Technological Forecasting & Social Change 188 (2023) 122306

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Responsible natural language processing: A principlist framework for social benefits

Rajat Kumar Behera [a,*], Pradip Kumar Bala [b], Nripendra P. Rana [c], Zahir Irani [d]

[a] School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University, KIIT Road, Patia, Bhubaneswar, Odisha 751024, India
[b] Area of Information Systems & Business Analytics, Indian Institute of Management Ranchi, Nayasarai Road, Prabandhan Nagar, Ranchi, Jharkhand 835303, India
[c] College of Business and Economics, Qatar University, Doha, PO Box 2713, Qatar
[d] University of Bradford, School of Management, Richmond Road, Bradford BD7 1DP, United Kingdom

ARTICLE INFO

Keywords:
Responsible NLP
Ethical work climate
Social benefits
Responsible decision making
Ethics

ABSTRACT

Businesses harness the power of natural language processing (NLP) to automate processes and make data-driven decisions. However, NLP raises concerns on a number of fronts due to its potential for disruption, which can be addressed with the assignment of responsibility. Therefore, responsible NLP (RNLP) can be designed as a principlist framework to ensure NLP systems are used in an ethical manner. The study proposes a principlist framework with the formulation of eight principlist ethical principles to ensure NLP is safe, secure and reliable for responsible decision making and subsequently results in social benefits. Using snowball sampling, data are collected from 15 informants, who represent senior-level positions in diversified industries. The analysis is performed with qualitative research methodology. The result produces two ethical practices. First is the adoption of RNLP as a disruptive technology for ethical decision making for social benefits and second is the creation of a culture of responsibility.

1. Introduction

Science and technology have offered various techniques in recent years to provide new understandings of the world, as well as more direct, broad, and less expensive answers to a wide range of human, technological, and social challenges (Perazzoli et al., 2022). Natural language processing (hereafter, NLP) is one of the technologies and (Chowdhury, 2003) is the process of understanding how computerized systems use texts, speech, and other resources, and how they are operated on computers. The major purpose of NLP is to achieve human-like language processing for a variety of jobs or applications, as well as to use computational approaches to examine the generated texts (Heylighen, 2008). Over the last decade, the amount of data and computing power has increased, resulting in the emergence of multiple new fronts in the field of NLP (Tripathy et al., 2021). Entity recognition, parts of speech tagging and semantic role labelling are examples of NLP applications that focus on text processing (Hutchinson, 2020) and many researchers have looked into the uses of NLP in various fields (Agarap, 2018; Lu et al., 2018). Business managers have been turning to NLP for a variety of organisational analysis, communications processing and textual analytics tasks amid the present rush to use artificial intelligence (hereafter, AI) for assisting enterprises in a number of tasks (Liu et al., 2021; Pandey and Pandey, 2019).

NLP, like any other technology, faces ethical challenges, both in terms of how it is conducted (i.e., reproducibility, transparency) and the societal implications of its findings (Fort and Couillault, 2016). Studies have shown how NLP technology may be abused to cause harm like repressing dissenters (Zhang et al., 2014), compromising privacy (Coavoux et al., 2018) or profiling (Wang et al., 2018). A lot more prevalent negative unforeseen consequences of NLP, such as unfairness resulting from demographic biases, varying levels of performance across a variety of user groups (Tatman, 2017), speakers misidentification and their expectations (Perez, 2019) or the propagation of negative stereotypes (Kiritchenko and Mohammad, 2018). These ethical challenges could be the result of the field's rapid growth and failure to adapt to new circumstances as any scepticism or concerns regarding NLP's power were purely theoretical.

In continuation, when NLP tools are trained on a demographic sample, it performs poorly on another sample (Garimella et al., 2019). This mismatch affects text and socio-demographic domains i.e., for instance, persons of different age groups are linguistically as different as a text from an online blog and a newspaper (Johannsen et al., 2015).

* Corresponding author.
E-mail addresses: rajat_behera@yahoo.com (R.K. Behera), pkbala@iimranchi.ac.in (P.K. Bala), z.irani@bradford.ac.uk (Z. Irani).

https://doi.org/10.1016/j.techfore.2022.122306
Received 14 May 2022; Received in revised form 18 September 2022; Accepted 20 December 2022
Available online 27 December 2022

Intriguingly, demographics such as age and text-domain are frequently correlated (Hovy, 2015). The advancement of NLP raises concerns regarding how to best include wellbeing, human-centred values, fairness, privacy and security, reliability, transparency, interrogation and accountability in NLP systems. Both the business and technology ethics literature have stressed and explored the importance of deliberative engagement in influencing responsible innovation (Brand and Blok, 2019; Lubberink et al., 2017). Literature has proposed AI ethical framework (i.e., Ashok et al., 2022; Celik, 2022; John-Mathews, 2022; Alabed et al., 2022), but none of these frameworks have covered the relationship between responsible innovation and information technology (hereafter, IT) and in addition, none of these explicitly discussed responsible decision making and social benefits.

In continuation, NLP is a sub-domain of AI, which deals with how computers understand and translate different natural languages, whereas AI includes systems that simulate cognitive capabilities, such as learning from examples and problem solving. Different languages have different grammar and different biases, which make particular individuals or groups feel misunderstood, alienated, or misrepresented. Hence, AI and NLP from ethical lens are different and hence, NLP requires a standalone framework. Therefore, responsible NLP (hereafter, RNLP) is a principlist framework that focuses on how an organisation addresses the challenges around NLP from an ethical point of view for responsible decision making and subsequently, witnessing the social benefits. A principlist framework is used to identify ethical problems and make decisions about what to do about these. Rather than resolving contentious substantive disagreements regarding specific cases, the purpose of the principlist framework is to demonstrate its use in building ethical sensitivity. Responsible decision making is the ability of NLP systems to make constructive choices about business behavior based on ethical standards. Social benefits are well-planned benefits and economically justified by the consumer of the NLP system.

The importance of the relationship between NLP and ethics is highlighted in several studies such as the fairness, outcomes, rights, responsibilities, liberty, empathy, authority and character are the lenses through which to examine challenging ethical scenarios using NLP (Mitri, 2020). Based on the foregoing discussions, the study could not identify literature that has categorically called for the core principlist ethical principles of NLP, structurally relates the NLP ethical principles to responsible decision making, defined the moderator role of ethical work climate (hereafter, EWC) between NLP ethical principles and responsible decision making. EWC is a collection of shared formal and informal perceptions of procedures and rules that impact ethical behavior expectations (Victor and Cullen, 1987, 1988). Concerning the importance, the attention of further research is highlighted by literature such as in the development and application of NLP technology, there is a need for NLP ethics and natural language ownership, privacy, filtering, intellectual property, and governance are all challenges that need to be addressed (Li et al., 2020). Concerning future research, the attention to the NLP-based ethical challenge is highlighted in literature such as understanding and mitigating bias in data and algorithms, recognising problematic information such as hate speech, stereotypes, and abusive language, and developing frameworks for improved system design and data management methods are all ethical challenges that NLP has to tackle (Prabhumoye et al., 2020). The discussions on literature gaps, future research and research challenges justify the rationale behind the motivation of this study.

The study is undertaken to highlight how RNLP results in social benefits. To achieve the motivation, the study aims to propose a principlist framework that focuses on ensuring NLP technologies that are used in an ethical, transparent and accountable manner for social benefits. To achieve the aim, the study has formulated the objectives i.e., to discuss responsible innovation and NLP, core principlist ethical principles of NLP, responsible decision making, social benefits and the conceptual model. The conceptual model aims to systematise the static coherence of NLP principlist ethical principles, responsible decision making, social benefits, EWC and represents a pillar to build the roadmap of this study. The theoretical implication intends to shed light on the framework based on NLP ethical principles. The managerial implication intends to witness social benefits for practitioners i.e., principles towards the achievement of sustainable development goals.

This paper is organised as follows: the principlist framework is presented in Section 2, followed by research methodology (i.e., Section 3). This is followed by results (i.e., Section 4) and the discussion in Section 5. Finally, in Section 6, the conclusion is presented.

2. Principlist framework

Every technical endeavour should be pervaded by ethical considerations, which should be systemic (Ocone, 2020). Therefore, a principlist framework is required to provide the broader knowledge on the interlocking of NLP with ethics for responsible decision making for social benefits. Chang (2021) proposed an ethical framework for big data to enhance data-driven decision making competency, but was limited to smart cities and public transportation systems. Vermanen et al. (2021) proposed an ethical framework for Internet of Things (IoT) deployment in medium-sized enterprises (SMEs), but lacks conducting expert interviews and applying the model to the test in real-world industry scenarios. Ashok et al. (2022) proposed an ethical framework illustrating the ethical use of AI in digital technologies, but lacks in-depth interviews and quantitative testing. Akrami et al. (2018) proposed a comprehensive and user-friendly framework for evaluating public health programmes in an ethical manner, but lacks discussion on responsible decision making and social benefits. Based on the above discussion, none of the framework adopts a principlist approach, as it is one of the most useful for elucidating the important ethical principles in a specific domain (Formosa et al., 2021), such as NLP.

Literature has criticised the principlist approach of ethical frameworks (i.e., John-Mathews et al., 2022; John-Mathews, 2022; John-Mathews, 2021; Mittelstadt, 2019), and highlighted the lack of interpretability of IT (i.e., AI) decisions. The study is in line with criticism provided by such literature that IT is unable to make constructive and respectful decisions based on the consideration of ethical standards and social norms (i.e., responsible decision making). Hence, an IT-enabled principlist approach must consider ethical standards, social and community norms in making decisions; apply and evaluate decision making skills to engage in a variety of situations. In specific, NLP-enabled principlist framework must make responsible decision making by understanding decisions that can have positive and negative effects on society; identifying ways that certain decisions effect short-term and long-term goals; identifying positive and negative consequences of decisions for society; monitoring how decision making affects progress towards achieving a goal, through reflection on past choices, social and community norms; devising stop, think and act strategy when making decisions; assessing lessons learned from past experiences and mistakes when making decisions; generate alternative solutions to identified problems; identifying and applying the steps of systematic decision making, using creativity and innovation; analysing and evaluating evidence, arguments, claims, and beliefs to inform effective decisions. Therefore, a principlist framework is proposed covering the relation between responsible innovation and NLP, core principlist ethical principles of NLP, responsible decision making, social benefits and the conceptual model to meet the motivation of the study.

