



ARTICLE



<https://doi.org/10.1057/s41599-024-02926-5>

OPEN

# Exploring excitement counterbalanced by concerns towards AI technology using a descriptive-prescriptive data processing method

Simona-Vasilica Oprea<sup>1</sup>✉ & Adela Bâra<sup>1</sup>

Given the current pace of technological advancement and its pervasive impact on society, understanding public sentiment is essential. The usage of AI in social media, facial recognition, and driverless cars has been scrutinized using the data collected by a complex survey. To extract insights from data, a descriptive-prescriptive hybrid data processing method is proposed. It includes graphical visualization, cross-tabulation to identify patterns and correlations, clustering using *K*-means, principal component analysis (PCA) enabling 3D cluster representation, analysis of variance (ANOVA) of clusters, and forecasting potential leveraged by Random Forest to predict clusters. Three well-separated clusters with a silhouette score of 0.828 provide the profile of the respondents. The affiliation of a respondent to a particular cluster is assessed by an F1 score of 0.99 for the test set and 0.98 for the out-of-sample set. With over 5000 respondents answering over 120 questions, the dataset reveals interesting opinions and concerns regarding AI technologies that have to be handled to facilitate AI acceptance and adoption. Its findings have the potential to shape meaningful dialog and policy, ensuring that the evolution of technology aligns with the values and needs of the people.

<sup>1</sup>Bucharest University of Economic Studies, Bucharest, Romania. ✉email: [simona.oprea@csie.ase.ro](mailto:simona.oprea@csie.ase.ro)

## Introduction

The current society increasingly depends on AI technologies. Most people possess at least one social media account, exposing them to diverse information sources. When traveling, we depend on security measures like facial recognition. Furthermore, automated transportation, such as driverless cars, has been in use for some time; for example, Narita Airport<sup>1</sup> has operated such commuting trains for over 20 years. Other assistants in our homes are tasked with performing mundane activities, gathering data, and consequently enhancing our quality of life. Developments in AI technologies have the potential to reshape and enhance society in the upcoming decades. Recently, the promise of these technologies in enhancing daily life and human capabilities has been acknowledged (Choung et al., 2023; Kopalle et al., 2022; Wang et al., 2023a; Yu and Helwig, 2022). Nonetheless, public perspectives are heavily influenced by considerations such as the intended use of these technologies, the regulation in place, and the beneficiaries or potential losses when AI advancements become implemented on a large scale. As expected, a sense of caution is prevalent in public attitudes toward AI and human enhancement applications, often revolving around concerns related to personal autonomy, lack of human connections, loss of jobs, unintended consequences, and the substantial societal changes that these innovations might entail (Querci et al., 2022; Lambert and Stevens, 2023). People might be apprehensive that economic inequalities may worsen as certain advancements emerge, and they also express concerns about the potential for increased surveillance through technologies like facial recognition software and driverless cars.

According to a recent Pew Research Center<sup>2</sup> (PRC) survey applied in February 2021 to over 10,000 Americans, it revealed that they perceived potential in how AI technologies could enhance life and society capabilities. The survey delves into a wide spectrum of scientific and technological advancements, testing the perceptions of respondents. It was mainly focused on public opinions regarding three AI developments that are frequently debated in social media: (1) The utilization of facial recognition technology (Smith and Miller, 2022), (Andrejevic and Selwyn, 2020); (2) The use of algorithms by social media companies to detect false information (Tseng et al., 2023); (3) The advancement of driverless passenger vehicles (Karmakar et al., 2021). The dataset is a subset of a larger survey that approached both technological and science-related issues. The investigated survey aims to gather opinions on the impact of technology on society, specifically looking at perceptions of fairness related to algorithms (Giovanola and Tiribelli, 2023) and the interplay between technology and issues of discrimination (Kim, 2022). It also seeks to understand how these views correlate with demographic factors such as income, political affiliation, personal ideologies, etc. The survey addresses several themes, including the overall impact of technology on daily life and societal well-being, concerns and excitement about future technological developments and their implications, views on the possibility of algorithms functioning in a fair and unbiased manner, etc. The distribution of responses suggests that the survey targeted a diverse adult population, likely with a broad age range, covering various income levels, political affiliations, and ideological beliefs. The findings extracted from answers could be used to inform policymakers and business leaders about public concerns and expectations regarding AI and its governance, guide the development of more equitable and fair technology policies and practice, enhance understanding of the relationship between socioeconomic status and perspectives on AI (van Noordt and Misuraca, 2022).

The survey has been structured to provide quantitative data through a series of closed-ended questions. This approach allows for a clear analysis of trends and patterns within the responses,

potentially supporting a wide range of statistical and comparative analyses. Given the analysis of the dataset and the prevalence of topics related to technology and algorithms, the survey may have sought to capture public sentiment on AI and its acceptance. The survey data we have analyzed offers a window into this complex landscape, capturing the multifaceted attitudes toward AI and technology as a whole. Respondents were asked to qualify the impact of AI technologies (question code: TECH1\_W99) as mostly positive, negative, or equally mixed. The distribution of responses indicates a general optimism towards technology, with many participants recognizing its benefits. However, a significant portion of the populace retains a level of skepticism or concern, highlighting the need for ongoing dialog about AI's role in society. The nuanced balance of concern and excitement regarding technological advances (CNCEXC\_W99) suggests that while the public is hopeful about the potential of AI, there is an underlying caution about its implications, ethical, privacy-related, and socioeconomic. This duality is a clear sign that acceptance of AI is not unconditional and is tempered by realistic appraisals of its potential downsides. Questions regarding the fairness of algorithms (ALGFAIR\_W99) delve into the heart of AI ethics. The responses reflect a divided opinion, with a notable portion of participants unsure about whether algorithms can be truly fair. This uncertainty underscores a pivotal aspect of AI acceptance: trust in the systems' design and the transparency of their decision-making processes. The dataset incorporates demographic variables like political affiliation (F\_PARTYSUM\_FINAL) and income (F\_INC\_SDT1), which could be correlated with AI acceptance levels. For instance, one might hypothesize that individuals with higher incomes, who may benefit more from AI services, could show greater acceptance, or that political leanings might influence trust in technology providers and regulatory approaches.

The columns that start with POSNEGAI relate to questions where respondents were asked about their levels of concern or excitement. These questions are designed to measure the respondents' attitudes towards certain issues, gauging their emotional responses on a spectrum from concern to excitement. The SMALG columns in the codification file correspond to questions about social media, algorithms, and their impact or perceptions among the respondents. The FACEREC prefix in these column names indicates that they are related to questions about facial recognition technology. These questions cover a range of issues related to facial recognition technology, including public sentiment, expected outcomes, societal impacts, and regulatory opinions. The available responses suggest that respondents can express a variety of opinions, from concerns about privacy and fairness to support for the potential benefits of the technology in areas such as law enforcement and crime prevention. The DCARS prefix in the column names indicates that these questions are related to the topic of driverless cars (autonomous vehicles) and the issues surrounding them. These questions cover a broad spectrum of issues related to autonomous vehicles (general feeling, societal and economic impact, regulation, and standards), from their potential societal and economic impacts to ethical considerations and personal comfort levels with the technology. This nuanced perspective on AI acceptance is important for stakeholders in the AI ecosystem. As AI continues to evolve, monitoring and understanding public sentiment is significant in navigating its integration into the fabric of society.

In the current research, we propose a combined descriptive and prescriptive model that integrates cross-tabulation, *K*-means clustering, 3D graphical visualization using PCA, analysis of variance (ANOVA) of clusters, and the predictive capabilities of Random Forest for cluster prediction. The proposed model

combines supervised and unsupervised Machine Learning (ML) algorithms and contributes to the state of the art in the field of survey data processing, gaining insights that are essential for promoting the acceptance and adoption of AI. Our primary objective is to conduct a comprehensive analysis of survey data on the excitement and concerns surrounding AI technology in areas like social media, facial recognition, and driverless cars, using a descriptive-prescriptive method. Our research motivation consists of understanding relevant groups of individuals, identifying their concerns, and predicting new instances of individuals into a cluster using prediction algorithms. This approach will enable us to delve deeply into the public perception and attitudes towards these technologies. Thus, the contribution of this paper lies in extracting meaningful insights from this data, which may lead to advancements in both humanity and social sciences.

Our research questions are “How do different groups of individuals perceive AI technologies such as social media, facial recognition, and driverless cars, and what are their primary concerns, and how can these perceptions and concerns be predicted for new instances of individuals?” This paper is structured as follows: in the “Literature review” section, a brief literature survey is presented focusing on similar research, the “Methodology” section is dedicated to the proposed methodology, and in the section “Results”, the results are presented. Section “Discussion” is designed for discussions, whereas section “Conclusion” is designed for conclusions.

## Literature review

The acceptance and adoption of AI have been investigated in many fields, such as: agriculture (Mohr and Kühl, 2021), manufacturing, health, education, transportation, social media (Chung et al., 2021), security, etc. Moreover, the ML and AI-induced technology-related stress in organizations was analyzed in (Kumar et al., 2023). The study on the practice of AI in manufacturing firms in Malaysia, focusing on its role in driving digital transformation, revealed significant insights (Ahmad et al., 2022). Integration of digital technologies into business processes is essential for digital transformation, a trend underscored by the emergence of “Industry 4.0”. In this context, AI is a key component in enhancing performance and stimulating demand and productivity in the manufacturing sector. The research also highlighted several barriers and challenges faced by manufacturing firms in Malaysia, primarily due to the nascent stage of AI development. These include a lack of skilled talent, insufficient incentives, and a dearth of innovation. To gather data, managers from various manufacturing companies were selected as respondents. From the 93 answers, descriptive analysis was employed to pinpoint the specific barriers and challenges. The findings indicate that the most significant barrier to AI implementation in these firms is the lack of talent. Furthermore, the primary challenge identified is the absence of in-house experts. This study underscores the need for focused strategies to overcome these issues, such as investing in education and training, fostering a culture of innovation, and creating incentives to attract and retain AI experts. The researchers focused on exploring the mediating role of perceptions on the relationship between technology readiness and the adoption of AI in the field of accounting (Damerji and Salimi, 2021). Specifically, it aimed to determine how Perceived Ease Of Use (PEOU) and Perceived Usefulness (PU) influence the correlation between the technology readiness of accounting students and their inclination to adopt AI technologies. This quantitative research involved assessing individual accounting students’ perceptions of their readiness for technology and their attitudes toward adopting it. To collect data, a questionnaire with 31 items was used. This questionnaire also

gathered demographic information. The findings of the study revealed that technology readiness significantly impacts technology adoption. However, more nuanced insights were obtained through mediation analysis using hierarchical regression. This analysis showed that the link between technology readiness and the adoption of AI is significantly influenced by PEOU and PU.

Patients’ views on the implementation of AI in radiology were investigated by a questionnaire (Ongena et al., 2020). The study aimed to gauge patients’ perspectives on the implementation of AI in radiology. A questionnaire was developed based on insights from previous qualitative research. The questionnaire was administered to 150 patients scheduled for various radiographies. The primary analytical tool used in this research was exploratory factor analysis (EFA) with principal axis factoring and oblique promax rotation. This approach helped to uncover latent variables underlying patients’ responses. To ensure the reliability of these factors, internal consistency was measured using Cronbach’s alpha and composite reliability. The EFA identified 5 distinct factors concerning patients’ views on AI in radiology: distrust and accountability, procedural knowledge, personal interaction, efficiency, and being informed. The internal consistency of the responses was good for the first three factors and acceptable for the last two. Moreover, in (Xuan et al., 2023), the readiness towards AI among undergraduate medical students in Malaysia was analyzed using frequency tables, percentages, standard deviation, unpaired *t*-tests, and ANOVA. Significant correlations were found between the age and academic year of participants and their readiness in the ability, vision, and ethics domains of medical AI. Additionally, a notable association was identified between prior training and all four domains of medical AI readiness. Additionally, other studies investigated AI impact and adoption factors in health systems (Kim et al., 2022; Kosan et al., 2022; Patrzyk et al., 2022; Vorisek et al., 2023).

Researchers explored the use of natural language processing (NLP) to analyze patient experiences (van Buchem et al., 2022). The data was primarily gathered through 5 open-ended questions, allowing for a more nuanced understanding of patient perspectives. An NLP pipeline, incorporating sentiment analysis and topic modeling, was employed to process the responses, and a visualization tool was developed to help physicians navigate the results. The research involved an iterative process of developing and validating both the questionnaire and the NLP pipeline within a clinical setting. The questionnaire focused on various aspects of patient care, including the information provided, personal approach, collaboration among healthcare professionals, organization of care, and other experiences. A total of 534 patients participated by responding to this questionnaire. The performance of the sentiment analysis model was notable, achieving an F1 score of 0.97 for identifying positive texts and 0.63 for negative texts. The use of NLP proved efficient in reducing the time required by healthcare professionals to assess and prioritize patient experiences. This approach offers a significant advantage over traditional methods that rely on closed-ended questions, as it captures a broader and more detailed range of patient feedback.

The researchers explored the implementation of AI in talent acquisition by human resource managers (Pillai and Sivathanu, 2020). A survey involving 562 managers was conducted. The collected data was analyzed through Partial Least Squares Structural Equation Modeling (PLS-SEM). The findings indicate that factors such as cost-effectiveness, relative advantage, support from top management, HR readiness, competitive pressure, and backing from AI vendors have a positive impact on the adoption of AI for talent acquisition. Conversely, concerns regarding security and privacy serve as deterrents. The research also identified that the characteristics of tasks and technology affect the

suitability of AI for talent acquisition tasks. Furthermore, the adoption and appropriateness of AI technology are linked to its actual usage in talent acquisition. Interestingly, the study uncovers that a preference for traditional methods of talent acquisition can weaken the relationship between AI adoption and its actual usage in this field.

Almarashda et al. (2021) focused on identifying factors influencing the implementation of AI in the United Arab Emirates' energy sector. A comprehensive analysis was conducted using data from 350 respondents. The study found that the primary drivers for adopting AI in organizations are its potential to foster effective business innovation, align with business strategies, and enhance production levels. Key aspects of AI, such as user-friendliness, the ability to improve work quality, and compatibility with existing tasks, were identified as critical factors for its adoption. Additionally, the research highlighted the importance of human resource management factors, including the need for a knowledgeable referral person when encountering difficulties with AI and expert support in utilizing AI technology.

The factors influencing the adoption of AI in banking services were examined (Rahman et al., 2022). A survey was conducted, yielding 302 opinions from Malaysian banking customers. This research aimed to identify the key determinants of customers' intentions to adopt AI in banking services. From the qualitative analysis, it was found that AI is essential for detecting fraud and preventing risks. However, concerns about data privacy and security, lack of regulatory frameworks, and deficiencies in relevant skills and IT infrastructure pose significant challenges to AI adoption. The quantitative analysis revealed that several factors significantly impact customers' intentions to adopt AI. These include their attitudes towards AI, PU, perceived risk, perceived trust, and subjective norms. Interestingly, factors like PEOU and awareness were not found to influence this intention. Moreover, the study highlighted the role of attitude towards AI as a mediator in the relationship between PU and the intention to adopt AI.

Another research presented an innovative approach to exploring the dynamics between digital technology and the adoption of AI in the context of electronic manufacturing enterprises (Binsaeed et al., 2023). Specifically, the study aimed to (a) investigate the direct relationship between digital technology and AI adoption; (b) examine the mediating role of Knowledge Sharing (KS) in this relationship; (c) assess how privacy and security considerations might moderate the link between digital technology and AI adoption. A questionnaire survey was distributed among electronic firms in Saudi Arabia. The analysis is based on data from 298 respondents. The research methods include multi-level correlation and regression analysis to test the hypotheses. The findings from this analysis are noteworthy. There is a positive direct influence of digital technology on AI adoption within electronics manufacturing enterprises. KS acts as a significant mediator in the relationship between digital technology and AI adoption. Privacy and security considerations play a moderating role in this relationship. Moreover, (Qiu et al., 2022) delved into the influence of AI-enabled services in the hospitality industry, particularly focusing on how AI technology supports Frontline Employees (FE) physically, mentally, and emotionally, thereby enhancing their ability to provide hospitable service. A total of 342 valid questionnaires were collected to examine various aspects related to AI-enabled services. The study employed factor analyses and measurement model evaluation to understand the construct of AI-enabled service better. Four key factors were identified as central to this construct: anthropomorphic attributes, entertainment attributes, functional attributes, and information attributes. While entertainment does not reduce the physical and mental fatigue experienced by FE, it has a notable positive impact on their emotional well-being.

