to online payments. Following an extensive review of the relevant literature, it is proposed to adopt a conceptual model termed a "participatory ecosystem" that emphasizes an inclusive environment.

II. LITERATURE REVIEW

II.1 AI, Social Media & Digital Business Models

Due to the necessary changes brought about by digital transformation, especially during the COVID-19 pandemic, traditional business methods have been transformed into a completely new model. The idea of online platforms for businesses began several years ago, although this concept is now being put into practice. For this reason, it is worth mentioning the study conducted by Parker in 2016 (Parker et al., 2016), in which he wrote about the concept of the “Platform Revolution”. This book examined the impact of digital platforms on the global economy. As very successful platforms of those years, the authors mention Uber, Airbnb, Facebook, and Alibaba. The book primarily pointed out the importance of platform architecture, its functional design, and its governance. The term governance, in this context, refers to the regulations and methods that govern online business operations and facilitate an appropriate balance between the enterprise and its clients. Furthermore, the study is detailed down to the strategies followed in financial transactions, subscriptions, marketing, etc. In conclusion, the authors emphasise the necessity of recognising the risks associated with using these platforms, including data exploitation and antitrust challenges. This ensures that the strategy for using these platforms is well thought out and technically maintained to avoid these problems.

Social networks today are widely used by all age groups and are the most widespread platforms due to their ease of use, entertaining and informative content, and minimum technical requirements. Soelaiman & Ekawati (2022) studied the factors that encourage businesses to be on social media and the expectations of their managers. The study was made possible by a questionnaire that asked managers about the expected benefits and perceived impact of using them, which was distributed to managers across various businesses. Moreover, the author analysed the acquired data to assert the risks associated with operating on social media. The research indicated that most companies used various social networks, with Instagram the most popular at 88.83%. In conclusion, the results of the data analysis once again confirmed the outstanding advantages that social media offers for promoting products or services, improving customer relationships, reducing marketing costs, and expanding market share. Bharadyia (2021) studied the applications of AI in business, with a focus on customer management. For this, the main applications of AI in customer relations have been discussed in detail, such as:

- improving customer service (through ML algorithms and NLP) (Coombs et al., 2020);
- providing services tailored to customers based on the analysis of their behavior on the online platform or the preferences shown about services or products, which, according to the author, is known as 'audience segmentation';
- figuring out the clients' satisfaction by using sentiment analysis with certain ML and DL classification algorithms, along with NLP methods for handling text;
- identification of fraud through real-time monitoring and reporting of transactions carried out;
- increasing efficiency in supply chain activities;

The figure below shows the cyclical relationship among data, users, and the products/services the business offers.

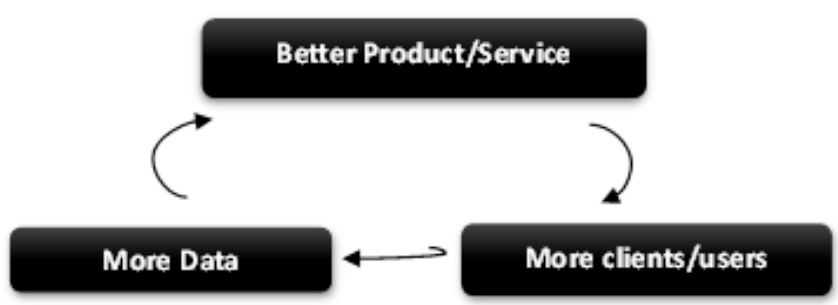


Figure 1. The virtuous cycle of AI

Source: Authors processing

Ruggieri at al. (2018) studied the impact of digital platforms on business models, specifically in businesses that base their activity on online platforms. These businesses are otherwise known as SMEs (Small and Medium-sized Enterprises), which refer to businesses that do not necessarily have a physical location, such as a store, but operate online. The study involved several Italian firms operating in different fields and several startups. Through well-structured interviews and mixed-method analyses of qualitative and quantitative data, the study compared SMEs' various strategies. These comparative analyses were also made possible by data collected through direct observation of secondary sources, such as websites, magazines, reports, and multiple articles, and by the application of the Canvas business model. This study highlighted the great importance of digital platforms in fostering innovative business models and in improving the supply-demand balance, thereby steadily driving a technology-based economy. The data analysis also confirmed the positive impacts of customer relationships and transparency, as well as of business scalability and adaptability to market and environmental changes.

II.2 AI in Business Decision-Making

In their study, Prasanth et al. (2023) show how AI has impacted business decision-making and how business models have changed as a result of this major technological shift. AI-based systems can analyse large volumes of data in a very short time thanks to sophisticated ML and DL mechanisms. This is also emphasised in this study, which points out that decision-making is a more efficient and accurate process thanks to the digital systems used. Furthermore, the author argues that the integration of AI in business, especially in its most important process, which is decision-making, significantly impacts managerial success and improves future practices.

Ogunbukola (2025) highlighted several important subtopics related to AI in business strategies. The first subtopic is AI and decision-making. In this context, he highlighted that AI can help identify business trends and generate predictive analytics and valuable insights for essential business operations. Data-driven decision-making is most important in businesses dealing with healthcare, economics, manufacturing, etc., where complex data need to be analysed to optimise resources and detect and predict faults in specific equipment. Another area explored by the author is the role of AI in process automation and productivity enhancement. Automation of various functionalities not only speeds up operations and increases productivity. For example, here we can mention the automation of repetitive emails, meeting scheduling, virtual assistants, etc. Furthermore, the study emphasises that the use of technology, specifically AI in customer relations, enables real-time support and creates a positive customer experience. AI, through ML mechanisms, enables analysis of customer behaviour and preferences, offering personalised services and immediate interactions with virtual assistants or chatbots. This increases customer trust in the business's products or services.

Gu (2024) analysed the role of AI in business decision-making by considering several case studies of concrete businesses with different characteristics. The study also compared AI implementation techniques across large and small businesses, focusing primarily on online work within SMEs. Gu examined in detail how certain technologies have been used in specific business contexts, enabling an accurate analysis of the significant impact that AI has on business strategies and success in the present and future. The companies for interviews were selected based on several criteria, such as:

- They should have AI at the core of the services or products they offer.
- They should have a specialised field in one of the types/applications of AI.

The author asserts that this selection was crucial in assessing the collected data to determine if organizations employing AI in their service or product operations utilized advanced AI processes, such as machine learning, natural language processing, or deep learning, for decision-making. Following an analysis of the chosen companies—those seeking to engage with AI, including those that marketed AI products—and a series of semi-structured online interviews, it was determined that 71% of these companies use AI for decision-making in their operations. In the extensive and detailed examination of firms that integrate AI at their core across several contexts, Gu presents several recommendations, including:

- To enable AI with its functionalities and benefits to be equally useful for small and large businesses (SMEs), a strategy or framework must be designed specifically for SMEs, which are businesses that face many challenges and limitations due to their way of operating in the market.
- As AI-based online platforms are being widely used in businesses, a general system that offers key functionalities in an easier and more advanced way should be developed to optimise operations and the decision-making process.
- In the automation of work processes, specifically through robotic process automation (RPA), it is essential to prioritise time-consuming, critical, and frequently occurring processes within the organisation to optimise technological utilisation and enhance overall workflow efficiency.

III. METHODOLOGY AND DATA

The methodology adopted in this study begins with a contextual definition of Artificial Intelligence (AI) and its current applications in business environments. The research is grounded in an extensive review of professional literature, including peer-reviewed articles, academic studies, and documented case analyses. Sources were selected based on their relevance to AI in strategic business processes and their citation frequency in recognised databases. The literature was categorised into two main groups:

1. Studies that focus on the role of digital platforms and the drivers of social media adoption in business.
2. Research that highlights the benefits of AI integration in business models, particularly through machine learning (ML) applications.

A conceptual analysis of these sources informed the development of a theoretical framework that illustrates the interrelationships among AI technologies, customer engagement, and business value creation.

IV. RESEARCH RESULTS

Studies have shown the benefits of using AI in business models. However, many researchers categorise the limitations and challenges that accompany this innovation as follows (Gu, 2024; Soelaiman & Ekawati, 2022):

- Challenges related to the technological aspect include problems that may be encountered with data. The quality of the collected data directly influences the training phase of machine learning models and, consequently, the outcomes. Moreover, resolving the issue of data security requires appropriate expertise, and in several regions globally, there is a deficiency of cybersecurity professionals (Shehu & Lezzerini, 2024). IT infrastructure may be another problem, particularly in developing countries, where a digital divide between urban and rural areas still exists. One challenge that could lead to the misuse or non-use of AI systems in business strategies is algorithmic transparency, which is linked to the need for a specialised technological sector.
- The organisational challenges primarily involve managing and adapting to change. In today's rapidly evolving world, driven by technology and AI, all organisations must adopt new tactics and strategies to thrive in the market. This necessitates training employees who can endure significant transformations. On the other hand, the technical costs of proper infrastructure and its maintenance are relatively high. The most important factor in eliminating these challenges is a clear vision for AI adoption, given the numerous benefits it offers businesses across all fields.
- Ethical and regulatory challenges regarding data concern the bias that may exist in data that may not equally represent the given context. This requires a thorough training phase on the data, using appropriate techniques to address any imbalance in the collected data. Another challenge in this regard is the lack of a clear regulatory framework for the use of customer data. The General Data Protection Regulation (GDPR) is an important step in this direction. However, challenges in its interpretation and implementation are particularly acute for companies that use highly advanced AI models.

The graph below illustrates the variation in AI utilisation across various workplaces in the US over the course of a year. The sectors related to information, technical services, finance, manufacturing, education, public administration, and agriculture have been analysed (Maslej, 2025). The graph clearly indicates that the information and data sector has expanded by almost 80% in 2024. The heightened interest in AI's functions across various domains has also led to a marked increase in the ratio within the professional services, banking, and manufacturing industries. Moreover, in areas deemed more conventional, such as agriculture, transportation, and waste management, there has been a notable increase. The only area that has seen a decrease is public administration, probably because there are not enough skills to use AI effectively, along with rules that limit its use and other reasons. In conclusion, the graph unequivocally illustrates the broad proliferation of AI utilisation across all industries.

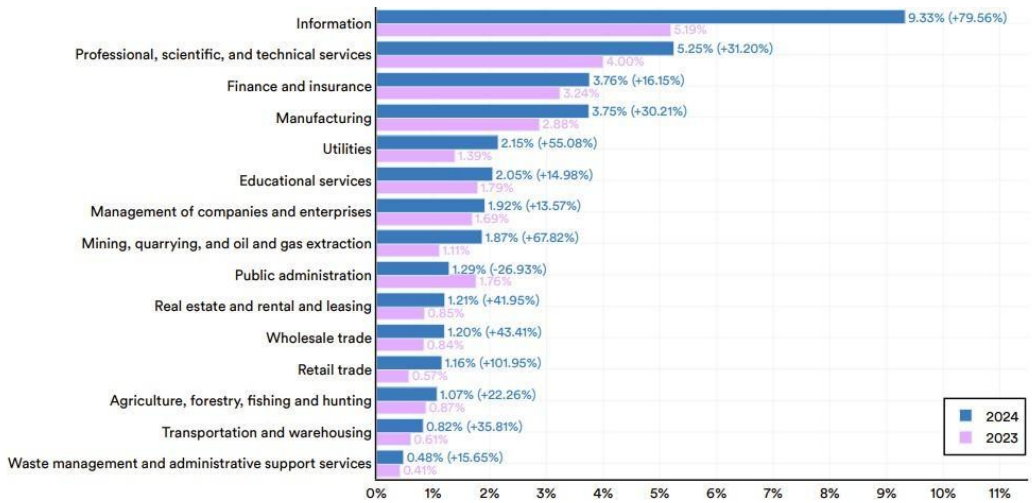


Figure 2. AI job shares by sector in the US, 2023 vs. 2024

Source: Authors processing

Therefore, it is critical to increase public awareness of the benefits AI offers to society, the economy, and other aspects. The goal is not only to train them to adapt easily to rapid changes but also to increase their trust in AI services. Global reports show that the proportion of customers who consider AI-based services or products beneficial and of added value has increased to 55% from 52% in 2022, reflecting a gradual increase in the acceptance and adoption of AI in the consumer experience (Maslej, 2025). However, trust in the transparency of businesses and companies regarding the use of their data has fallen to 47%, down from 50% in 2023. This decrease reflects a lack of transparency in how businesses use, store, and manage customers' personal information, resulting in uncertainty and scepticism among users. This situation emphasises the necessity of taking citizens' voices into consideration in the decision-making process of businesses, in the process of establishing policies or rules on how their data will be used, etc. (Gonesh et al., 2023).

Only through such a process, where the citizen's voice is truly considered, can long-term trust be strengthened and a healthier relationship between consumers and the advanced technologies that serve them be ensured (Shehu &Luarasi, 2024).

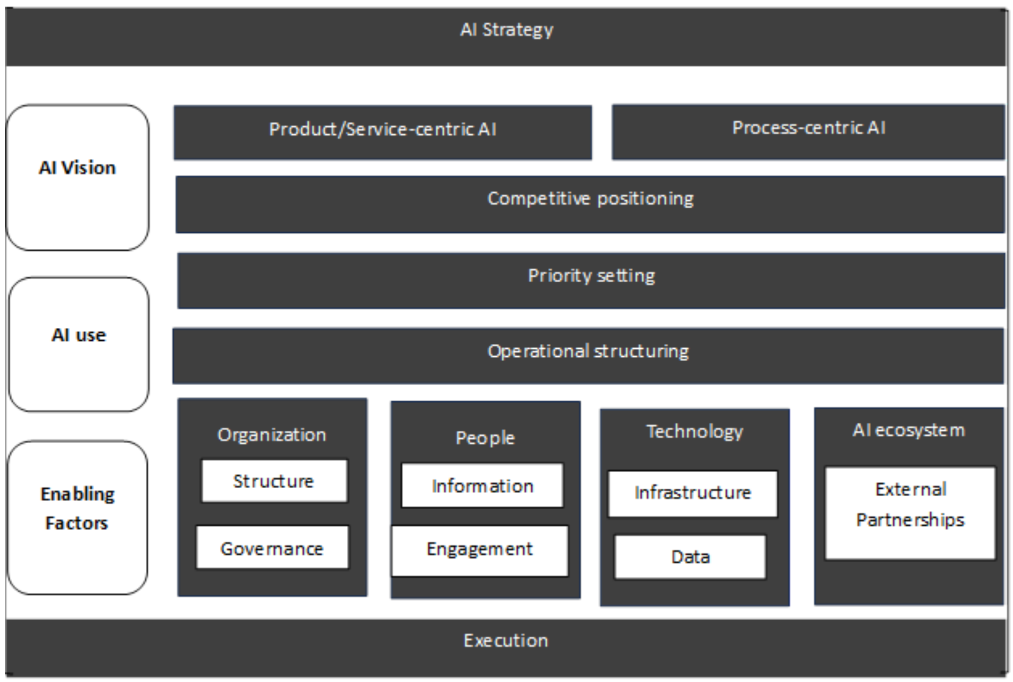


Figure 3. Conceptual model of AI in business models

Source: Authors processing

The following figure illustrates a conceptual framework for incorporating AI into a business. In a structured form, we see that in the first layer lies the strategy which should be in the same line as the business strategy. This layer is followed by a vision that presents the focus, purpose, and objectives for using AI to optimise business operational processes and to create new products/services creatively and competitively in the market. The penultimate layer encompasses the key elements that enable the proper integration of AI, such as leadership structures, citizens through cooperation and expertise, technology and third-party collaborations. At the end of the model lies the execution of these defined strategies. The scheme summarises the key concepts to keep in mind when designing a business mechanism—essentially, the voices of citizens and customers.

V. CONCLUSIONS

The recent advances in artificial intelligence have presented numerous obstacles for businesses, alongside various benefits. As the world progresses rapidly, there is an increasing necessity for adaptation and training in this domain. The capacity of AI, utilising ML, NLP, and DL techniques, to analyse vast quantities of data rapidly and produce highly accurate models and predictive analytics renders it an essential asset for organisations to maintain strength and competitiveness in the marketplace. Decision-making processes for them have been significantly facilitated, and citizens are increasingly becoming part of this ecosystem every day. Digital platforms and continuous public reports for customers and citizens in general demonstrate transparency, which directly impacts the trust they place in the services/products businesses offer. One of the key factors affecting the accuracy of analyses is the quality of the collected data; therefore, organisations must use regulatory models that ensure accurate decisions by minimising data-related losses and risks, and preventing biased decisions that arise when the data fails to accurately represent reality. The study thus recommends establishing a specific business model that aligns with business objectives while prioritising customers in the value-creation process. The success of this paradigm ultimately depends on the ethical application of AI, with genuine involvement from customers and citizens, who may both contribute to and directly profit from this transformation. Definitive objectives, robust data governance, transparency, adaptability through innovation, and ongoing monitoring and evaluation of AI systems enhance a business's safety, competitiveness, and efficiency in the marketplace.