2.1. Responsible innovation and NLP

Innovation is a multi-stage process through which organisations transforms ideas into new or improved products, services, or processes to advance, compete and differentiate themselves successfully in the marketplace (Baregheh et al., 2009). There are three core components to responsible innovation i.e., the responsibility to prevent harm, the obligation to do good and the responsibility to govern (Buhmann and

Fieseler, 2021). First, 'responsibility to prevent harm' includes risk management strategies aimed at preventing potentially negative outcomes. Second, 'obligation to do good' refers to improving living situations, such as in accordance with sustainable development goals. Finally, 'governance responsibility' entails the creation and support of global governance mechanisms that can help with the former two obligations. The mindset that underpins responsible innovation, with its focus on social benefits, involves prioritising not only inclusive involvement but also meaningful procedures of inclusive reflection and deliberative democracy (Owen et al., 2012). Therefore, NLP innovation refers to an organisation's effort to train algorithms to replicate human-like tasks including language processing, learning and problem-solving and to scale these functions through software for responsible decision making.

The first dimension (i.e., the responsibility to prevent harm) of responsible innovation holds true for NLP i.e., much of the work on NLP for social good focuses on detecting or preventing harm, such as classifying hate speech, moderating bias, or recognising depression symptoms and on the other hand, NLP research has the promise of positive proactive applications that can increase user and public well-being or stimulate constructive dialogue (Field et al., 2021). The second dimension (i.e., the obligation to do good) of responsible innovation holds true for NLP, i.e., the use of NLP is particularly useful in the context of contracts that are less prone to costly arbitrations and litigations (Agrawal et al., 2021). The third dimension (i.e., governance responsibility) of responsible innovation holds true for NLP i.e., text messages and emails include extremely sensitive data that is commonly utilised in NLP applications and the adoption of a federated learning technique in the domain of NLP requires sensitive data (Prabhu et al., 2021). Since RNLP innovation cannot do well without ethics, therefore, systematic attention to core principlist ethical principles of NLP is important for RNLP innovation practices to avoid breakthroughs that have a detrimental societal impact.

2.2. Core principlist ethical principles of NLP

NLP models and algorithms, in particular, have been widely used in a number of decision making scenarios, including criminal justice (Curreen, 1995), traffic control (Li et al., 2018), financial lending (Chen et al., 2020) and medical diagnostics (Carchiolo et al., 2019). In many ways, the growing use of NLP-based decision making systems poses a concern. Therefore, the core principlist ethical principles of NLP emphasise the need to weigh factors while making responsible decisions. The core principlist ethical principles demand not just realistic and thorough technical frameworks and methods to impose ethical limitations on NLP models, algorithms and products, but also to reflect society's goodwill and moral beliefs. The study proposes eight principlist core ethical principles of NLP, i.e., wellbeing, human-centred values, fairness, privacy and security, reliability, transparency, interrogation and accountability. Each principle is further elaborated.

2.2.1. NLP principlist ethical principle 1 (wellbeing)

Any activity that improves everyone's wellbeing while making no one worse off is regarded as having the greatest overall benefit to society and hence being the most morally desirable (Choy, 2018). The wellbeing principlist ethical principle ensures individuals, society and the environment should benefit from NLP systems throughout its operational lifecycle. This principle states unequivocally that NLP systems should be used for the benefit of humans, society and the environment from the inception and the goals of NLP systems should be explicitly defined and justified. NLP systems created for legitimate internal business reasons, such as enhancing efficiency, can have far-reaching consequences for human, social and environmental wellbeing. Deep-learning algorithms based on neural networks are driving the biggest breakthroughs in NLP and to ensure wellbeing, such methods are designed to imitate the function of neurons in the human brain (Marulli et al., 2021).

2.2.2. NLP principlist ethical principle 2 (human-centred values)

Human-centred tacit knowledge is valued and the human is empowered by technology and automation at disposal, the power structures linked with technology can operate as barriers to ethical behavior (O'Neill and Stapleton, 2020). The human-centred values principlist ethical principle ensures individual autonomy, diversity and human rights should be respected by NLP systems throughout its operational lifecycle. The objective of this principle is to ensure that NLP systems are relevant to human values. By honouring, defending and promoting human rights, respecting human freedom and conserving the environment, NLP systems should enable an equitable and democratic society. Individuals who interact with NLP systems should be able to govern them completely and effectively. Machine intelligence of NLP can be enabled as a cognitive function that a machine replicates, such as reasoning, learning, and problem solving, perception, rationality and thought process could impede ethical and moral standards handled by human knowledge and responsibility (Shahriari and Shahriari, 2017).

2.2.3. NLP principlist ethical principle 3 (fairness)

Good character is based on the ethical premise of human equality and is defined in terms of values such as fairness (Kazim and Koshiyama, 2021). The fairness principlist ethical principle ensures NLP systems should be inclusive and accessible to every individual, community and groups, and prejudice towards individuals, communities or groups is not involved or resulted in. This principle aims to ensure that NLP systems are equitable and inclusive over the entire lifespan. NLP systems should be user-centric, with access to associated products and services available to everyone who interacts with them. This is crucial in light of concerns that the NLP may perpetuate societal injustices and unfairly affect marginalised and vulnerable groups, including those based on age, disability, race, sex, intersex status, gender identity, and sexual orientation. Newer deep learning models are delivering breakthrough results in a variety of NLP applications, such as machine translation, video captioning, and speech recognition, while stressing algorithmic fairness (Kim et al., 2020).

2.2.4. NLP principlist ethical principle 4 (privacy and security)

In the creation and use of innovations in and around information and communication technology, ethical challenges and concerns regarding privacy, security and other human rights and social values prevail (Raab, 2020). The privacy and security principlist ethical principle ensures NLP systems should uphold and respect private rights and data protection and also ensure data security. When using NLP systems, this principle strives to secure privacy and data security, which includes providing effective data governance and management for all data that the NLP system uses and generates during its lifecycle. Additionally, the link between data and the inferences provided by NLP systems should be robust and assessed in a timely manner. Additionally, this principle aims to ensure that adequate data security and NLP system security are in place. To achieve privacy in NLP, a typical method is to use homomorphic encryption, which allows encrypted data to undergo arithmetic operations without having to decrypt the encrypted data or expose the underlying data and multi-party computation is method for achieving security, in which computations are done on secret inputs from many parties such as service providers, users and so on (Feng et al., 2020).

2.2.5. NLP principlist ethical principle 5 (reliability)

Reliability is an important issue in ethics (Gordon et al., 2021) and is regarded as the organisation's guarantee to deliver services on time, resolve difficulties, and provide error-free transactions (Westbrook, 1981). The reliability principlist ethical principle ensures NLP systems should operate in a consistent manner in accordance with the intended purpose. This principle aims to ensure that NLP systems consistently perform their intended task for the duration of existence. This includes making sure NLP systems are as precise, dependable, and repeatable as feasible. NLP systems should not pose unreasonably high safety risks,

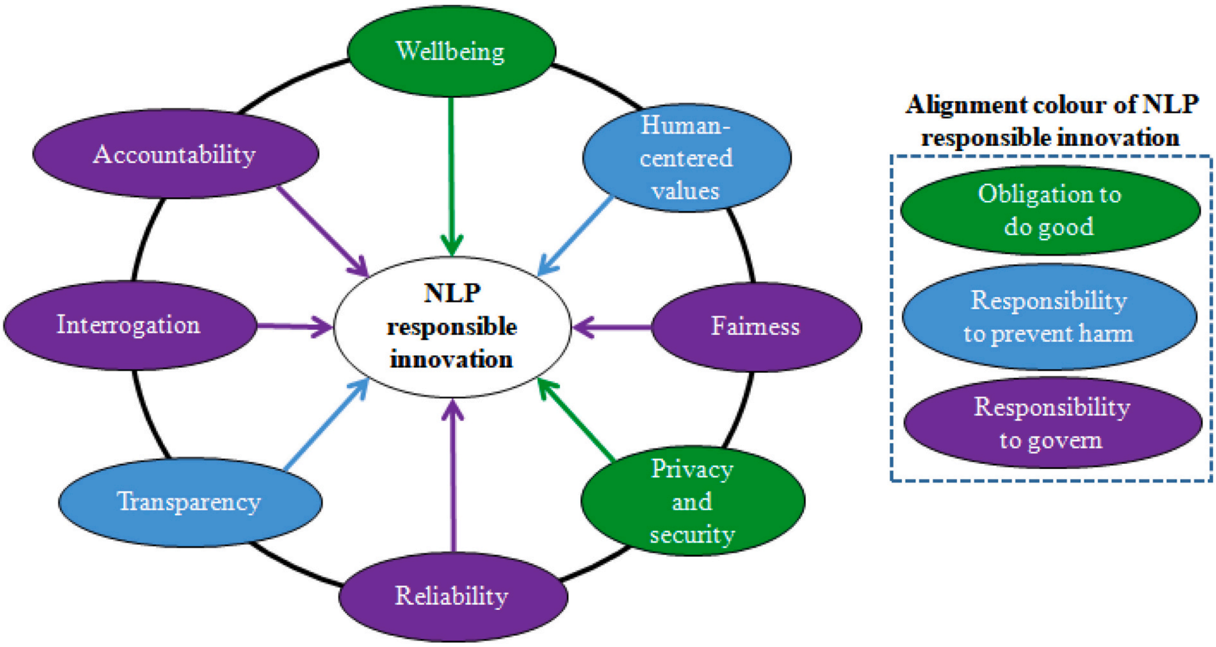

Fig. 1. Alignment diagram.

and safety measures should be appropriate to the size of potential threats. Monitoring and testing should be done on NLP systems to ensure that it continues to serve its original function and any issues should be addressed with appropriate risk management. To ensure that an NLP system is robust and secure, responsibility should be explicitly and appropriately allocated. The most prevalent attack on reliability is a poisoned data set and to create a reliable system, the NLP model must be trained using a large volume of data (Verde et al., 2021).

2.2.6. NLP principlist ethical principle 6 (transparency)

Corruption scandals (Harrison and Sayogo, 2014), poor decision making (Guillamón et al., 2016), lack of accountability (Lourenço, 2015) and dysfunctional governance of organisations (Kosack and Fung, 2014) are all linked to a lack of transparency in operations and decision making processes. The transparency principlist ethical principle ensures NLP systems should ensure transparency and responsible disclosure so that people can recognise when NLP is having a substantial influence on them and when the NLP system is interacting with them. This principle aims to ensure appropriate disclosure when an NLP system has a significant impact on a person's life. The notion of a 'substantial impact' criterion is determined by the NLP system's context, impact and application. Responsible disclosures should be made in a timely fashion, with reasonable justifications for the outcomes of NLP systems. Recent legislative and regulatory trends have increased the demand for algorithm transparency, making explanations more necessary than desired, and another important reason to advocate for algorithm transparency in real-world decision making is due to inherent biases in data and data quality issues (Kim et al., 2020).