The adoption of AI and related technologies by public and academic librarians in North America was investigated (Yoon et al., 2022). Quantitative analysis of the survey responses was conducted using various chi-square tests and crosstab analyses. It was observed that academic librarians reported higher usage and awareness of AI and related technologies. Conversely, public librarians exhibited more positive attitudes towards AI technologies. Overall, 67% of the participants believed that AI would significantly change library functions, and 68% of the librarians expressed interest in receiving training in AI. These findings suggest a growing recognition of the transformative potential of AI in the library sector, coupled with a strong interest in training and development to leverage these technologies effectively.

Another research contributed to the field of AI education by introducing a novel AI curriculum framework (Chiu et al., 2022). The effectiveness of this curriculum in enhancing AI learning was established through a multifactorial pre-post-test evaluation, focusing on students' perceptions of AI learning. The curriculum's development involved a collaborative co-creation process, with a team of 14 professors working alongside 17 principals and teachers from 6 secondary schools. The study participants included 335 students and 8 teachers from these schools. The research employed a mixed-methods approach, combining quantitative data from pre- and post-questionnaires with qualitative insights highlighting teachers' views on the co-creation process. Data analysis was conducted using paired t-tests and ANCOVAs for quantitative data and thematic analysis for qualitative data. Post-intervention, students reported an increased sense of competence in AI and a more positive attitude towards learning AI. Additionally, other studies investigated AI in education (Rauf et al., 2021; Wang et al., 2023b; Nouraldeen, 2023) in Lebanese universities. Another study investigated an integrated model designed to predict elementary school students' acceptance of AI robots serving as teachers (Chen et al., 2023). This research area has garnered significant interest due to AI teachers' potential to address the global teacher shortage. Participants were students from Chinese elementary schools. The study utilized descriptive statistics and SEM. The research identified several key factors influencing the acceptance of AI teachers, including Robot Use Anxiety (RUA), PU, PEOU, and Robot Instructional Task Difficulty (RITD). The findings indicated that students had positive attitudes towards AI teachers, and these attitudes could be predicted based on their perceptions of the AI teachers' PU, PEOU, and the difficulty of the tasks they were programmed to perform. Furthermore, the study discovered that the relationship between RITD and the acceptance of AI teachers was mediated by factors such as RUA, PEOU, and PU. This indicated that the difficulty of tasks assigned to AI teachers, the anxiety students feel about using robots, and their PEOU of AI teachers all play important roles in determining how well these AI teachers are accepted by students.

The study (Amichai-Hamburger et al., 2020) focused on the adoption of personal autonomous cars, particularly how tailoring the vehicle's operations to the user's personality could improve user experience. A questionnaire was created, targeting various features of a proposed information system for autonomous cars. Analysis of the responses identified two key factors influencing user preferences: the willingness to share personal information and the desire to maintain control over the vehicle. Additionally, a regression analysis examining the relationship between preferences for autonomous car features and factors such as personality traits, gender, and age revealed that traits like openness, conscientiousness, and age significantly influenced these preferences. Another study (Rowthorn, 2019) investigated the ethical decision-making processes in autonomous vehicles, examining aspects of machine ethics, AI in driving, and algorithms used during crash scenarios. Utilizing data from multiple sources,

including the Pew Research Center, the researcher conducted a series of analyses and estimations. These included assessing the percentage of U.S. adults willing or not willing to travel in a driverless car, evaluating how safe these adults feel as pedestrians in cities where Autonomous Vehicles (AVs) operate, determining which countries are best equipped for the advent of autonomous vehicles, and identifying the key data infrastructure needs in smart cities that are essential for the testing and implementation of AVs. More studies focused on identifying the behavioral factors influencing the adoption of AVs (Acheampong and Cugurullo, 2019). The primary objective was to develop conceptual frameworks and measurement models to forecast trends in public transportation, sharing, and ownership of self-driving cars. These proposed frameworks incorporated socio-demographic variables along with key latent behavioral factors. They included PEOU of AVs, public concerns and apprehensions about AVs, subjective norms, perceived behavioral control, and attitudes related to the environment, technology, shared consumption, public transport, and car ownership. Utilizing survey data with CFA, the research aimed to establish and validate the reliability of scale indicators. This process involved confirming both the convergent and discriminant validity of the relationships among the latent variables. As a result of this comprehensive analysis, four distinct measurement models were developed and presented.

Passengers' perceptions and satisfaction levels with the digital technology implemented by airlines during the COVID-19 pandemic, focusing specifically on the Chinese market, were explored (Shiwakoti et al., 2022). An online survey was conducted to assess passengers' views on 11 different digital technology-based services provided by airlines, including facial recognition. A total of 365 valid responses were collected and analyzed using ANOVA tests and stepwise multiple linear regression analysis. The findings revealed a generally positive attitude among passengers towards the adoption of new technologies by airlines. From the regression analysis, 6 technologies were identified as having a statistically significant impact on passenger satisfaction. These include AI customer service, electronic luggage tags, cleaning robots, ultraviolet light and antimicrobial cabin cleaning, an app-controlled in-flight entertainment system, and an electronic library. Conversely, three technologies were found to be less favorable among passengers: facial recognition, digital documentation, and AI customer service.

Other recent research focused on public sentiment extracted from texts from social platforms such as Twitter (Liu, Zhou, et al., 2023b) and emotions classification involving ML algorithms and complex data processing pipelines (Liu et al., 2023a). Autonomous cars, driving, and traffic issues were debated in (Xiao et al., 2023), addressing the safety challenges posed by occlusions in autonomous driving by proposing a multi-tier perception task offloading framework that leverages collaborative computing between autonomous vehicles and roadside units. Federated learning in autonomous cars (Fu et al., 2023), technology in of moving trajectory of autonomous cars (Ding et al., 2023), and human activity recognition (Jannat et al., 2023) were further examined by exploring the use of Wi-Fi Channel State Information (CSI) signals for recognizing human activities, leveraging the insight that human movement alters Wi-Fi signal propagation, evident in CSI signal variations. Other intriguing aspects such as management in traffic with autonomous vehicles (Yue et al., 2023) and policies related to technological innovations in vehicles and their manufacturers (Jiang and Xu, 2023) were investigated.

Comparing the brief findings extracted in Table 1, we can conclude that AI acceptance and adoption were investigated in various countries (Malaysia, U.S.A., Netherlands, Malaysia, India,

United Arab Emirates, and Saudi Arabia. North America, China (Hong-Kong), China-other regions, Israel, Ireland (Dublin). The number of respondents varied from 93 to 5,400, showing a wide range of cases. Furthermore, the methods used in these studies vary widely, ranging from descriptive analysis, quantitative research, and hierarchical regression in the Malaysian industry ([17]) and U.S. accounting fields ([18]), to EFA and CFA in health-related studies in the Netherlands ([19], [25]), PLS-SEM in human resources and banking system research in India and Malaysia ([26], [28]), and more advanced techniques like NLP and multi-level regression analysis in health and manufacturing sectors in the Netherlands and Saudi Arabia ([25], [29]). Other methods include OLS regression in transportation studies in Israel ([37]), and a mix of ANOVA, chi-square tests, and thematic analysis in various fields across different countries. This diversity in methods reflects the unique requirements and complexities of each field and research question.

## Methodology

**Proposed data processing method.** Survey data can be analyzed in various ways to gain insights and make informed decisions: descriptive analysis, cross-tabulation (Warwas et al., 2022), hypothesis testing (*t*-tests and ANOVA, Chi-square tests, regression analysis) (Ko et al., 2021), (Sampurna et al., 2023), factor analysis (Băra and Oprea, 2023), (Oprea and Băra, 2022), cluster analysis, sentiment analysis (Lepelaar et al., 2022), principal component analysis (PCA) to reduce the dimensionality of the data and identify the most important components; latent class analysis (LCA), Bayesian analysis, text mining and topic modeling like latent dirichlet allocation (LDA) or word cloud visualizations. To investigate the responses, we propose a descriptive-prescriptive hybrid model that includes cross-tabulation, clustering using *K*-means, analysis of variance (ANOVA), and prediction ensemble algorithm (Random Forest) to estimate clusters. The process flow of the methodology is described in Fig. 1.

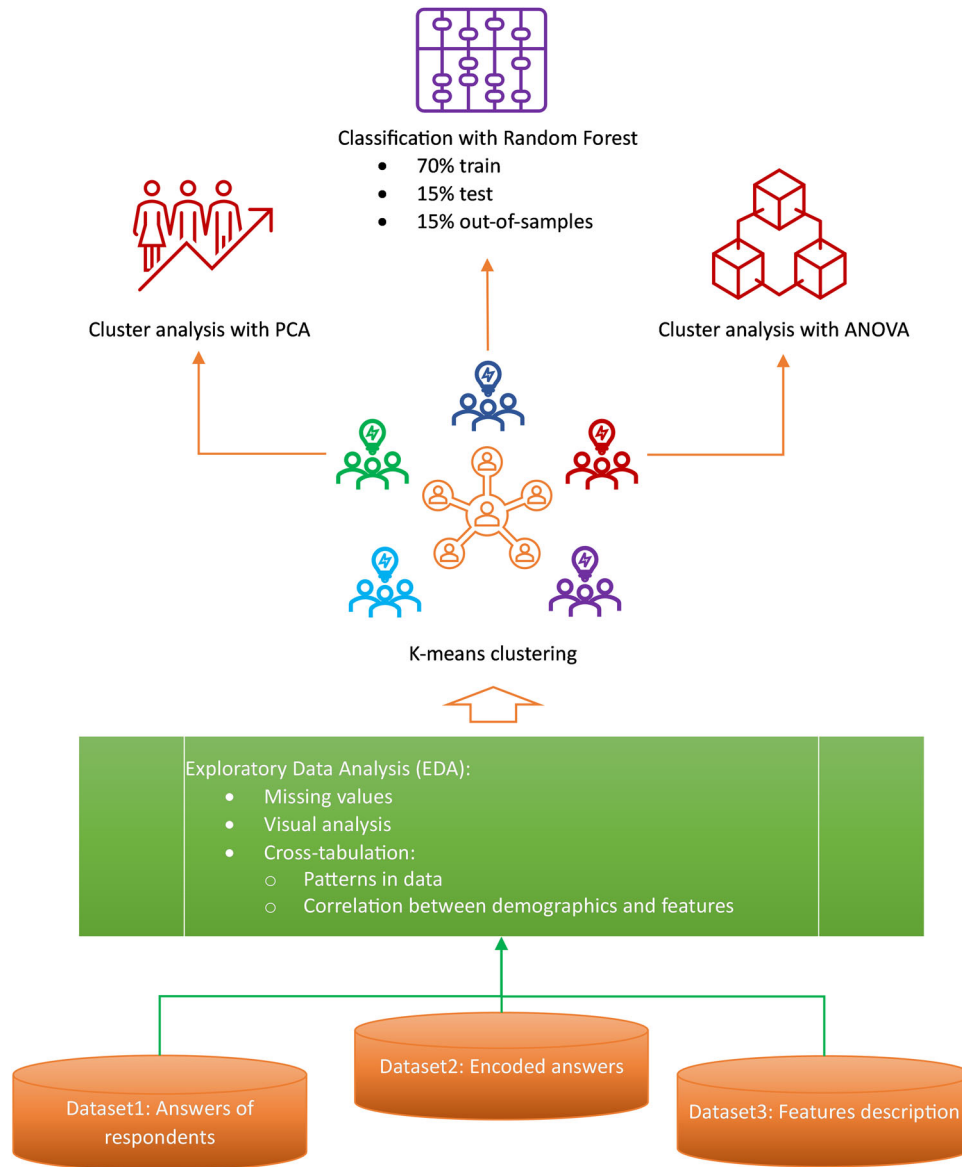
**Algorithms and statistical approach background.** *K*-means is an unsupervised ML algorithm used for clustering similar data points into groups or clusters. The algorithm aims to partition a dataset into *K* distinct, non-overlapping clusters, with each data point belonging to the cluster with the nearest mean value. The number of clusters is the most important parameter in *K*-Means. Thus, it is important to determine the optimal number of clusters using the elbow method or silhouette score. It represents the number of centroids (cluster centers) to be generated. *K*-means starts by randomly initializing *K* cluster centroids. These centroids serve as the initial cluster centers. Each data point is assigned to the nearest centroid based on a distance metric, commonly Euclidean distance. The data points are grouped into clusters based on their proximity to the centroids. After assigning data points to clusters, the algorithm calculates new centroids for each cluster. The new centroids are computed as the mean of all data points within the cluster. The assignment and update steps are repeated iteratively until one of the stopping criteria is met. *K*-Means is guaranteed to converge to a solution, but it may not always find the global optimum. The final clustering result depends on the initial random centroid selection. It has several advantages, such as: simple and easy to understand, scales well to large datasets, and can be used for various types of data, including numerical and categorical. Nonetheless, it has also disadvantages, such as: requires specifying the number of clusters (*K*) in advance; being sensitive to the initial placement of centroids, which leads to different results; assuming that clusters are spherical and equally sized, which may not be the case in real complex datasets. There are several applications where *K*-means can be applied

**Table 1 Brief comparative analysis of previous research.**

Ref.	Country	Field	No. of respondents	Methods	Findings summary
(Ahmad et al., 2022)	Malaysia	Industry	93 (managers)	Descriptive analysis	Absence of in-house experts
(Damerji and Salimi, 2021)	U.S.A.	Accounting	824 students	Quantitative research, mediation analysis using hierarchical regression	Technology readiness significantly impacts technology adoption, the link being significantly influenced by both PEOU and PU
(Ongena et al., 2020)	Netherlands	Health (radiology)	105 (patients)	EFA, suggestions for CFA	5 distinct factors concerning patients' views on AI in radiology: distrust and accountability, procedural knowledge, personal interaction, efficiency, being informed
(Xuan et al., 2023)	Malaysia	Health	105 participants	Frequency tables, percentages, standard deviation, unpaired t-test, and ANOVA	Significant correlations were found between the age and academic year of participants and their readiness in the ability, vision, and ethics domains of medical AI
(van Buchem et al., 2022)	Netherlands	Health	534 patients	NLP pipeline: sentiment analysis (multilingual BERT model) and topic modeling (LDA)	NLP proved efficient in reducing the time required by healthcare professionals to assess and prioritize patient experiences
(Pillai and Sivathanu, 2020)	India	Human resources	562 managers	PLS-SEM	Cost-effectiveness, relative advantage, support from top management, HR readiness, competitive pressure, and backing from AI vendors have a positive impact on the adoption of AI
(Almarashda et al., 2021)	United Arab Emirates	Energy	350	Cronbach's alpha value, RII technique to prioritize parameters	Primary drivers for adopting AI in organizations are its potential to foster effective business innovation, align with business strategies, and enhance production levels
(Rahman et al., 2022)	Malaysia	Banking system	302 bank customers	PLS-SEM	Attitudes towards AI, perceived usefulness, perceived risk, and perceived trust influence the intention to adopt AI
(Binsaeed et al., 2023)	Saudi Arabia	Manufacturing	298 respondents from electronic firms	Multi-level correlation and regression analysis	Digital technology has a positive influence on AI adoption. Knowledge sharing mediates the linkage between digital technology and AI adoption
(Qiu et al., 2022)	N.a.	Hospitality industry	342 from 15 hotels employees	Factor analyses and measurement model	4 factors: anthropomorphic, entertainment, functional, and information attributes - explain the construct of AI-enabled service
(Yoon et al., 2022)	North America	Library	242 public and academic librarians	Multiple chi-square tests and crosstab analyses	Academic librarians reported higher usage and awareness of AI, while public librarians exhibited more positive attitudes toward these technologies

**Table 1 (continued)**

Ref.	Country	Field	No. of respondents	Methods	Findings summary
(Chiu et al., 2022)	China (Hong-Kong)	Education	335 students and 8 teachers	Paired t-tests and ANCOVAs for quantitative data and thematic analysis for qualitative data PLS-SEM	Post-intervention, students reported an increased sense of competence in AI and a more positive attitude toward learning AI It confirms the new explanatory power of resource-based theory in technological practices.
(Wang et al., 2023b)	China	Education	561 valid questionnaires		
(Chen et al., 2023)	China	Education	665	SEM, descriptive statistics	Openness, consciousness, and age were related to different preferences
(Amichai-Hamburger et al., 2020)	Israel	Transportation (driverless cars)	155 students	Ordinary Least Square (OLS) regression analysis	
(Rowthorn, 2019)	U.S.A.	Transportation (driverless cars)	5400	SEM	A multifaceted analysis of the current state and challenges of integrating autonomous vehicles into society, focusing on ethical considerations, public perception, infrastructural needs, etc.
(Acheampong and Cugurullo, 2019)	Ireland	Transportation (driverless cars)	507	CFA	Confirming both the convergent and discriminant validity of the relationships among the latent variables
(Shiwakoti et al., 2022)	China	Security (facial recognition)	365	ANOVA tests and stepwise multiple linear regression analysis	Three technologies were found to be less favorable among passengers: facial recognition, digital documentation, and AI customer service



**Fig. 1** Methodology flow diagram.

alone or in combination with other ML algorithms providing solutions for customer segmentation for marketing, image compression, anomaly detection, document classification, genomic data analysis, recommendation systems, etc.