REFERENCES

Bharadyia. (2021). *Rise of artificial intelligence in business and industry*. ResearchGate. https://www.researchgate.net/publication/371307024_Rise_of_Artificial_Intelligence_in_Business_and_Industry

Choi, J., & Ozkan, B. (2019). Innovation and disruption: Industry practices and conceptual bases (pp. 3–13). https://doi.org/10.1108/S1569-376720190000020003

Coombs, C., Hislop, D., Taneva, S. K., & Barnard, S. (2020). The strategic impacts of intelligent automation for knowledge and service work: An interdisciplinary review. *The Journal of Strategic Information Systems, 29*(4), 101600. https://doi.org/10.1016/j.jsis.2020.101600

Gonesh, C., Saha, G., Menon, R., Paulin, S., Yerasuri, S. S., Saha, H., & Lalit, S. (2023). *The impact of artificial intelligence on business strategy and decision-making processes*. https://doi.org/10.52783/eel.v13i3.386

Gu, S. (2024). *Exploring the role of AI in business decision-making and process automation*.

Lexalytics. (2020, September 29). *Machine learning (ML) for natural language processing (NLP)*. https://www.lexalytics.com/blog/machine-learning-natural-language-processing/

Maslej, N. (2025). *Artificial Intelligence Index Report 2025*. Artificial Intelligence.

McKinsey. (2024). *What is machine learning (ML)?* https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-machine-learning

Ogunbukola, M. (2025). *Artificial intelligence implications for business strategy*.

Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. W. W. Norton.

Prasanth, A., Vadakkan, D. J., Surendran, P., & Thomas, B. (2023). Role of artificial intelligence and business decision making. *International Journal of Advanced Computer Science and Applications, 14*(6). https://doi.org/10.14569/IJACSA.2023.01406103

Ruggieri, R., Savastano, M., Scalingi, A., Bala, D., & D'Ascenzo, F. (2018). The impact of digital platforms on business models: An empirical investigation on innovative start-ups. *Management & Marketing, 13*, 1210–1225. https://doi.org/10.2478/mmcks-2018-0032

Shehu, D., Luarasi, T., & Kyratsis, P. (2025). Designing a natural language processing-driven communication system for urban planning: A case study. *Journal of Chinese Architecture and* Urbanism, 0(0), Article 0. https://doi.org/10.36922/jcau.8417

Shehu, D., & Lezzerini, L. (2024). *The challenges of applying big data in the urban planning practices for the developing countries: Case study in Albania.* https://press.universitetipolis.edu.al/book/the-challenges-of-applying-big-data-in-the-urban-planning-practices-for-the-developing-countries-case-study-in-albania-dhurata-shehudr-lucca-lezzerini/

Shehu, D., & Luarasi, T. (2024, July 23). *A review on citizen motivation analyses for participating on urban planning process.* https://press.universitetipolis.edu.al/book/a-review-on-citizen-motivation-analyses-for-participating-on-urban-planning-process/

Soelaiman, L., & Ekawati, S. (2022). The role of social media in enhancing business performance. https://doi.org/10.2991/aebmr.k.220501.060

Stryker, & Kavakoglu. (n.d.). *What is artificial intelligence (AI)?* Retrieved June 4, 2025, from https://www.ibm.com/think/topics/artificial-intelligence

Velu, P., & B, V. (2020). Role of artificial intelligence in business transformation. *International Journal of Advanced Science and Technology, 29*, 392–400.

19

AI AND IMAGE PROCESSING. SOME KEY MOMENTS IN THE IMPLEMENTATION OF THESE METHODS

DOI: 10.37199/c41000319

Melisa TUFA

POLIS University, Tirana, Albania
melisa_tufa@universitetipolis.edu.al
ORCID 0009-0006-6376-9775

Tamara LUARASI

POLIS University, Tirana, Albania
tamara_luarasi@universitetipolis.edu.al
ORCID 0009-0002-7449-5491

Abstract

The integration of artificial intelligence (AI) into architecture and urban planning represents a significant shift in how professionals approach design, analysis, and decision-making. This article aims to underscore the relevance and applicability of various AI methodologies in the daily practices of architects and urban planners.

This article will serve as a starting point for a more detailed study of transforming existing architecture and urbanism curricula to include basic knowledge of modern technologies involving AI methods. This paper focuses on two common issues in implementing the methods mentioned above. The creation of custom datasets to solve a specific problem, and the CNN model that serves as the basis for many AI methods in image processing.

Through a complete example of the implementation of one of the AI methodologies, which is image classification, a working practice is illustrated.

Keywords: AI, Image Processing, Architecture, Urban Planning

I. INTRODUCTION

The purpose of this article is to sensitise architects and urban planners to the importance of using AI's various methodologies in their work, and to provide essential practices for these methodologies. All this supports the idea of revising curricula for architects and urbanists, which would constitute a second phase of the study.

Only a limited number of articles, across a wide range, address the role AI has taken on in the work of architects or urbanists. The authors in Stojanovski et al. 2021 present a combination of urban visions of the future, digital tools, and AI techniques to inspire architects, engineers, programmers, and computer scientists to explore AI and improve architectural and urban practices.

In Jin et al (2024), it is mentioned that higher education has undergone major transformations with the development of artificial intelligence. According to the authors, this transformation consists not only in the use of new technologies but also in the way work is conceptualised. Undoubtedly, the introduction of these new methods into the daily work of architects and urbanists requires knowledge and skills to understand and implement them.

By possessing AI tools such as programming and data management, students would be able to handle datasets and use them in software for drawing that includes AI elements. Such software includes Rhino, Grasshopper, and Revit, which architects widely use.
There are two ways the architect can successfully use AI. Architectonic programming is a preliminary step in the design process, where architects and design professionals collect, analyse, and determine clients' desires and requirements to inform the design. Technologies such as Information and Communication Technology (ICT), Building Information Modelling (BIM), Virtual Reality (VR), and Augmented Reality (AR) are gradually changing the way architectural programming is conducted. AI in design further expands these experiences by creating highly efficient 3-dimensional interactive environments for communicating with customers (Stas, 2025)

The second direction is architectural design itself. AI algorithms, such as Generative Adversarial Networks (GANs) and other generative AI algorithms, are further expanding the boundaries of traditional architectural drawing.

A Generative Adversarial Network, or GAN, is a machine learning approach used for generative modelling designed by Ian Goodfellow and his colleagues in 2014 (Idiot Developer, 2020).

Another example is Foster + Partners (2021), where it is stated that if we train a machine with a set of hundreds of public spaces and extract those that are successful (it is easier to identify an existing successful space than to define the conditions for a space to be successful), the system in the machine can be used to generate other spaces with similar characteristics.

However, Nasir (2024) argues for a balance between traditional architectural methods and AI applications. Here, it is emphasised that, despite the use of AI to solve complex problems, hand drawing and even hand calculations remain very important for architects and cannot be replaced by AI.

Some image processing methodologies. Design is the most important element in the work of an architect or an interior design professional. An architectural design is an image for AI. In Mohit (2005), a coloured image is defined as a data holder that represents a combination of RGB values. The colour image is stored in a 3D array on the computer. The first two dimensions correspond to the image's dimensions (in pixels), and the third to the colour intensity (red, green, and blue) in each pixel.

In a program (in any programming language), such a structure can be represented as a 2-dimensional matrix of pixels, where the rows and columns represent the image's dimensions, and the value of each pixel is a composition of the colour intensities in that pixel. The image is the object of processing of many tasks that AI can do, and which can be categorised as follows:

- Image Classification;
- Object detection and localization;
- Image segmentation;
- Image retrieval;
- Image denoising/restoration;
- Image super-resolution.

In the last decade, a big progress has been made in:

- Image generation.

The field to which these methodologies belong is described in Restack.io. as: "Computer vision is a field of artificial intelligence (AI) that enables computers to interpret and extract meaningful information from digital images and videos". It involves using machine learning and neural networks to analyse visual data, allowing systems to make decisions or recommendations based on what they "see". This interdisciplinary field combines aspects of AI and image processing to help machines gain a high-level understanding of visual inputs".

To clarify the meaning of each AI methodology, a concrete example of application in architecture or urban planning is provided for each.

Image classification. A restoration researcher would be interested in classifying a building, or a part of it, or another object based on a photograph of it. This classification would relate to the time of its construction, the style used, and other factors that together create a labelling of each object. An architect would also be interested in classifying his collection of drawings into different categories and, in turn, find for each design, which classification category it would belong to. Machine Learning, part of AI, provides a solution to this problem. The idea is that the machine has learned to make this classification through a model (composed of algorithms) and a considerable number of already classified data points, and then, for each unclassified data point, it can classify it with a certain certainty or predict a classification.

Object detection and localisation. Localisation of an object refers to the identification of its position within an image. As an example of the application of this process, the Unetiq project in Munich is worth mentioning, where architects design the layout of the buildings' land and hand it over to subcontractors. Electrical engineers perform specialised planning tasks. The earth plans are designed with Computer-Aided-Design (CAD) software. They are transformed into a separate file format comprising basic geometric shapes such as lines, dots, and rectangles. The project aims to identify and locate objects in a plan.

Moreover, to enable algorithms to understand the semantic structure of a given plan, contextual information is required for different rooms and objects, for example, the position and size of a kitchen sink, as well as the use of Deep Learning (AI) techniques such as Region Convolutional Neural Networks (R-CNN) [10].

Another example of the application of object localisation is the analysis of "patterns" in pedestrian and car movements to inform city planning decisions that shape infrastructure development and meet community needs.

The implementation of this analysis requires data that, in this case, is video data from many sources, including cameras, to create datasets for training models as well as AI techniques to perform these models, which are the YOLO (You Only Look Once) and SSD (Single Shot MultiBox Detector), the basis of which is the CNN algorithm.

Image segmentation. Image segmentation is the separation of objects, contours, or structures within an image for further analysis. Here, semantic segmentation assigns a class label to each pixel in an image, whereas instance segmentation goes beyond this by identifying and delineating each object. It distinguishes between separate objects of the same class. This is where deep learning algorithms are used.

Regarding the architecture, in Li et al 2023, it is shown that segmentation is the basis of so-called generational AI, and specifically related to several categories of work: 1) generating facade images based on semantic segmentation maps; 2) use of AI generating models for transferring facade image style; 3) generating semantic segmentation maps of the facade based on images. UNET is one of the algorithms the AI uses to segment an image, but its core is the CNN.

Image retrieval. A study by Condorelli (2023) presents a methodology for using photogrammetry-derived imaging data to create 3D models. A combination of AI and photogrammetry has been used specifically for cultural heritage images. Photogrammetry includes multiple photos of a building or structure to create a 3D model. This study aims to expand imaging datasets by increasing the quantity and quality of photogrammetric imagery, resulting in more accurate and detailed 3D models. AI algorithms are used to generate additional images from existing ones.

Some deep learning AI algorithms have been used here, including the visual search engine, which again uses Convolutional Neural Networks (CNNs). They are used for image classification and similarity comparison.

*Image denoising/restoratio**n.*** Removing noise from an image means improving its quality by eliminating noise sources, which, according to Goyal (2024), arise from various natural causes and are difficult to identify and avoid.

According to the same article, the CNN algorithm has achieved excellent results in image recognition, and Chiang and Sullivan were the first to use CNN (deep learning, a specialised subset of ML) to denoise the image.

Image super-resolution. It is stated that the abstraction of buildings from images is extremely important for numerous applications, including urban planning and management, 3D modelling of cities, and change detection. In this context, satellite and aerial images are widely used. Despite numerous successful studies in the literature, this task remains challenging and complex due to several factors that complicate imaging. The U-Net algorithm is applied to the project described in this paper in several variants. This algorithm is a variant of the CNN algorithm.

Image generation. The architect or urban player has a series of building sketches, photos, and photos of the real buildings. This set of images (sketch, photo) can be used to train the machine to predict the view of a building from a given sketch. This approach, called pixel-to-pixel, is explained in detail in TensorFlow. Another approach is to generate image versions from a set of images provided to the machine for training.

I.1 What is Common in Image Processing Methodologies?

DL (deep learning) is a subset of ML (machine learning), which itself is a subset of AI. It is DL that is applied in image processing. It is called deep learning because it involves neural networks with many layers. The data and images are supplied to algorithms that learn from the most informative features, enabling further predictions.

The paper focuses on two aspects that it considers essential for a beginner to start applying image processing methodologies.

- Some ways to create a customised dataset of images,
- The Convolutional Neural Network (CNN) model, used to extract the most expressive features of an image, has the same base logic as many other DL algorithms for image processing.

As a last note, for the Python implementations, Keras is used, a high-level neural network API that runs on top of TensorFlow, one of the most popular deep learning frameworks used in both academic research and commercial applications.

II. DATASETS AND DATA PREPROCESSING

II.1 Use of a dataset of images provided by Python libraries

Python is a programming language specialised in the application of algorithms and machine learning methodologies, as it includes libraries of specialised programs for various operations.

For these methodologies, Python provides ready-made datasets, and users must build and train the model. Datasets are structured data collections organised in table form, where each column represents a variable and each row a record. They also support a variety of methods that perform different actions on datasets, such as uploading data and selecting data for training, validation, and testing. Training, validation, and test data are three categories of data, usually used by ML and DL algorithms.

Among the datasets offered by Python, there are image datasets such as MNIST in K**eras.datasets** package [18] of Python. This dataset is a collection of 60,000 handwritten digit images, with dimensions 28x28, and also includes 10,000 test images. These datasets are very useful for people to learn different image processing techniques at the beginning, before using their customised datasets. The following code shows how such a dataset, provided by the Python K**eras** package, can be used.

```
#import the necessary libraries
import matplotlib.pyplot as plt
import numpy as np
from keras.datasets import mnist
from keras.models import Sequential
from keras.layers import Dense, Dropout, Flatten, Conv2D, MaxPooling2D
from tensorflow.keras import utils

#data loading from mnist dataset
(training_data,train_labels), (validation_data, validation_labels)= mnist.load_data()
'''
Each image have size of 28*28 pixels; that is 28 pixels height and 28 pixels width and hence (1, 28, 28), 1 in first part is to specify color depth of the pixel. 1 is for greyscale image(black and white image).
(60000, 1, 28, 28)- is the shape of matrix the data have since we have 60000 images in which 28*28 pixels and each pixel have a value in this matrix. astype('float32') convert data in float format; this makes computation more convenient.
```

```
'''
training_data = training_data.reshape(training_data.shape[0],28,28).astype('float32')
validation_data = validation_data.reshape(validation_data.shape[0],
'''
Normalization of the data. The most common pixel format is the byte image, where this number is stored as an 8-bit integer giving a range of possible values from 0 to 255. Typically zero is taken to be black, and 255 is taken to be white. Values in between make up the different shades of gray. Dividing an image by 255 simply rescales the image from 0-255 to 0-1
'''
training_data = training_data /255
validation_data = validation_data /255
```

Of course, the architect or urban planner's interest would be to create a customised dataset with his images. In the following are some ways to create customised datasets.

II.2 Some methodologies to create a customised dataset

A dataset is the main component of machine learning methods. The quality of the images during their creation and a preliminary treatment of them substantially affects the outcome we receive. These steps are detailed in Kokorin (2023). One of the recommendations given here is that if the dataset is over 100K images with many classification classes, it is recommended to use a structured database where the main data could be:

- file name;
- file path;
- annotation data (coordinates of the bounding box position, and some relative coordinates);
- class data.

The case of database use is not the object of this study. It is worth noting an important element in this preliminary image treatment: "Data Augmentation," a technique for creating new images from existing ones by applying several transformations. This way, the dataset is expanded with new images, leading to more accurate results. These transformations, according to GeeksforGeeks (2023 are:

- Rotation (rotates the image by a certain specified degree).
- Shearing (transforms the image's orientation).
- Zooming (zoom in or zoom out).
- Cropping (crops the image or selects a particular area from an image).
- Flipping (flips the orientation of the image).
- Changing the brightness level (helps to combat illumination changes).

There are different ways to realise a dataset, and some of them are presented here.

Keras ImageDataGenerator with flow_from_directory(). According to StudyMachineLearning, using Keras' ImageDataGenerator class and flow_from_directory method allows loading images from a hierarchy of folders, where the folder structure determines the image classification. The following image depicts the classification structure we will use for machine training.

```
train/
...classA/
......a_image_1.jpg
......a_image_2.jpg
. . . . .
...classB/
......b_image_1.jpg
......b_image_2.jpg
. . . . .
testing/
```

The ImageDataGenerator class provides several parameters to control scaling, the proportions of the training and validation sets, and data augmentation. In TensorFlow there are parameters that the constructor of this class can use:

flow_from_directory(): Takes the path to a directory and generates batches of augmented data. The full set of parameters of this method is given in AI Learner (2019.

The preprocessing of the data until the preparation of the training, validation and testing sets would be presented with these lines, where only a part of the possible parameters were used:

```
#imports
from keras.preprocessing.image import ImageDataGenerator
src_path_train = '.../train/'  # see the folder hierarchy above
src_path_test =' ...testing/'
img_width, img_height = 100, 100 #image dimensions
batch_size=8 #number of images in each batch
train_datagen = ImageDataGenerator(
    rescale=1 / 255.0,  #normalization of the data
    # Data Augmentation parameters
    rotation=20, zoom=0.05,width_ shift=0.05, height_shift =0.05,
      shear_range=0.05, horizontal_flip=True,
    validation_split=0.20
         #proportion of data split into training and validation data
)
train_generator = train_datagen.flow_from_directory(
    directory=src_path_train, #path of the training data folder
    target_size=( img_width, img_height),  color_mode="rgb",
           #imazhe features, color_mode=3 channels
    batch_size=batch_size
    class_mode="categorical", subset='training',  shuffle=True)
      #shuffle = True, the dataset will be randomly shuffled to avoid
      #any overfitting in training. It is True only for training and validation set
validation_generator = train_datagen.flow_from_directory(
   . . . the same parameters like train_generator )
test_datagen = ImageDataGenerator(rescale=1 / 255.0)
```

```
test_generator = test_datagen.flow_from_directory(
   directory=src_path_test, #path of the test data folder
   #the same parameters for the image
   batch_size=1,    class_mode=None,  shuffle=False,)
```

flow_from_directory(directory), description: Takes the path to a directory, and generates batches (sets) of augmented/normalised data.