2.2.7. NLP principlist ethical principle 7 (interrogation)

The interrogation ethical principle aims to ensure the provision of effective, accessible methods that enable individuals to challenge the use or output of the NLP system when it has a significant impact on a person, community, group or environment. The notion of a 'substantial impact' criterion is determined by the NLP system's context, impact and application. To make interrogation effective, there should be sufficient access to the information available to the algorithm and the inferences drawn. In the case of judgments that have a considerable impact on rights, an effective system of oversight should be in place that makes appropriate use of human judgment. However, in some circumstances, the dataset's complex diversity and dimensionality characteristics make implementation difficult. NLP must incorporate various different algorithms for analysing the human native language because of text and voice-based data, as well as their practical applications. Statistical and machine learning methods, as well as rules-based and algorithmic methods, could be used. Pre-processing, tokenization, normalization, typographical error correction, named entity reorganisation and dependency parsing are all part of such a process.

2.2.8. NLP principlist ethical principle 8 (accountability)

The relationship between accountability and ethics has long been a source of discussion, with accountability typically being viewed as a tool for controlling and directing administrative behavior by requiring answerability to some external authority (Dubnick, 2003). The accountability principlist ethical principle ensures NLP systems should enable human oversight of NLP systems and those responsible for the various phases of the NLP system lifecycle should be identified and held accountable for the outcomes of the NLP systems. This principle aims to recognise the responsibility of relevant organisations and individuals for the consequences of NLP systems that they design, build, deploy and operate. The application of legal principles relating to NLP system accountability has to continually evolve. More human-interpretable and relevant explanations should be provided in the deployment stage to improve the NLP model's accountability, taking into account various factors such as the balance between predictive and descriptive accuracy, end-user experience and the model's fairness and accountability (Kim et al., 2020).

Fig. 1 depicts the alignment diagram showing the alignment of core principlist ethical principles of NLP to NLP responsible innovation. In the figure, the principlist ethical principles of wellbeing and 'privacy and security' are aligned to the 'obligation to do good' component of NLP responsible innovation (depicted in green). The principlist ethical principles 'human-centred values' and transparency are aligned to the 'responsibility to prevent harm' component of NLP responsible innovation (depicted in light blue). The principlist ethical principles of fairness, reliability, interrogation and accountability are aligned to the 'responsibility to govern' component of NLP responsible innovation (depicted in purple). The alignment is made based on the matching criteria of the purpose of 'NLP responsible innovation' against the motivation of 'NLP ethical principle' e.g., the purpose of 'responsibility to prevent harm' should be fulfilled by the motivation of 'human-centred values' and transparency ethical principles.

2.3. Responsible decision making

A choice between two or more alternatives is referred to as a decision (Sugumaran and Degroote, 2010). Decisions can only be effective if they are based on factual data and accuracy is achieved through interpretation (Cronbach, 1988). While it is impossible to make responsible decisions without quality data, it is well acknowledged that supplying correct data does not guarantee that it will be acted on responsibly (Dean et al., 2020). Therefore, responsible decisions are based on social morals, ethics and are socially, ethically appropriate, and take into account the effects on society. In the context of NLP, responsible decision making is the ability of the NLP system to make constructive and respectful choices about social interaction on consideration of ethical standards, social norms, the realistic evaluation of consequences of various actions, and the wellbeing of society. The goal of responsible decision making by the NLP system is twofold, i.e., first, while making decisions, ethical standards, social and community norms be taken into account and second, in a range of scenarios, applies and evaluates decision making skills.

2.4. Social benefits

A growing number of studies are promoting the purported advantages of technology to society (Chopik, 2016), such as NLP. For instance, technology is frequently used by older adults to avoid loneliness (Vošner et al., 2016). Social benefit is the total benefit to society owing to NLP outcome or from consuming NLP services. For example, in the face of chatbots, NLP is on the rise, forming appropriate responses to questions by analysing the language typed into text fields. Chatbots not only reply to customer or prospect inquiries, but also allow society to gain access to new information. This has a positive impact on both businesses and society. NLP has progressed well beyond machine translation, revealing previously unseen patterns in datasets, automating processes, and allowing people to focus on higher-value, more creative tasks. NLP is becoming more successful at detecting fraud in insurance. The paper defines RNLP as a governance framework which is the integration of NLP responsible innovation, core ethical principles of NLP that results in responsible decision making for social benefits.

2.5. Conceptual model

The conceptual model or principles based on ethical decision making has been proposed in the literature. Clarke (2019) proposed principles for responsible AI and its purpose is to offer organisations and individual practitioners' guidance on how to fulfil their responsibilities in relation to AI technology and AI-based artifacts and systems. The proposed principles are a logical fit for multi-stakeholder risk assessment and management needs, which include asset identification, value, and threats, vulnerabilities, and safeguards. However, it has not considered how such principles are structurally related to responsible decision making and do not demonstrate the effect of social benefits. Buhmann and Fieseler (2021) proposed communicative principles for responsible AI and discussed AI's lack of transparency, explainability and accountability stifles informed discussion. However, it lacks discussion on principlist conceptual framework and the structural alignment of independent variables and dependent variable of social benefit. Ashok et al. (2022) proposed a conceptual model for ethical assessment for the use of AI in digital technologies. The impact of digital ethics implications on societal impact is highlighted in this study, which is moderated by digital technology archetypes and mediated by organisational impact. However, it lacks qualitative synthesis of NLP technologies.

Table 1
Construct definition and reference.

Construct	Definition	Reference
Wellbeing expectancy	Degree to which the NLP ethical principle 'wellbeing' benefits individuals, society and the environment.	Duncan et al., 2020
Human-centred values expectancy	Degree to which the NLP ethical principle 'human-centred values' respect human rights, diversity and the autonomy of individuals.	Bannon, 2011
Fairness expectancy	Degree to which the NLP ethical principle 'fairness' should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.	Lotz et al., 2013
Privacy and security expectancy	Degree to which the NLP ethical principle 'privacy and security' should respect and uphold privacy rights and data protection and ensure the security of data.	Allhoff et al., 2011
Reliability expectancy	Degree to which the NLP ethical principle 'reliability' should reliably operate in accordance with intended purpose.	Ratten, 2015
Transparency expectancy	Degree to which the NLP ethical principle 'transparency' should be transparent and responsible so people can understand when they are being significantly impacted by NLP and can find out when the NLP system is engaging with them.	Brunner and Ostermaier, 2019
Interrogation Expectancy	Degree to which the NLP ethical principle 'interrogation' should significantly impact a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the NLP system.	Schrecker, 2013
Accountability expectancy	Degree to which the NLP ethical principle 'accountability' should be identifiable and accountable for the outcomes of the NLP systems and human oversight of NLP systems should be enabled.	Beu and Buckley, 2001
Responsible decision making	The ability of the NLP system to make constructive decision making and social interactions in an ethically responsible way for the benefits and consequences of various actions for personal, social and collective wellbeing.	Clegg et al., 2007
Ethical work climate	Moral climate of the workplace and the amount of organisational ethics practice.	Victor and Cullen, 1988
Social benefits	Total benefit gained by the individual and the society owing to NLP outcome and consumption of NLP services.	Humphreys, 2015

In a similar lens, Maass and Storey (2021) investigated how conceptual modeling can be used in machine learning and present a framework for incorporating conceptual modeling into data science initiatives. A healthcare application is used to demonstrate the framework. However, the study discussed challenges such as model architecture development, model training, model testing, model optimization, model deployment and model maintenance in information systems. To sum up, none of the studies discussed the principlist ethical principles from NLP perceptive, defined structural alignment of core principlist ethical principles of NLP with responsible decision making, proposed NLP based principlist conceptual framework for social benefits and discussed the alignment of NLP responsible innovation to ethical principles. Therefore, the study has proposed responsible decision making NLP conceptual model for social benefits.

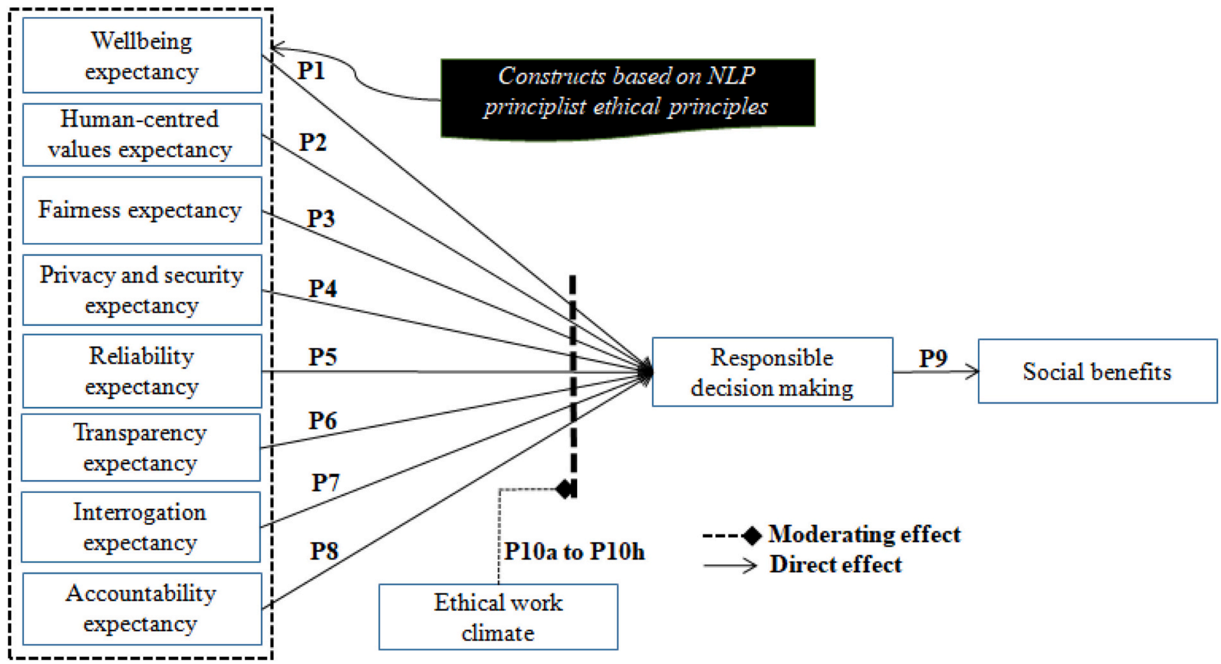

Fig. 2. Conceptual model (adapted from Duncan et al., 2020; Bannon, 2011; Lotz et al., 2013; Allhoff et al., 2011; Ratten, 2015; Brunner and Ostermaier, 2019; Schrecker, 2013; Beu and Buckley, 2001; Clegg et al., 2007; Victor and Cullen, 1988; Humphreys, 2015).