The first step consists of random initialization of  $K$  cluster centroids  $\mu_1, \mu_2, \dots, \mu_K$ , where  $K$  is the number of clusters. Then, the assignment of the data points to clusters is performed. For each data point  $x_i$ , where  $i = 1, 2, \dots, N$  ( $N$  is the number of data points), the clusters are selected based on Eq. (1):

$$c_i = \arg \min_k \|x_i - \mu_k\|^2 \tag{1}$$

Each data point  $x_i$  is assigned to the cluster with the nearest centroid, where  $c_i$  is the cluster assignment for  $x_i$ . After each assignment, the cluster centroids are updated. For each cluster  $k = 1, 2, \dots, K$ , the cluster centroids are calculated as in Eq. (2):

$$\mu_k = \frac{1}{|C_k|} \sum_{i \in C_k} x_i \tag{2}$$

The centroid  $\mu_k$  is updated as the mean of all data points  $x_i$  assigned to cluster  $k$ , where  $|C_k|$  is the number of data points in

cluster  $k$ . The assignment and update steps are repeated until a stopping criterion is met. Common stopping criteria include a maximum number of iterations reached or centroids no longer changing significantly (leading to convergence).

Finally, the clusters are evaluated in terms of size, centroids, distributions, and other characteristics. Therefore, the  $K$ -means algorithm involves iteratively assigning data points to clusters and updating cluster centroids until convergence. The algorithm aims to minimize the sum of squared distances between data points and their assigned centroids. The choice of distance metric and the initialization method (e.g., random initialization or  $K$ -means) are essential aspects of the  $K$ -means algorithm. Additionally, the algorithm's performance is sensitive to the choice of  $K$ , the number of clusters, which requires evaluation techniques.

ANOVA is a statistical technique used to analyze and compare the means of two or more groups or populations to determine whether there are statistically significant differences among them. ANOVA is a tool for understanding the variability within and between groups. It can be combined with  $K$ -means to further analyze the quality of clusters. Two-way ANOVA extends the one-way ANOVA to handle situations with two independent



categorical variables or factors. It allows us to simultaneously investigate the effects of two factors on a response variable. To formalize the two-way ANOVA, let's consider two independent factors A and B, where  $I$  is the number of levels or categories in the first factor (Factor A) and  $J$  is the number of levels or categories in the second factor (Factor B),  $N$  is the total number of observations,  $n_{ij}$  is the number of observations in the  $i$ -th level of Factor A and the  $j$ -th level of Factor B.  $\bar{x}_{ij}$  is the mean of observations in the  $i$ -th level of Factor A and the  $j$ -th level of Factor B.  $\bar{x}$  is the overall mean,  $SS_T$  is the total sum of squares (variation of all data points from the overall mean),  $SS_A$  is the sum of squares for Factor A (variation due to Factor A),  $SS_B$  is the sum of squares for Factor B (variation due to Factor B),  $SS_{AB}$  is the sum of squares for the interaction between Factors A and B (joint effect of A and B) and  $SS_W$  is the within-group sum of squares (residual or error variation). The total sum of squares is defined in Eq. (3):

$$SS_T = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^{n_{ij}} (x_{ijk} - \bar{x})^2 \tag{3}$$

It calculates the total variability in the data, measuring how far individual data points are from the overall mean. The sum of squares for Factor A is formulated in Eq. (4):

$$SS_A = \sum_{i=1}^I n_i (\bar{x}_i - \bar{x})^2 \tag{4}$$

The variation due to Factor A, measuring how much the group means (across levels of Factor A) differ from the overall mean, is calculated. The sum of squares for Factor B is defined in Eq. (5):

$$SS_B = \sum_{j=1}^J n_j (\bar{x}_j - \bar{x})^2 \tag{5}$$

The variation due to Factor B, measuring how much the group means (across levels of Factor B) differ from the overall mean, is calculated. The sum of squares for interaction is provided in Eq. (6):

$$SS_{AB} = \sum_{i=1}^I \sum_{j=1}^J n_{ij} (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j - \bar{x})^2 \tag{6}$$

Equation (6) calculates the variation due to the interaction between Factors A and B, measuring the combined effect of A and B. Within-group sum of squares is provided in Eq. (7):

$$SS_W = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^{n_{ij}} (x_{ijk} - \bar{x}_{ij})^2 \tag{7}$$

Equation (7) calculates the variation within each group, measuring how far individual data points are from their respective group means. The total degrees of freedom is defined in Eq. (8):

$$df_T = N - 1 \tag{8}$$

The degrees of freedom for Factor A and for Factor B are formulated in Eq. (9):

$$df_A = I - 1 \quad df_B = J - 1 \tag{9}$$

The degrees of freedom for the interaction and within-group are defined as in Eq. (10):

$$df_{AB} = (I - 1) \times (J - 1) \quad df_W = N - I \times J \tag{10}$$

The mean squares for Factor A and for Factor B are provided in Eq. (11):

$$MS_A = \frac{SS_A}{df_A} \quad MS_B = \frac{SS_B}{df_B} \tag{11}$$

The mean squares for the interaction and within groups are formulated in Eq. (12):

$$MS_{AB} = \frac{SS_{AB}}{df_{AB}} \quad MS_W = \frac{SS_W}{df_W} \tag{12}$$

F-statistics for Factor A, for Factor B, and for the interaction are provided in Eq. (13):

$$F_A = \frac{MS_A}{MS_W} \quad F_B = \frac{MS_B}{MS_W} \quad F_{AB} = \frac{MS_{AB}}{MS_W} \tag{13}$$

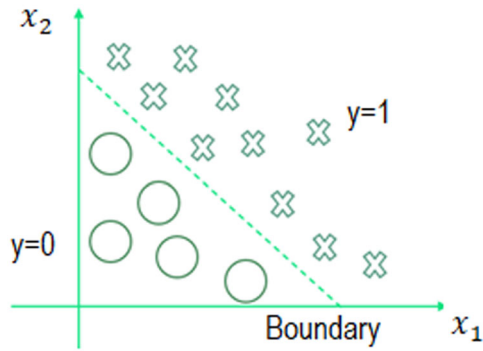
The  $p$ -values are calculated based on the respective F-statistics and the F-distribution. They indicate the probability of observing such extreme F-statistic values under the null hypothesis for each factor and the interaction. Null Hypotheses ( $H_0$ ) implies that  $H_{0A}$ : No significant effect on Factor A,  $H_{0B}$ : No significant effect of Factor B, and  $H_{0AB}$ : No significant interaction between Factors A and B. Alternative Hypotheses ( $H_1$ ): implies that  $H_{1A}$  Significant effect on Factor A,  $H_{1A}$  a significant effect on Factor B and  $H_{1AB}$ : Significant interaction between Factors A and B. If the  $p$ -value for a factor or interaction is smaller than a chosen significance level (e.g., 0.05), we reject the null hypothesis for that factor or interaction. Otherwise, we fail to reject it. These equations formalize the process of conducting a two-way ANOVA to analyze the effects of two categorical factors (Factor A and Factor B) and their interaction on a response variable. Two-way ANOVA helps determine whether variations in the response variable are influenced by one or both factors, as well as their interaction. Variation analysis is applied to the groups obtained with K-means to estimate the performance of the clustering process.

The aim of cluster prediction is to categorize a new instance within a dataset into a predefined cluster. For instance, when an individual responds to survey questions, they can be assigned to a specific cluster based on their answers. While it is acknowledged that perceptions and attitudes in technology and innovation are subject to change, such allocation can still be effectively done using existing classifications. However, to ensure accuracy and relevance in understanding technology adoption, it is necessary to periodically repeat the survey, thereby updating the responses and refining the calibration of adoption measures.

Classification methods focus on identifying variables that take on discrete values, which are represented as categories or classes (in our particular case cluster number). These classes can be binary, such as  $y \in \{0, 1\}$ , or multi-class, such as  $y \in \{0, 1, 2, 3, 4, \dots, p\}$ , where  $p$  represents the total number of distinct classes. In a binary classification scenario, the value 1 typically signifies the "positive" class, while 0 indicates the "negative" class. The core of classification techniques lies in the construction of a function that predicts the probability that a given instance belongs to a particular class. This is achieved by modeling a function based on independent variables  $x$  and computing the likelihood of belonging to a specific class. For instance, in binary classification, if the function  $h_\theta(x)$  yields a value equal to or greater than 0.5, it is predicted that  $\hat{y} = 1$ , meaning the instance is classified into the positive class. Conversely, if  $h_\theta(x)$  is less than 0.5,  $\hat{y} = 0$  is predicted, placing the instance in the negative class. The function  $h_\theta(x)$ , which ranges from 0 to 1 (i.e.,  $0 \leq h_\theta(x) \leq 1$ ), is used to model this classification decision. To effectively model the function  $h_\theta(x)$ , specific notations and methods are employed as in Eq. (14):

$$q = \theta^T \times X$$

$$q = [\theta_0 \quad \dots \quad \theta_n] \times \begin{bmatrix} x_1^{(1)} & \dots & x_n^{(1)} \\ \dots & \dots & \dots \\ x_1^{(m)} & \dots & x_n^{(m)} \end{bmatrix} \tag{14}$$



**Fig. 2** Boundary between the two classes ( $p=2$ ).

Where  $\theta = [\theta_0 \dots \theta_n]$ - the vector of weights that is multiplied by  $X$ ;  $X$ -the matrix of input variables;  $n$ -number of variables;  $m$ -number of observations;  $x = [x_1^{(i)} \dots x_n^{(i)}]$ -an observation. The following probability function is calculated as in Eq. (15):

$$h_{\theta}(x) = \sigma(q) = \frac{1}{1 + e^{-q}} \tag{15}$$

The sigmoid function approximates:  $\hat{y} = 1$ , if  $h_{\theta}(x) \geq 0.5$  (or  $q \geq 0$ ) and  $\hat{y} = 0$ , if  $h_{\theta}(x) < 0.5$  (or  $q < 0$ ), whereas the function  $h_{\theta}(x)$  estimates the probability that  $y = 1$  as in Eq. (16):

$$h_{\theta}(x) = P(y = 1|x;\theta) \tag{16}$$

The relation between the two probabilities  $P(y = 1|x;\theta)$  and  $P(y = 0|x;\theta)$  is defined in Eq. (17):

$$P(y = 1|x;\theta) + P(y = 0|x;\theta) = 1 \quad P(y = 0|x;\theta) = 1 - P(y = 1|x;\theta) \tag{17}$$

Let's consider  $q = \theta_0 + \theta_1 x_1 + \theta_2 x_2$ , where  $\theta_0 = -2, \theta_1 = 1$  and  $\theta_2 = 1$ . Thus,  $q = -2 + x_1 + x_2$  and  $h_{\theta}(x) = \sigma(q) = \frac{1}{1 + e^{-(-2+x_1+x_2)}}$ .  $y = 1$ , if  $-2 + x_1 + x_2 \geq 0$  or  $x_1 + x_2 \geq 2$ , creating a decision boundary (as in Fig. 2).

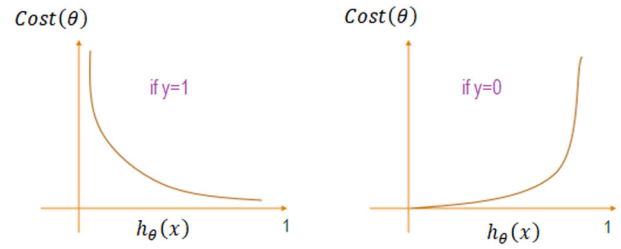
The cost function for classification depends on the dependent variable  $y$  as it can be either  $y = 1$  or  $y = 0$  as in Eq. (18):

$$Cost(\theta) = \frac{1}{m} \sum_{i=1}^m -\log(h_{\theta}(x^{(i)}))$$

$$Cost(\theta) = \frac{1}{m} \sum_{i=1}^m -\log(1 - h_{\theta}(x^{(i)})) \tag{18}$$

If  $y = 1$ , when  $h_{\theta}(x)$  tends 1, the  $Cost(\theta)$  is very small and tends to 0, whereas when  $h_{\theta}(x)$  tends 0, the cost is very high  $Cost(\theta)$  and tends to infinite. Furthermore, if  $y = 0$ , when  $h_{\theta}(x)$  is closer to 0, the cost tends to 0, then  $Cost(\theta) \rightarrow 0$ , whereas if  $h_{\theta}(x)$  tends to 1, then the cost is rocketing to infinite:  $Cost(\theta) \rightarrow \infty$  as the estimation is far from the target and it has to be significantly penalized (as in Fig. 3).

In the same way as with linear regression, where overfitting is a potential issue, a regularization term denoted by  $\gamma$  can be incorporated into the cost function. This addition aims to mitigate the problem of overfitting by decreasing the variance. Consequently, the cost function is modified to include this



**Fig. 3** Cost function.

regularization term as in eq. (19).

$$Cost(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) (\log(1 - h_{\theta}(x^{(i)})))] + \frac{\gamma}{2} \sum_{j=1}^n \theta_j^2 \tag{19}$$

The goal is to fine-tune the model's weights to achieve the lowest possible value of the cost function. The gradient descent algorithm is a commonly used method for tackling this optimization challenge. By iteratively updating the weights  $\theta$ , the algorithm works towards minimizing the overall cost, defined in Eq. (20).

$$\theta_j := \theta_j - \alpha \frac{\partial Cost(\theta)}{\partial \theta_j} \tag{20}$$

The partial derivative of the cost function is calculated to obtain Eq. (21):

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (\forall) j = 0, n \text{ and } x_0 = 1 \tag{21}$$

If the classification has multiple classes,  $y \in \{0, 1, 2, \dots, p\}$ , then the classification problem can be decomposed in  $p + 1$  problems as in Eq. (22):

$$h_{\theta}^0(x) = P(y = 0|x;\theta) \quad h_{\theta}^1(x) = P(y = 1|x;\theta) \dots$$

$$h_{\theta}^p(x) = P(y = p|x;\theta) \tag{22}$$

To evaluate the performance of classification, Accuracy (A), Precision (P), Recall (R), F1 score, and AUC are calculated as in Eq. (23):

$$A(y, \hat{y}) = \frac{1}{m} \sum_{i=1}^m 1(\hat{y} = y) \tag{23}$$

Where  $y$  - real value,  $\hat{y}$  - estimated value.