Labelling the images, writing them into a CSV file and using this .csv file as a dataset. Another way to create a customised dataset is to convert images to numerical data and save them in a .csv file. In Cswah (2018), there is a description of the steps necessary to read the images from a folder and to write them into a .csv file, associating them with a classification label:

- Gather images for the dataset in different folders corresponding to their classes.
- Rename the pictures according to their classes.
- Merge them into one folder.
- Resize the pictures.
- Convert all images into the same file format.
- Labelling the images.
- Convert images into a CSV file.
- A few tweaks to the CSV file.
- Load the CSV (BONUS).

This sequence of steps is kept as in the original material of Baheti (2023). However, the manual steps 5, 6, 7, and 8 are converted to code, and another step is added: a preprocessing step that produces training, validation, and test data for the machine learning algorithm. The following steps represent what is mentioned:

a) Data preparation

- Gather images for your dataset; assume all images are .jpeg files. The images are stored in different folders based on their classification.
- Rename the images according to their classes:
    - Select all the images in the folder
    - Rename the images, for example, with the name class_A. The names will be the same, but associated with (1), (2).
- Merge all the files into one folder

b) Image format unification

The next steps are done by code: image resizing, image labelling, writing to a CSV file, and preprocessing that results in training data, validation data, and test data.

```
#Resize all the images with the same format, this function is used in writetocsv():
```

```
def resizeimage(imgstart): #imgstart original image
    # size of original image
    width, height = imgstart.size
    # Setting the points for cropped image
    left = 6
    top = height / 4
    right = 174
    bottom = 3 * height / 4
    # Cropped image, the original is not changed
    imgend = imgstart.crop((left, top, right, bottom))
    newsize = (48, 48)
    imgend = imgend.resize(newsize)
    return imgend
```

c) Data Labeling

```
#Labeling the images, this function is used in writetocsv():
import os
def createFileList(myDir, format='.jpg'):
  fileList = []
  labels = []
  names = []
  # classify the images by a common letter of image name
  # the names here are classA(#).jpeg, classB(#).jpeg
  keywords = {"A": "1", "B": "0", }
  # os.walk() generates the file names in the path specified
  for root, dirs, files in
      os.walk(src_path_train, topdown=True):
      for name in files:
          if name.endswith('.jpeg'):
            fullName = os.path.join(root, name)
            fileList.append(fullName)
        for keyword in keywords:
            if keyword in name:
              labels.append(keywords[keyword])
            else:
              continue
        names.append(name)
  return fileList, labels, names
```

d) Writing into .csv file

```
import csv
```

```
from PIL import Image
import pandas as pd
import numpy as np
# Write image data into a CSV file
#Create three lists with images,labels and their names
def writetocsv():
       myFileList, labels, names = createFileList(src_path_train)
        i = 0
       for file in myFileList:
         # for each image
              img_file = Image.open(file)
             # image is resized
               img_file=resizeimage(img_file)
             # the original parameters
               width, height = img_file.size
               format = img_file.format
               mode = img_file.mode
             # Make image Greyscale
               img_grey = img_file.convert('RGB')
             # convert input(pixels) to an array
               value = np.asarray(img_grey.getdata(),dtype=int).
                               reshape((width, height, 3))
             # convert array value into one dimensional array
               value = value.flatten()
             # join all pixels into one string with space in between
               s = ' '.join(str(x) for x in value)
             # create two columns: label, pixels
               value = np.append(labels[i], s)
               i += 1
             #filename=src_path_train+'classescsv.csv'
              filename="/classescsv.csv"
              with open(filename, 'a') as f:
                  writer = csv.writer(f)
                  writer.writerow(value)
              #add some header
              headerList = ['Labels', 'Pixels']
              data = pd.read_csv(filename)
              # converting data frame to csv  assing a header
              data.to_csv(filename, header=headerList, index=False)
```

e) Transform the Image information from a sequence of numbers into an array of float numbers

```
import numpy as np
```

```
# Image preprocessing  used in gettrainvalidationdata():
def imagesetting(data):
    image_size = (48, 48, 3)# RGB image
    width, height, color = 48, 48, 3 #RGB image
    imgspixels = []
    for pixel_sequence in data:
        # get the pixel data from pixel column in .csv
      I mgpixels = [int(pixel) for pixel in  pixel_sequence.split(' ')]
        # create an array of pixels for the image
        imgpixels =  np.asarray(imgpixels).reshape(width, height, color)
        # Convert the pixel values of an image to uint8
        imgpixels =  np.resize(imgpixels.astype('uint8'), image_size)
        # pixel values of an image are converted this type
        # and added into an array
        imgspixels.append(imgpixels.astype('float32'))
  imgspixels = np.asarray(imgspixels)
  return imgspixels
```

Note: imgpixels.astype('uint8') is used for memory efficiency (because this format uses 1 byte per pixel), for standardisation (because image processing libraries use this format), and for performance of different operations.

f) Reading the .csv file and creating three sets of data: training data, validation data and test data

For some image processing tasks, such as image classification, three data sets are used: the training set, the evaluation set, and the test set. Training data is what the model uses to learn patterns. During this process, another set of data, the validation set, is used instead of the training set and serves to monitor performance during training. After training, the model is tested using another set, called the test set. In practice, all the data can be divided into 80% for training and 20% for validation. The test data can be data outside of the dataset.

The following Figure represents the role of each of these sets:

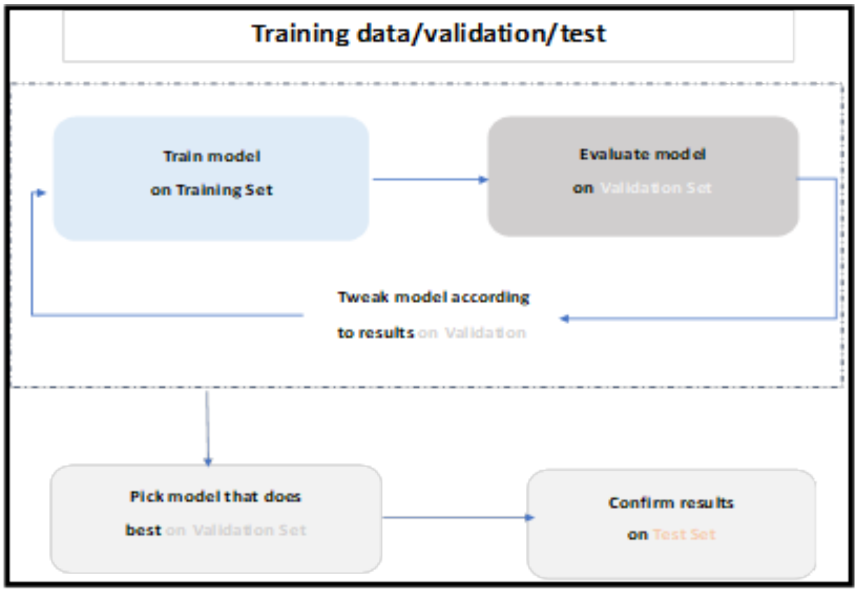


Figure 11. Training, validation, test data

Source: https://www.v7labs.com/blog/train-validation-test-set

```
# definition of the training, validation and test sets
 import pandas as pd
 from sklearn import model_selection
 def gettrainvalidationdata():
        Corpus = pd.read_csv(r"\classescsv.csv")
        X_train, rest_data,y_train, y_rest_data =
          model_selection.train_test_split(Corpus['Pixels'],
             Corpus['Labels'], train_size=0.8, shuffle=False)
        X_validation, X_test, y_validation, y_test =
          model_selection.train_test_split(rest_data,y_rest_data,
             test_size=0.5, shuffle=False)
        X_train = imagesetting(X_train)
        X_test = imagesetting(X_test)
        X_validation = imagesetting(X_validation)
        y_train = utils.to_categorical(y_train)
        y_test = utils.to_categorical(y_test)
        y_validation = utils.to_categorical(y_validation)
        print('y_train.shape[1]', y_train.shape[1])
        num_classes = y_train.shape[1]
        return X_train, X_test, y_train,
```

```
y_test,X_validation,y_validation, num_classes
```

Use images from a folder and add labels programmatically. In some cases, it is not possible to use Excel to create a dataset for problems that require images with associated labels. Excel has its own limitations; for example, the maximum number of columns is 16,384, which is not enough to store images of size 64x64. In this case, the images could be labelled programmatically. A practical way is to keep all the images in one folder, with regular filenames, and have a common part of each filename for each class. This part would be used by a program to label the image, as shown in the code below.

```
def createLabels():
    fileList = []
    labels = []
    names = []
    keywords = {"me": 1, "pa": 0, }  # keys and values to be changed as needed
    for root, dirs, files in os.walk(". . ./images", topdown=True):
        for name in files:
            if name.endswith('.jpg'):
                fullName = os.path.join(root, name)
                fileList.append(fullName)
            for keyword in keywords:
                if keyword in name:
                    labels.append(keywords[keyword])
                else:
                    continue
            names.append(name)
    return labels

def load_image(image_path):
    img = tf.io.read_file(image_path)
    img = tf.io.decode_jpeg(img)
    img = tf.image.resize_with_crop_or_pad(img, img_size, img_size)
    img = tf.cast(img, tf.float32)
    img = (img - 127.5) / 127.5
    return img

def tf_dataset(images_path):
    dataset = tf.data.Dataset.from_tensor_slices(images_path)
    dataset = dataset.shuffle(buffer_size=10240)
    dataset = dataset.map(load_image)
    return dataset
```

```
images_path = glob("C:\Polis\Kenti\images\Imagemeetikete\*")
X_train = tf_dataset(images_path)
print('images_dataset',len(X_train))
y_train = createLabels()
print('y_dataset',len(y_train))
y_train = tf.data.Dataset.from_tensor_slices(y_train)
print('X_train',X_train.take(0))
dataset = tf.data.Dataset.zip(X_train, y_train)
dataset = dataset.shuffle(buffer_size=1000).batch(batch_size)
```

Use images from a folder when labels are not needed (for image generation). As we mentioned above, unsupervised approaches, such as image generation, do not require associating images with labels. The following code provides a dataset of images for further processing.

```
#Import TensorFlow and Other Libraries
from glob import glob
import tensorflow as tf
from tensorflow.keras.layers import *
from matplotlib import pyplot as plt
IMG_H = 64
IMG_W = 64
IMG_C = 3
batch_size = 128
latent_dim = 128
 images_path = glob(". . ./train/*")

def load_image(image_path):
    img = tf.io.read_file(image_path)
    img = tf.io.decode_jpeg(img)
    img = tf.image.resize_with_crop_or_pad(img, IMG_H, IMG_W)
    img = tf.cast(img, tf.float32)
    img = (img - 127.5) / 127.5
    return img

def tf_dataset(images_path, batch_size):
    dataset = tf.data.Dataset.from_tensor_slices(images_path)
    dataset = dataset.shuffle(buffer_size=10240)
    dataset = dataset.map(load_image,
              num_parallel_calls=tf.data.experimental.AUTOTUNE)
    dataset = dataset.batch(batch_size)
    dataset = dataset.prefetch(buffer_size=tf.data.experimental.AUTOTUNE)
    return dataset
```

III. CONVOLUTIONAL NEURAL NETWORK (CNN) MODEL AND ITS IMPLEMENTATION.

There are many papers and materials on the Convolutional Neural Network, which is used specifically in image processing and is known for its high efficiency. In recent years, supervised and unsupervised learning with convolutional networks (CNNs) has seen widespread adoption in computer vision applications.

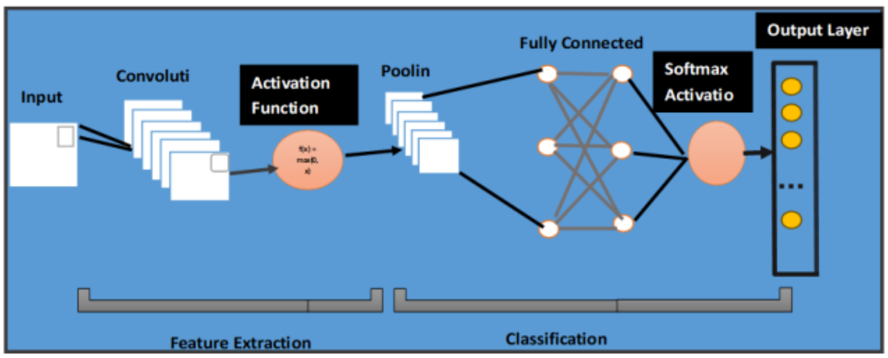


Figure 2. Block diagram of the CNN model

Source. https://blog.talent500.co/wp-content/uploads/2024/03/images-19-1.png

The idea of this model is to extract image features from images that will serve to build a law that defines image patterns when this model (features and law) is trained on a large number of images, so-called training images.

To clarify how this model works, we take as an example one of the machine learning methodologies that uses it: image classification into several classes or categories. Machine learning classification methods aim to create a model or function that, using a large dataset of images, learns a pattern for each image category. Then, each new image is compared with these patterns, and the closest one determines its category. This function or model is not derived from a prior assumption; rather, it is built on the training data and some operations that define its laws. Therefore, such a model belongs to the deep learning family of methods.

In CodezUp, several papers present this model by building it step by step, creating a so-called layer at each step, where each layer has an input and produces an output as a result of an algorithm. The basic layers are:

1. Data Preprocessing
2. Convolutional Layers
3. Pooling Layers
4. Flatten Layers

5. Fully Connected Layers

Below are the steps of this model, illustrated in Fig. 1, implemented in Python.

Input and Preprocessing. For images, the input data are pixels. Therefore, an image is presented as a matrix of numbers representing the colour intensity at each coordinate. However, the number of these values is extremely large as the input variable. The CNN model, using several algorithms, extracts the most determining qualities of an image by reducing the large number of qualities.

Convolutional Layer and Function: **ReLU (Rectified Linear Unit).** The defining qualities in relation to images are the image's most significant colour intensities. The methodology for determining these defining qualities consists of applying a matrix, called a kernel or filter, to the square of the image's pixels, with the same dimensions, element-by-element. This kernel matrix is initialised to random values, but over time, its values are learned. This matrix operates by sliding in steps across the entire pixel space of the image.

The result will again be a matrix with smaller dimensions than the initial pixel matrix, and the change in size depends on the sliding window step. The result is called a feature map, and its components are neurons. For simplicity of calculations, we are considering Figure 2 which shows the pixels of a three-color image, with dimension (5x5x3) and a kernel with dimension (3,3,3), where the third dimension is presented into three 3x3 planes, where each of them, operates with one of the colors of the image (red, green, blue), which in the language of the algorithm are called channels. So, there are 3 of them. The results are summarised , and the final result presents, in a summarised way, the representative quality of all the qualities of the squares in the three channels of the image pixels.

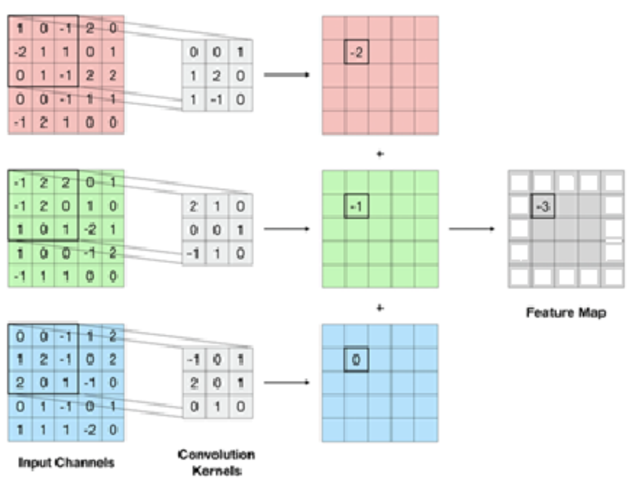


Figure 3. Convolution operation demonstration

Source: https://setosa.io/ev/image-kernels/

This is how the value -3 is calculated

The result from first kernel is:

$$1 * 0 + 0 * 0 + (-1) * 1 + (-2) * 1 + 1 * 2 + 1 * 0 + 0 * 1 + 1 * (-1) + (-1) * 0 = -2$$

Form the second one the calculation gives the result:

$$(-1) * 2 + 2 * 1 + 2 * 0 + (-1) * 0 + 2 * 0 + 0 * 1 + 1 * (-1) + 0 * 1 + 1 * 0 = -1$$

And from the third one the result is:

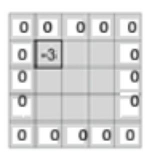
$$0 * (-1) + 0 * 0 + (-1) * 1 + 1 * 2 + 2 * 0 + (-1) * 1 + 2 * 0 + 0 * 1 + 1 * 0 = 0$$

Summarising the above results it is taken: $-2 + (-1) + 0 = -3$

Then a (3x3x3) kernel or filter defines a feature map with dimension (3x3) for a 5x5x3 image.