The conceptual model aims to propose constructs that enable responsible decision making for social benefits. Therefore, the study introduced new constructs, namely wellbeing expectancy, human-centred values expectancy, fairness expectancy, privacy and security expectancy, reliability expectancy, transparency expectancy, interrogation expectancy, accountability expectancy, responsible decision making and adopted EWC and social benefits from existing studies. The construct definition and reference are listed in Table 1.

The rationale behind consideration of each construct is further justified. Improving and maintaining wellbeing of society is critical because having healthy people at work and in society in general is crucial, and a healthy workforce is linked to effective organisational performance (Day and Nielsen, 2017). Individuals who are subjected to discrimination, exclusion, oppression, disempowerment, and disadvantage are listed in the provisions of human rights treaties forbidding discrimination (Krupiy, 2020). Individual autonomy emphasises how people are socially entrenched and how social and relational forces form their identities and preferences (Sylvestre et al., 2021). According to empirical research on ethics programme effectiveness, fairness evaluations are a crucial influence on ethics-related outcomes (Treviño et al., 1999). Skaug (2022) emphasises the ethical aspects of the trade-off between privacy and security, which are crucial for policy formation in modern liberal democracies. In the context of robustness, reliability is critical since an unreliable system that does not reproduce results would contribute to the system's untrustworthiness (Kazim and Koshiyama, 2021). For relevance, robustness and societal acceptance, concerns about transparency in the analytical model must be addressed (Mathrani et al., 2021). Interrogation is both ethical and effective, and has been used by excellent individuals in extreme circumstances (Balfe, 2021). Accountability is a moral phenomenon that should be ethically considered (Shearer, 2002). Integrating goal-directed effortful decision making and norm-activation processes into a theoretical framework can provide a more complete and precise understanding of individuals' responsible decision making processes or behavior (Han and Hwang, 2016). The social benefits have proven to be difficult to quantify, particularly because their objectives are vague (Yang et al., 2021). Employee perceptions of ethically suitable acts and rules witnessed in the workplace are reflected in the EWC (Schwepker et al., 1997).

Fig. 2 depicts the conceptual model based on the constructs defined in Table 1 and the propositions. In the figure, propositions P1 to P9 represent the direct effect and P10a to P10h represent the moderating effect.

3. Methodology

For researchers investigating novel phenomena and theorising new constructs, qualitative research methodology has been regarded as the preferred option (Eisenhardt and Graebner, 2007; Lee and Lee, 1999). As a result, the study used key informant interview and thematic analysis as the method of preference (Eisenhardt et al., 2016). For maximal information elicitation, the study used a semi-structured open-ended interview format. We were unable to identify a suitable five-point or seven-point Likert scale questionnaire for the measurement items (i.e., wellbeing expectancy, human-centred values expectancy, interrogation expectancy, accountability expectancy) in the context of this study, hence qualitative research methodology was used.

The data were collected from India, a developing economy with the fastest growth (Sohal et al., 2022). The data were collected by ten interviewers who were well-versed in conducting personal interviews in English. The respondents who agreed to participate in the survey were given a consent form and an information sheet narrating the objective of the study. For more than two decades, India's expansion has been fuelled by its IT industry, which is regarded as one of the best in the world (Jha et al., 2020). Since India is a diverse country with many people speak and write many languages, NLP plays an important role in multilingual settings i.e., an ordinary person should be able to understand the syntax (what the words imply) and semantics (what the words mean) when utilising NLP (Miner et al., 2012). The study conducted a total of 15 interviews. It is recommended that a minimum sample size of 12 be employed in qualitative studies to ensure data saturation (Braun and Clarke, 2016; Fugard and Potts, 2015). As a consequence, 15 key informants are regarded as a sufficient sample size for the data collection, which avoids/reduce design bias. We argue that the study through natural enquiry achieved a sampling data saturation at 15, where further sampling sources expect to add little value to the study, which coincides with (Braun and Clarke, 2016; Fugard and Potts, 2015) that profess data

Table 2
Responder's profile.

Serial No.	Gender	Firm size	Age	Industry	Experience	Interview length
1	Male	>1000	48	Healthcare	<5	45
2	Female	>1000	52	Retail and e-commerce	>10	35
3	Male	>1000	49	Healthcare	5–10	55
4	Male	>1000	51	Retail and e-commerce	>10	45
5	Male	1–100	52	Finance	<5	40
6	Male	>1000	55	Retail and e-commerce	>10	55
7	Male	>1000	53	Healthcare	>10	55
8	Male	1–100	49	Healthcare	<5	45
9	Male	>1000	50	Retail and e-commerce	5–10	60
10	Male	100–1000	53	Healthcare	<5	35
11	Male	1–100	54	Finance	<5	60
12	Male	100–1000	55	Healthcare	<5	65
13	Female	100–1000	55	Healthcare	<5	65
14	Female	>1000	49	Retail and e-commerce	>10	50
15	Male	>1000	55	Healthcare	>10	40

saturation with an adequate sample design size of 12.

The study used the data triangulation approach (Jick, 1979) to improve the validity and credibility of findings. Thus, the 15 informants represented senior-level positions in diversified industries i.e., retail and e-commerce, healthcare and finance. In India, there is a growing interest in the continuous usage of NLP in such industries, wherein the senior-level informants reflect industry-level perspectives as well as personal experience with the ethical issues of NLP. NLP is quickly becoming one of the most popular and important technologies for businesses today (Bhattacharjee et al., 2022), and hence, it can be concluded that mainstream businesses across the globe are now showing major signs of NLP adoption. Since generalisation is an essential element of the scientific process, the above argument ensures that the findings of the study are generalisable to other cultural contexts. Studies focusing on novel phenomena with inclusion of one or few industries may lack support for findings, therefore, consideration of multiple industries are worthy and justified. Furthermore, the data were acquired in a semi-structured fashion to gather information about a topic (i.e., RNLP) that was so complex that a structured interview or survey would not have yielded the same results. The semi-structured interview has the advantage of allowing for a deeper knowledge of the respondent's opinions and the reasons for them (Carruthers, 1990). Third-party newsletters and reports of such industries were categorised according to different use cases to ensure the consistency and dependability of the collected data. Senior-level interviewees (aged between 48 and 55) were chosen based on their extensive experience in the fields of NLP and ethics, as well as their ability to interact with IT products or services, which avoids/reduces the sampling bias. The interviewers were contacted through personal networks and digital channels, which were spread across geography.

The firms considered for this study had an industrial research ethics committee (IREC), which is responsible for ensuring the research's ethical safety and scientific merit. It has the power to reject, authorise, or terminate the research, as well as to require changes to the research protocol. The IREC's primary mission is to ensure the safety of potential research subjects (i.e., RNLP) and to assess the risks and benefits to the community. As a result, researchers consulted the IREC and submitted an application stating the study's nature and purpose. Before approaching the informants, the IREC grants formal approval to guarantee that the research methodology is evaluated for cultural and legal suitability.

During the interview, the respondents were informed about the purpose of the study to avoid/reduce response bias. They were briefed on the importance of NLP in business (e.g., sentiment analysis, market intelligence, customer-centring service etc.), the benefits of NLP, the relationship between responsible innovation and NLP, the ethical principles of NLP, alignment of ethical principles of NLP to responsible innovation, responsible decision making in the context of business, ethical models, what ethical problems NLP can solved, how to promote ethical practices, what NLP has to do in ethics, why business should invest in NLP from the perspective of ethics and social benefits. The participants were also requested to report the extent to which they have adopted NLP tools and technologies in their operational systems, as well as any ethical concerns they have encountered. Follow-up questions were only asked if the researchers needed more information on certain ethical concerns or if they needed clarification on a particular ethical point.

The snowball sampling method assists in understanding a topic that has been obscured by the veil of ignorance (Yadav et al., 2022), and is ideally suited to the investigation of a specific phenomenon (Cabeza-Ramírez et al., 2022). Therefore, after each interview, the study employed snowball sampling to collect information on prospective future informants who satisfied the study inclusion criteria. To guarantee that most information could be acquired, all participants were given anonymity and the interviews were held in a semi-structured manner. For internal control, the researchers created a broad list of thematic questions along which the interview would be conducted. The interview was free flowing, as the researchers attempted to ensure that the respondents explored all of the broad questions. Table 2 represents the responder's profile, wherein firm size is measured in terms of employees, experience in years and interview length in approximate minutes. As per the table, 3 females and 12 males senior functionalities participated with the firm size distribution of 1–100, 100–1000 and >1000 employee strength; experience of the firm with the distribution of <5 years, between 5 and 10 years and >10 years and the age of informants is between 48 and 55. The average length of the interview was 50 min. The paucity of female respondents (three out of fifteen) indicates that women are underrepresented in such positions in different industries. In India, organised leadership and coaching programmes aimed specifically at females are lacking (Messner, 2017), which could be one of the reasons why men are holding senior-level positions in industries considered for this study.

The data were analysed and coded using the MAXQDA software tool, which was also used to record and transcribe the interviews. For the sake of anonymity, the names of interviewees and their firms have been omitted. The researchers alternated between conducting interviews and analysing the data. There were instances where theoretical insights in data analysis were used to determine further data requirements. The data gathering period lasted from June 2021 through March 2022. Between January and March 2022, the final transcription and coding were completed. As a methodological framework, the study followed a five step approach for qualitative data analysis i.e., Step 1: collection of data, Step 2: organisation of data, Step 3: theoretical coding, Step 4: analysis of the data for insights and Step 5: reporting on the insights from the analysis. The prime objective of the analysis was to examine the qualitative patterns through theoretical coding for the identification of themes in the form of core category, subcategory, and subsequently, to

Table 3
Category of codes with sample quotes.