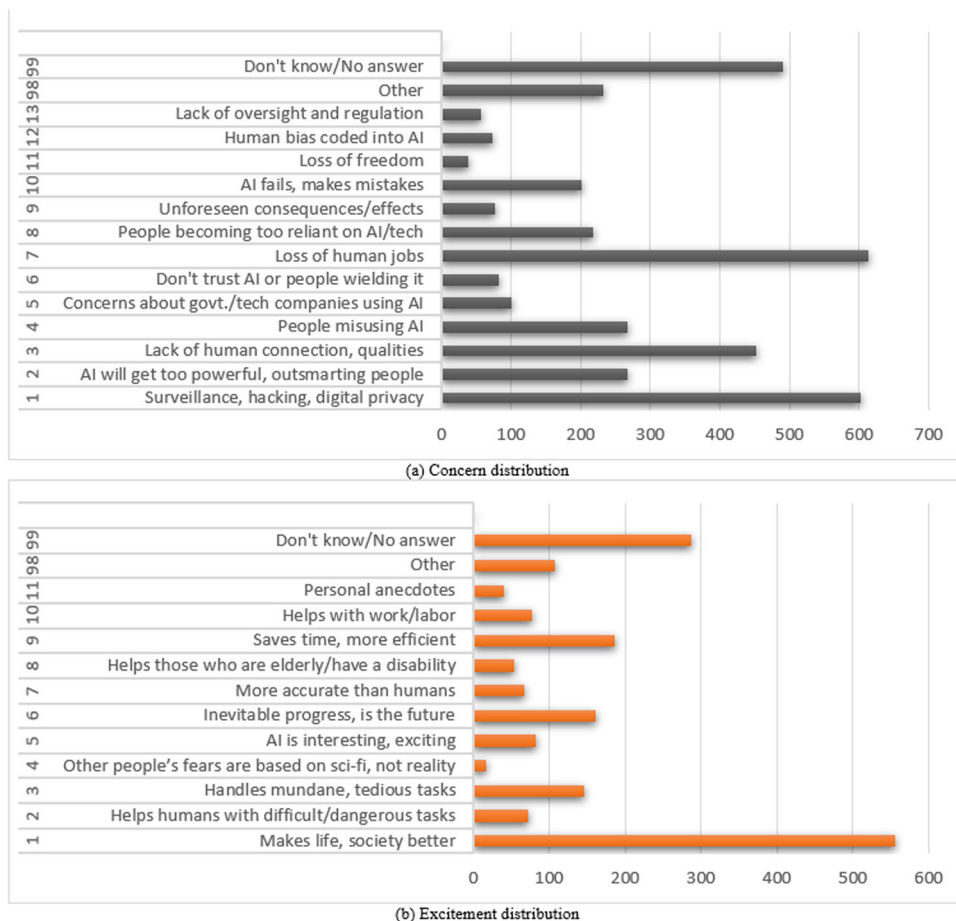
Recall (R) or True Positive Rate (TPR) is the ratio between True Positive (TP) and the sum of TP and False negative (FN). Precision (P) is the ratio between TP and the sum of TP and False Positive (FP). Usually, F1score is calculated mostly as it represents a combination of R and P. They are defined in Eq. (24):

$$R = \frac{TP}{TP+FN} \quad P = \frac{TP}{TP+FP} \quad Fscore = \frac{P \times R}{P+R} \tag{24}$$

Area under the curve (AUC) represents the performance metric of a classifier. This metric ranges from 0.5 to 1, where 1 represents the ideal score for excellent model performance. An AUC significantly greater than 0.5, typically above 0.7, indicates strong predictive accuracy.

**Results**

**Input data.** This research examines a subset of a comprehensive, publicly accessible survey, the details of which are available in a GitHub repository and on the Pew Research Center (PRC) website. The analyzed data originates from a survey panel that



**Fig. 4 Concerns and excitement distributions.** **a** Concern distribution—includes the list of concerns related to AI technologies and their frequencies; **b** excitement distribution—includes the list of excitements related to AI technologies and their frequencies.

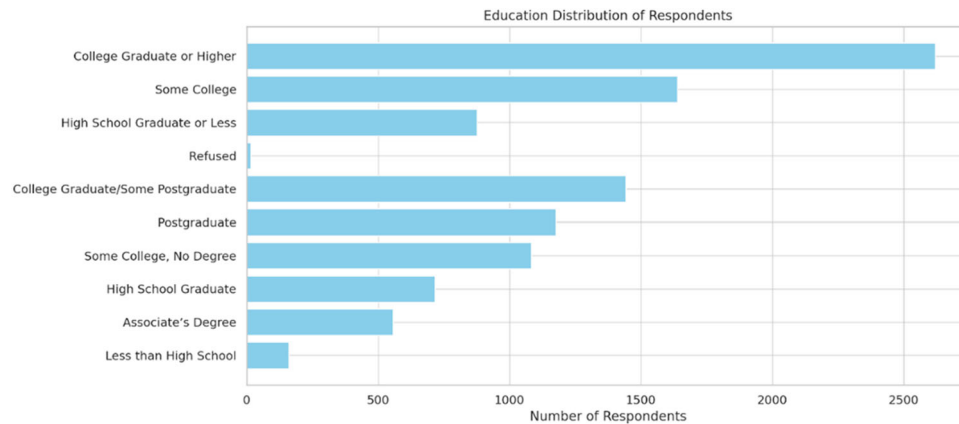
took place between November 1st and November 7th, 2021. Out of the 11,492 individuals selected for the survey, 10,260 completed their responses, resulting in an 89% participation rate. Therefore, the survey was designed mainly by PCR in consultation with Ipsos and responses were collected online and pre-processed (including the data quality checks) by the same institution. According to PRC methodology, when considering non-responses to initial recruitment efforts and subsequent dropouts, the overall response rate was 3%. Among those panelists who began the survey and answered at least one question, the break-off rate was 1%. For the entire sample of 10,260 respondents, the margin of error in sampling was estimated to be  $\pm 1.6$  percentage points.

This survey aimed to gather data from non-institutionalized adults aged 18 and over who reside in the United States, encompassing those living in Alaska and Hawaii. Within each household, the survey methodology involved selecting the adult who had the nearest upcoming birthday to participate by completing the survey online. Each participant was provided with a pre-arranged reward for their involvement in the survey.

While the original survey explored both technological and scientific impacts, our analysis in this paper is confined to the technological aspects (that refer to 5,153 out of the 10,260 records of the dataset). This decision is due to the dataset’s complexity and its extensive range of variables. For reader convenience and to support data transparency, we included a link to the dataset on GitHub in the data availability section. This allows for the replication of our analysis.

Addressing potential biases in survey data is essential for the validity of a study’s outcomes. Concerns such as response bias, where respondents may provide socially desirable answers, can be mitigated by ensuring anonymity and using neutral, non-leading questions. Sampling bias, which occurs when the sample is not representative of the population, can be minimized through randomized sampling methods and, if necessary, stratified sampling to include underrepresented groups. We used data from the PRC. It is an organization renowned for conducting surveys, sampling, data analysis, and pre-processing. We considered their pre-processed survey data, which comes with an array of explanatory read-me files. PRC employs a rigorous survey methodology and provides a wealth of resources for each survey, including files in both CSV and PDF formats. Their surveys feature questions labeled with full descriptions and answers presented in various formats, ranging from all numeric to mixed formats. Additionally, the PRC offers detailed information on its methodology and pre-processing stages. The answers were not linked to the respondents’ names being pre-processed by the PRC.

The exploratory data analysis (EDA) provides some initial insights into the dataset sample: The QKEY column is a unique identifier for respondents, ranging from 2 to 10,260, with a total count of 5,153 entries. This indicates that the dataset is a subset of a larger survey that approached both technological and science-related issues. TECH1\_W99: Responses suggest that the majority of participants (2627 out of 5153) perceive the impact of technology as “Mostly positive”, with a significant number (2036) seeing equal positive and negative effects. A smaller group



**Fig. 5** Education distribution of respondents.

(482) views it as “Mostly negative”. CNCEXC\_W99: The sentiment about concern vs. excitement is fairly even, with 2,267 respondents being “Equally concerned and excited” and 1,981 being “More concerned than excited”. A smaller portion of 889 is “More excited than concerned”. ALGFAIR\_W99: A plurality of respondents (2046) is “Not sure” about the fairness of algorithms, with almost equal numbers finding it “Not possible” (1580) and “Possible” (1480) for algorithms to be fair. F\_PARTYSUM\_FINAL: Political affiliations are split closely between “Dem/Lean Dem” (2589) and “Rep/Lean Rep” (2445), with a small number (119) not leaning towards any party or refusing to disclose. F\_INC\_SDT1: Income distribution varies, with the highest number of respondents (1427) reporting an income of “\$100,000 or more”. The next largest group earns “Less than \$30,000” (812), followed by a fairly even distribution across other income ranges. F\_IDEO: Ideological self-identification is mostly “Moderate” (1,853), with “Conservative” (1,344) and “Liberal” (913) also represented. Fewer respondents identify as “Very conservative” (541) or “Very liberal” (424).

This dataset is designed to capture a diverse array of opinions on technology, societal concerns, perceptions of fairness, and personal beliefs, alongside demographic information such as political affiliation and income. The dataset in focus offers a compiled array of survey responses, designed to capture a snapshot of public opinion on a spectrum of contemporary issues. The data is organized into 126 columns, each representing a unique question or data point, and spans 5153 rows, corresponding to individual participants. The survey’s breadth covers topics as varied as technological impact (TECH1\_W99), concerns and excitement about future developments (CNCEXC\_W99), and perceptions of algorithmic fairness (ALGFAIR\_W99). Additionally, it delves into personal beliefs and demographic details, including political leanings (F\_PARTYSUM\_FINAL), income brackets (F\_INC\_SDT1), and self-identified ideology (F\_IDEO). The columns cover responses to questions regarding technology, fairness of algorithms, discrimination, and personal beliefs. Moreover, the dataset includes demographic information such as political affiliation, income brackets, and frequency of internet use, among other variables. Given the presence of certain columns such as TECH1\_W99 and CNCEXC\_W99, it can be inferred that the survey explores attitudes towards technological advancements and their perceived impact on society. Additionally, columns like DISCRIM1\_a\_W99 to DISCRIM1\_f\_W99 suggest an assessment of perceived discrimination across various facets of life. The survey data is rich with individual viewpoints that are categorized to facilitate a nuanced analysis of public opinion. The diversity of topics within the dataset indicates a multifaceted approach to understanding complex societal issues,

ranging from technology and privacy to socio-political dynamics and personal ideology.

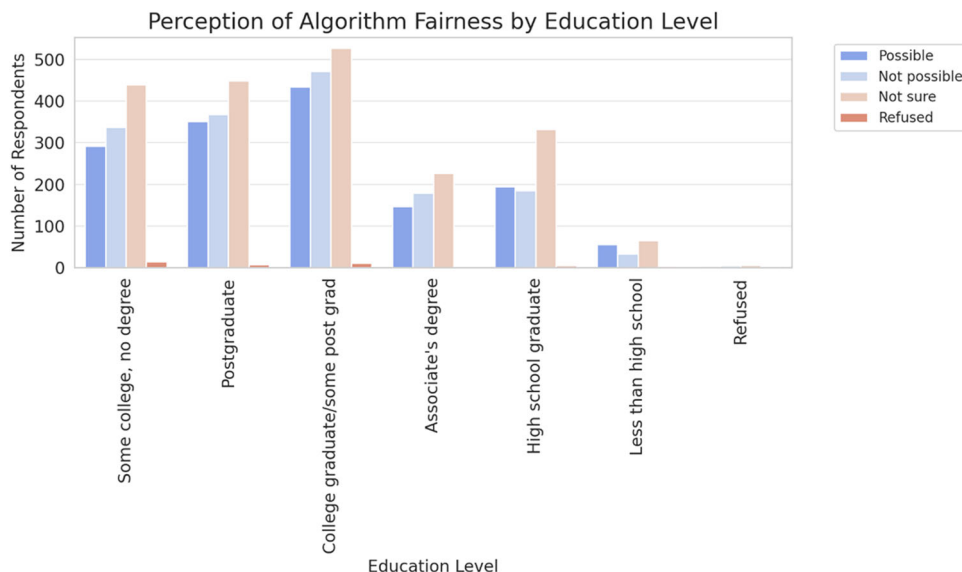
An EDA was performed, and it revealed a dataset that is complete and free from common data entry errors such as duplicates or missing values. This level of data integrity allows for a robust analysis of the survey responses. The data showcases a diversity of viewpoints, with no single perspective dominating the responses. For instance, opinions on technology’s impact range broadly from “Mostly positive” to “Mostly negative,” reflecting a spectrum of individual experiences and attitudes. The demographic information suggests a balanced representation of political affiliations and income levels, providing a rich ground for examining correlations between socioeconomic status and opinions on the surveyed topics. Similarly, the distribution of ideological self-identification points to a nuanced political landscape, which could be pivotal in interpreting the subtleties of the participant’s views on the fairness of algorithms and other technology-related issues.

Americans’ attitudes toward the large-scale use of facial recognition technology by law enforcement to monitor crowds and search for potential crime suspects lean more toward the positive side. Specifically, 46% of U.S. adults view this as a beneficial concept for society, while 27% consider it unfavorable, and another 27% are uncertain about it. Regarding the use of computer algorithms by social media companies to detect false information on their platforms, public opinion slightly favors it being a good idea for society (38%) rather than a bad one (31%). Ambivalence is another notable theme in the survey data: 45% indicate they feel both excitement and concern about the increased use of AI programs in daily life, while 37% express more concern than excitement, and 18% report feeling more excited than concerned.

The greatest concern relates to the loss of jobs, followed by the loss of privacy (as in Fig. 4). Top of FormBottom of Form This discussion contextualizes the survey data within the broader topic of AI acceptance, highlighting the complexity of public sentiment and the factors that influence it. The dataset encompasses a comprehensive collection of survey responses, potentially aimed at gauging public sentiment on a variety of topics.

### Distributions and cross-tabulation

*Education and AI technologies.* To focus on the education level of the respondents and analyze its relationship with the acceptance of AI and technology, the dataset contains two columns that relate to the education of the respondents: ‘F\_EDUCCAT’ and ‘F\_EDUCCAT2’. These columns contain categorical data reflecting the highest level of education attained by the survey participants. We analyze the responses in these columns and



**Fig. 6** Algorithm fairness and education.

cross-reference them with other columns related to attitudes toward technology and AI. This helps us understand if there are any discernible patterns or correlations between education levels and opinions on these topics. We start by examining the distribution of educational attainment among the respondents and then look at the relationship with their views on AI technology.

The analysis of the dataset with respect to the education of the respondents reveals the following education distribution: College Graduate or Higher 2620 respondents; Some College 1640 respondents; High School Graduate or Less 877 respondents; Refused 16 respondents. When broken down further ('F\_EDUC-CAT2'), we notice: College Graduate/Some Postgraduate 1443 respondents; Postgraduate 1177 respondents; Some College, No Degree 1083 respondents; High School Graduate 717 respondents; Associate's Degree 557 respondents; Less than High School 160 respondents; Refused 16 respondents (as in Fig. 5).