In CNN commands, a number of filters is defined as 16, 32, 64, etc. If 32 such filters were to operate as above, they would generate 32 feature maps with dimensions 3x3 in this case, which would be considered the input for further operations, called channels. So initially, we had 3 colour channels, and then the components of the created feature map were used as channels.

The dimension of the feature map, as we see, is smaller than that of the input matrix. Since we want to keep the feature map at the same dimension as the initial image, we complete the pixel matrix to the same dimension by padding with zeros.



Using an input matrix with a contour consisting mostly of zeros allows the example above to maintain the dimension of the input image matrix. In algorithm terminology, this is called padding, and considering this element, the CNN uses the following formula to determine the dimension of the feature map:

The output feature map size is calculated as follows:

$$O = \lfloor \frac{n - f + 2 \cdot p}{s} \rfloor + 1$$

were,

- O: Output size
- n: Input size (height or width)

- f: kernel size
- p: Padding
- s: Stride

Padding based on is calculated by the formula

$$if\ n\%s = 0:$$

$$p = max(f - s, 0)p$$

$$else:$$

$$p = max(Fh - (n\%s), 0)$$

In this case, p=3-1=2

Applied to Fig 2, it would be

$$O = \lfloor \frac{5 - 3 + 2 \cdot 1}{1} \rfloor + 1 = \lfloor 4 \rfloor + 1 = 5$$

A so-called activation function is applied to this result. Its goal is to provide a nonlinear relationship between the input and output that better reflects the reality of complexity. There are different activation functions, such as Sigmoid, Tanh, Softmax, ReLU, and Softplus.

In this implementation, the ReLU function is chosen, and the graph is shown in Figure 3.

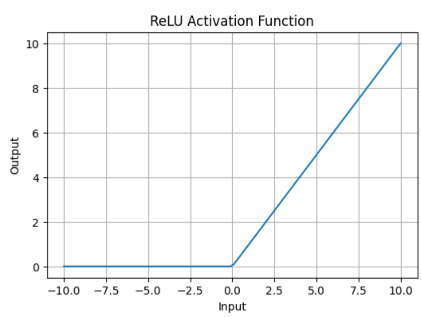


Figure 4: Visualisation of the Rectified Linear Unit (ReLU) Activation Function

Source: Author processing

Expressed by a formula, this functionality is $f(x) = max(0, x)$. So, for the value of the feature map in Figure 2, the result is f(-3)=0.

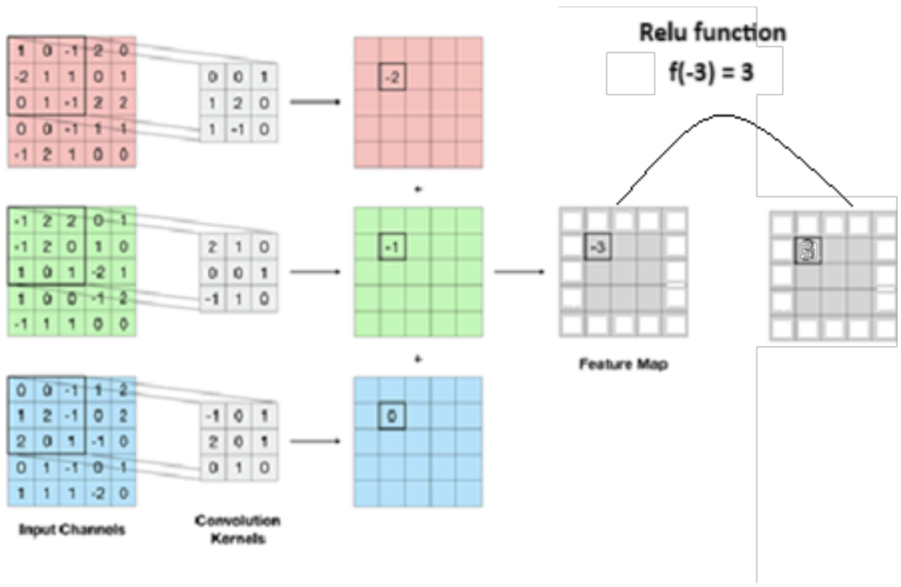


Figure 5. Convolution Process Across Input Channels Followed by ReLU Activation in a CNN

Source: Author's processing

Convolutional Layer Implementation. What is mentioned is implemented in Python by the following lines.

```
model=Sequential()
model.add(Conv2D(32,3,1, padding='same', input_shape=(48,48,3),
activation='relu'))
```

Here, the feature map is created using 32 3x3x3 kernels, with padding to maintain the same output dimension as the input. Conv2D is used for the 3-dimensional images (RGB). The output dimension is = $\lfloor\frac{48-3+2*1}{1}\rfloor + 1 = 48$ and there are 32 channels. A formula in GeeksforGeeks calculates the number of parameters.

Parameters=(kernel_w x kernel_h x channels+1)× filters=(3 x 3 x 3+1) x 32= 896

Where:

- kernel_w = kernel width
- kernel_h = kernel height
- channels = number of input channels
- filters= number of filters (output channels)

This is the result that model.summary() gives as well.

Model: "sequential"

Layer (type)	*Output Shape*	*Param #*
conv2d (Conv2D)	*(None, 48, 48, 32)*	*896* . . .

Pooling Layer: **Max Pooling/Average Pooling.** After creating a convolutional layer, an operation is applied to the feature maps to reduce their dimensions. The following Figure shows how it works. The square of values in feature maps is replaced with the maximum or average of their values. An example is shown in Figure 6.

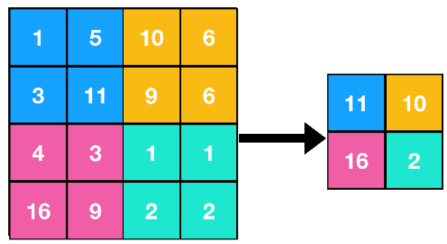


Figure 6. Fully connected layers & SoftMax function output

Source: Authors processing

Then the result's dimensions are reduced. The formula in[34] calculates the output feature map size:

$$O = \left(\frac{h-f}{s} + 1\right) X \left(\frac{w-f}{s} + 1\right) x\ c$$

where,

- O: Output size
- h: height of feature map
- w: width of feature map
- f: size of pooling filter
- c: channels, or the depth of the output
- s: Stride length

Python implementation

```
model.add(MaxPooling2D(pool_size=(2,2), padding='same'))
```

Note:

In the case of odd dimensions, padding can be used by adding a column on the right and a row on the bottom.

This case would be implemented by

model.add(MaxPooling2D(pool_size=(2,2), padding='same', ceil_mode=true))

In this implementation, the formula gives $\frac{h-f}{s}+1=\frac{48-2}{2}+1=24$, which can be shown by model.summary()

Model: "sequential"

Layer (type)	*Output Shape*	*Param #*
conv2d (Conv2D)	*(None, 48, 48, 32)*	*896*
max_pooling2d (MaxPooling2D)	*(None, 24, 24, 32)*	*0*

Flatten Layer

In this step, the feature maps are converted into a 1D array, and the implementation in Python is:
model.add(Flatten())
The result shown by the model summary()

Model: "sequential"

Layer (type)	*Output Shape*	*Param #*
conv2d (Conv2D)	*(None, 48, 48, 32)*	*896*
max_pooling2d (MaxPooling2D)	*(None, 24, 24, 32)*	*0*
flatten (Flatten)	*(None, 18432)*	*0* . . .

Fully Connected Layer, or Dense Layer. The 1D array that comes from the Flatten function step represents the most important features that define the image and is used as input for functionalities that predict each image's class label. The following Figure shows the operations in this step:

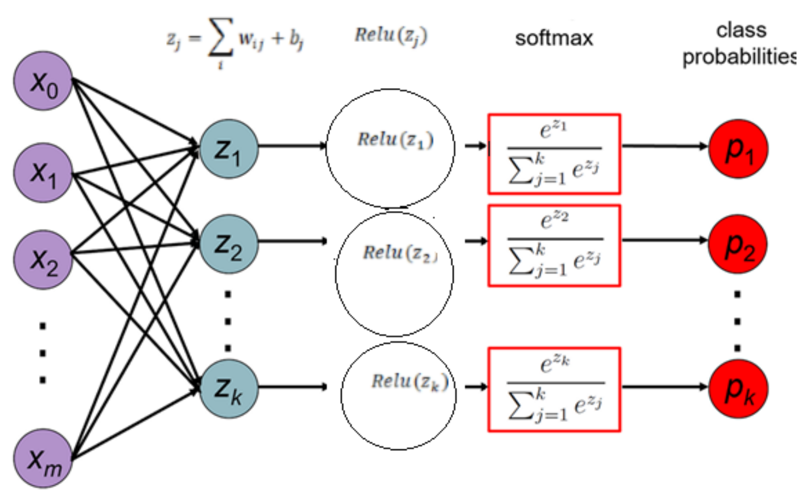


Figure 7. Operations in steps

Source: Adapted from

https://purwadhika.com/blog/neural-network-implementation-for-image-classification-using-cnn

Figure 7 shows a very defining element in the neural network. Here, there are two layers of neurons, where each element of the first layer from the previous step is connected to each element of the second layer, the number of which equals the number of classes that will categorise the images. The equation expresses this connection.

$$z_j = \sum_i \quad w_i \cdot x_i + b_j$$

where the input neurons $x_1, x_2, \ldots$ the output neurons $z_1, z_2, \ldots$ are determined by means of the so-called weights w_i and biases b_j. The weights show the influence of each input neuron on that output and the bias defines a type of regulator regarding the dependence $z_j = \sum_i \quad w_i \cdot x_i$.

Then it operates the Relu function that returns all z_j values to positive and the softmax function that determines the probability for each output neuron in the form of an expression

$$Softmax(z_j) = \frac{e^{z_j}}{\sum_{l=1}^{k} \quad 1e^{z_l}}$$

The following two lines represent the implementation of the above operations:

```
model.add(Dense(128, activation='relu'))
model.add(Dense(num_classes, activation='softmax'))
```

A very clear explanation of the Dense layer is given in [37]. The weights and biases are learnable parameters that the neural network adjusts during training to minimise the loss function.

In this case, the dimensionality reduction happens because the weight matrix in the Dense layer has a shape of (*18432*, 128). Therefore, there are *18432* neurons (input_dim), reduced to 128 neurons (output_dim). The weight matrix would have a shape of (*18432*, 128). The result would be a 128-dimensional vector, effectively reducing the input dimensionality.

After the linear operation, in this layer an activation function is typically applied to introduce non-linearity into the model, allowing the network to learn complex patterns. Here are two Dense layers. The first reduces the output dimension to 128, and the second gives an output dimension equal to the number of classes. For each class, the Softmax activation function computes the probability.

During the process described above, the concept of learning is mentioned twice: once regarding the values of weights and biases related to the filter values in the convolutional layer, and once regarding the values of weights and biases related to the dense layer, which are called model trainable parameters. The process of learning is based on a function, called a loss function, that measures the error between the true distribution of class values in the training set and the distribution of class values the model outputs.

Based on the values returned by the loss function, the above process is repeated a number of times, and before each repetition, the trainable parameters are updated. One way to define this loss function and some other parameters for the model is the following

```
model.compile(loss='categorical_crossentropy', optimizer='adam',  metrics=['accuracy'])
```

Shortly:

loss='categorical_crossentropy' measures the difference between the predicted probability distribution and actual (true) class distribution. Categorical Cross-Entropy (CCE), also known as softmax loss or log loss, is one of the most commonly used loss functions in machine learning, particularly for classification problems.

optimizer='adam' activates a method that adjusts the learning rates of parameters during training.

metrics=['accuracy'] allows the use of several accuracy metrics, in this case the Keras Accuracy metric, to evaluate the performance of machine learning models.

The completed information given by model.summary() would be

```
Model: "sequential"
Layer (type)                    Output Shape              Param #
=================================================================
 conv2d (Conv2D)                (None, 48, 48, 32)        896
 max_pooling2d (MaxPooling2D)   (None, 24, 24, 32)        0
 flatten (Flatten)              (None, 18432)             0
 dense (Dense)                  (None, 128)               2359424
 dense_1 (Dense)                (None, 2)                 258
=================================================================
Total params: 2360578 (9.00 MB)
Trainable params: 2360578 (9.00 MB)
Non-trainable params: 0 (0.00 Byte)
```

All that is mentioned above defines the construction of the CNN model. The model creates patterns using the training and validation data, and it can be tested on test data. The lines implement these two moments:

```
cnnmodel.fit(X_train, y_train, validation_data=(X_validation,y_validation),epochs=16, batch_size=10, verbose=2)

score= cnnmodel.evaluate(X_test, y_test, verbose=0)
```

It is the function fit() that performs the training, and to do so, the data are divided into "batches" of size 10. This process is repeated 16 times, defined by the number of epochs. It is just these iterations that help learn better which parameters work inside the model, improving them at each iteration. The model is trained on the training set, and evaluation is performed on the validation set after every epoch.

Some notes about improving the model. The CNN model's presented layers are the main ones it uses to extract images and create patterns. Nevertheless, the studies and practical experiences show that the same ways can improve this model:

- Using some convolutional layers (implemented by conv2d())
- Adding the dropout layer (implemented by dropout())
- Using some fully connected layers (implemented by dense())

Dropout is a technique that helps avoid overfitting, making the model more general and better adapted to reality. Moreover, this is achieved by randomly dropping some outputs during training. The figure below shows this trick.

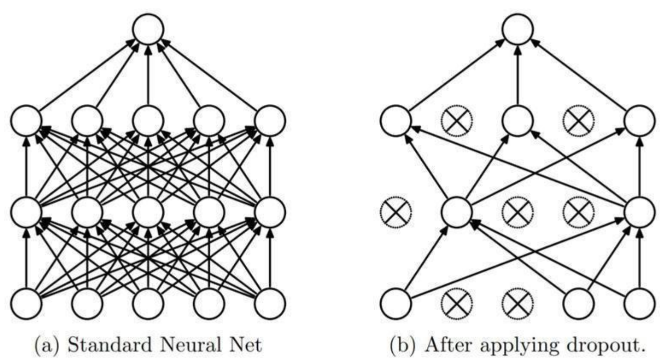


Figure 8. Effect of Dropout on Neural Network Architecture

Source: Authors processing based on LearnDataSc

In this paper, several models and evaluation metrics are used, and the following results are shown: accuracy, classification report, and confusion matrix.

The models:

```
#Model1

model = Sequential()
model.add(Conv2D(32, 3,  padding='same', input_shape=(48, 48, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2), padding='same'))
model.add(Flatten())
model.add(Dense(128, activation='relu'))
model.add(Dense(num_classes, activation='softmax'))
```

```
#Model 2

model = Sequential()
model.add(Conv2D(32, (3, 3), input_shape=(48, 48, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(128, (3, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Flatten())
model.add(Dense(128, activation='relu')),
model.add(Dense(num_classes, activation='softmax'))
```

```
#Model 3

model = Sequential()
model.add(Conv2D(32, (3, 3), input_shape=(48, 48, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(128, (3, 3), activation='relu'))
model.add(MaxPooling2D((2, 2)))
model.add(Dropout(0.2))
model.add(Flatten())
model.add(Dense(128, activation='relu')),
model.add(Dense(num_classes, activation='softmax'))
```

III. SOME RESULTS AND INTERPRETATIONS

These models are trained on a dataset of road images. The goal is to classify the data into irregular and regular roads. The labelling is done with 1 for irregular roads and 0 for regular roads. There are different metrics to evaluate classification results, such as accuracy, classification report, and confusion matrix. Accuracy is included in the classification report.

Three indicators are elements of a classification report: precision, recall, and F1 score for each class in a classification problem. Each of these indicators helps provide a complete understanding of classification performance.