Core category	Sub-category	Sample quote
NLP principlist ethical principles	Wellbeing expectancy	The most significant benefit of RNLP is that it frees people up to accomplish the more creative, vital, and innovative things that humans would rather be doing by automating some of the most monotonous, time-consuming jobs in many industries.
	Human-centred values expectancy	At its best, technology can help to reduce bias and benchmark diversity. It can also be argued that technology has the potential to improve human autonomy and respect human rights. So, while making decisions, the RNLP system should take responsibility and respect the autonomy, human rights and diversity of individuals.
	Fairness expectancy	Since RNLP impacts on how individuals react to the decision, therefore the RNLP system should be inclusive and accessible to every individual and should not lead to chauvinism against individuals.
	Privacy and security expectancy	Trust is required for privacy protection and security provisions. Violations of these privacy provisions put an individual's security at risk, as well as breaching the law. Law offers a foundation for businesses, whereas ethics can assist a business in filling up the gaps while making various decisions. Simply stated, when data privacy is violated and security is compromised, customers lose faith in the business. The business is then exposed to legal liabilities and security risks, as well as a breach of ethical values. The net outcome is bad news for the business. Therefore, RNLP system should be more privacy-conscious and tackle security problems by identifying malicious domains, determining patterns associated with known vulnerabilities, potentially vulnerable code segment and identifying phishing attempt.
	Reliability expectancy	Since businesses rely on their reputations, it is critical for the RNLP systems to set clear and consistent expectations for ethical norms to assist decision makers.
	Transparency expectancy	When done correctly, transparency can increase society's motivation and satisfaction, as well as marketing, traffic and many other areas of business. In addition, a transparent decision making process prevents misalignment. Therefore, RNLP system should enhance transparency.
	Interrogation expectancy	If there is one thing society can count on in RNLP, it is that nothing remains the same for long i.e., RNLP algorithms need to be upgraded, if not replaced, from year to year. Industries must keep up with this in order to remain competitive. Investing in a principle that tries to ensure the provision of efficient, accessible methods that allow individuals to question or challenge the use or output of RNLP systems will offer the business an advantage.
	Accountability expectancy	The industry must demonstrate that it is more than just a generator of output by holding the RNLP system accountable when it makes poor decisions. This requires inculcating a culture of continuous learning by the people responsible for the various phases of the RNLP system lifecycle development, the expansion of their knowledge. Therefore, the decision making by RNLP systems must reflect in the legal framework of the business.
Responsible decision making	None	RNLP enabled responsible decision making identifies the constructive options in the decision, build confidence and reflect the ethical result of the decision and social interactions for the wellbeing of the business and its stakeholder.
Ethical work climate	None	EWC of an organisation is crucial because it can strengthen organisational commitment, and generate a more engaged and loyal information technology workforce such as RNLP systems.

identify the ethical practices followed by the formulation of propositions.

4. Results

Table 3 depicts the result of the analysis and theoretical coding. The top-level actionable variables are acted upon by management to make up the core category of thematic codes. The MAXQDA sub-category codes enable to undertake investigation of the core theme's component to generate the propositions.

The result conduce two ethical practices. First is the adoption of RNLP as a disruptive technology for ethical decision making for social benefits. Digital expansion is a business necessity and disruptive technology is frequently the only way to break even, not only to keep up with the competition. Embracing RNLP does not have to be solely accompanied by ethical issues; rather, it is how RNLP is applied that determines whether or not that technical advancement will result in a better world. Respect for wellbeing, human-centred values, fairness, privacy and security, reliability, transparency, interrogation and accountability are essential for building and maintaining a healthy relationship between RNLP and the public for the achievement of social benefit. This practice is in line with (Scuotto et al., 2022) i.e., several factors can impact a user's decision to adopt disruptive technology, and employees will be encouraged to accept and use disruptive technology.

Second is the creation of a culture of responsibility. The industry must foster a culture of technological responsibility such as RNLP. More governance, ethical attitude and fair use of data can be seen if the IT workforce and industry titans believe they are responsible for the safe and ethical use of NLP. Although professional excellence is still not fully or regularly integrated into technical practice, ethical reflection and practice are necessary in technology. Individual practitioners and industry alike must take active measures to make ethical contemplation and practice mainstream and pervasive. To be effective, ethical contemplation and practice must be conducted iteratively. Since technologies are constantly evolving, the industry must treat NLP ethics as an active and ongoing learning cycle in which individuals continuously observe the ethical outcomes of practices, learn from errors, gather more information, hone ethical and technical expertise, and then update and improve operating practices accordingly. This practice is in line with (Chen, 2022) i.e., a culture of responsibility emerges over time and can be influenced and changed significantly by its members.

The results are also presented as a set of propositions from the theoretical code and data analysis. The propositions shown in the conceptual model are elaborated further to establish the relationship between the constructs.

4.1. Wellbeing expectancy and responsible decision making

RNLP has a good impact on society by making advances that have made life easier for humans, from being able to properly store and analyse data across many industries to improve daily routines with virtual and home assistants. The study advocates the ability of RNLP systems to make constructive choices that benefit individuals, society based on ethical standards can positively realise the wellbeing of oneself and others. Therefore, the proposition is postulated as:

P1. Wellbeing expectancy of NLP systems will positively influence responsible decision making.

4.2. Human-centred values expectancy and responsible decision making

With the rapid development of technology, full-fledged automation without the consideration of ethics can raise serious questions about potential adverse impacts on the rights, autonomy and diversity of individuals. The study advocates that the ability of RNLP systems that focus on creating better human experiences can positively balance the need for human-centred values by achieving the desired impact. Therefore, the proposition is postulated as:

P2. Human-centred values expectancy of NLP systems will positively influence responsible decision making.

4.3. Fairness expectancy and responsible decision making

The requirements of fairness in decision making by the RNLP system can be accomplished in a variety of ways, which includes a decision's significance and its type to the individual and the method employed in making the decision. The paper advocates that the good use of NLP systems improves fairness among individuals by demonstrating open-mindedness, making a reasoned judgment after analysing information, data, and facts, and evaluating personal, interpersonal, community, and institutional impacts. Therefore, the proposition is postulated as:

P3. Fairness expectancy of NLP systems will positively influence responsible decision making.

4.4. Privacy and security expectancy, and responsible decision making

Securing data privacy is a difficult task. When a business acquires personal information from an individual, it has several responsibilities to that individual, including trust, security, legal, and, most importantly, ethical responsibility. The study advocates that the moral principles and behaviours of those participating, as well as the norms of the profession, will form the foundation of RNLP systems and such systems should be secure and respect privacy. Therefore, the proposition is postulated as:

P4. Privacy and security expectancy of NLP systems will positively influence responsible decision making.

4.5. Reliability expectancy and responsible decision making

Reliability is sometimes referred to as the hallmark of ethics. Decision making requires reliability in the sense that the moral standards, actions and values of the RNLP system should not be contradictory. The study advocates that reliability is surely not sufficient for responsible decision making, it is at least necessary for making such decisions in NLP systems. Therefore, the proposition is postulated as:

P5. Reliability expectancy of NLP systems will positively influence responsible decision making.

4.6. Transparency expectancy and responsible decision making

Transparency is the cornerstone of ethics and is an attribute of society. From the decision making perspective, transparency infers into the RNLP system's clear, complete, and intuitive reporting, as well as the atmosphere in which information is understandable, accessible, and visible to all consumers of society. The study advocates that transparency infers the way RNLP systems offer decision making outcomes that have been free of hidden conditions for free and open exchange of information, with rules and reasoning that have been fair and apparent to all consumers. Therefore, the proposition is postulated as:

P6. Transparency expectancy of NLP systems will positively influence responsible decision making.

4.7. Interrogation expectancy and responsible decision making

Technology is at the heart of today's businesses; therefore, RNLP systems necessitate feedback, which is output that is returned to a person, community, group in order to assist them in evaluating or correcting the input data. The study advocates that each level of decision making is served by a different type of RNLP system, i.e., operational, tactical, and strategic and the business should resolve any implementation related challenges of NLP systems and learn from the past failures by allowing individuals to challenge the use of such systems. Therefore, the proposition is postulated as:

P7. Interrogation expectancy of NLP systems will positively influence responsible decision making.

4.8. Accountability expectancy and responsible decision making

Ethics and accountability go hand in hand in business. To inculcate ethical responsibility in society, a list of ethical responsibilities is not enough. The study advocates that the roles that RNLP systems take on imply that they are responsible for performing certain functions connected with assigned responsibilities and its decision making should be judged by society. Therefore, the proposition is postulated as:

P8. Accountability expectancy of NLP systems will positively influence responsible decision making.

4.9. Responsible decision making and social benefits

The development of responsible decision making abilities is critical for society. Businesses should prepare for real-world challenges that will affect their long-term sustainability by building responsible decision making abilities through RNLP systems. The study advocates that RNLP systems should be built with the capability to make responsible decision making based on principlist ethical principles and social norms for the realistic assessment of the repercussions of various actions and the consideration of society's wellbeing. For social benefits, responsible decision making necessitates a high level of ethical discipline in NLP systems. Therefore, the proposition is postulated as:

P9. Responsible decision making of NLP systems will positively benefit society.

4.10. Moderating role of EWC

NLP systems must act as role model for society to build an organisational culture that fosters ethical behavior by augmenting responsible decision making. EWC enables a business to run its operations in an honest, respectable and proper manner. The study advocates that every organisation operates in an ethical work climate and when such organisation's information systems (e.g., RNLP systems) demonstrate and produce highly responsible decision making, it is likely that the organisation has an ethical atmosphere. In a similar notation, when information systems make unethical decisions on a routine basis, an unethical work climate emerges. Therefore, the propositions are postulated as:

P10a. EWC moderates the relationship between wellbeing expectancy and responsible decision making.

P10b. EWC moderates the relationship between human-centred values expectancy and responsible decision making.

P10c. EWC moderates the relationship between fairness expectancy and responsible decision making.

P10d. EWC moderates the relationship between privacy and security expectancy and responsible decision making.

P10e. EWC moderates the relationship between reliability expectancy and responsible decision making.

P10f. EWC moderates the relationship between transparency expectancy and responsible decision making.

P10g. EWC moderates the relationship between interrogation expectancy and responsible decision making.

P10h. EWC moderates the relationship between accountability expectancy and responsible decision making.

5. Discussion

For the propositions, the findings are presented as the synthesis of the result of data analysis and theoretical code into the literature. In addition, it includes theoretical implications, managerial implications and limitations and future research.