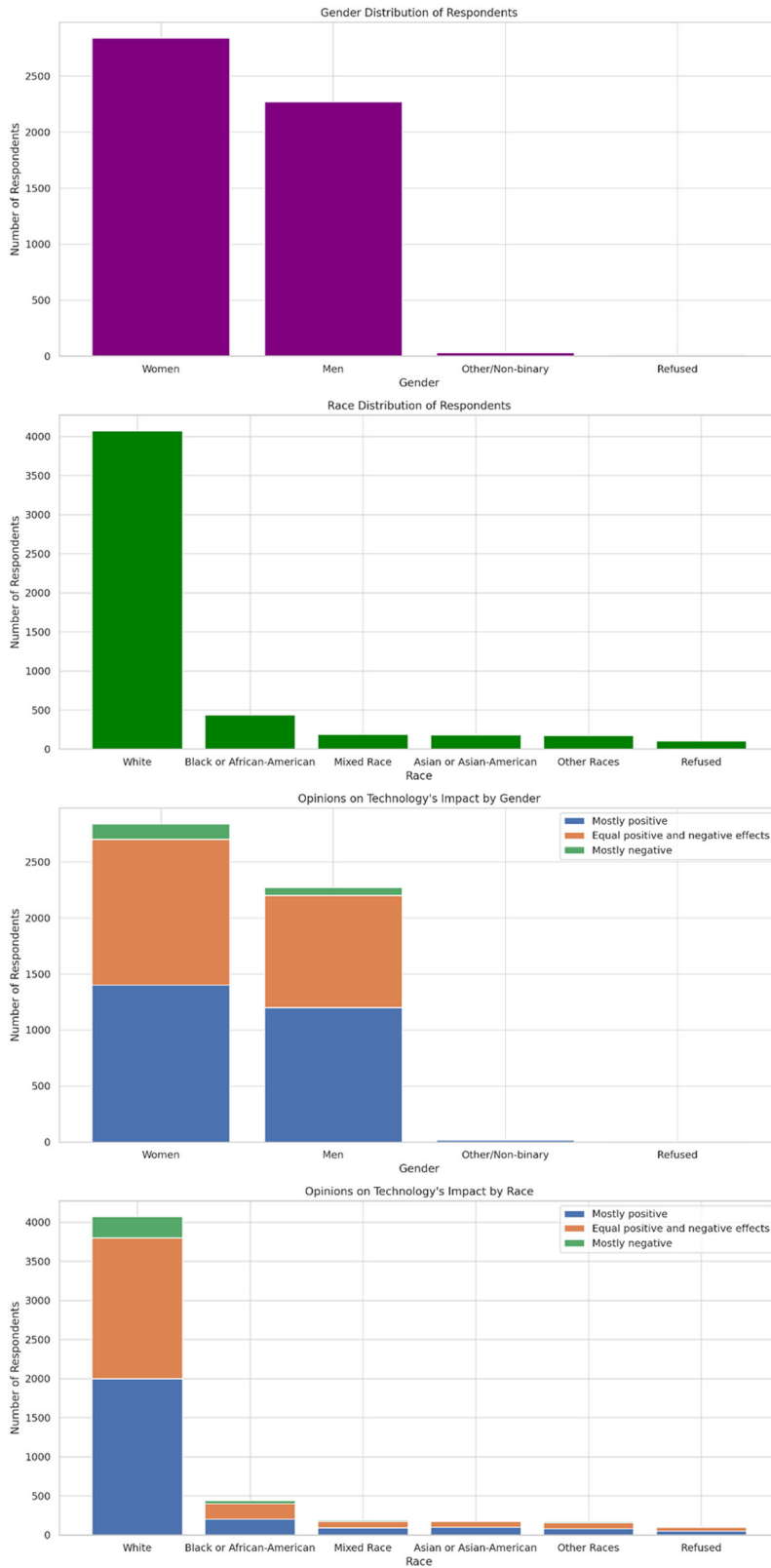
To determine opinion on technology's impact (TECH1\_W99), we cross-tabulate education levels with opinions on the impact of technology. Respondents with a higher education level (College Graduate+) tend to have a more positive view of technology, with 1,527 considering it "Mostly positive". Those with a high school education or less have a greater tendency to see technology's effects as "Equally positive and negative" (359) or "Mostly negative" (124). Some college education respondents also exhibit a balanced perspective, with significant numbers viewing technology's impact as both positive (736) and mixed (724). Looking at perceptions of algorithm fairness (ALGFAIR\_W99), college graduates and postgraduates show a higher degree of skepticism or uncertainty, with 976 "Not sure" about the fairness of algorithms and 839 finding it "Not possible" for algorithms to be fair. Those with less education also show uncertainty but with fewer respondents (398 "Not sure" for high school education or less). Across all education levels, a significant number of respondents are unsure about the fairness of algorithms, indicating a general ambiguity regarding this aspect of AI. As education often plays a role in technological acceptance and understanding, a bar plot for the opinion on algorithm fairness by the highest level of education attained is showcased in Fig. 6.

Figure 6 shows the perception of algorithm fairness by the respondents' education level. The education levels are on the x-axis, and the number of respondents is on the y-axis. The colors differentiate the respondents' views on whether they believe it is possible for algorithms to be fair, not possible, or if they are not

sure. These provide an image of how education correlates with opinions on technology's impact and perceptions of algorithm fairness, which are key indicators of AI acceptance.

*Age and AI technologies.* To focus on the age of the respondents, we analyze the distribution of respondents' ages and examine how different age groups' opinions vary on technology's impact, excitement and concerns about future developments, and perceptions of algorithmic fairness. The column 'F\_AGE-CAT' represents the age categories of the respondents. We further analyze the distribution of respondents' ages and then look at how these age groups relate to their views on technology's impact (TECH1\_W99), excitement and concerns about future developments (CNCEXC\_W99) and perceptions of algorithmic fairness (ALGFAIR\_W99). Age distribution is as follows: Ages 30–49 1,642 respondents; Ages 65+ 1,527 respondents; Ages 50–64 1,509 respondents; Ages 18–29 457 respondents; Refused to specify 18 respondents. Analyzing these age segments and opinions on technology's impact (TECH1\_W99), younger respondents (ages 18–29) show a more positive view of technology's impact, with the majority (237) considering it "Mostly positive". The 30–49 age group also leans toward a "Mostly positive" view (774) but with a substantial number seeing equal positive and negative effects (702). Older age groups (50–64 and 65+) show a balanced perspective, with a significant number viewing technology's impact as both positive (754 and 857, respectively) and mixed (603 and 546, respectively). Furthermore, analyzing the age groups and the concern vs. excitement about future developments (CNCEXC\_W99), the youngest cohort (18–29) exhibits a higher level of excitement than concern, with a notable proportion equally concerned and excited (215). As age increases, the level of concern seems to rise, with the 50–64 and 65+ groups showing the highest number of respondents more concerned than excited (643 and 651, respectively). Moreover, considering the perception of algorithm fairness (ALGFAIR\_W99), the following aspects were extracted. There is a fair degree of uncertainty across all age groups regarding the fairness of algorithms, with "Not sure" being a common response. The perception of algorithms as "Not possible" to be fair is consistently high across all age groups, indicating a widespread skepticism about AI fairness.

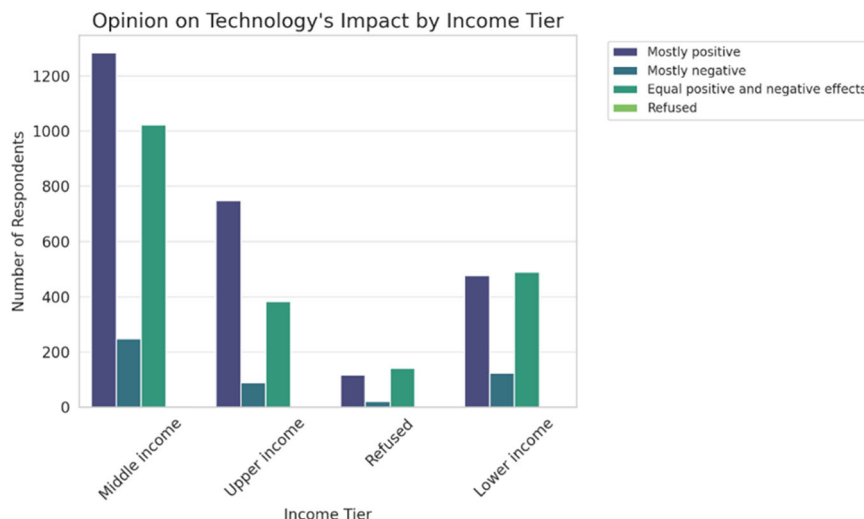
*Gender, race, and AI technologies.* To focus on the gender and race of the respondents and their relationship with attitudes



**Fig. 7** AI technology's impact on gender and race.

toward AI and technology, we analyze how opinions on technology's impact, excitement and concerns about future developments, and perceptions of algorithmic fairness vary by these demographic factors. The columns 'F\_GENDER', 'F\_RACECMB', and 'F\_RACETHNMOD' correspond to the respondents' gender

and race/ethnicity. First, we analyze the distribution of respondents' gender and race and then look at how these demographic groups relate to their views on technology's impact (TECH1\_W99), excitement and concerns about future developments (CNCEXC\_W99), and perceptions of algorithmic fairness



**Fig. 8** Opinions on technology’s impact categorized by income tier.

(ALGFAIR\_W99). The gender distribution reveals: Women 2,840 respondents; Men: 2,270 respondents; Other/Non-binary: 31 respondents; Refused: 12 respondents, while the race distribution distinguished among White: 4,072 respondents; Black or African-American 436 respondents; Mixed Race: 186 respondents; Asian or Asian-American 180 respondents; Other Races 177 respondents; Refused: 102 respondents.

Their opinions on technology’s impact (TECH1\_W99) reflect that both men and women predominantly see technology’s impact as either “Mostly positive” or “Equal positive and negative effects.” However, women are slightly more likely than men to see the impact as “Mostly negative”. Respondents who identify as Asian or Asian-American are most likely to view technology’s impact as “Mostly positive,” followed by White respondents. Regarding the concern vs. excitement about future developments (CNCEXC\_W99), women are slightly more likely than men to be “More concerned than excited” about future technological developments. Black or African-American respondents are the most likely to be “More concerned than excited,” whereas Asian or Asian-American respondents are more likely to be “More excited than concerned.” The perception of algorithm fairness (ALGFAIR\_W99) that a higher proportion of women than men believe it is “Not possible” for algorithms to be fair, whereas men are more likely to believe it is “Possible”. White respondents show the greatest skepticism about the possibility of fair algorithms, with the highest count of respondents who believe it is “Not possible” (as in Fig. 7).

*Religion, ideology, and AI technologies.* To analyze the survey data based on the religion and ideology of the respondents, we explore how different religious and ideological groups perceive the impact of technology, their concerns and excitement about future developments, and their views on the fairness of algorithms. The columns ‘F\_RELIG’ and ‘F\_RELIGCAT1’ correspond to the respondents’ religious affiliations, and ‘F\_PARTYSUMIDEO\_FINAL’ and ‘F\_IDEO’ to their ideological leanings. We analyze the distribution of respondents’ religion and ideology and then look at how these demographic groups relate to their views on technology’s impact (TECH1\_W99), excitement and concerns about future developments (CNCEXC\_W99), and perceptions of algorithmic fairness (ALGFAIR\_W99).

Examining the distribution of religion among the respondents, we find that the largest group of respondents identified as Protestant (2,204), followed by Roman Catholic (1,075), and

those with no particular religious affiliation (826). Smaller religious groups included Agnostics (307), Atheists (279), and other faiths like Jewish (108), Mormon (100), Buddhist (34), Hindu (26), and Muslim (24). Regarding ideology distribution, Moderates represent the largest ideological group (1,853), followed by Conservatives (1,344), Liberals (913), Very Conservatives (541), and Very Liberals (424). A small number refused to specify their ideology (78). Their opinions on technology’s impact (TECH1\_W99) reveal that across most religious groups, there’s a balance between seeing technology’s effects as equally positive and negative. However, Atheists and Agnostics tend to have a more positive view. Ideologically, Moderates and Conservatives are more inclined to see technology’s impact as equally positive and negative, while Liberals and Very Liberals are more likely to see it as mostly positive. Concerns and excitement about future technological developments (CNCEXC\_W99) are distributed relatively evenly across religious affiliations, with no significant outliers. Ideologically, Moderates and Conservatives tend to be more concerned than excited, while Liberals and Very Liberals show more excitement than concern. The perception of algorithm fairness (ALGFAIR\_W99) reflects that skepticism about the possibility of algorithms being fair is high across all religious groups, with Protestants and Roman Catholics having the highest number of respondents who think it is not possible. Ideologically, Conservatives and Very Conservatives are more likely to believe that it is not possible for algorithms to be fair, while Liberals and Very Liberals are more divided between “Not possible” and “Not sure.”

*Income and AI technologies.* To analyze the survey data based on the income levels of the respondents, represented by the columns ‘F\_INC\_SDT1’ and ‘F\_INC\_TIER2’, we look into how income correlates with attitudes toward technology, concerns about future developments, and views on the fairness of algorithms.

A summary of the analyses focusing on the income levels (‘F\_INC\_SDT1’ and ‘F\_INC\_TIER2’) of the respondents is provided. The highest number of respondents have an income of \$100,000 or more (1427). The next largest group earns less than \$30,000 (812). Other income ranges have smaller counts, with the least being those earning \$80,000 to less than \$90,000 (277). Some respondents refused to disclose their income (244). Middle income is the largest group (2560). Upper income is next (1222), followed by lower income (1090). A number of respondents refused to disclose their income tier (281).

Their opinions on technology’s impact (TECH1\_W99) show that in the ‘F\_INC\_SDT1’ category, those earning \$100,000 or more have the highest number of respondents viewing technology’s impact as “Mostly positive” (846), while also having a considerable portion seeing it as “Equal positive and negative effects” (481). The ‘F\_INC\_TIER2’ category shows similar patterns, with upper-income respondents more likely to view technology’s impact as “Mostly positive” compared to other income tiers. Regarding concern vs. excitement about future developments (CNCEXC\_W99), for ‘F\_INC\_SDT1’, lower-income respondents have a higher count of being “More concerned than excited” (445) compared to those excited (176), whereas upper-income individuals are more balanced with a considerable number being “More excited than concerned” (268). ‘F\_INC\_TIER2’ analysis shows that middle-income respondents are more evenly split between concern and excitement, with a notable number being “Equally concerned and excited” (1119). Concerning the perception of algorithm fairness (ALGFAIR\_W99), respondents across all income levels in the ‘F\_INC\_SDT1’ category show skepticism about the possibility of algorithms being fair, with high-income earners being the most doubtful (440 believe it’s “Not possible”). The ‘F\_INC\_TIER2’ category also reflects skepticism across all tiers, with middle-income individuals showing the highest number of respondents who think it is “Not possible” for algorithms to be fair (812).

In Fig. 8, the opinion on AI technology’s impact on income is showcased.

The income tiers are represented on the x-axis, and the number of respondents is on the y-axis. Different colors indicate whether respondents view technology’s impact as mostly positive, mostly negative, or equally positive and negative.

**Clustering using K-means.** The columns QKEY, SMALG3\_W99, F\_REG, and F\_INTFREQ that contain identifiers or strings are removed and the values are scaled using a standard scaler, which centers the data around zero with a standard deviation of one. The majority of data points are grouped into Cluster 4 (using 5 clusters setting) and Cluster 1 (using 3 clusters), while the other clusters contain significantly fewer points.

After rerunning the K-means clustering with 5 and 3 clusters, the respondents’ distribution is presented in Table 2.

With 3 clusters, the majority of the respondents are in Cluster 1, while Clusters 0 and 2 are much smaller. To visualize the clusters, a dimensionality reduction technique such as Principal Component Analysis (PCA) is applied to reduce the dataset to 2 dimensions for a 2D plot, or to 3 dimensions for a 3D plot (as in Fig. 9). The silhouette score, with the clusters assigned by K-Means, is approximately 0.828. This score suggests a good separation between the clusters, indicating that the clustering configuration is well-defined for the sampled data. In our research, we explored various datasets that contained both numeric and non-numeric variables. We applied multiple clustering techniques, including K-prototypes, DBSCAN, and hierarchical and spectral clustering. However, we found that the

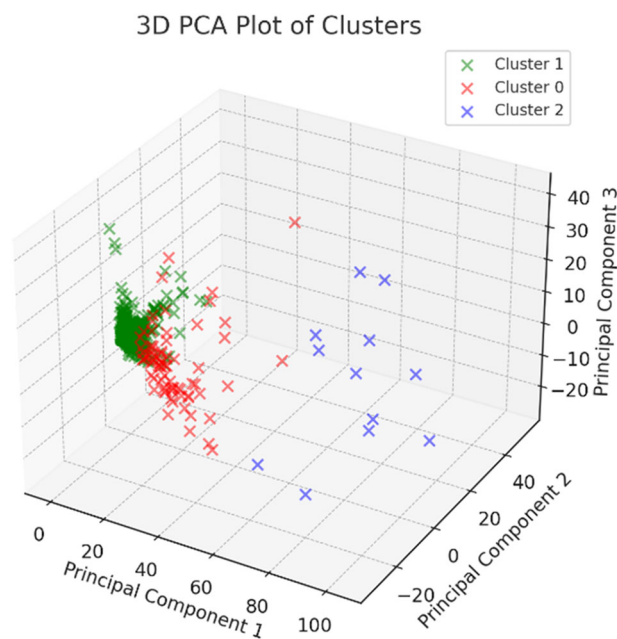
K-means clustering algorithm yielded the most effective results in terms of segmentation quality.

Each color represents one of the three clusters: Red: Cluster 0; Green: Cluster 1; Blue: Cluster 2. ANOVA is performed for each feature across the three clusters. For most features, the *p*-values are significantly low (often below 0.05), suggesting that there are statistically significant differences in the mean values of those features across the clusters. This indicates that the clustering has effectively separated the respondents into groups with distinct characteristics. ANOVA is useful for identifying features that contribute significantly to clustering, which can help in interpreting the clusters or in feature selection for further modeling.