The classification report for the database of road images and for the labelling done represents the following indicators:

$$precision_{class0} = \frac{correctly\ predicted\ broken\ roads}{total\ \ broken\ roads}$$

$$recall_{class0} = \frac{correctly\ predicted\ broken\ roads}{predicted\ as\ broken\ (broken+not\ broken)}$$

$$f1 - score_{class0} = 2 * \frac{precision*recall}{precision+recall}$$

The confusion matrix shows four outcomes. For the dataset used in this application, the confusion matrix gives this information:

$$r_{00=}regular\ road_{classified\ as\ not\ regular}$$

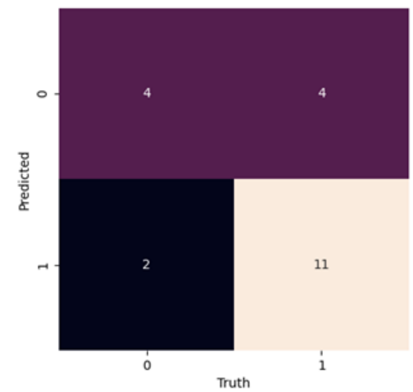
$$r_{01=}iregular\ roads_{classified\ as\ regular}$$

$$r_{10=}regular\ roads_{classified\ as\ iregular}$$

$$r_{11} = iregular\ road\ _{classified\ as\ iregular}$$

Model 1

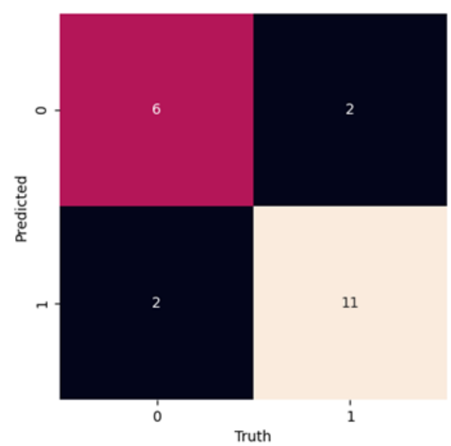
Classification	Report			
	precision	recall	f1 score	support
0	0.67	0.5	0.57	8
1	0.73	0.85	0.79	13
accuracy			0.71	21
macro avg	0.7	0.67	0.68	21
weighted avg	0.71	0.71	0.7	21



Confusion Matrix

Model 2

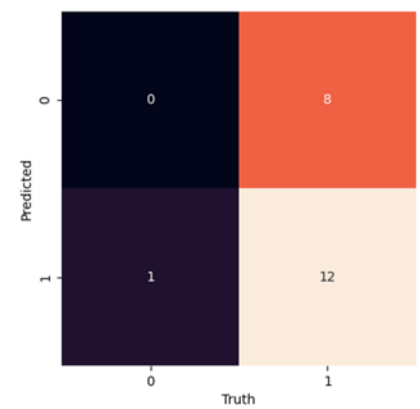
Classification	Report			
	precision	recall	f1-score	support
0	0.75	0.75	0.75	8
1	0.85	0.85	0.85	13
accuracy			0.81	21
macro avg	0.8	0.8	0.8	21
weighted avg	0.81	0.81	0.81	21



Confusion Matrix

Model 3

Classification	Report			
	precision	recall	f1-score	support
0	0	0	0	8
1	0.6	0.92	0.73	13
accuracy			0.57	21
macro avg	0.3	0.46	0.36	21
weighted avg	0.37	0.57	0.45	21



Confusion Matrix

Where the calculations (for instance, for class 1, for Naive Bays, in the first model) are as follows:

$$precision_{class0} = \frac{correctly\ predicted\ broken\ roads}{total\ \ broken\ roads} = \frac{11}{15} = 0.73$$

$$recall_{class0} = \frac{correctly\ predicted\ broken\ roads}{predicted\ as\ broken\ (broken+not\ broken)} = \frac{11}{13} = 0.85$$

The f1-score is the harmonic mean of precision and recall:

$$f1 - score_{class0} = 2 * \frac{precision*recall}{precision+recall} = 2 * \frac{0.73*0.85}{0.73+0.85} = 0.79$$

Support is the number of actual occurrences, which is 21.

The macro average is the average of the precision, recall, and F1-scores:

macro avg: $\frac{0.67+0.73}{2} = 0.7$

The weighted average considers the number of samples in each class:

$$weighted\ avg = \frac{0.67*8+0.73*13}{21} = 0.71$$

The idea of using three models here is to understand and compare their performances on this dataset and with the number of epochs used for the three models, based on the information provided by the classification report and confusion matrix.

In the table below, the averages and weighted averages of the indicators:

	Model 1	Model 2	Model 3
Accuracy	0.71	0.81	0.57
Macro precision	0.7	0.8	0.3

Weighted precision	0.71	0.81	0.37
Macro recall	0.67	0.8	0.46
Weighted recall	0.71	0.81	0.57
Macro f1-score	0.68	0.8	0.36
Weighted f1-score	0.7	0.81	0.45
performance		The best model	

- Accuracy: The number of correctly classified roads (the broken ones and the regular ones) divided by the total number of roads. It seems the second model has a better result for this indicator, but this is not enough information; it is a general one.
- Precision: the formula $\frac{correctly\ predicted\ broken\ roads}{total\ \ broken\ roads}$ consider only the broken roads and count how many are correctly predicted among them. Ideally, it should be 1; therefore, the closer to 1, the better the performance. In that aspect, the second model is the best.
- F1-score: the formula $2 * \frac{precision*recall}{precision+recall}$ Becomes when both precision and recall are high, and we see that the second model has the highest value of this indicator.

Therefore, the second model is the best for classifying roads.

IV. CONCLUSIONS

This paper serves as a sensitisation for architects and urban planners on the use of AI methodologies, specifically DL methodologies, in their work.

For a beginner in this field, it is helpful to know how to start using them. Therefore, this paper focuses on two aspects common to the implementation of image processing methodologies, specifically important for architects and urban planners.

The first aspect is the organisation of the data into a dataset. Therefore, this article presents a summary of alternatives for creating image datasets to address various problems. Given that the convolutional neural network (CNN) model is a key component of many other Deep models for image processing, the CNN is presented in detail, and some alternatives are applied to a dataset of images of regular and irregular roads.

An interpretation and comparison of the results are provided, giving a beginner in this field practical experience. After all, this paper will, for a start, help review the teaching programs related to the inclusion of AI elements.

REFERENCES

Arat, M. M. (2019, January 17). *Implementing "SAME" and "VALID" padding of TensorFlow in Python*. Mustafa Murat ARAT. https://mmuratarat.github.io/2019-01-17/implementing-padding-schemes-of-tensorflow-in-python

Baheti, P. (2021, September 13). *Train test validation split: How to & best practices*. V7 Labs. https://www.v7labs.com/blog/train-validation-test-set

Condorelli, F. (2023). Image retrieval for 3D modelling of architecture using AI and photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-M-2*, 441–446. https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-441-2023

GeeksforGeeks. (2023, August 3). *Python | Data augmentation*. GeeksforGeeks. https://www.geeksforgeeks.org/machine-learning/python-data-augmentation

GeeksforGeeks. (2024). *Categorical cross-entropy in multi-class classification*. GeeksforGeeks. https://www.geeksforgeeks.org/deep-learning/categorical-cross-entropy-in-multi-class-classification

GeeksforGeeks. (2025, June 20). *How to calculate the number of parameters in CNN?* GeeksforGeeks. https://www.geeksforgeeks.org/deep-learning/how-to-calculate-the-number-of-parameters-in-cnn/

GeeksforGeeks. (2025, July 15). *What is Adam optimizer?* GeeksforGeeks. https://www.geeksforgeeks.org/deep-learning/adam-optimizer/

Goyal, C. (2024, November 12). *Image denoising using autoencoders – A beginner's guide to deep learning project.* Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/07/image-denoising-using-autoencoders-a-beginners-guide-to-deep-learning-project

Idiot Developer. (2020, July 24). *DCGAN — Implementing deep convolutional generative adversarial network in TensorFlow*. Idiot Developer. https://idiotdeveloper.com/dcgan-implementing-deep-convolutional-generative-adversarial-network-in-tensorflow/

Jin, S., Tu, H., Li, J., Fang, Y., Qu, Z., Xu, F., Liu, K., & Lin, Y. (2024). Enhancing architectural education through artificial intelligence: A case study of an AI-assisted architectural programming and design course. *Buildings, 14*(6), Article 1613. https://doi.org/10.3390/buildings14061613

Kohir, V. (2020, June 25). *Calculating output dimensions in a CNN for convolution and pooling layers with Keras*. Medium. https://kvirajdatt.medium.com/calculating-output-dimensions-in-a-cnn-for-convolution-and-pooling-layers-with-keras-682960c73870

Kokorin, O. (2023, September 5). *7 steps to prepare a dataset for an image-based AI project*. HackerNoon. https://hackernoon.com/7-steps-to-prepare-a-dataset-for-an-image-based-ai-project

LearnOpenCV. (2023, January 18). *Convolutional neural networks (CNN): A complete guide*. LearnOpenCV. https://learnopencv.com/understanding-convolutional-neural-networks-cnn

Li, C., Zhang, T., Du, X., Zhang, Y., & Xie, H. (2025). Generative AI models for different steps in architectural design: A literature review. *Frontiers of Architectural Research*. https://doi.org/10.1016/j.foar.2024.10.001

Mohit. (2025, May 1). *Image processing using CNN: A beginner's guide.* Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/06/image-processing-using-cnn-a-beginners-guide/

Nasir, O. (2024, October 31). *AI education in architecture: Revolutionising learning*. Parametric Architecture. https://parametric-architecture.com/ai-education-in-architecture

Prasanth, A., Vadakkan, D. J., Surendran, P., & Thomas, B. (2023). Role of artificial intelligence and business decision making. *International Journal of Advanced Computer Science and Applications, 14*(6). https://doi.org/10.14569/IJACSA.2023.01406103

Purwadhika. (2024, March 21). *Neural network implementation for image classification using CNN.* Purwadhika. https://purwadhika.com/blog/neural-network-implementation-for-image-classification-using-cnn

Sai, U. (2020, August 24). *How to create a new custom dataset from images.* Towards AI. https://pub.towardsai.net/how-to-create-a-new-custom-dataset-from-images-9b95977964ab

Stas. (2025, January 14). *AI in architecture rendering: Benefits and examples.* Pixready. https://www.pixready.com/blog/ai-in-architecture-rendering-benefits-and-examples

Stojanovski, T., Zhang, H., Peters, C., Frid, E., Lefosse, D., & Chattre, K. (2021, April). *Architecture, urban design and artificial intelligence (AI): Intersection of practices and approaches.* Society for Modeling & Simulation International (SCS). SimAUD 2021. https://www.diva-portal.org/smash/record.jsf?pid=diva2:1901303

TUM-DI-LAB. (2019). *Object localisation on building plans using machine learning techniques*. Munich Data Science Institute. https://www.mdsi.tum.de/en/di-lab/past-projects/unetiq-object-localization-on-building-plans-using-machine-learning-techniques

20

AI IMAGE GENERATION AND ITS POSSIBLE CONTRIBUTIONS IN ARCHITECTURAL LANGUAGE

DOI: 10.37199/c41000320

Andia VLLAMASI
POLIS University (Tirana, Albania)
andia_vllamasi@universitetipolis.edu.al
ORCID 0009-0007-1730-4130

Skender LUARASI
POLIS University (Tirana, Albania)
skender_luarasi@universitetipolis.edu.al
ORCID 0000-0002-0967-2482

Tamara LUARASI
POLIS University (Tirana, Albania)
tamara_luarasi@universitetipolis.edu.al
ORCID 0009-0002-7449-5491

Abstract

AI, and specifically Machine Learning methodologies, can help the architect's imagination in the design process or the urban planner to develop ideas about urban planning. This paper focuses on generative ML methodologies that generate images from input datasets. The machine is trained to create patterns that help generate new images.

In all of these methodologies, the Generative Adversarial Networks (GANs) model is used, with specific details for each case. The idea is to highlight their specifics, both in concept and in implementation, and to test different metrics for assessing the accuracy of the generative process.

The input datasets are facade images, their sketches, or the combinations of both, and as a result, new images can be generated. The machine learning techniques are used to help us interpret architectural historical concepts, such as the relationship between the natural and the customary. Traditionally, the natural is represented by geometry, while the customary has stood for inherited stylistic languages.

Keywords*:* Design, pattern, generative, algorithm, architectural

I. INTRODUCTION

The development of generative artificial intelligence technologies has opened new horizons in the way architectural forms are conceived and presented (Li, 2024). Formal generation is being revolutionised by the generation of realistic images using sophisticated models like Generative Adversarial Networks (GANs) and other machine learning approaches. The AI-generated images undoubtedly daunt one. By plugging a bunch of natural keywords in the command prompt of AI models like Midjourney, DALL-E, or Stable Diffusion, a myriad of otherworldly images emerge as if out of thin air. The AI-generated images are outputs of neural algorithms "diving" deep into large datasets of 'observations' and intelligently reading and synthesising them (Chaillou, 2022, p. 65). This approach not only enables the formal analysis of complex visual data and the visualisation of projects before physical realisation, but also introduces new, previously unthought-of or unimagined images. This has led both scientists and AI apologists to claim that the latter can generate semantic content 'entirely autonomously' from humans (del Campo, 2022, p. 142). The question, of course, is how we - the designers, in general, can make sense of such output.

AI describes the statistical rather than the epistemological being of a phenomenon. It 'learns' to recognise images by categorising and labelling pixels, or regions in an image, into several groups' (Parente et al, 2023, p. 2). A large amount of information is processed through a series of computational layers comprised of 'neural' parameters or 'dials,' each of which 'filters' the input and 'feeds' it the following layer. Changing these dials will change the probability that the network model gives for the next informational pattern on a given input (IBM Think Topics). Unlike parametric modelling, where the user formulates the parameters, in AI it is the model itself that 'learns' them, and the user guides the general direction of learning through 'high-level settings' or 'hyperparameters' (Cahillou, 2022, p. 65). Out of this 'semantic training' AI 'learns' not only to recognize but also predict and generate new images.

Matias Del Campo has likened AI's ability to learn patterns to the architect's analysis of 'a building's style. The architect would break down the building into its basic components, looking for features such as columns, arches, curves, colours, materials that are associated with specific time periods, school of thoughts, or that can differentiate the image content in a *unique* and *meaningful way*' (del Campo, 2022, 72). Del Campo's claim, however, conflates terms that, historically, have stood for different categories. For example, the curve, as a geometrical entity, would fall under the 'natural,' while the columns and arches, as semantic or stylistic components, would fall under the 'customary.' The aim of such categorization was precisely to match and calibrate formal and content, generative epistemic systems (which are always technological) with meaning.
What is critical in such generative large language models, however, is precisely the very im-possibility of such a distinction. AI tends to erase such a distinction by being used like a black box

that does not account for the architect's intentions and overflows the latter with infinite patterns. This paper aims to use machine learning in such a way as to revitalize and reconceptualize the distinction between form and content, syntax and semantics, the natural and customary. By analysing three GAN extensions, this paper aims to explore the potential of AI in creating visual images of facades of buildings through a generator, on the one hand, and then in matching these images with geometrical sketches through a pixel-to pixel algorithm.

This material focuses on three generative machine learning methodologies that have as their objective the generation of images based on a dataset that is used to train the machine learning algorithm. The machine is trained, creating patterns that help in the generation of new images.

These methodologies consist of:

- Generating new views as a result of training the algorithm with facade views and their details
- Generating a facade view or producing an object detail as a result of training machine learning algorithm with real sketch-photo pairs
- Generating new views similar to those that the machine is trained. In this case algorithm is trained on facade views that are classified and this classification helps in the generation process.

In all three of these methodologies, the Generative Adversarial Networks (GANs) model is used with some specifics for each case. The idea is simply to highlight their specifics, in concept and implementation, and to test some different metrics in assessing the accuracy of the generative process.

II. LITERATURE REVIEW

II.1 Related Work

A Generative Adversarial Network, or GAN, was designed by Ian Goodfellow and his colleagues in 2014, and is a machine learning approach to generative modelling that has served as a basis for many other generative modelling approaches. The original GAN is often called a Vanilla GAN. Its two main components are neural networks: a generator G and a discriminator D. D takes images generated by G as input, along with real images, and is trained to distinguish the generated images from the real ones. G is trained to synthesise images as close to reality as possible. (Goodfellow, 2014)

While the GAN model has a wide range of applications, other generation models are more focused on image generation, which is directly related to the architectural design process. The following models are as follows.

The Progressive Growing GAN model is an extension of the GAN model adapted for large and high-quality images. Its idea is to add new generator and discriminator layers during the training process, going from a low resolution to a higher one and thus accelerating the execution time. (Karras, 2018)

An alternative of GAN model is proposed Kaneko and Harada (2020), which is called noise robust GANs (NR-GANs). In this alternative, the generator is trained with a noise input in addition to training a generator with a clean image input, which solves the problem when there is a lack of information in the input noise, such as the type of distribution, amount of noise, or noise-signal relationship.

While the GAN model uses a pair of images, one real and one synthesized to perform image generation, the CycleGAN (Zhu, 2017) model does not use a pair of images, which is often difficult to obtain. For this, this model uses two generators that pass, the first, an image from one domain to another and the second, returns the transformed image back by comparing it with the first in an attempt to get it as close as possible to the first. A model of a robust noise-generation generative adversarial network (NG-GAN) is proposed by Hossain and Lee (2023). The idea of the proposed model is how to deal with old images, how to obtain noise distribution of the degraded old images inspired by the CycleGAN model.

Deep Convolutional Generative Adversarial Network (DCGAN) were introduced by Alec Radford & Luke Metz & Soumith Chintala (2016) , where an architectural topology of Convolutional GANs were proposed. It has been utilized by artists and designers to create unique and visually captivating artworks, logos, and graphics.

Conditional GAN (cGAN) allows us to condition the network with additional information such as class labels (Mirza, 2014). It means that beside the images, some associative labels are the inputs of this model (facades with roof, facades without roofs).

In their study, Isola et al. (2017) presented Image-to-Image Translation with Conditional Adversarial Networks (PixToPix). This model It is a special case of the conditional GAN where the label is image. The two networks learn the mapping of two images, but also a loss function to train this mapping.

Based on a study, Convolutional Generative Adversarial Network (CGAN-Pix2pix) is the method used to provide a quality assessment of dehazing images.

II.2 GAN Overview

There are two models included in the proposed GAN model, that cooperate and oppose each other, and each is a so-called a multilayer perceptron. (Goodfellow, 2014)

Multilayer perceptron (MLP) are artificial neural networks (X. Li, 2021). Such networks consist of several fully connected layers that transform input data from one dimension to another. There is an input layer, one or more hidden layers, and the output layer.