5.1. Wellbeing expectancy and responsible decision making (P1)

Literature suggests that by modernising information and knowledge management that supports human rights monitoring in conflict zones, a human rights organisation has seen demonstrable gains through the use of state-of-the-art NLP methodologies (Alhelbawy et al., 2020). Justifying decisions in the interests of individuals, society and the environment is a prerequisite for responsible decision making (Arrieta et al., 2020) and such justifications will become increasingly important in the future (Biran and Cotton, 2017). Depending on the user's background knowledge, different explanations for the reasons behind a forecasting model can be deduced and the NLP system must deliver enough information to the user so that the user can clearly comprehend the models' logic and make responsible decisions (Doran et al., 2017).

5.2. Human-centred values expectancy and responsible decision making (P2)

Literature suggests that human rights encompass a range of obligations and technology provides enormous potential for improving many aspects of human life to get unique insights and improve decision making in a specific context (Nersessian, 2018). It is essential to take into account the diversity of individuals and technology is one type of input or resource that can be used to fulfil such capabilities to improve the fit between decision making proclivities (Ryan, 2022). Autonomy is an ethical factor, and technologies should be employed in a way that respects human autonomy (Formosa et al., 2021).

5.3. Fairness expectancy and responsible decision making (P3)

Literature suggests that the ethical principle of human equality underpins fairness (Kazim and Koshiyama, 2021). The emerging research community is putting effort into developing fairness frameworks (Chouldechova, 2017). However, due to the complexities of fairness as a concept, it is important to understand how individuals perceive algorithmic-based fairness (Kleanthous et al., 2022). Fairness is a complex concept that is interpreted differently by different actors and individuals understand it differently depending on the context in which the system is used (Hutchinson and Mitchell, 2019). Fairness has been proven as a behavioural factor by industry observations as well as behavioural research and it has a significant impact on an individual's decision making (Kang et al., 2021).

5.4. Privacy and security expectancy, and responsible decision making (P4)

Literature suggests that the ethical and privacy issues surrounding the use of technology have piqued the curiosity of a diverse range of audiences (Zhang et al., 2021). In the creation and implementation of advancements in and around information and communication technology, ethical issues and concerns about privacy and social values abound (Raab, 2020). Unauthorized access, disclosure or misuses of sensitive private information are all instances of privacy breaches (Chen et al., 2021). As consumers are more sensitive to privacy and security issues, the digital business privacy and security policy can have a substantial impact on its reputation and value (Lanfang and Peng, 2014). The issues related to security and privacy need to be addressed from the perspective of ethics (Arpaci et al., 2015). Technology has significant implications for privacy and security (Blakesley and Yallop, 2019). Security threats may act as cost-increasing fear mechanisms (Chen and Li, 2017).

5.5. Reliability expectancy and responsible decision making (P5)

Literature suggests that reliability is only one aspect to consider while deciding whether or not to trust an information system (O'Neill, 2002). Improving the output of the decision making process is regarded as a direct effect of improving the information system reliability (Dahmani et al., 2016). Subjective insights improve the decision making process's reliability and an automated information system can facilitate it (Egan and Haynes, 2018).

5.6. Transparency expectancy and responsible decision making (P6)

Literature suggests that transparency given by timely information generation and transmission can help to speed up the process and encourage efforts to identify effective solutions to complex challenges (Archibald et al., 2021). Increasing transparency has a positive impact on business (Khosroshahi et al., 2021) and consumer satisfaction (Simintiras et al., 2015). Lack of transparency in information technology-based decision making is frequently at the root of disputes (Zhdanov et al., 2021). Legislative and regulatory trends are progressively requiring algorithm transparency, making explanations more necessary than desirable (Kim et al., 2020). Regardless of the differences in transparency criteria among different sources of law, the ultimate result is the same i.e., decision makers must understand why a particular conclusion is proposed by a decision model in order to make an adequately informed decision (Mourby et al., 2021).

5.7. Interrogation expectancy and responsible decision making (P7)

Literature suggests that information about data use, bias, accuracy, and the diagnostic role of information technology is required for effective interrogation (Ploug and Holm, 2020). Reframing fallible companion technology opens up the possibility of challenging the underlying logic of the algorithms and starts to understand how decisions are made regarding them (Gal et al., 2020). The question of whether only the wealthy elderly will profit from autonomous systems, or whether the less fortunate elderly will be disproportionately affected by the unintended consequences of autonomous systems, has been posed (Tan et al., 2021).

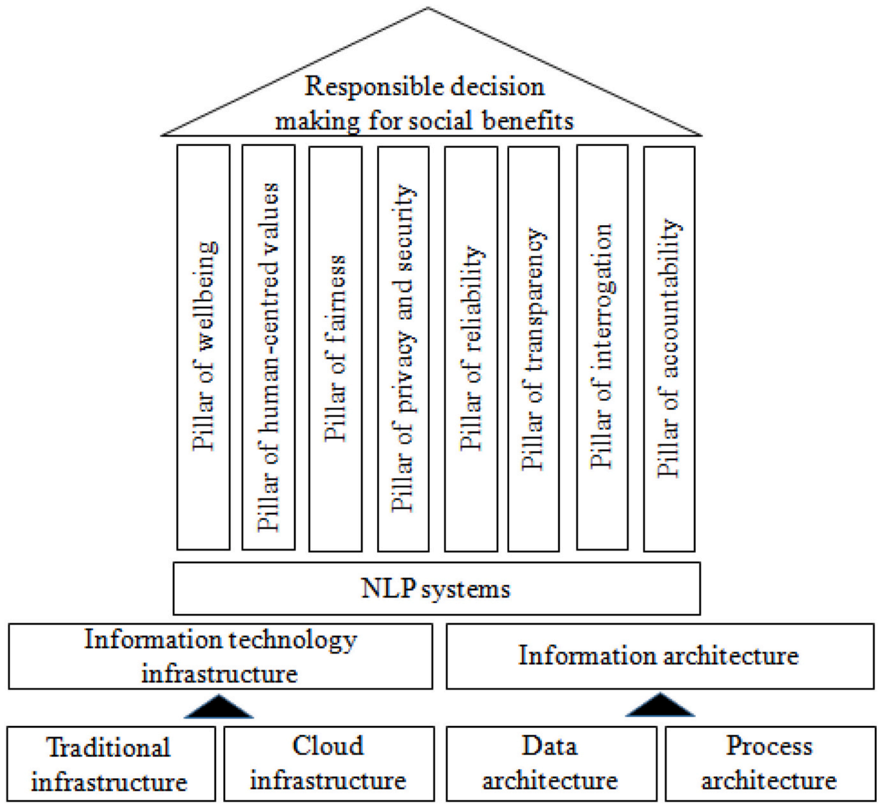

Fig. 3. Pillars of NLP based principlist ethical principles.

5.8. *Accountability expectancy and responsible decision making (P8)*

Literature suggests that accountability is regarded as a crucial enabler of ethical information systems (Srinivasan and González, 2022). For information systems, the need to build effective accountability mechanisms is a fundamental challenge due to the involvement of multiple actors. Therefore, the goal of developing effective accountability mechanisms should be to limit the power of those who have the potential to harm (Maas, 2022). Achieving a strong objectivity for the ethical outcomes of the information systems may be difficult, but establishing more recorded results can assure meaningful responsibility (Wagner, 2020).

5.9. *Responsible decision making and social benefits (P9)*

Literature suggests that responsible decision making is a sub-domain of social competence that emerges from the social learning process (Corcoran et al., 2020). Misuse of technology is linked to a lack of ability to make responsible decisions and threatens society (Nasaescu et al., 2018). To understand social responsibility, a multi-layered analysis of responsible decision making is required (Aguinis and Glavas, 2012).

5.10. *Moderating role of EWC (P10a to P10h)*

Literature suggests that building long-term relationships with its customers requires an ethical work climate (Schwepker and Hartline, 2005) and responsible decision making boosts long-term relationships (Ferreira et al., 2015). There has been the focus of countless research in the ethical work climate and its impact on sales and marketing practices and outcomes (Mulki and Lassk, 2019). The ethical climate has a substantial influence on behavioural decisions (Forte, 2004) and especially crucial for sales organisations (DeConinck, 2011). In the decision making process, ethical sensitivity is critical (Deshpande et al., 2000; Kotler and Zaltman, 1971). There is a growing need to uncover the intuitive factors that underpin ethical decision-making processes and frameworks (Valentine and Godkin, 2019). By defining policies and procedures and creating behavioural expectations, an organisation's ethical climate can provide structure (DeConinck, 2010).

5.11. *Theoretical implications*

The study capitulates radical theoretical implications. First, the study proposed an NLP based responsible decision making model for social benefits. The proposed model is grounded on NLP-based eight principlist ethical principles, which are the independent variables, responsible decision making as the mediating variable, social benefits as the dependent variable and finally, EWC as the moderator variable. In doing so, it extends the conceptual framework of Victor and Cullen (1987) from the viewpoint of the EWC that is highly linked to job performance and positive attitudes and behaviours among employees (Pagliaro et al., 2018) of a business to build principlist ethical principles in NLP systems. In addition, it also extends the conceptual framework of Garrigan et al. (2018) from the viewpoint of moral decision making that guides moral development studies for a typical and atypical society.

Second, the study integrates responsible innovation and NLP, which promotes creativity and opportunities for NLP in the ethical dimension, which business can undertake in the interest of society. Since its inception, responsible innovation has received traction (Owen et al., 2020) and several businesses have effectively implemented NLP (Perboli et al., 2021). For businesses, the integration of responsible innovation and NLP entails thorough assessment of, and action to address, the potential consequences of releasing a new product, service, process and business model to the market. In addition, it can look at the impact on innovating business, the relationship between customers and suppliers, investors and wider society. When the advantages to society can be realised and any potential drawbacks can be avoided, NLP based new innovative products, services, processes, and business models are preferable.

Third, the study discussed eight pillars of NLP based principlist ethical principles (see Fig. 3). Ethical principles are regarded as a crucial component in decision making and the general rule is that if there is a risk, strict governance standards must be applied to the design, training, implementation and operation of information systems (Behera et al., 2022). The study argues that since search engines, chatbots and other NLP algorithms are not humans, businesses must apply principlist ethical principles on a wide scale, i.e. on a global scale, to NLP systems. Since NLP systems are widespread and utilised by a wide range of individuals in a range of situations, businesses must ensure that ethical principles are objective and unbiased. Even if implemented with the best of intentions, the business must ensure that decision making of NLP systems should not have a negative influence on marginalised groups. As depicted in Fig. 3, principlist ethical principles are built on the NLP system with the implementation of appropriate algorithms which results in responsible decision making for social benefits.