The centroids of the clusters represent the mean value for each feature within the cluster. These are the centroids in the original feature space for each of the three clusters: Cluster 0: This cluster has higher average values for several features compared to Cluster 1 but is generally lower than Cluster 2. Features like TECH1\_W99, CNCEXC\_W99, and ALGFAIR\_W99 are notably higher than Cluster 1 but less than Cluster 2. This cluster might represent a middle ground between the other two clusters in terms of features; Cluster 1: This cluster tends to have the lowest average values across most features. It suggests that this cluster may represent a group with less pronounced characteristics in terms of the variables measured or a baseline group in comparison to the other clusters; Cluster 2: This cluster has the highest average values for the features, particularly for TECH1\_W99, CNCEXC\_W99, ALGFAIR\_W99, DISCRI-M1\_a\_W99, and DISCRIM1\_b\_W99.

This indicates that this cluster’s characteristics are quite distinct and more pronounced in comparison to the other two clusters. These interpretations are general and based on the average values of the features; individual cluster members may vary. The centroids can help in understanding the profiles of the respondents in each cluster, which could be useful for targeted analysis, marketing, or further study. The feature with the most variation across the centroids of the three clusters is DCARS11\_b\_W99: A feature with high variance among the clusters, suggesting that respondents have significantly different values for this feature across clusters; DCARS11\_a\_W99: Similar

No. of clusters/cluster characteristics	5 clusters	3 clusters
Cluster 0	246 respondents	72 respondents
Cluster 1	104 respondents	5069 respondents
Cluster 2	10 respondents	12 respondents
Cluster 3	28 respondents	
Cluster 4	4765 respondents	



**Fig. 9** Clusters 3D visualization after using PCA.



**Table 3 Feature-specific analysis by cluster.**

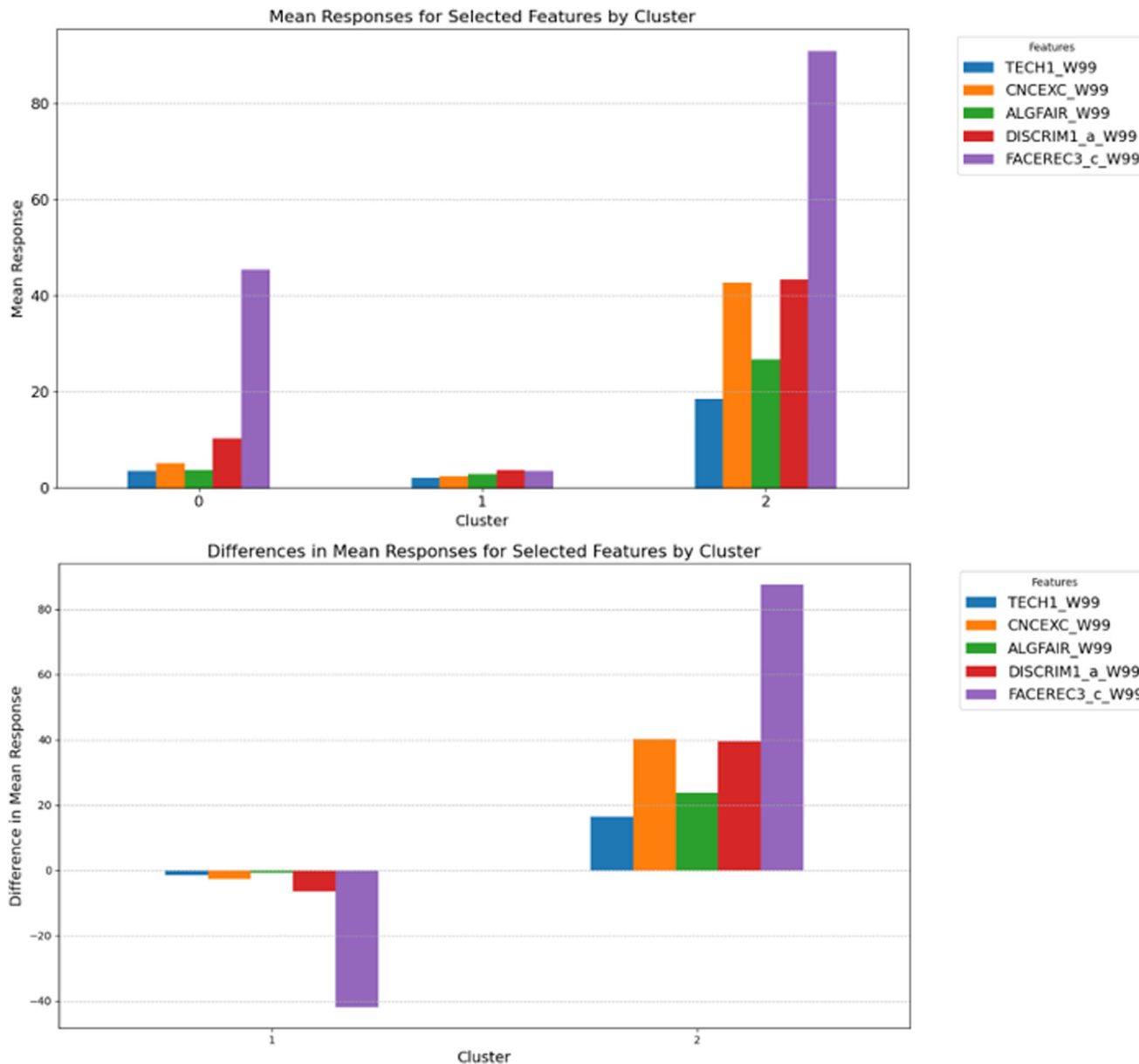
Feature	Cluster
TECH1_W99 (Opinions on technology):	Cluster 0 has a moderate mean value of approximately 3.5. Cluster 1 has a lower mean value of about 2, suggesting less enthusiasm or engagement with technology. Cluster 2 has a significantly higher mean value of over 18, indicating a strong sentiment towards technology, which could be positive or negative.
CNCEXC_W99 (Views on the concept of excellence):	Cluster 0 has an average mean value of around 5. Cluster 1 has a lower mean value of about 2.4, which could indicate a different standard or priority regarding excellence. Cluster 2 has a high mean value of approximately 42.7, suggesting that the concept of excellence is particularly significant to this group.
ALGFAIR_W99 (Views on the fairness of algorithms):	Cluster 0 has a mean value of around 3.7, indicating moderate views on algorithmic fairness. Cluster 1 has a lower mean value of about 2.9, suggesting less concern or awareness of algorithmic fairness. Cluster 2 shows a much higher mean value of about 26.8, which could reflect a strong opinion on the fairness of algorithms, possibly seeing them as often unfair.
DISCRIM1_a_W99 (Views on discrimination):	Cluster 0 shows a mean value of approximately 10.3, suggesting some recognition of discrimination issues. Cluster 1 has a lower mean value of about 3.7, indicating that discrimination might not be perceived as a significant issue among this group. Cluster 2 has a high mean value of 43.25, which could indicate a heightened awareness or concern regarding discrimination.
FACEREC3_c_W99 (Opinions on potential racial bias in the use of facial recognition):	Cluster 0 has a mean value of around 45.3, possibly acknowledging the potential for racial bias. Cluster 1 has a much lower mean value of about 3.5, which could suggest less concern about this issue. Cluster 2 shows a very high mean value of approximately 90.9, indicating strong concerns about racial bias in facial recognition technology.

to the previous feature, indicating that this feature strongly differentiates the clusters. The feature-specific analysis for the selected features shows the following mean responses for each cluster are presented in Table 3.

From these observations, we notice that Cluster 2 is characterized by a strong engagement with the survey’s topics, particularly concerning technology and societal issues. Cluster 0 appears to have moderate views, while Cluster 1 seems less engaged or concerned about these topics. To understand the profiles of the respondents in each cluster, we can look at the centroids for each cluster and interpret them based on the features that stand out. A general profile summary based on the centroids provided earlier is provided: Cluster 0 - The Moderates. It has moderate average values for most features, neither extremely high nor low. This could indicate a more balanced or moderate perspective on various issues. Respondents in this cluster might not hold extreme views and could represent a general population in terms of the survey’s topics; Cluster 1—The Baseline. It shows the lowest average values for most of the features. This suggests that respondents in this cluster may be less engaged or have less intense opinions on the surveyed topics. This could be the most common profile, possibly representing the average respondent with no strong inclinations toward the topics surveyed; Cluster 2—The Intensives. It exhibits the highest values across nearly all features, indicating strong opinions, higher engagement, or more pronounced behaviors in the topics covered by the survey. Respondents in this group might be more passionate, involved, or affected by the survey’s topics, which could make them stand out in specific areas.

For a more detailed profile, we map the questions’ code with their description. With that information, we can interpret the clusters’ profiles more precisely in the context of the survey and refine our understanding of the three clusters: Cluster 0 - The

Cautious Moderates. This group appears to be moderate in their acceptance of new technologies like driverless vehicles and facial recognition, possibly favoring some regulations like dedicated lanes for driverless cars (‘DCARS11\_a\_W99’) and identifiable labels (‘DCARS11\_b\_W99’). They might hold moderate views on the role of companies and federal agencies in setting standards for facial recognition technology used by the police (‘FACEREC6\_b\_W99’ and ‘FACEREC6\_a\_W99’). Their opinions on social media’s role in filtering false information (‘SMALG4\_d\_W99’ and ‘SMALG12\_W99’) might be balanced, suggesting they see both benefits and drawbacks in technology’s role in public discourse. Cluster 1—The Baseline or Technology Skeptics. Members of this cluster could be less inclined towards the adoption of new technologies, perhaps less excited about the prospect of enhancements like advanced vision (‘POSNEGHE\_e\_W99’), or less accepting of driverless vehicles under various conditions (‘DCARS11’ series). They might also be more skeptical of the effectiveness of computer programs in filtering false information on social media (‘SMALG4\_d\_W99’ and ‘SMALG12\_W99’). This cluster could be more conservative in its views on how facial recognition technology is regulated and used by law enforcement (‘FACEREC6’ series). Cluster 2—The Technologically Enthusiastic or Concerned. This group likely has strong opinions about technology, possibly showing a high level of acceptance for driverless vehicles (‘DCARS11’ series) and new human abilities through technological advancements (‘POSNEGHE\_e\_W99’). They may advocate for a significant role of both companies and government agencies in setting standards for the use of facial recognition technology by the police (‘FACEREC6\_b\_W99’ and ‘FACEREC6\_a\_W99’). Concerns about racial bias in the use of facial recognition technology by the police (‘FACEREC3\_c\_W99’) and strong views on the role of technology in combating false information



**Fig. 10 Mean response for selected features by clusters (up).** Differences in mean response for selected features by clusters (down).

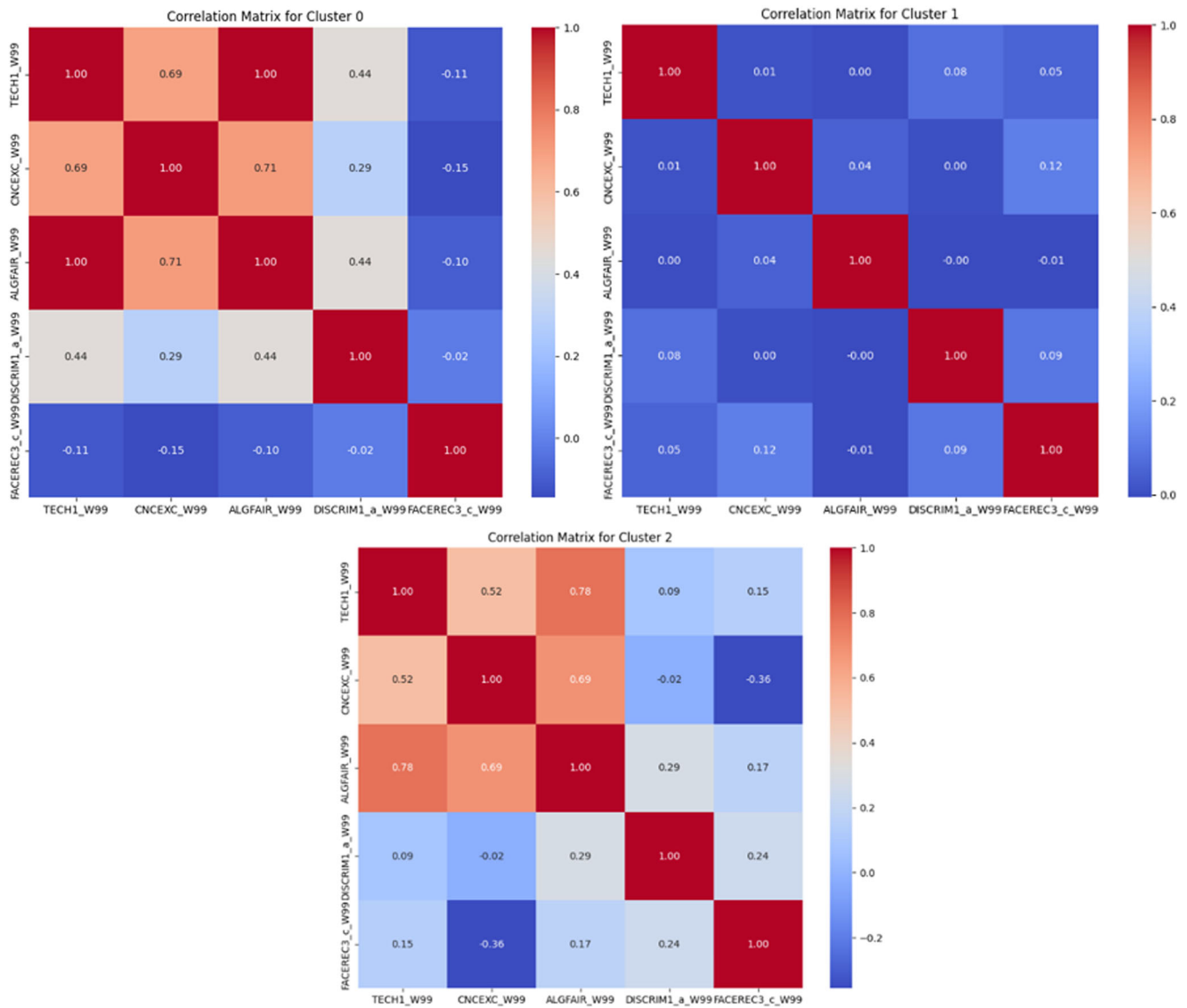
on social media ('SMALG4\_d\_W99' and 'SMALG12\_W99') could also be defining characteristics.

These profiles are inferred from the centroids, which represent the average values for each cluster, and are based on the assumption that higher values indicate greater acceptance or concern, as suggested by the descriptions. Each cluster seems to have a unique stance on technology and its implications for society. Corroborating the below features specifically with their description in the survey, significant differences suggest that they are the most influential in distinguishing the groups defined by the clustering algorithm.

- FACEREC6\_b\_W99: Another distinctive feature among the clusters, possibly relating to a respondent's recognition or attitude towards something specific.
- FACEREC6\_a\_W99: As with the "b" variant of this feature, it seems to be a strong differentiator between clusters.
- FACEREC3\_c\_W99: This feature also shows considerable variance and may represent a unique aspect of the respondent's profile.

- SMALG4\_d\_W99: Indicates a distinctive characteristic or opinion among the clusters.
- POSNEGHE\_e\_W99: Could represent a particular positive or negative health aspect that differs significantly among clusters.
- DCARS11\_c\_W99: Like the other "DCARS11" features, this one is also a strong differentiator.
- SMALG12\_W99: Another feature that significantly varies among clusters.

The features with the most variation across clusters relate to opinions on emerging technologies such as driverless vehicles, facial recognition, and computer programs to detect false information. They can be used to understand the clusters' general attitudes toward technology and regulation. For example, a cluster with high values in DCARS11\_b\_W99 and DCARS11\_a\_W99 may consist of respondents who are more accepting of driverless vehicles under certain conditions. Conversely, a cluster with low values in FACEREC3\_c\_W99 might be less concerned about potential racial biases in the use of facial



**Fig. 11 Correlation matrices for the three clusters.** Correlation matrix for each cluster.

recognition technology by the police. For a more nuanced understanding, it is useful to visualize the distribution of these key features within clusters to see how the responses vary (as in Fig. 10). Figure 10 and Table 3 show the differences in mean responses between the clusters for the selected features. As of TECH1\_W99 (Opinions on technology), the mean response for Cluster 1 is approximately 1.5 points lower than that of Cluster 0, indicating less enthusiasm or engagement with technology. Cluster 2's mean response is significantly higher than Cluster 1's by over 16 points, reflecting a much stronger sentiment towards technology among respondents in Cluster 2. Regarding CNCEXC\_W99 (Views on the concept of excellence), the mean response for Cluster 1 is about 2.6 points lower than Cluster 0's, suggesting different standards or priorities regarding excellence. Cluster 2's mean response is dramatically higher than Cluster 1's by about 40 points, indicating that the concept of excellence is significantly more important to respondents in Cluster 2. In terms of ALGFAIR\_W99 (Views on the fairness of algorithms), Cluster 1 has a mean response that is about 0.78 points lower than Cluster 0, suggesting less concern or awareness of algorithmic fairness. The mean response for Cluster 2 is about 23.8 points higher than Cluster 1, which could indicate a strong

opinion on the fairness of algorithms, possibly perceiving them as often unfair.