The figure below illustrates such a network, where each node (or neuron) in the input layer represents a feature of the input data, in our case of an image, and has a full connection between a neuron of one layer and all neurons of the previous layer.

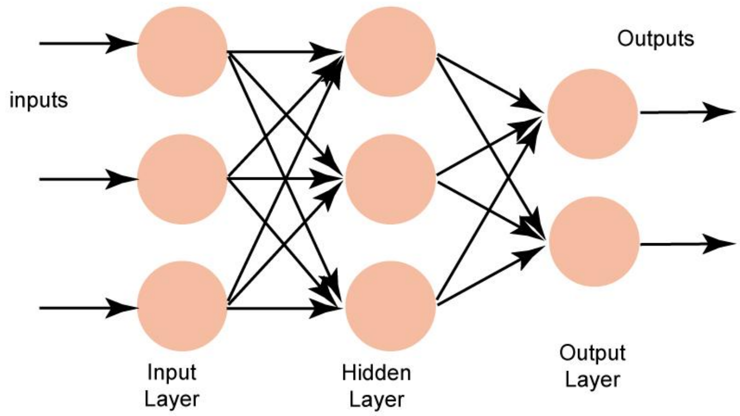


Figure 1. Fully connected neural network
Source: Authors processing

It is important to mention some key point in the paper of the authors. Many authors have used and commented on this paper. Between them in this work (Arjovsky, 2017) is a detailed interpretation of these key points is presented. It is necessary to mention these key points, to have a better understanding of the implementation of GAN model.

According to the authors of the GAN model (Goodfellow, 2014), the GAN model consists of two multi-layers neural networks: Discriminator and Generator.

As input data for the generator, are used some fixed-length vectors of randomly selected values representing the qualities of an image, referred to as noise, with values chosen within the possible bounds that suggest concrete data for their qualities, or sampling these values from a probabilistic distribution of the qualities of the data that is taken a priori. As a result of this input, the generator produces an image that is called a fake image.: noise(z), parameters(g)-> image: $G(z, param_g)$, and this image that will serve as an input for the discriminator.

The discriminator, on the other hand, is also a multi-layer neural network $D(x, param_d)$ that takes as input real images and those generated by the generator, and produces a single numerical output that is the labelling of these images with predicted labels.

The labelling of the images, coming from real data, should be a probability $D(x)$ as close to 1 as possible, and the labeling of a generated image $G(z)$ should be a probability as small as possible D(G(z)), which translates into a value as high as possible of $1 - D(G(z))$. Therefore, the goal of the discriminator is to provide a value as high as possible to $D(x) * (1 - D(G(z)))$ or to the expression $log(D(x) + log(1 - D(G(z)))$.

In order for the discriminator to become capable of this, it needs to be trained, which means improving (or learning) the parameters of the network. Thus, we have an autonomous training of the discriminator that is carried out through the backpropagation process, where in each step of this process we have an improvement of the network parameters that results in increasing the discriminator's ability to avoid incorrect evaluations.

Backpropagation is conditioned by the mistakes that the discriminator makes and they are calculated. There are two types of errors, discriminator loss and generator loss. The first is related to the error against real images, so the image is real and its labelling is close to 0, while the second is related to the error against generated images, meaning the image is generated and its labelling is close to 1. The sum of these errors, which is based on the differences between the expected labels and those that occur, forces the discriminator to repeat the process from the beginning now with modified parameters.

On the other hand, the generator is interested in generating images as close to reality as possible, thus minimizing the expression $D(G(z)$ or $maximizing$ $(1 - D\big(G(z)\big)$ or $log(1 - D\big(G(z)\big)$, and this also requires training of the generator, which is carried out through a backpropagation process, a process that modifies the parameters of the neural network of the generator. Everything said is illustrated with the figure 2.

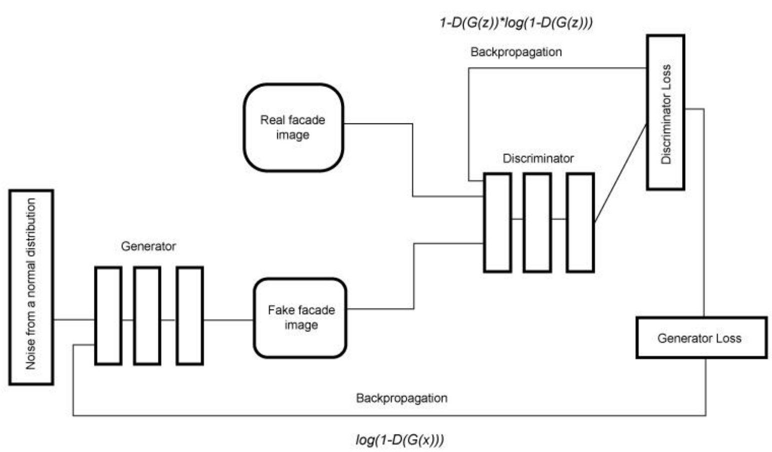


Figure 2. Generative Adversarial Networks (GANs) Model
Source: Authors processing

The achievement of these two objectives by the discriminator and generator has been expressed by the authors finding of:

$$V(D,G) = E_x(x)[logD(x)] + E_z\left[log\ log\ \left(1 - D\big(G(z)\big)\right)\right] \quad (1)$$

Where$E_x, dhe\ E_z$ represent the mathematical expectations of the respective expressions.

According to the author (Goodfellow) of the model, the modifications of the parameters of the discriminator and the generator are made through stochastic gradient, which is an iterative method of modifying the parameters like it is represented in the Figure 3.

Algorithm 1 Minibatch stochastic gradient descent training of generative adversarial nets. The number of steps to apply to the discriminator, k, is a hyperparameter. We used $k = 1$, the least expensive option, in our experiments.

for number of training iterations **do**
  for k steps **do**
  - Sample minibatch of m noise samples $\{z^{(1)}, \dots, z^{(m)}\}$ from noise prior $p_g(z)$.
  - Sample minibatch of m examples $\{x^{(1)}, \dots, x^{(m)}\}$ from data generating distribution $p_{data}(x)$.
  - Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(x^{(i)}\right) + \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right)\right].$$

  end for
  - Sample minibatch of m noise samples $\{z^{(1)}, \dots, z^{(m)}\}$ from noise prior $p_g(z)$.
  - Update the generator by descending its stochastic gradient:

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right).$$

end for
The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.

Figure 3. Algorithm of the iterative modifications of Discriminator and Generator
Source: Authors processing

In this figure an iterative independent cycle of parameter modifications of the discriminator is observed, and this cycle is included in another cycle where the generator is modified.
For fixed G, the $V(D,G)$ maximum will be achieved when
$D^*(x) = realdataDistribution(x) / [realdataDistribution(x) + generateddataDistribution(x)]$ *(2)*
The global minimum in (1) is achieved when the total error calculated (when the real images are predicted with low probabilities and when generated images are predicted with high probabilities) gets its smallest value.
To measure this error it is used **Jensen-Shannon (JS) divergence**, which measures the differences between two distributions of real data and generated data.
Jensen-Shannon (JS) Divergence Formula is:

$$JS(D||G) = 1/2 * KL(P||\,M) + 1/2 * KL(Q||M) \ (3)$$

where M is the average distribution:

$$M = (D + G)/2$$

and the **Kullback-Leibler (KL) divergence** is:

$$KL(D||G) = \sum D(x)\, log[D(x)/G(x)]$$

D(x) is the distribution of the predictions of the real images and G(x) the distribution of predictions of the generated images. Using this method, it is concluded that the global minimum is achieved when two distributions are identical.

III. METODOLOGY AND DATA

III.1 Implementation of the GAN model

III.1.1 The internal structure of the generator and discriminator and its implementation

A fully connected neural network is composed by some layers, where each layer is the input of the next layer. A layer is composed by neurons and every neuron of a layer is connected with each neuron of the its input layer like the figure 1 shows.

Each neuron is in fact a non-linear function, called perceptron neuron as well, and performs two steps:

- Finds the sum of the outputs of previous layer multiplied by the weights of previous layer, that can be presented as: $z_j = \sum_i \quad w_i \cdot x_i + b_j$
- Applies the activation function, which is a non-linear function:

The Figure 4 shows the layer created after the input layer by these two steps.

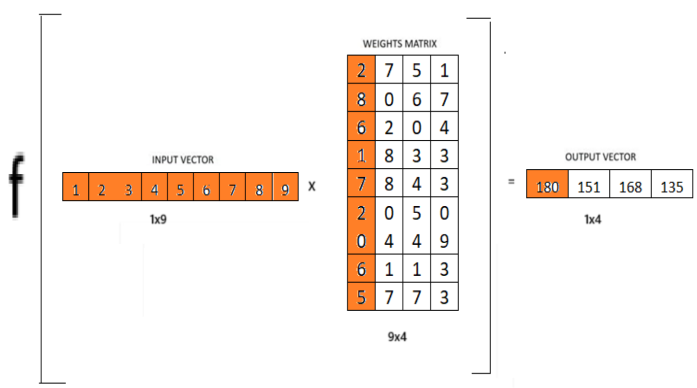


Figure 4. The steps of perceptron neuron (reprinted from *Fully Connected Layer vs. Convolutional Layer: Explained*, Built In, n.d., https://built-in.com/…).

Source: Author's processing

Where the activation function f can be one of the functions: Sigmoid, Tanh, ReLU (Rectified Linear Unit) function, etc. (Shiv Ram Dubey, 2021). The implementation code of a fully connected layer depends on the Python framework. For all the examples below, it is selected the TensorFlow framework of python and Keras. Layers library on it. In this framework the placeholders of the data are so called tensors. "In TensorFlow, tensors are the basic building blocks used to represent data.

A tensor can be thought of as a multi-dimensional array, similar to a matrix but with an arbitrary number of dimensions. Tensors can hold various data types, including integers, floating-point numbers, and strings." (Bader, 2009)

The implementation of fully connected layer for the GAN model would be:

Create a dense layer with 128 units

layer = tf.keras.layers.Dense(units=128, activation='relu'), and the generator and discriminator would be composed of some lines of this type.

Because there are generation models more focused on image generation than GAN model, some implementation details are provided related to some specifics of them in the following sections.

III.1.2 The implementation of GAN model extensions related to some specifics

The common characteristic of these modes is the use of convolutional layers in their model. In difference with the fully connected layer, in a convolutional layer, each neuron of a layer applies a convolutional operation to the input layer. The neuron uses a sliding window (called kernel), and represents the space occupied by this window in the input layer as a sum of a dot product between the values in the input layer and values in the window. In this type of layer (convolutional one), a neuron is not connected with all the neurons of the input layer. Figure 5 represent a convolutional operation of a neuron. The neuron with value 48, uses the dot multiplication of the space in orange colour of the input layer and the kernel to get this value. Output is called a feature map too.

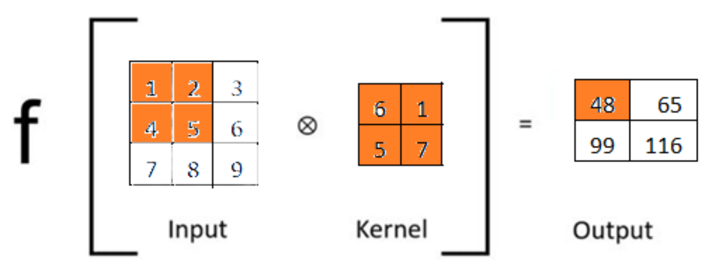


Figure 5. Convolutional operation (adapted from *Conditional Generative Adversarial Network*

Source:https://www.geeksforgeeks.org/conditional-generative-adversarial-network/).

Some key moments will be mentioned for the models DCGAN, CGAN and PixToPix:

- Model structure: Discriminator and Generator
- Definition of the discriminator loss and generator loss
- Training process of discriminator and generator

III.2 DCGAN model structure

DCGAN model generates new images. The generator starts with a noise, and the discriminator compares the generated image from the noise by the generator, and a real image. The process progresses like in GAN model and the result will be some new images.

- *Discriminator* structure in tensorflow keras.layers:

```
#inputs are colored images
def discriminator
 image = tf.keras.layers.Input(shape=(img_size, img_size, 3)) # [(None, 32, 32, 3)]
 i_output = tf.keras.layers.Conv2D(128,(3,3 ),strides=(2, 2), padding='same')(image)
                                        # (None, 16,18,128)
 i_output = tf.keras.layers.LeakyReLU(alpha=0.2)( i_output)   # (None, 16,18,128)
 i_output = tf.keras.layers.Conv2D(128, (3, 3),strides=(2,
2),padding='same')(i_output)
                                        # (None, 8,8,128)
 i_output = tf.keras.layers.LeakyReLU(alpha=0.2)(i_output)    # (None, 8,8,128)
 i_output = tf.keras.layers.Flatten()(i_output)          # (None, 1,8192)
 i_output = tf.keras.layers.Dropout(0.4)(i_output)         # (None,1,8192)
 i_output = tf.keras.layers.Dense(1, activation='sigmoid')(i_output)# (None,1)
 model = Model(image, i_output)
 return model
d_model = build_discriminator()
d_model.summary()
```

Some notes about discriminator structure:

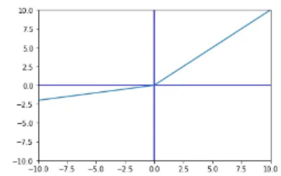
1. Input (shape= (img_size, img_size, 3): creates the tensor for the image with dimensions [None, 32, 32, 3] (here the image_size=32), None represent the size of data
2. Conv2D (128, (3,3), strides= (2, 2), padding='same') (image):
convolutional layer creates 128 features maps with dimensions (16, 16), or an output of dimensions (None, 16, 16, 128); the formula is applied

$$O = \lfloor \frac{n - f + 2 \cdot p}{s} \rfloor + 1$$

where, O: Output size, n: Input size (height or width) =32, f: kernel size =3, p: Padding =1, s: Stride =2

The second Conv2 produces the output of (None, 8, 8,128), based on the same formula, but now n: Input size (height or width) =16

3. LeakyReLU(alpha=0.2) (i_output)
It is applied the formula to avoid linearity
f(x) = alpha * x if x < 0
f(x) = x if x >= 0



4. Flatten () (i_output)
The feature maps are converted into a 1D array (1, 8192) where 8192=8x8x128

5. Dropout (0.4) (i_output)
The Dropout layer with rate=0.4 ensures that 40% of the neurons in the preceding layer are randomly deactivated during training to enhance the robustness of neural network model.

6. Dense (1, activation='sigmoid') (i_output)
This line provides one dimensional tensor and the activation function is applied. The mathematical formula for a sigmoid function is given by: f(x) = 1 / (1 + exp(-x)). It transforms numbers to values typically between 0 and 1.

- *Generator* structure in tensorflow keras.layers:

```
#Inputs are noises defined randomly from a normal distribution
    noise = tf.keras.layers.Input(shape=(noise_dim,))   # [(None, 100)], noise_dim=100
    n_nodes = 128 * 8 * 8
    g_output = tf.keras.layers.Dense(n_nodes)(noise)     # (None, 8192), 8192=128x8x8
    g_output = tf.keras.layers.LeakyReLU(alpha=0.2)( g_output)# (None, 8192)
    g_output = tf.keras.layers.Reshape((8, 8, 128))( g_output)   #(None, 8, 8, 128)
    g_output = tf.keras.layers.Conv2DTranspose(128,4,4),strides=(2,2),padding='same')
                                  (g_output) # (None, 16, 16, 128)
    g_output = tf.keras.layers.LeakyReLU(alpha=0.2)(gen)       #(None, 16, 16, 128)
    g_output = tf.keras.layers.Conv2DTranspose(128,(4,4),strides=(2,2), padding='same')
                                  (g_output) #(None, 32, 32, 128)
    n_output = tf.keras.layers.LeakyReLU(alpha=0.2)( g_output)  #(None, 32, 32, 128)
    g_output = tf.keras.layers.Conv2D(3,(8,8),activation='tanh',padding='same')(g_output)
                                       # (None, 32, 32, 3)
    model = Model (noise, g_output)
```

Some notes about generator structure:

1. Input(shape=(noise_dim,)
   creates the tensor for a fixed vector with size noise_dim (here noise_dim=100)
2. n_nodes = 128 * 8 * 8
   defines the dimension of the generated image at the starting point
3. Dense(n_nodes)(noise) provides one dimensional tensor for the noise

4. Reshape ((8, 8, 128))
   Reshape the tensor in correspondence with the tensors used by discriminator
5. Conv2DTranspose (128,4,4), strides= (2,2), padding='same')(l_output)

The need for transposed convolutions generally arises from the need to use a transformation going in the opposite direction of a normal convolution, i.e., from something that has the shape of the output of some convolution to something that has the shape of its input while maintaining a connectivity pattern that is compatible with said convolution. (Odena, 2016) he following formula calculates the output dimension in this layer based on some parameters mentioned below: (Contributors, In PyTorch 2.7. , 2025)

$$n_{out} = (n_{in} - 1)s - 2p + (f - 1) + p + 1 = (n_{in} - 1)s - p + f$$
$$where$$

- n_{out} *is the output width or height (same in both dimensions)*
- n_{in} *is the input width or height (same in both dimensions)*
- s *is the stride (same in both dimensions)*
- p *is the padding (same in both dimensions)*
- f *is the filter dimensions (same in both dimensions)*

Padding is calculated: (Liu, 2018)

$$if \;\; n_{in}\%s == 0$$
$$p = max(f - s, 0)$$
$$else$$
$$p=max(f-(n_{in}\%S),0$$

for the above case:

n_{in}%S=8%2=0 then p=4-2=2

And $n_{out} = (n_{in} - 1)s - p + f = (8 - 1) * 2 - 2 + 4 = 16$

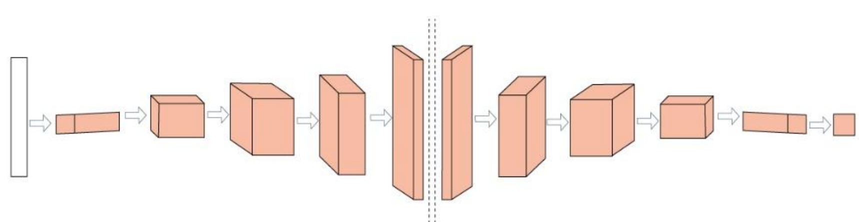


Figure 6. Illustration of GAN structure
Source: Authors processing

III.3 Definition of the discriminator loss and generator loss

An important element of the GAN model is the error calculation during the training process. These errors dictate the updating of the parameters in each step and backpropagation of this process.