The NLP systems are grounded on IT infrastructure and information architecture. The IT infrastructure comprises hardware (i.e., computers, servers, routers, switches etc.) and software (i.e., operating systems, web servers, customer relationship management, content management systems etc.). The two types of infrastructure recommended namely the traditional infrastructure or cloud infrastructure. The information architecture captures and stores the information needs and formalises it, and is composed of two models, i.e., data architecture and process architecture. The data architecture outlines an organisation's information needs, whereas the process architecture specifies its operations (Kim and Lee, 1996).

5.12. *Managerial implications*

The study has substantial managerial implications and will be particularly valuable to practitioners. First, decision making is crucial because it evaluates if feasible solutions will solve significant problems for a given scenario (Liu et al., 2022). Therefore, the paper advocates that the responsible decision making model is at the heart of developing scenario-based NLP systems for making better decisions. Making responsible decisions can be established in five steps. In Step 1, the problem to be identified, Step 2 to analyse the scenario, Step 3 is the development of NLP solutions or algorithms and subsequently building the system to solve the problem matching to the given scenario, Step 4 is the adoption of ethical responsibilities with the consideration of principlist ethical principles of NLP and the final step (i.e., Step 5) is to evaluate and reflect. Multi-facet solutions have different strengths and challenges when it comes to making responsible decisions, and these challenges can be best addressed by focusing on the strengths of the business. The steps progress in a sequential manner, which can help the business to ease the walk of complicated process of making responsible decisions. With the adoption of such a stepwise decision making model, society will become open-minded and will show confidence in the responsible decision making choice of NLP systems.

Second, for a full-fledged implementation of the NLP systems supporting the eight principlist ethical principles, an optimal information technology infrastructure is recommended for the business (Radkov, 2018), as it offers businesses high-performance storage, a low-latency network, security infrastructures, efficient wide area network (WAN), virtualisation and zero downtime. High-performance storage systems back up and store data, as well as providing a data recovery solution in the event of a disaster. Low-latency networks reduce data flow delay by utilising enterprise-level infrastructure components. Secure infrastructures protect businesses from data breaches and cyber-attacks no matter where the data is stored, ensuring that consumer's trust is maintained. WANs control the network by prioritising traffic and allocating more or less bandwidth to certain applications as needed. Virtualisation speeds server deployment, improves uptime, enhances disaster recovery, and reduces energy use. Zero downtime strives to minimise business disruptions and eliminate system downtime in order to keep costs low and profits high.

Table 4
RNLP use cases of healthcare industry for social benefits.

Use case	Description	Reference
Dictation	Physicians dictate clinical procedures and outcomes to a voice recorder, who then transcribe the processes into texts and input them into electronic medical records (hereafter, EMR) and electronic health record (hereafter, EHR) systems. NLP is used to analyse speech recordings and transform them to text so that they can be fed into EMRs and patient files.	Hoffer et al., 2012
Clinical documentation	NLP is coupled with optical character recognition (hereafter, OCR) to extract healthcare data from EHRs, notes from physicians or medical forms so that it can be input into data entry software. This dramatically decreases data entry time and improves data quality by eliminating human errors.	Stocking et al., 2020
Clinical diagnosis	Medical models that recognise disease criteria based on standard clinical language and medical word usage are built using NLP.	Remes et al., 1991

Third, by discriminating against specific social groups and forming the biased associations of individuals through the media they are exposed to, biased NLP algorithms have an immediate harmful impact on society. Therefore, Hovy and Prabhumoye (2021) proposed sources of bias in NLP systems, i.e., data and models. Data bias in NLP and model bias is a serious problem that needs to be addressed immediately. For data bias, the early stage would be debiasing the data and for model bias, the early stage would be debiasing the model. For debiasing the data, the successful approach is to remove existing features or observations that are filled with social biases including racist, misogynistic and demeaning labels. This does not rule out the use of input data, but it does mean that it should be removed or edited to account for bias. In the same way, new data must be auto examined for bias. Debiasing the model is accomplished by altering the vector representations of words to remove stereotypes while retaining useful information.

5.13. *Limitations and future research directions*

There are limitations to this study which pave the avenues for future research. First and foremost, the goal of this study was theoretical saturation rather than comprehensive empirical research. As a result, any conclusions on NLP based principlist ethical principles should be approached with caution and in the context of responsible decision making for social benefits. Hence, future research can be instituted through in-depth interviews and quantitative research methodology to validate and evolve the proposed conceptual model with instruments in developing countries like India (Behera et al., 2021b) and developed countries.

Second, the proposed principlist ethical principles may not be feasible or practicable for a few industries. In addition, implementation of such principles requires a good understanding of NLP technology and ethics. Therefore, knowledgeable resources are required for the correct implementation and failure to do so can result in a huge drain on business. It can cost businesses millions of dollars in productivity, expense of technology and infrastructure. Hence, future research can be instituted to experiment the following use cases.

5.13.1. *Use cases of healthcare industry*

The RNLP use cases for the healthcare industry for social benefits are captured in Table 4.

5.13.2. *Use cases of finance industry*

The RNLP use cases for the finance industry for social benefits are

Table 5
RNLP use cases of finance industry for social benefits.

Use case	Description	Reference
Insurance claims management	To assess insurance claims, NLP is combined with OCR.	Comuzzi et al., 2012
Financial auditing	By scanning financial records of a business and recognising document similarities and differences, NLP enables the automation of financial audits. As a result, irregularities and anomalies in financial statements might be detected.	Knechel et al., 2020
Fraud detection	To detect fraud from financial records, NLP is integrated with machine learning and predictive analytics.	Khan et al., 2022
Stock prices prediction	To assess current financial news on the web, NLP is combined with KNN classification algorithms to support 'news-based trading,' in which analysts aim to isolate financial news that influences stock prices and market activity.	Jing et al., 2021

Table 6
RNLP use cases of retail industry for social benefits.

Use case	Description	Reference
Customer service cognitive chatbot	Use NLP to provide automated answers to repetitive questions.	Behera et al., 2021a
Semantic search	Semantic search is a search approach that seeks to comprehend the context of the search query and provide appropriate responses in addition to finding keywords. To use long-tail search strings, understand the shopper's intent, and increase the visibility of a variety of products, many retailers rely on NLP-powered semantic search engines.	Malhotra and Rishi, 2021

captured in Table 5.

5.13.3. Use cases of retail industry

The RNLP use cases for the retail industry for social benefits are captured in Table 6.

6. Conclusions

Industries can use RNLP to define important objectives and develop governance strategies, resulting in solutions that allow NLP systems and businesses to thrive. Therefore, industries must address the major challenge of converting NLP principlist ethical principles into practical and measurable metrics that should operationally work. In addition, they also require the appropriate organisational, technological, operational, and reputational framework to integrate these into day-to-day processes. Organisations may set core objectives and build governance strategies with RNLP, resulting in systems that enable NLP and business to flourish. Future payoffs in RNLP will give early adopters a competitive advantage that competitors may never be able to overtake. The result conduces two ethical practices i.e., adoption of RNLP as a disruptive technology for ethical decision making for social benefits and creation of a culture of responsibility. Organisations can become smarter, more agile, more adaptable and more responsive as a result of the application of disruptive technologies in ethical decision making. The difference between high and low performing organisations is attributable to a culture of responsibility. In particular, the study contributes to the domain area of information systems and ethics.

Data availability

Data will be made available on request.

References

Agarap, A.F., 2018. Statistical analysis on E-commerce reviews, with sentiment classification using bidirectional recurrent neural network (RNN) arXiv preprint arXiv:1805.03687.

Agrawal, A.K., Jagannathan, M., Delhi, V.S.K., 2021. Control focus in standard forms: an assessment through text mining and NLP. J. Leg. Aff. Disput. Resolut. Eng. Constr. 13 (1), 04520040.

Aguinis, H., Glavas, A., 2012. What we know and don't know about corporate social responsibility: a review and research agenda. J. Manag. 38 (4), 932–968.

Akrami, F., Zali, A., Abbasi, M., Majdzadeh, R., Karimi, A., Fadavi, M., Bahar, A.M., 2018. An ethical framework for evaluation of public health plans: a systematic process for legitimate and fair decision-making. Public Health 164, 30–38.

Alabed, A., Javornik, A., Gregory-Smith, D., 2022. AI anthropomorphism and its effect on users' self-congruence and self–AI integration: a theoretical framework and research agenda. Technol. Forecast. Soc. Chang. 182, 121786.

Alhelbawy, A., Lattimer, M., Kruschwitz, U., Fox, C., Poesio, M., 2020. An NLP-powered human rights monitoring platform. Expert Syst. Appl. 153, 113365.

Allhoff, F., Lin, P., Steinberg, J., 2011. Ethics of human enhancement: an executive summary. Sci. Eng. Ethics 17 (2), 201–212.

Archibald, D.W., McIver, R., Rangeley, R., 2021. Untimely publications: delayed Canadian fisheries science advice limits transparency of decision-making. Mar. Policy 132, 104690.

Arpaci, I., Kilicer, K., Bardakci, S., 2015. Effects of security and privacy concerns on educational use of cloud services. Comput. Hum. Behav. 45, 93–98.

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, S., Benjamin, R., Chatila, R., Herrera, F., 2020. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf.Fusion 58, 82–115.

Ashok, M., Madan, R., Joha, A., Sivarajah, U., 2022. Ethical framework for artificial intelligence and digital technologies. Int. J. Inf. Manag. 62, 102433.

Balfe, M., 2021. Violence and violations: perspectives of non-professional audiences on enhanced interrogation. EthicsMed.Public Health 19, 100700.

Bannon, L., 2011. Reimagining HCI: toward a more human-centered perspective. Interactions 18 (4), 50–57.

Baregheh, A., Rowley, J., Sambrook, S., 2009. Towards a multidisciplinary definition of innovation. Manag. Decis. 47 (8), 1323–1339.

Behera, R.K., Bala, P.K., Ray, A., 2021a. Cognitive chatbot for personalised contextual customer service: behind the scene and beyond the hype. Inf. Syst. Front. 1–21.

Behera, R.K., Bala, P.K., Rana, N.P., Kizgin, H., 2022. Cognitive computing based ethical principles for improving organisational reputation: a B2B digital marketing perspective. J. Bus. Res. 141, 685–701.

Behera, R.K., Bala, P.K., Tata, S.V., Rana, N.P., 2021b. Retail atmospherics effect on store performance and personalised shopper behaviour: a cognitive computing approach. Int. J. Emerg. Mark. https://doi.org/10.1108/IJOEM-03-2021-0433.