As of DISCRIM1\_a\_W99 (Views on discrimination), Cluster 1 shows a mean response that is about 6.6 points lower than Cluster 0, indicating that discrimination might not be perceived as a significant issue among this group. Cluster 2 has a mean response that is about 39.6 points higher than Cluster 1, which could signify a heightened awareness or concern regarding discrimination. Related to FACEREC3\_c\_W99 (Opinions on potential racial bias in the use of facial recognition), the mean response for Cluster 1 is about 41.9 points lower than Cluster 0, suggesting significantly less concern about potential racial bias.

Cluster 2's mean response is about 87.4 points higher than Cluster 1, indicating strong concerns about racial bias in facial recognition technology. These differences underline the distinct profiles of each cluster: Cluster 0: Moderate views on technology and social issues; Cluster 1: Generally lower concern or engagement with the issues. Cluster 2: Strongly engaged and often with pronounced concern or optimism regarding technology and its societal implications.

The correlation matrices for each cluster are displayed above in Fig. 11. Regarding Cluster 0, there is a moderate positive

correlation between TECH1\_W99 (Opinions on technology) and ALGFAIR\_W99 (Views on the fairness of algorithms). This suggests that within this cluster, individuals who have positive sentiments toward technology may also perceive algorithms as fair. A moderate positive correlation is also observed between ALGFAIR\_W99 and FACEREC3\_c\_W99 (Opinions on potential racial bias in the use of facial recognition), indicating that those concerned with algorithm fairness might also be concerned about racial bias in facial recognition. Cluster 1 shows a strong positive correlation between TECH1\_W99 and ALGFAIR\_W99, suggesting that attitudes towards technology are strongly linked to perceptions of algorithm fairness within this group. DISCRIM1\_a\_W99 (Views on discrimination) also shows a moderate correlation with ALGFAIR\_W99, indicating that those who are concerned about discrimination also tend to be concerned about the fairness of algorithms. As for Cluster 2, a significant positive correlation exists between TECH1\_W99 and ALGFAIR\_W99, and between TECH1\_W99 and DISCRIM1\_a\_W99, which suggests that in this cluster, individuals who are more optimistic or concerned about technology are also more likely to be concerned about algorithm fairness and discrimination. FACEREC3\_c\_W99 shows moderate to strong correlations with all other features, highlighting that concerns about racial bias in facial recognition are associated with more general concerns about technology, discrimination, and the fairness of algorithms.

These correlations provide insights into the shared sentiments and concerns within each cluster, which can be especially valuable for understanding the nuances of each group's attitudes toward technology and social issues.

**ANOVA.** We proceed with the ANOVA test, considering the gender variable. The 'DISCRIM1\_a\_W99' through 'DISCRIM1\_f\_W99' columns contain ordinal responses that are suitable for an ANOVA analysis when encoded numerically. These columns represent perceptions on various aspects of discrimination and can be encoded as follows: 'Very well' = 4, 'Somewhat well' = 3, 'Not too well' = 1, 'Not at all well' = 0. We select one of these columns, encode the responses, and then perform the ANOVA test considering gender. We encode the 'DISCRIM1\_a\_W99' responses. The responses in the 'DISCRIM1\_a\_W99' variable are investigated by comparing across the different gender groups as indicated by the 'F\_GENDER' column. The result of the ANOVA test yielded an *F*-statistic of approximately 6.017 and a *p*-value of approximately 0.000439.

The *p*-value is less than 0.05, which is the common threshold for statistical significance. This suggests that there are statistically significant differences in the perceptions of discrimination (as indicated by 'DISCRIM1\_a\_W99') across different gender groups in the dataset. The *F*-statistic from the ANOVA test indicates the ratio of the variance between the group means to the variance within the groups. A higher *F*-statistic generally suggests a greater disparity between the groups. In our case, the *F*-statistic of approximately 6.017 suggests that there is a difference between the means of the groups. The *p*-value tells us about the significance of the differences that we observe between the groups' means. A *p*-value less than 0.05 (which is commonly used as a threshold for statistical significance) suggests that it is very unlikely that the observed differences in group means are due to random chance alone. In the context of our analysis, with a *p*-value of approximately 0.000439, we can conclude with high confidence that there is a significant difference in the way different genders perceive discrimination ('DISCRIM1\_a\_W99').

To perform ANOVA considering the 'TECH1\_W99' variable, which refers to the respondents' views on whether technology has had a mostly positive or mostly negative effect on society,

alongside the ideology of respondents, we perform an ANOVA to see if there is a significant difference in the views on technology's impact on society across different ideological groups. The ANOVA test reflects the respondents' views on technology's impact on society, considering their ideological orientation ('F\_IDEO'). The results yield an *F*-statistic of approximately 15.674 and a *p*-value of approximately  $9.43 \times 10^{-13}$ . The *p*-value is much lower than the standard threshold of 0.05, indicating that there is a statistically significant difference in how respondents of different ideological backgrounds perceive the impact of technology on society. With such a low *p*-value, we can confidently reject the null hypothesis that there are no differences among the ideological groups' opinions regarding technology's impact. In other words, people's political ideology appears to have a significant association with their perceptions of whether technology has a mostly positive or mostly negative effect on society.

The results of the Tukey's HSD test comparing the means of the 'TECH1\_W99\_encoded' variable across different ideological groups ('F\_IDEO') provide the following insights:

1. Very Conservative vs. Conservative: There is a statistically significant difference, with a mean difference of 0.1078. The confidence interval does not include zero (0.0171 to 0.1984), indicating that Very Conservative respondents are slightly more likely to view technology's impact on society as more negative compared to Conservative respondents.
2. Very Conservative vs. Moderate: The difference is larger and significant, with a mean difference of 0.1473. This suggests that Moderate respondents are more likely to view technology's impact more positively than Very Conservative respondents.
3. Very Conservative vs. Liberal and Very Conservative vs. Very Liberal: Both comparisons show significant differences with Liberals and Very Liberals viewing technology's impact even more positively than Very Conservative respondents.
4. Conservative vs. Moderate: There is no statistically significant difference in their views.
5. Conservative vs. Liberal and Conservative vs. Very Liberal: Conservatives view technology's impact as more negative compared to Liberal and Very Liberal respondents, with significant mean differences.
6. Moderate vs. Liberal and Moderate vs. Very Liberal: Moderates are slightly more negative in their views compared to Liberals and Very Liberals, with significant mean differences.
7. Liberal vs. Very Liberal: There is no statistically significant difference between these two groups in their views of technology's impact.

From these results, it shows that ideology is a strong factor in how respondents perceive the impact of technology on society. Generally, as we move from Very Conservative to Very Liberal, the perception of technology's impact becomes increasingly more positive.

To perform ANOVA with the variables 'CNCEXC\_W99' (which relate to a specific concept or concern/excitement), 'F\_IDEO' (ideology), and income features 'F\_INC\_SDT1' and 'F\_INC\_TIER2', we grouped the variables and obtained the following results: for the 'F\_INC\_SDT1' income variable combined with 'F\_IDEO': *F*-statistic: 4.9488, *p*-value:  $3.63 \times 10^{-24}$ . For the 'F\_INC\_TIER2' income tier variable combined with 'F\_IDEO': *F*-statistic: 13.3624, *p*-value:  $1.13 \times 10^{-31}$ . The *p*-values are significantly below the 0.05 threshold, indicating that there are statistically significant differences in respondents' excitement or concern regarding the issue in question across different combinations of income levels

**Table 4 Performance metrics.**

Set/metric	Accuracy [%]	F1 score	AUC score
Testing set	99.48	0.994	0.9998
Out-of-sample set	99.09	0.989	0.9988

and ideological orientations. This suggests that both a person's income and their ideological beliefs have a significant association with their level of excitement or concern regarding technology. Given the magnitude of the p-values, this effect is very unlikely to be due to chance.

**Predicting clusters.** The clustering process introduced an additional column in the dataset, which serves as the target variable for predictive analysis. To perform clustering, we utilized the original survey dataset. Following this, for the predictive modeling, we incorporated the newly added column and then executed a splitting procedure. This procedure divided the dataset into three subsets: training, testing, and validation. These subsets facilitate the development and evaluation of predictive models. The Random Forest classifier was trained on the training set (70%) and evaluated on both the testing (15%) and out-of-sample (15%) sets. The accuracy, F1 score, and AUC score results are showcased in Table 4. These results show that the classifier performs very well on both unseen parts of the original dataset, indicating good generalization. The F1 scores for the Random Forest classifier on the testing and out-of-sample sets are consistent with the high accuracy we observed, indicating a strong performance of the classifier in predicting the clusters across both sets.

The F1 score considers both precision and recall, so these high values suggest that the classifier has a good balance between these metrics. These AUC scores are very close to 1, which indicates excellent performance in distinguishing between the different clusters. The model has strong classification capabilities across all classes in both the testing and out-of-sample sets.

For the purpose of predicting clusters, we opted for the Random Forest algorithm due to its renowned efficacy and the impressive results it delivered for our classification issue. Although we also experimented with Linear Regression (F1 score 0.87 for out-of-sample set) and Decision Tree algorithms (F1 score 0.89 for out-of-sample set), the F1 score achieved with RF was superior to those obtained with these algorithms.

Given the excellent performance of the Random Forest model, we chose not to explore additional advanced algorithms like XGBoost (XGB) or LightGBM (LGBM), which are known to potentially yield good results as well as our primary goal was to derive insights from the data and achieve accurate cluster predictions, which we successfully accomplished with Random Forest.

Addressing the limitations and challenges of using Random Forest in predictive modeling is significant for maintaining a robust analytical approach. One of the primary concerns with Random Forest, as with any machine learning algorithm, is the risk of overfitting. This occurs when the model becomes too complex, capturing noise in the data rather than underlying patterns. It performs well on the training data but poorly on unseen data. To mitigate overfitting, we assessed the model's performance on different subsets of the data, ensuring that it generalizes well to new data.

## Discussion

Coming back to our research questions "How do different groups of individuals perceive AI technologies such as social media, facial

recognition, and driverless cars, and what are their primary concerns, and how can these perceptions and concerns be predicted for new instances of individuals?", we discovered that the primary worries among respondents centered on job loss and privacy breaches due to heightened surveillance. Conversely, the most significant source of enthusiasm was the potential for technology to enhance the quality of life. From cross-tabulation the following aspects are relevant. There is a correlation between higher education levels and a more positive perception of technology's impact. This could suggest that education plays a role in how individuals understand and appreciate the complexities and benefits of technology. Education also seems to correlate with skepticism about the fairness of algorithms, with more educated respondents showing higher levels of doubt or uncertainty. This may be due to a greater awareness of the issues surrounding AI and technology among those with more education. Across the board, there is a notable level of uncertainty about the fairness of algorithms, which might reflect the public's overall wariness of AI decision-making processes and the need for greater transparency and education on this topic.

The younger respondents are more optimistic about the impact of technology, while there is a more cautious view among older respondents. This could be due to generational differences in technology adoption and familiarity. Concerns about technological developments increase with age. This could reflect the greater potential impact of technology on the livelihoods and lifestyles of older generations, or a more conservative approach to new developments. Skepticism about algorithmic fairness is prevalent across all age groups. However, younger respondents appear slightly more optimistic about the potential for algorithms to be fair. This age-based analysis of the survey data indicates that while there is general optimism about technology among the younger demographic, there is also a tangible concern about the implications of technological advancements, particularly regarding the fairness of algorithms. The data also suggests that education efforts around technology and AI might need to be tailored to address the specific concerns and levels of familiarity within different age groups.

When analyzing technology concern levels based on gender, our findings indicated a modest gender disparity in perceptions towards technology. Women, in particular, tended to express greater concern about the adverse effects of technology and the fairness of algorithms. This variation in perspective could be attributed to different experiences with technology between genders, or a heightened awareness among women about issues such as algorithmic bias, which might impact women disproportionately.

Racial differences in perceptions of technology are also evident. Asian or Asian-American respondents tend to have a more positive view of technology's impact and future developments, while Black or African-American respondents show more concern. This may be influenced by socioeconomic factors or different lived experiences with technology. Across genders, there is a level of skepticism about the fairness of algorithms, with a higher proportion of women expressing doubt. When broken down by race, White respondents are the most skeptical, which might suggest a heightened awareness or concern about the ethical implications of AI within this demographic.

Regarding religion and AI technologies, religious affiliation does not show a strong correlation with views on technology's impact, but there is a tendency among non-religious groups (Atheists and Agnostics) to view technology more positively, while ideological orientation appears to influence perceptions of technology, with Moderates displaying a balanced view, and Liberals leaning towards a more positive outlook. Conservatives are more skeptical about the fairness of algorithms, which might

reflect a broader caution towards technology governance and ethical considerations. Across both religious and ideological lines, there's a notable skepticism about the fairness of algorithms, suggesting a common concern about AI ethics and the potential for bias. These findings indicate that demographic factors such as gender and race can influence perceptions of technology. Tailoring communication to address the unique concerns of different religious and ideological groups could foster greater understanding and acceptance of AI technologies.

Regarding income and views on technology, higher-income respondents tend to have a more positive view of technology's impact. This could be related to greater access to technology and its benefits among those with higher incomes. Concerns about future developments tend to decrease as income increases, with lower-income individuals more worried than their higher-income counterparts. Financial stability might influence the level of concern regarding the implications of technological changes. Across all income levels, there is skepticism about the fairness of algorithms. However, this skepticism is most pronounced among the middle and upper-income tiers, which might indicate a higher awareness of algorithmic biases and fairness issues among these groups. Income level is a significant factor in shaping individuals' perceptions of technology. Those with higher incomes may feel more secure and optimistic about the benefits of technology, while those with lower incomes may have more immediate concerns related to financial security and job displacement due to technological advancements. Addressing these disparities could be key to fostering a more inclusive technological future.

## Conclusion

In this paper, we propose a descriptive-predictive hybrid model that includes a) EDA, checking and handling missing values, graphically visualizing, cross-tabulation to understand patterns in data and correlations between demographics and the main features, identifying the most significant features, b) clustering, analyzing clusters and respondents' profile with PCA and ANOVA including statistical tests (Tukey's HSD test), and finally c) classification to predict clusters.

The proposed model was applied to a complex survey that consists of 5153 records and 126 columns that represent the encoded questions with their answers provided by people from U.S.A. Questions refer to the AI technology impact, concerns, and excitement regarding AI, positive and negative perceptions, fairness of algorithms, discrimination issues, face recognition, and social media usage of AI, and driverless cars. The findings from this survey (presented at length in the results and discussion sections) resonate with the current discourse on AI, where enthusiasm for technological progress contends with valid concerns over privacy, job displacement, and the opacity of AI systems. The data signifies a society that is at a crossroads with AI acceptance, eager to harness its potential while acutely aware of the need for ethical guardrails.

The practical implications derived from the cluster analysis of survey respondents' engagement with technology issues may inform targeted communication strategies, policy development, and educational initiatives. The following four practical applications are identified: (1) Communication strategies—understanding the distinct profiles of Cluster 0 (Moderates), Cluster 1 (Baseline), and Cluster 2 (Intensives) allows for the development of customized messages that resonate with each group's level of engagement and concern. For instance, awareness campaigns about AI's impact could be nuanced to address the specific interests and concerns of each cluster, thereby increasing the relevance and effectiveness of the communication; (2) Policy development—insights into the different clusters guide

policy-makers in crafting policies and initiatives that reflect the diverse perspectives and engagement levels within the population. For example, policies aimed at addressing privacy concerns and job displacement due to AI could be prioritized, considering the strong engagement of Cluster 2 (Intensives) with these issues; (3) Educational programs - the findings of this analysis help in designing educational and outreach programs tailored to the varying levels of interest and concern among the clusters. Programs aimed at Cluster 2 (Intensives) might delve deeper into the nuances of AI technology and ethical considerations, while those targeting Cluster 1 (Baseline) might focus on raising awareness and providing basic information to increase engagement; (4) Market segmentation for technology products—companies developing AI-related products or services may use these insights for market segmentation, creating offerings that cater to the specific needs and concerns of each cluster. For example, products designed with enhanced privacy features might appeal more to Cluster 2 (Intensives), while user-friendly and introductory AI tools could be more suitable for Cluster 1 (Baseline).