- *JS Divergence*

As it is mentioned above this approach finds an "average" distribution between distribution of the discriminator predictions and generator prediction, and uses this distribution as a reference distribution against the two first, finding the differences with them like it is shown in the figure.

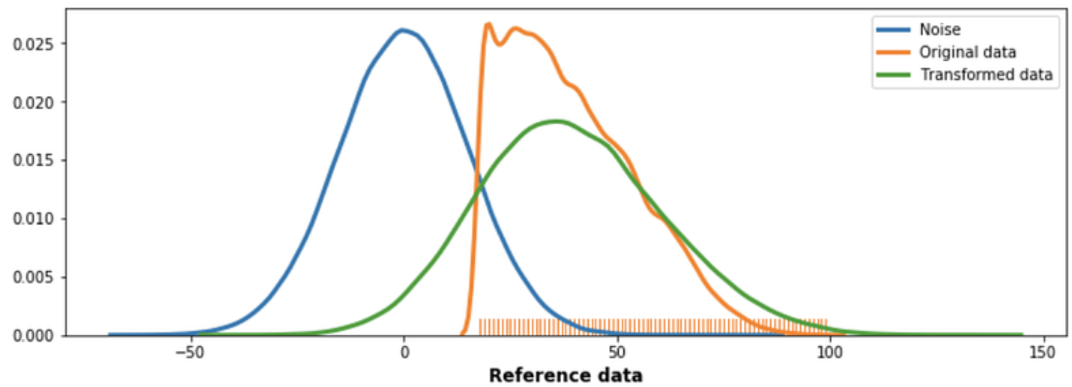


Figure 7. Jensen–Shannon divergence - Search Images
Source: Authors processing

The implementation of this approach is:

```
def jensen_shannon_divergence(d, g):
  # The values outside (0,1] are eliminated
  d = tf.clip_by_value(p, 1e-10, 1.0)
  g = tf.clip_by_value(q, 1e-10, 1.0)
  # the "average" distribution
  m = 0.5 * (d + g)
  # Compute the KL divergence for each part
  kl_dm = tf.reduce_sum(p * tf.math.log(d / m), axis=-1)
  kl_gm = tf.reduce_sum(q * tf.math.log(g / m), axis=-1)

  # JS divergence as average of the two Kullback-Leibler divergences
  js_divergence = 0.5 * (kl_dm + kl_gm)

  return js_divergence

def discriminator_loss_js(real, fake):
  real_loss = jensen_shannon_divergence(tf.ones_like(real), real)
  fake_loss = jensen_shannon_divergence(tf.zeros_like(fake), fake)
  total_loss = real_loss + fake_loss
  return total_loss
```

```
def generator_loss_js(preds):
  return jensen_shannon_divergence(tf.ones_like(preds), preds)
```

- *Binary cross-entropy*

It is considered the binary cross-entropy approach too for the error calculation, which calculates the error by the formula as follows:

$$BCE = -\frac{1}{N}\sum_{i=1}^{n} \quad [y_i log(p_i) + (1 - y_i)log(1 - p_i)] \quad (4)$$

where: N is the data size, y_i is the actual binary label (0 or 1) of the i^{th} real image, p_i is the predicted probability of the i^{th} real images. It quantifies the uncertainty of the prediction distribution measuring the average of the differences between the real labelling of the images and the predicted probabilities for these images to be real or fake. (Goodfellow I. B., 2016).

The implementation of this formula in Python is:

```
bce_loss = tf.keras.losses.BinaryCrossentropy()
def discriminator_loss(real, fake): #real->labeling of a real image,
                          #fake->labeling of a generated image
  real_loss = bce_loss(tf.ones_like(real), real)
                #error against real image
  fake_loss = bce_loss(tf.zeros_like(fake), fake)
                #error against generated image
  total_loss = real_loss + fake_loss #sum of both errors
  return total_loss

def generator_loss(preds):
  return bce_loss(tf.ones_like(preds), preds)
             #sum of differences (1-prediction)
```

This formula is based on this logic:

We have errors in discriminator if it labels with 0 the real images, and with 1 the generated images. Then the expression $p_i^{y_i} * (1 - p_i)^{1-y_i}$ represents how close or far are the predictions from the correct labeling.

Truly,

If an image is real, $y_i = 1$, the expression $p_i^{y_i} * (1 - p_i)^{1-y_i} = p_i$

If an image is generated, $y_i = 0$, the expression $p_i^{y_i} * (1 - p_i)^{1-y_i} = (1 - p_i)$

Instead of the above expression the logarithm of it with negative sign is considered:

$$-(y_i log(p_i) + (1 - y_i) log(1 - p_i))$$

Therefore:

If an image is real, $y_i = 1$, the expression $-(y_i log(p_i) + (1 - y_i) log(1 - p_i)) = -log(p_i)$

If an image is generated, $y_i = 0$, the expression $-(y_i log(p_i) + (1 - y_i) log(1 - p_i)) = -log(1 - p_i)$

In both cases if the prediction is far from the correct label (close to 0 for y=1 and close to 1 for y=0), the value of the expression is high and when the prediction is close to the correct label the value of the expression is low. The figure 5 represent the graph in both cases.

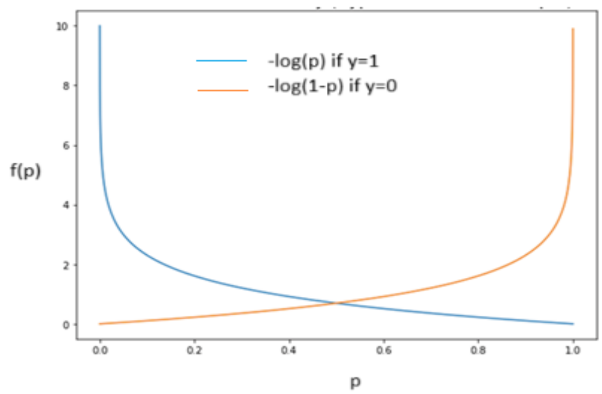


Figure 8. the Graph of $-(y_i log(p_i) + (1-y_i) log(1-p_i))$ for y=0 and y=1, the blue line represents the distribution of the real images and the orange line represents the distribution of the generated images, the optimal value is achieved when both distributions are equal. Source: Binary cross-entropy graph - Search Images

The value of the expression (4), which in fact represents the average of the errors for each image, dictates the backpropagation of the process, firstly for the discriminator and after for the generator to achieve a discriminator loss and generator loss as small as possible until the optimal value for the prediction is achieved, as shown in Figure 5.

Training process of the discriminator and generator code:

```
def train_step(dataset):
  real_images= dataset #real images
  # Sample random noise represented by a vector.
  random_vectors = tf.random.normal(shape=(batch_size, noise_dim))
  generated_images = g_model(random_vectors)

  # Train the discriminator.
  with tf.GradientTape() as tape:
    pred_fake = d_model(generated_images)
    pred_real = d_model(real_images)
    d_loss = discriminator_loss(pred_real, pred_fake)

  #gradient approach to modify the parameters(trainable_variables)
  grads = tape.gradient(d_loss, d_model.trainable_variables)
  d_optimizer.apply_gradients(zip(grads, d_model.trainable_variables))
```

```
# Sample random noise.
random_vectors = tf.random.normal(shape=(batch_size, noise_dim))

# Train the generator
with tf.GradientTape() as tape:
   fake_images = g_model(random_vectors)
   predictions = d_model(fake_images)
   g_loss = generator_loss(predictions)

#gradient approach to modify the parameters(trainable_variables)
grads = tape.gradient(g_loss, g_model.trainable_variables)
g_optimizer.apply_gradients(zip(grads, g_model.trainable_variables))

return d_loss, g_loss
```

III.4 Conditional GAN (CGAN) and the differences from the GAN model

The GAN model described above it is named as unconditioned GAN, because there is not control over the data generated. In conditional GANs, additional information is provided, for example in preliminarily the images provided could be associated by some labels.

Often the input data are images that can be preliminary labelled, and this helps in the generation process in the speed of it and in the quality of the image generated, when it is needed the generated images to keep their categorization.

The figure 9 represent such a model. The images are labelled.

The formula (1) based on the first definition of Conditional GANs by Mirza and Osindero (Osindero, 2014) can be presented as:

$$\min_{G} \max_{D} V(D, G) = \mathbb{E}_{x}[\log D(x/y)] + \mathbb{E}_{z}[\log(1 - D(G(z/y)))] \quad (5)$$

As this research describes this model, in GAN model the data are generated randomly from a distribution chosen a priory, but in CGAN the generator creates images by considering *specific conditions.*

In the example considered in this paper, as dataset is considered a collection of 400 facades and for the GAN model they are used without any label. For the CGAN model the facades are classified in two classes: facades with roofs and facades without roofs.

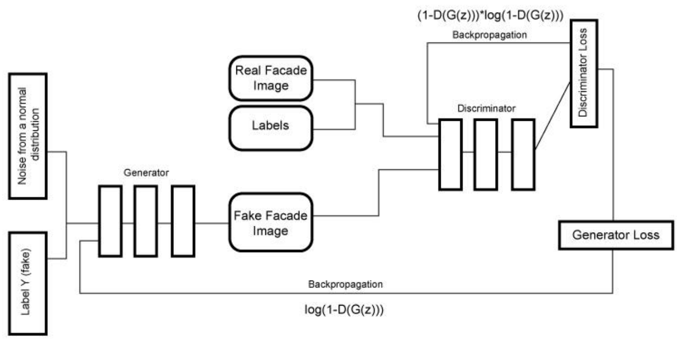


Figure 9. The CGAN model
Source: Authors processing

This difference takes place discriminator model and generator model, where ae add some lines to include labels. To keep everything the same lake for the GAN model:

- *In discriminator:*

```
# label input
label = tf.keras.layers.Input(shape=(1,))
#create a vector of size 50
l_output = tf.keras.layers.Embedding(n_class, 50)(label) #n_class=2
n_nodes = img_size * img_size
l_output = tf.keras.layers.Dense(n_nodes)( l_output)
l_output = tf.keras.layers.Reshape((img_size, img_size, 1))( l_output)
#image input
image = tf.keras.layers.Input(shape=(img_size, img_size, 3))
                                  #this line was before
i_output = tf.keras.layers.Concatenate()([image, l_output])
 . . .
model = Model([image, label], i_output)
```

- *In generator:*

```
 # label input
label = tf.keras.layers.Input(shape=(1,))
# create an embedding layer for 2 classes in the form of a vector
# of size 50
l_output = tf.keras.layers.Embedding(n_class, 50)(label)
n_nodes = 8 * 8
l_output = tf.keras.layers.Dense(n_nodes)( l_output)
# reshape the layer
l_output = tf.keras.layers.Reshape((8, 8, 1))( l_output)
# image generator input
noise = tf.keras.layers.Input(shape=(noise_dim,)) #this line was before
```

```
. . .
g_output = tf.keras.layers.Concatenate()([noise, l_output])
. . .
model = Model([noise, label], g_output)
```

III.5 PixelToPixel Model

The Image generation model Pix2Pix is as well a conditional GAN (cGAN). Phillip Isola developed it 2017. The characteristic of this model is it uses a pair of images as input, for example a sketch and the corresponding facade and a new image is generated. The architect is interested to generate a facade, then the sketch is the input of the generator. If the architect wants a generated sketch than the facade will be the input of the generator.
These specifics of the pixel-to-pixel model are presented below and illustrated with figures and code presented in this paper (Isola, 2017).

III.5.1 Input data

A pair of images: a sketch or a blueprint of an object, and a real image of this object are the **input images** for the discriminator, for example the sketch of a facade and a real photo of the facade

For the generator the sketch of the object is the **Input image**, like we see in the figure. It is this sketch that is transformed to generate a new image and it is the discriminator that controls this process until it decides to accept the generated image as a real one.

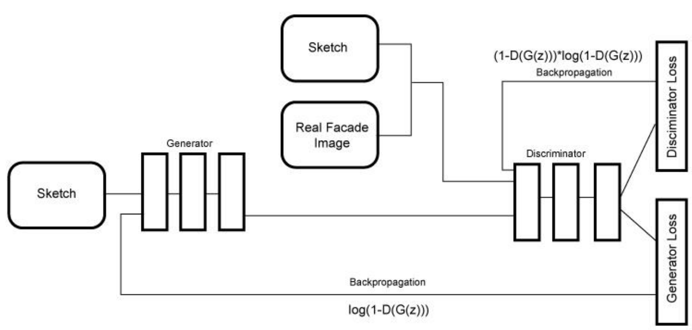


Figure 10. PixelToPixel Model
Source: Authors processing

III.5.2 Discriminator and Generator Structures

- *Generator*

To generate something new from a sketch, so to translate an image to another one, some transformations are performed to the input image by the generator using some layers.

The neural network architecture called UNET is used for the generator to generate new images from input images. The generator with this architecture uses some layers (encoder layers) to reduce the image into a smaller one and some layers (decoder layers) to enlarge it again but during the second part of transformation the image is mixed with the image of the parallel step of the first part of the process, like it is shown in the figure:

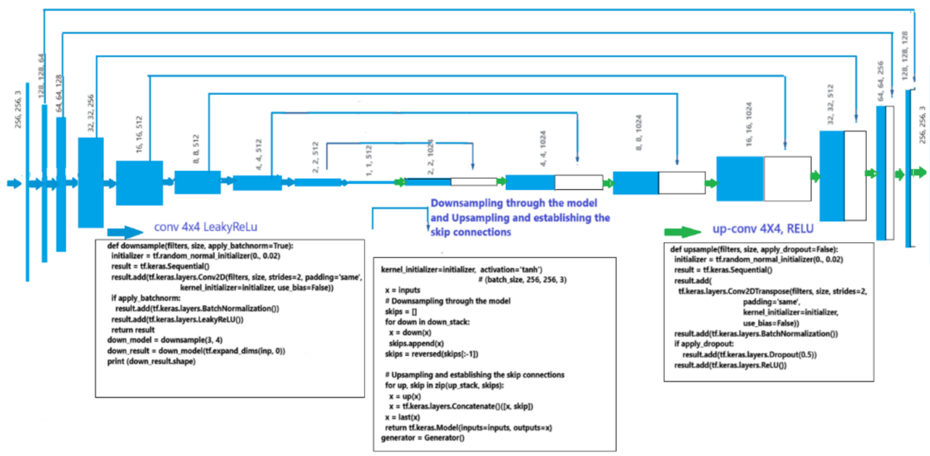


Figure 11. Generator Structure

Source: Authors processing

- *Discriminator*

As well as in the above models GAN and CGAN, the task of the discriminator is to distinguish between a real image and a generated one. The Discriminator specific in Pixel-to-Pixel is that this comparison is done for patches of the images and not for the entirely image. Therefore, it is called a patchGAN classifier. The inputs are the sketch (real or a generated one) image of dimensions 256x256x3, and the facade photo image, the two images (one sketch and facade photo) are concatenate before going in the other layers.

```
inp = tf.keras.layers.Input(shape=[256, 256, 3], name='input_image')
target = tf.keras.layers.Input(shape=[256, 256, 3], name='target_image')
x = tf.keras.layers.concatenate([inp, target])
```

These layers are the usual layers that a cGAN discriminator like Conv2D layer, BatchNormalization, LeakyReLU, etc. to achieve to a patch of a shape 30, 30, 1. This provide a better analysis of the image.

- *Training Process*

Regarding the training process, in this model the discriminator and generator are trained simultaneously, and, as in the other models, they consider different object as goal of training. Discriminator training aims to distinguish as much as possible the errors in the classification of generated images as real ones and of real images as generated ones, meanwhile the generator training aims to diminish the difference between the generated and real images. The simultaneously training is presented with a loop like below:

```
with tf.GradientTape() as gen_tape, tf.GradientTape() as disc_tape:
  gen_output = generator(input_image, training=True)
  disc_real_output = discriminator([input_image, target], training=True)
  disc_generated_output = discriminator([input_image, gen_output],
                                        training=True)
  gen_total_loss, gen_gan_loss, gen_l1_loss =
      generator_loss(disc_generated_output, gen_output, target)
  disc_loss = discriminator_loss(disc_real_output, disc_generated_output)
```

IV. RESEARCH RESULTS

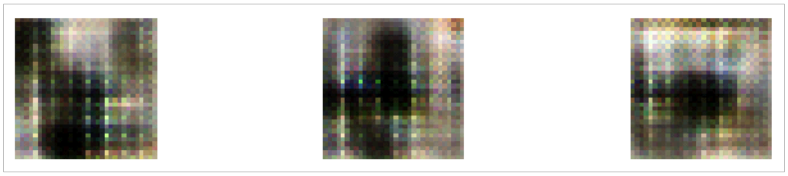
The dataset used for all the cases is dataset of 400 facades and sketches taken from www.kaggle.com/datasets. They are considered as two separated datasets. A series of experiments are done, considering different values for the parameters batch size (number of samples processed together in one forward and backward pass) and epoch number (The number of times the entire training dataset processed through the model).

In our practice there are used different structures of the generator and discriminator.

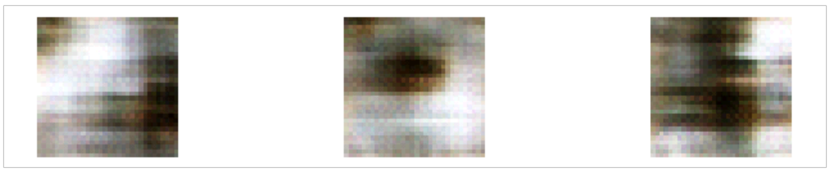
1. The first five results, that means, the result for batch size=16 and epoch=50,100,200,500,1000, are taken using a simple generator and discriminator, where:
   - generator composed by layers: Input, Dense, Reshape, LeakyReLU, Conv2DTranspose, LeakyReLU Conv2DTranspose, LeakyReLU , Conv2D
   - discriminator composed by layers: Input, Conv2D, LeakyReLU, Conv2D, LeakyReLU, Flatten, Dropout, Dense

It is shown that the improvement of the quality needs a big epoch number, which is 1000. The graphics of generator and discriminator loss also show that in the case of 1000 epoch there is a closer relation between them.

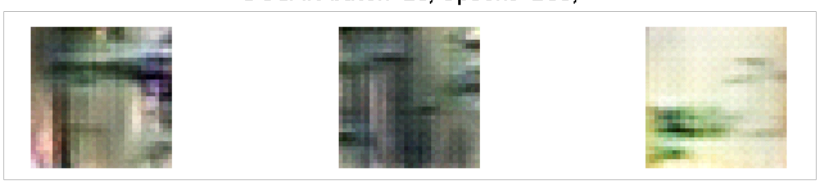
DCGAN batch=16, epochs=50, kernel initialization by uniform distribution



DCGAN batch=16, epochs=100,

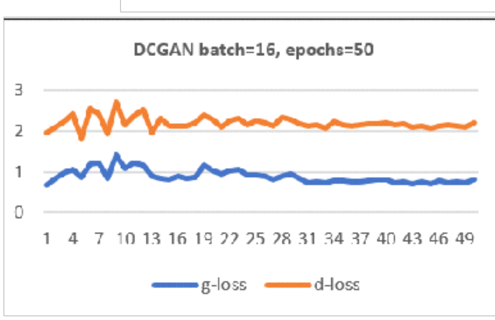


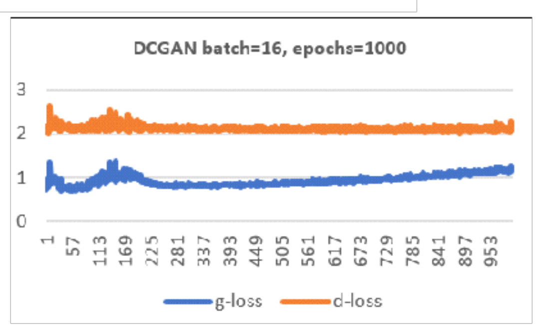
DCGAN batch=16, epochs=200,



DCGAN batch=16, epochs=500,

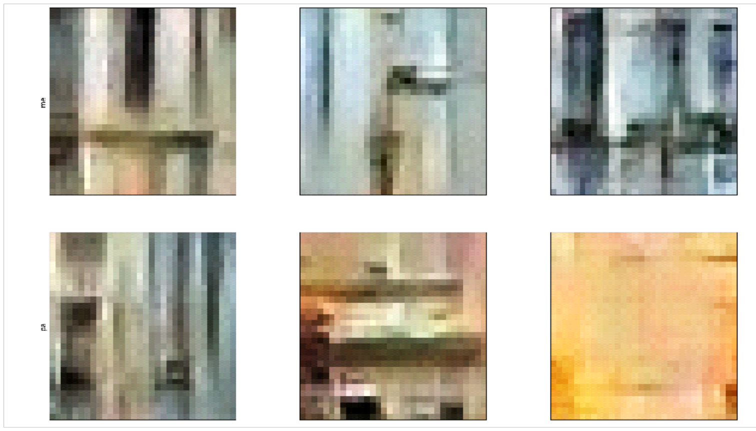
DCGAN batch=16, epochs=1000, kernel initialization by uniform distribution





2. The following is the result of the conditional GAN model but with the structure of the generator and discriminator. The images are labelled preliminary as facades with roof and without roof.

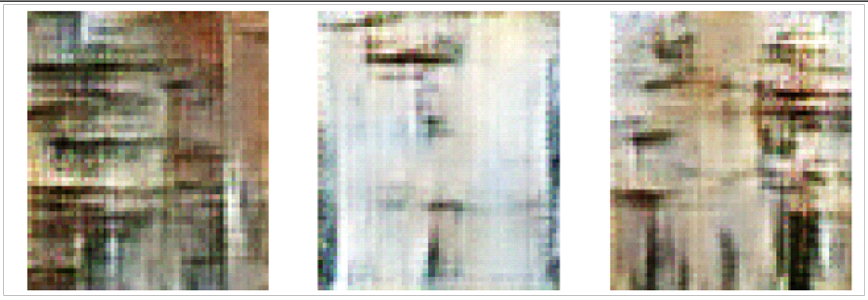
CGAN batch=16, epochs=1000, kernel initialization by normal distribution



3. The two following results shows correspond to some richer structures of generator and discriminator where:
   - generator composed by layers: Input, Dense, Reshape, LeakyReLU , Conv2DTranspose, LeakyReLU Conv2DTranspose, LeakyReLU , Conv2DTranspose, LeakyReLU , Conv2D
   - discriminator composed by layers: Input, Conv2D, LeakyReLU, Conv2D, LeakyReLU, Dropout, Conv2D, LeakyReLU, Dropout, Conv2D, LeakyReLU, Dropout, Conv2D, Flatten, Dense

It is shown that the clarity of the images needs a smaller number of epochs. For the initialization of the network parameters (kernel initialization), there are used two ways: by uniform distribution, and normal distribution.

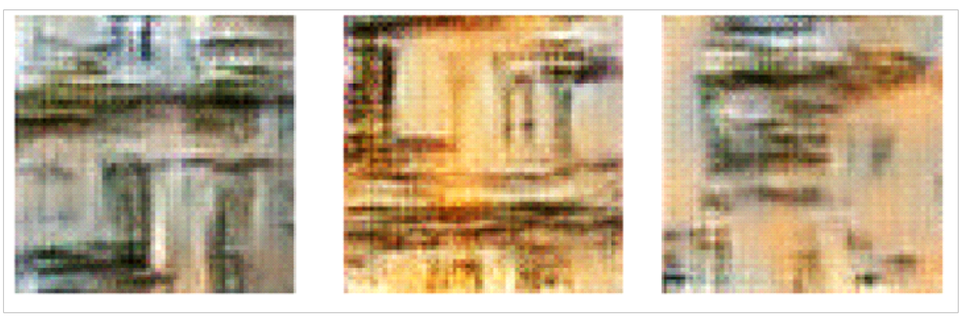
DCGAN batch=16, epochs=100, kernel initialization by uniform distribution

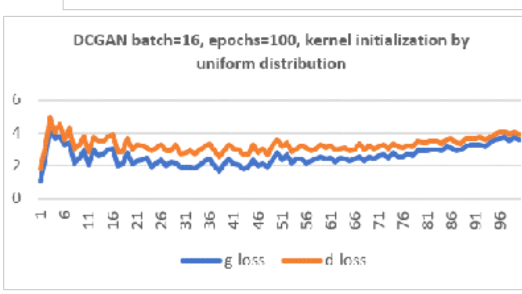


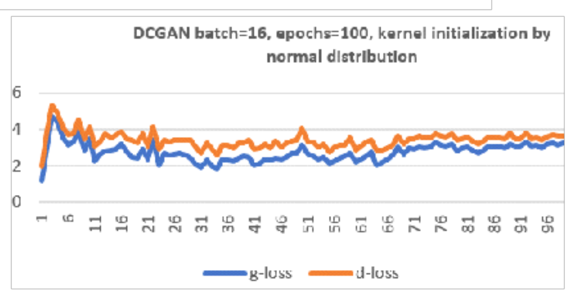
DCGAN batch=16, epochs=100, kernel initialization by normal distribution

DCGAN batch=16, epochs=100, kernel initialization by uniform distribution

DCGAN batch=16, epochs=100, kernel initialization by normal distribution

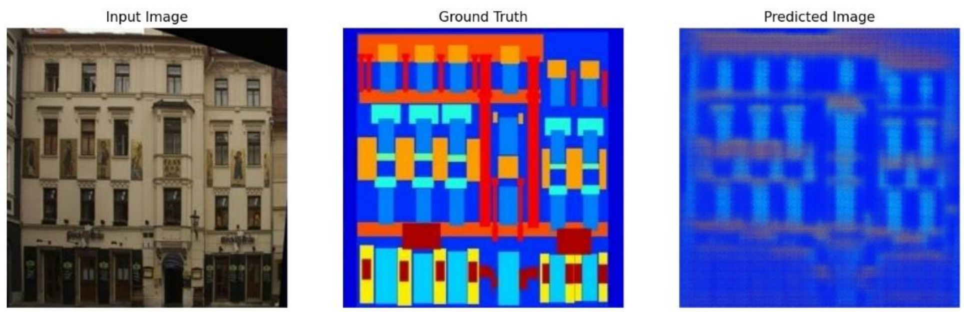




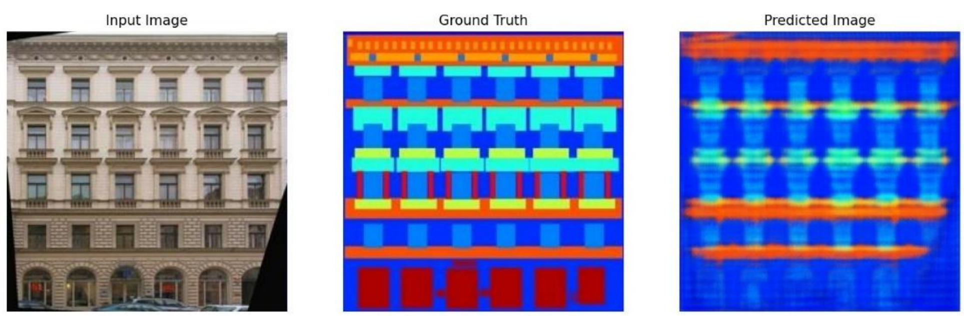


4. The following are the result of model PixelToPixel used in two ways:
   - The input for the generator is the sketch, and it is generated the fasade
   - The input of the generator is the façade and the sketch is generated

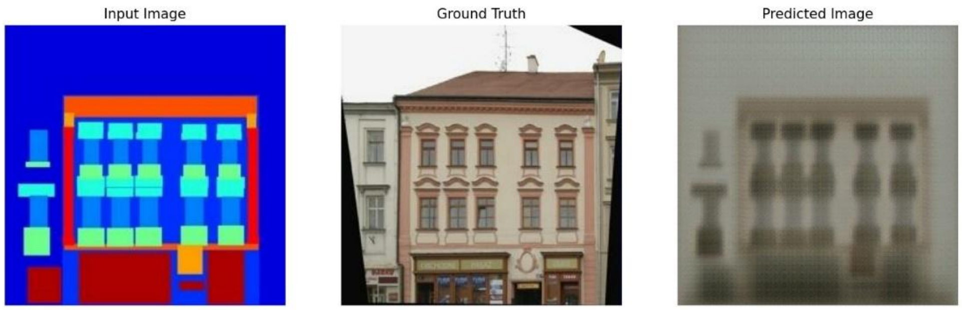
Input: the facade, generated: the sketch number of epochs=100



Input: the facade, generated: the sketch number of epochs=800



Input: the sketch, generated: the facade number of epochs=100



Input: the sketch, generated: the facade number of epochs=800

V. CONCLUSIONS

This paper discusses Machine Learning methodologies of image generation. These computational methodologies can be utilized in the architectural design process by furnishing a wealth of images in a short amount of time. At the same time the paper attempts to use these methodologies in such a way that the distinction between the geometrical syntax and architectural form is maintained.

The GAN model and some extension of it, like DCGAN, cGAN and PixToPix model are discussed. The paper is focused in some key parts of these models, like their structures and the training process of them, pointing out the differences between them. The experiments have used some facades and sketches datasets provided by the site www.kaggle.com.

The experiments are based on two elements batch size and number of epochs and also, in different structures of the models like generator and discriminator. There are not definitive values for the batch size and epochs, so an experience is given to avoid underfitting and overfitting phenomena for a dataset of the size around 400. This experience shows that for a such small database, to get a clear view of a generated facade are needed at least 1000 epochs, but with a richer structure of layers of the two models Generator and Discriminator, the clarity comes faster, starting from 100 epochs. Regarding PixToPix model, it is used in two ways. The architect has a sketch, and needs a generated facade, or he has a facade and needs to generate a sketch. The results are better for a bigger number of epochs. The metric for the error calculation used in the tests, mainly was binary cross-entropy approach, because it was shown not a big change in results when we used JS-Divergence metric.

From both perspective as an architect or data scientist, the comparative analysis of the three GAN-based models — DCGAN, CGAN, and Pix2Pix — reveals notable differences in structural design and output capabilities. Given these architectural advancements and the empirical results observed, Pix2Pix emerges as the most promising model for further exploration, particularly in domains such as urban design or architectural visualization. Its ability to translate different geometric inputs or sketches into realistic façade images highlights its potential for practical applications where structure-to-image generation is essential.

REFERENCES

Arjovsky, M., & Bottou, L. (2017). *Towards principled methods for training generative adversarial networks*. In Proceedings of the 5th International Conference on Learning Representations (ICLR).

Bader, B. W., & Kolda, T. G. (2009). Tensor decompositions and applications. *SIAM Review, 51*(3), 455–500. https://doi.org/10.1137/07070111X

Chaillou, S. (2022). *Artificial intelligence and architecture: From research to practice*. Birkhäuser.

Del Campo, M. (2022). *Neural architecture: Design and artificial intelligence*. ORO Editions.

Goodfellow, I. (2014). Generative adversarial nets. In Advances in neural information processing systems (NeurIPS) (pp. 2672–2680).

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Hossain, S., & Lee, D. (2023). NG-GAN: A robust noise-generation generative adversarial network for generating old image noise. *Sensors, 23*(2), Article 251. https://doi.org/10.3390/s23020251

IBM. (2025). *Semantic segmentation*. IBM Think Topics. https://www.ibm.com/think/topics/semantic-segmentation

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 5967–5976). IEEE. https://doi.org/10.1109/CVPR.2017.632

Kaneko, T., & Harada, T. (2020). Noise-robust generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 9994–10003). IEEE.

Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2018). Progressive growing of GANs for improved quality, stability, and variation. In Proceedings of the International Conference on Learning Representations (ICLR).

Li, C., Zhang, Y., & Wang, J. (2024). Generative AI models for different steps in architectural design: A literature review. *Frontiers of Architectural Research*. https://doi.org/10.1016/j.foar.2024.xx.xxx

Liu, G., Reda, F. A., Shih, K. J., Wang, T.-C., Tao, A., & Catanzaro, B. (2018). Image inpainting for irregular holes using partial convolutions. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 85–100). Springer.

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

Odena, A., Dumoulin, V., & Olah, C. (2016). Deconvolution and checkerboard artifacts. *Distill*. https://doi.org/10.23915/distill.00003

Parente, J., et al. (2023). Integration of convolutional and adversarial networks into building design: A review. *Journal of Building Engineering, 76*, 107155. https://doi.org/10.1016/j.jobe.2023.107155

Radford, A., Metz, L., & Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. In Proceedings of the International Conference on Learning Representations (ICLR).

Sadiku, M. N. O., Musa, S. M., & Ashaolu, T. J. (2025). Generative design in architecture. *International Journal of Trend in Scientific Research and Development*.

Shiv Ram Dubey, S., Singh, S. K., & Chaudhuri, B. B. (2021). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing, 450*, 92–120. https://doi.org/10.1016/j.neucom.2021.03.061

X. Li, Y. Zhang, & Z. Wang. (2021). Research on multi-layer perceptron based on TensorFlow. In *Proceedings of SPIE – The International Society for Optics and Photonics*. https://doi.org/10.1117/12.258xxxx

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 2223–2232). IEEE. https://doi.org/10.1109/ICCV.2017.244

ISBN 9789928347213