One of the limitations of our study is that it draws its conclusions from a survey conducted among U.S. participants, potentially leading to a skewed representation of worldwide attitudes toward AI technology. Despite this limitation, it is important to acknowledge that the United States is at the forefront of AI technology development, making it a potentially representative sample for understanding individual opinions on this matter. The advanced state of AI technology in the U.S. offers insights into how perceptions and attitudes could evolve in other countries as they progress in AI adoption and integration. This context suggests that, while the study's findings are primarily reflective of U.S. perspectives, they could still provide valuable foresight into emerging global trends and attitudes towards AI technology. This narrow demographic focus may limit the broader applicability of our findings. Thus, it is important to consider the consequences of relying on this U.S.-centric dataset, as it might not encompass the diverse perspectives present in a global context. To enhance the relevance and applicability of our research on a global scale, future studies will aim to incorporate a more varied demographic, thereby broadening the scope and improving the generalizability of the results.

## Data availability

The data is available on the GitHub repository: [https://github.com/simonavoprea/Human\\_enhancements](https://github.com/simonavoprea/Human_enhancements).

Received: 13 November 2023; Accepted: 5 March 2024;

Published online: 11 March 2024

## Notes

1 <https://www.naa.jp/en/airnarita/automation.html>

2 <https://www.pewresearch.org/>

## References

- Acheampong RA, Cugurullo F (2019) Capturing the behavioural determinants behind the adoption of autonomous vehicles: Conceptual frameworks and measurement models to predict public transport, sharing and ownership trends of self-driving cars. *Transp Res Part F Traffic Psychol Behav* 62:349–375. <https://doi.org/10.1016/j.trf.2019.01.009>
- Ahmad MF, Husin NAA, Ahmad ANA, Abdullah H, Wei CS, Nawi MNM (2022) Digital transformation: an exploring barriers and challenges practice of artificial intelligence in manufacturing firms in Malaysia. *J Adv Res Appl Sci Eng Technol* 29:110–117. <https://doi.org/10.37934/araset.29.1.110117>
- Almarashda HAHA, Baba IB, Ramli AA, Memon AH, Rahman IA (2021) Human resource management and technology development in artificial intelligence

- adoption in the UAE energy sector. *J Appl Eng Sci* 11:69–76. <https://doi.org/10.2478/jaes-2021-0010>
- Amichai-Hamburger Y, Mor Y, Wellington T, Landesman T, Ophir Y (2020) The personal autonomous car: personality and the driverless car. *Cyberpsychol Behav Soc Netw* 23:242–245. <https://doi.org/10.1089/cyber.2019.0544>
- Andrejevic M, Selwyn N (2020) Facial recognition technology in schools: critical questions and concerns. *Learn Media Technol* 45:115–128. <https://doi.org/10.1080/17439884.2020.1686014>
- Băra A, Oprea SV (2023) What makes electricity consumers change their behavior? Influence of attitude and perceived impact of DR programs on awareness. *Kybernetes*. <https://doi.org/10.1108/K-01-2023-0032>
- Binsaeed RH, Yousaf Z, Grigorescu A, Samoila A, Chitescu RI, Nassani AA (2023) Knowledge sharing key issue for digital technology and artificial intelligence adoption. *Systems* 11:316. <https://doi.org/10.3390/systems11070316>
- Chen S, Qiu S, Li H, Zhang J, Wu X, Zeng W, Huang F (2023) An integrated model for predicting pupils' acceptance of artificially intelligent robots as teachers. *Educ Inf Technol* 28:11631–11654. <https://doi.org/10.1007/s10639-023-11601-2>
- Chiu TKF, Meng H, Chai CS, King I, Wong S, Yam Y (2022) Creation and evaluation of a pretertiary artificial intelligence (AI) curriculum. *IEEE Trans Educ* 65:30–39. <https://doi.org/10.1109/TE.2021.3085878>
- Choung H, David P, Ross A (2023) Trust in AI and its role in the acceptance of AI technologies. *Int J Hum-Comput Interact* 39:1727–1739. <https://doi.org/10.1080/10447318.2022.2050543>
- Chung KC, Chen CH, Tsai HH, Chuang YH (2021) Social media privacy management strategies: a SEM analysis of user privacy behaviors. *Comput Commun* 174:122–130. <https://doi.org/10.1016/j.comcom.2021.04.012>
- Damerji H, Salimi A (2021) Mediating effect of use perceptions on technology readiness and adoption of artificial intelligence in accounting. *Account Educ* 30:107–130. <https://doi.org/10.1080/09639284.2021.1872035>
- Ding C, Li C, Xiong Z, Li Z, Liang Q (2024) Intelligent identification of moving trajectory of autonomous vehicle based on friction nano-generator. *IEEE Trans Intell Trans Syst* 1–8. <https://doi.org/10.1109/TITS.2023.3303267>
- Fu Y, Li C, Yu FR, Luan TH, Zhao P (2023) An incentive mechanism of incorporating supervision game for federated learning in autonomous driving. *IEEE Trans Intell Trans Syst* 24:14800–14812. <https://doi.org/10.1109/TITS.2023.3297996>
- Giovanola B, Tiribelli S (2023) Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms. *AI Soc* 38:549–563. <https://doi.org/10.1007/s00146-022-01455-6>
- Jannat MKA, Islam MS, Yang SH, Liu H (2023) Efficient wi-fi-based human activity recognition using adaptive antenna elimination. *IEEE Access* 11:105440–105454. <https://doi.org/10.1109/ACCESS.2023.3320069>
- Jiang Z, Xu C (2023) Policy incentives, government subsidies, and technological innovation in new energy vehicle enterprises: evidence from China. *Energy Policy* 177:113527. <https://doi.org/10.1016/j.enpol.2023.113527>
- Karmakar G, Chowdhury A, Das R, Kamruzzaman J, Islam S (2021) Assessing trust level of a driverless car using deep learning. *IEEE Trans Intell Trans Syst* 22:4457–4466. <https://doi.org/10.1109/TITS.2021.3059261>
- Kim KH, Hong KJ, Shin SD, Ro YS, Song KJ, Kim TH, Park JH (2022) How do people think about the implementation of speech and video recognition technology in emergency medical practice? *PLoS ONE* 17:e0275280. <https://doi.org/10.1371/journal.pone.0275280>
- Kim PT (2022) Race-aware algorithms: fairness, nondiscrimination and affirmative action. *Calif Law Rev*. <https://doi.org/10.15779/Z387P8TF1W>
- Ko E, Kim H, Lee J (2021) Survey data analysis on intention to use shared mobility services. *J Adv Trans* 2021:1–10. <https://doi.org/10.1155/2021/5585542>
- Kopalle PK, Gangwar M, Kaplan A, Ramachandran D, Reinartz W, Rindfleisch A (2022) Examining artificial intelligence (AI) technologies in marketing via a global lens: current trends and future research opportunities. *Int J Res Market* 39:522–540. <https://doi.org/10.1016/j.ijresmar.2021.11.002>
- Kosan E, Krois J, Wingenfeld K, Deuter CE, Gaudin R, Schwendicke F (2022) Patients' perspectives on artificial intelligence in dentistry: a controlled study. *J Clin Med* 11:2143. <https://doi.org/10.3390/jcm11082143>
- Kumar A, Krishnamoorthy B, Bhattacharyya SS (2023) Machine learning and artificial intelligence-induced technostress in organizations: a study on automation-augmentation paradox with socio-technical systems as coping mechanisms. *Int J Organ Anal*. <https://doi.org/10.1108/IJOA-01-2023-3581>
- Lambert J, Stevens M (2023) ChatGPT and generative AI technology: a mixed bag of concerns and new opportunities. *Comput School*. 1–25 <https://doi.org/10.1080/07380569.2023.2256710>
- Lepelaar M, Wahby A, Rossouw M, Nikitin L, Tibble K, Ryan PJ, Watson RB (2022) Sentiment analysis of social survey data for local city councils. *J Sens Actuat Netw* 11:7. <https://doi.org/10.3390/jsan11010007>
- Liu X, Shi T, Zhou G, Liu M, Yin Z, Yin L, Zheng W (2023a) Emotion classification for short texts: an improved multi-label method. *Humanit Soc Sci Commun*. <https://doi.org/10.1057/s41599-023-01816-6>
- Liu X, Zhou G, Kong M, Yin Z, Li X, Yin L, Zheng W (2023b) Developing multi-labelled corpus of Twitter short texts: a semi-automatic method. *Systems*. <https://doi.org/10.3390/systems11080390>
- Mohr S, Kühl R (2021) Acceptance of artificial intelligence in German agriculture: an application of the technology acceptance model and the theory of planned behavior. *Precis Agri* 22:1816–1844. <https://doi.org/10.1007/s11119-021-09814-x>
- Nouraldeen RM (2023) The impact of technology readiness and use perceptions on students' adoption of artificial intelligence: the moderating role of gender. *Dev Learn Organ* 37:7–10. <https://doi.org/10.1108/DLO-07-2022-0133>
- Ongena YP, Haan M, Yakar D, Kwee TC (2020) Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. *Eur Radiol* 30:1033–1040. <https://doi.org/10.1007/s00330-019-06486-0>
- Oprea S-V, Băra A (2022) A measurement model for electricity Consumers' awareness with covariance structure Analyses. A solid pillar for boosting demand response programs. *Sustain Energy Technol Assess* 53:102738. <https://doi.org/10.1016/j.seta.2022.102738>
- Patrzyk S, Bielecki W, Woźniacka A (2022) A study of attitudes among polish dermatologists and dermatology trainees regarding modern technologies in medicine. *Postepy Dermatologii i Alergologii* 39:531–537. <https://doi.org/10.5114/ada.2022.117738>
- Pillai R, Sivathanu B (2020) Adoption of artificial intelligence (AI) for talent acquisition in IT/ITeS organizations. *Benchmarking* 27:2599–2629. <https://doi.org/10.1108/BIJ-04-2020-0186>
- Qiu H, Li M, Bai B, Wang N, Li Y (2022) The impact of AI-enabled service attributes on service hospitableness: the role of employee physical and psychological workload. *Int J Contemp Hosp Manag* 34:1374–1398. <https://doi.org/10.1108/IJCHM-08-2021-0960>
- Querci I, Barbarossa C, Romani S, Ricotta F (2022) Explaining how algorithms work reduces consumers' concerns regarding the collection of personal data and promotes AI technology adoption. *Psychol Market* 39:1888–1901. <https://doi.org/10.1002/mar.21705>
- Rahman M, Ming TH, Baigh TA, Sarker M (2021) Adoption of artificial intelligence in banking services: an empirical analysis. *Int J Emerg Market* 18:4270–4300. <https://doi.org/10.1108/IJOEM-06-2020-0724>
- Rauf MA, Ashfaq M, Hasan R, Manju MA (2021) A comparative study on the impact of artificial intelligence on employment opportunities for university graduates in Germany and the Netherlands: AI opportunities and risks. *Int J Environ Workplace Employment* 6:185. <https://doi.org/10.1504/IJEW.2021.119679>
- Rowthorn M (2019) How should autonomous vehicles make moral decisions? Machine ethics, artificial driving intelligence, and crash algorithms. *Contemp Read Law Soc Just* 11:9. <https://doi.org/10.22381/CRLSJ1120191>
- Sampurna MTA, Handayani KD, Utomo MT, Angelika D, Etika R, Harianto A, Mapindra MP (2023) Determinants of neonatal deaths in Indonesia: a national survey data analysis of 10,838 newborns. *Heliyon* 9:e12980. <https://doi.org/10.1016/j.heliyon.2023.e12980>
- Shiwakoti N, Hu Q, Pang MK, Cheung TM, Xu Z, Jiang H (2022) Passengers' perceptions and satisfaction with digital technology adopted by airlines during COVID-19 pandemic. *Future Transp* 2:988–1009. <https://doi.org/10.3390/futuretransp2040055>
- Smith M, Miller S (2022) The ethical application of biometric facial recognition technology. *AI Soc* 37:167–175. <https://doi.org/10.1007/s00146-021-01199-9>
- Tseng FH, Zeng JY, Cho HH, Yeh KH, Chen, CY (2024) Detecting adversarial examples of fake news via the neurons activation state. *IEEE Trans Comput Soc Syst* 1–11. <https://doi.org/10.1109/TCSS.2023.3293718>
- van Buchem MM, Neve OM, Kant IMJ, Steyerberg EW, Boosman H, Hensen EF (2022) Analyzing patient experiences using natural language processing: development and validation of the artificial intelligence patient reported experience measure (AI-PREM). *BMC Med Inf Decis Making* 22:183. <https://doi.org/10.1186/s12911-022-01923-5>
- van Noordt C, Misuraca G (2022) Exploratory insights on artificial intelligence for government in Europe. *Soc Sci Comput Rev* 40:426–444. <https://doi.org/10.1177/0894439320980449>
- Vorisek CN, Stellmach C, Mayer PJ, Klopfenstein SAI, Bures DM, Diehl A, Henningsen M (2023) Artificial intelligence bias in health care: web-based survey. *J Med Intern Res* 25:e41089. <https://doi.org/10.2196/41089>
- Wang J, Xing Z, Zhang R (2023a) AI technology application and employee responsibility. *Humanit Soc Sci Commun* <https://doi.org/10.1057/s41599-023-01843-3>
- Wang S, Sun Z, Chen Y (2023b) Effects of higher education institutes' artificial intelligence capability on students' self-efficacy, creativity and learning performance. *Educ Inf Technol*. <https://doi.org/10.1007/s10639-022-11338-4>
- Warwas I, Podgórnjak-Krzykacz A, Wiktorowicz J, Górniak J (2022) Demographic and generational determinants of Poles' participation in the sharing economy: findings from a survey data analysis. *PLoS ONE* 17:e0265341. <https://doi.org/10.1371/journal.pone.0265341>
- Xiao Z, Shu J, Jiang H, Min G, Chen H, Han Z (2023) Perception task offloading with collaborative computation for autonomous driving. *IEEE J Selected Areas Commun* 41:457–473. <https://doi.org/10.1109/JSAC.2022.3227027>

- Xuan PY, Fahumida MIF, Al Nazir Hussain MI, Jayathilake NT, Khobragade S, Soe HHK, Moe S (2023) Readiness towards artificial intelligence among undergraduate medical students in Malaysia. *Educ Med J* 15:49–60. <https://doi.org/10.21315/eimj2023.15.2.4>
- Yoon J, Andrews JE, Ward HL (2022) Perceptions on adopting artificial intelligence and related technologies in libraries: public and academic librarians in North America. *Library Hi Tech* 40:1893–1915. <https://doi.org/10.1108/LHT-07-2021-0229>
- Yu C, Helwig EJ (2022) The role of AI technology in prediction, diagnosis and treatment of colorectal cancer. *Artif Intell Rev* 55:323–343. <https://doi.org/10.1007/s10462-021-10034-y>
- Yue W, Li C, Wang S, Xue N, Wu J (2023) Cooperative incident management in mixed traffic of CAVs and human-driven vehicles. *IEEE Trans Intell Trans Syst* 24:12462–12476. <https://doi.org/10.1109/TITS.2023.3289983>

### Acknowledgements

This work was supported by a grant from the Ministry of Research, Innovation and Digitization, CNCS- UEFISCDI, project number PN-III-P4-PCE-2021-0334, within PNCDI III.

### Author contributions

Simona-Vasilica Oprea: methodology, formal analysis, investigation, writing—original draft, writing—review and editing, visualization, project administration. Adela Băra: Conceptualization, validation, formal analysis, investigation, resources, data curation, writing—original draft, writing—review and editing, visualization, supervision.

### Ethical approval

Ethical approval was not required as the study did not involve designated human participants.

### Informed consent

Informed consent was not required as this research did not involve designated human participants.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to Simona-Vasilica Oprea.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024