Beu, D., Buckley, M.R., 2001. The hypothesized relationship between accountability and ethical behavior. J. Bus. Ethics 34 (1), 57–73.

Bhattacharjee, S., Delen, D., Ghasemaghaei, M., Kumar, A., Ngai, E.W., 2022. Business and government applications of text mining & Natural Language Processing (NLP) for societal benefit: introduction to the special issue on "text mining & NLP". Decis. Support Syst. 162 (C), 113867 https://doi.org/10.1016/j.dss.2022.113867.

Biran, O., Cotton, C., 2017. Explanation and justification in machine learning: a survey. In: IJCAI-17 Workshop on Explainable AI, 8(1), pp. 8–13.

Blakesley, I.R., Yallop, A.C., 2019. What do you know about me? Digital privacy and online data sharing in the UK insurance sector. J. Inf. Commun. Ethics Soc. 18 (2), 281–303.

Brand, T., Blok, V., 2019. Responsible innovation in business: a critical reflection on deliberative engagement as a central governance mechanism. J.Responsib.Innov. 6 (1), 4–24.

Braun, V., Clarke, V., 2016. (Mis) conceptualising themes, thematic analysis, and other problems with Fugard and Potts'(2015) sample-size tool for thematic analysis. Int. J. Soc. Res. Methodol. 19 (6), 739–743.

Brunner, M., Ostermaier, A., 2019. Peer influence on managerial honesty: the role of transparency and expectations. J. Bus. Ethics 154 (1), 127–145.

Buhmann, A., Fieseler, C., 2021. Towards a deliberative framework for responsible innovation in artificial intelligence. Technol. Soc. 64, 101475.

Cabeza-Ramírez, L.J., Sánchez-Cañizares, S.M., Santos-Roldán, L.M., Fuentes-García, F. J., 2022. Impact of the perceived risk in influencers' product recommendations on their followers' purchase attitudes and intention. Technol. Forecast. Soc. Chang. 184, 121997.

Carchiolo, V., Longheu, A., Reitano, G., Zagarella, L., 2019. Medical prescription classification: a NLP-based approach. In: In 2019 Federated Conference on Computer Science And Information Systems. IEEE, pp. 605–609.

Carruthers, J., 1990. A rationale for the use of semi-structured interviews. J. Educ. Adm. 28 (1).

Celik, I., 2022. Towards intelligent-TPACK: an empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. Comput. Hum. Behav. 107468.

Chang, V., 2021. An ethical framework for big data and smart cities. Technol. Forecast. Soc. Chang. 165, 120559.

Chen, C.C., Huang, H.H., Chen, H.H., 2020. NLP in FinTech applications: past, present and future arXiv preprint arXiv:2005.01320.

Chen, C.H., 2022. The mediating effect of corporate culture on the relationship between business model innovation and corporate social responsibility: a perspective from small-and medium-sized enterprises. Asia Pac. Manag. Rev. https://doi.org/10.1016/j.apmrv.2022.01.001.

Chen, H., Li, W., 2017. Mobile device users' privacy security assurance behavior: a technology threat avoidance perspective. Inf.Comput.Secur. 25 (3), 330–344.
Chen, R., Kim, D.J., Rao, H.R., 2021. A study of social networking site use from a three-pronged security and privacy threat assessment perspective. Inf. Manag. 58 (5), 103486.
Chopik, W.J., 2016. The benefits of social technology use among older adults are mediated by reduced loneliness. Cyberpsychol. Behav. Soc. Netw. 19 (9), 551–556.
Chouldechova, A., 2017. Fair prediction with disparate impact: a study of bias in recidivism prediction instruments. Big Data 5 (2), 153–163.
Chowdhury, G.G., 2003. Natural language processing. Annu. Rev. Inf. Sci. Technol. 37 (1), 51–89.
Choy, Y.K., 2018. Cost-benefit analysis, values, wellbeing and ethics: an indigenous worldview analysis. Ecol. Econ. 145, 1–9.
Clarke, R., 2019. Principles and business processes for responsible AI. Comput.LawSecur. Rev. 35 (4), 410–422.
Clegg, S., Kornberger, M., Rhodes, C., 2007. Organizational ethics, decision making, undecidability. Sociol. Rev. 55 (2), 393–409.
Coavoux, M., Narayan, S., Cohen, S.B., 2018. Privacy-preserving neural representations of text. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 1–10. Association for Computational Linguistics, Brussels, Belgium.
Comuzzi, M., Vonk, J., Grefen, P., 2012. Measures and mechanisms for process monitoring in evolving business networks. Data Knowl. Eng. 71 (1), 1–28.
Corcoran, R.P., O'Flaherty, J., Xie, C., Cheung, A.C., 2020. Conceptualizing and measuring social and emotional learning: a systematic review and meta-analysis of moral reasoning and academic achievement, religiosity, political orientation, personality. Educ.Res.Rev. 30, 100285.
Cronbach, L.J., 1988. Five perspectives on validity argument. In: Test Validity, pp. 3–17.
Curreen, M.P., 1995. A simple hypnotically based NLP technique used with two clients in criminal justice settings. Aust.J.Clin.Exp.Hypn. 23 (1), 51–57.
Dahmani, S., Boucher, X., Peillon, S., Besombes, B., 2016. A reliability diagnosis to support servitization decision-making process. J. Manuf. Technol. Manag. 27 (4), 502–534.
Day, A., Nielsen, K., 2017. What does our organization do 16 to help our well-being? Creating healthy workplaces and workers. In: An Introduction to Work And Organizational Psychology: An International Perspective, 295.
Dean, A., Shubita, M., Claxton, J., 2020. What type of learning journey do students value most? Understanding enduring factors from the NSS leading to responsible decision-making. J.Glob.Responsib. 11 (4), 347–362.
DeConinck, J.B., 2010. The influence of ethical climate on marketing employees' job attitudes and behaviors. J. Bus. Res. 63 (4), 384–391.
DeConinck, J.B., 2011. The effects of ethical climate on organizational identification, supervisory trust, and turnover among salespeople. J. Bus. Res. 64 (6), 617–624.
Deshpande, S.P., George, E., Joseph, J., 2000. Ethical climates and managerial success in Russian organizations. J. Bus. Ethics 23 (2), 211–217.
Doran, D., Schulz, S., Besold, T.R., 2017. What does explainable AI really mean? A new conceptualization of perspectives arXiv preprint arXiv:1710.00794.
Dubnick, M.J., 2003. Accountability and ethics: reconsidering the relationships. Int.J. Org.TheoryBehav. 6 (3), 405–441.
Duncan, N., Field, R., Strevens, C., 2020. Ethical imperatives for legal educators to promote law student wellbeing. Legal Ethics 23 (1–2), 65–88.
Egan, D., Haynes, N.C., 2018. Manager perceptions of big data reliability in hotel revenue management decision making. Int.J.Qual.Reliab.Manag. 36 (1), 25–39.
Eisenhardt, K.M., Graebner, M.E., 2007. Theory building from cases: opportunities and challenges. Acad. Manag. J. 50 (1), 25–32.
Eisenhardt, K.M., Graebner, M.E., Sonenshein, S., 2016. Grand challenges and inductive methods: rigor without rigor mortis. Acad. Manag. J. 59 (4), 1113–1123.
Feng, Q., He, D., Liu, Z., Wang, H., Choo, K.K.R., 2020. SecureNLP: a system for multi-party privacy-preserving natural language processing. IEEE Trans.Inf. ForensicsSecur. 15, 3709–3721.
Ferreira, F.A., Jalali, M.S., Meidutė-Kavaliauskienė, I., Viana, B.A., 2015. A metacognitive decision making based-framework for bank customer loyalty measurement and management. Technol. Econ. Dev. Econ. 21 (2), 280–300.
Field, A., Prabhumoye, S., Sap, M., Jin, Z., Zhao, J., Brockett, C., 2021. Proceedings of the 1st Workshop on NLP for Positive Impact. In: Proceedings of the 1st Workshop on NLP for Positive Impact.
Formosa, P., Wilson, M., Richards, D., 2021. A principlist framework for cybersecurity ethics. Comput.Secur. 109, 102382.
Fort, K., Couillault, A., 2016. Yes, we care! results of the ethics and natural language processing surveys. In: International Language Resources and Evaluation Conference (LREC).
Forte, A., 2004. Business ethics: a study of the moral reasoning of selected business managers and the influence of organizational ethical climate. J. Bus. Ethics 51 (2), 167–173.
Fugard, A.J., Potts, H.W., 2015. Supporting thinking on sample sizes for thematic analyses: a quantitative tool. Int. J. Soc. Res. Methodol. 18 (6), 669–684.
Gal, U., Jensen, T.B., Stein, M.K., 2020. Breaking the vicious cycle of algorithmic management: a virtue ethics approach to people analytics. Inf. Organ. 30 (2), 100301.
Garimella, A., Banea, C., Hovy, D., Mihalcea, R., 2019. Women's syntactic resilience and men's grammatical luck: gender-bias in part-of-speech tagging and dependency parsing. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Florence, Italy, pp. 3493–3498.
Garrigan, B., Adlam, A.L., Langdon, P.E., 2018. Moral decision-making and moral development: toward an integrative framework. Dev. Rev. 49, 80–100.
Gordon, D., Stavrakakis, I., Gibson, J.P., Tierney, B., Becevel, A., Curley, A., Collins, M., O'Mahony, W., O'sullivan, D., 2021. Perspectives on computing ethics: a multi-stakeholder analysis. J. Inf. Commun. Ethics Soc. https://doi.org/10.1108/JICES-12-2020-0127.
Guillamón, M.D., Ríos, A.M., Gesuele, B., Metallo, C., 2016. Factors influencing social media use in local governments: the case of Italy and Spain. Gov. Inf. Q. 33 (3), 460–471.
Han, H., Hwang, J., 2016. Cruise travelers' environmentally responsible decision-making: an integrative framework of goal-directed behavior and norm activation process. Int. J. Hosp. Manag. 53, 94–105.
Harrison, T.M., Sayogo, D.S., 2014. Transparency, participation, and accountability practices in open government: a comparative study. Gov. Inf. Q. 31 (4), 513–525.
Heylighen, F., 2008. Complexity and self-organization. In: Encyclopedia of Library And Information Sciences, 3, pp. 1215–1224.
Hoffer, D.N., Finelli, A., Chow, R., Liu, J., Truong, T., Lane, K., Punnen, S., Knox, J.J., Legere, L., Kurban, G., Gallie, B., Jewett, M.A., 2012. Structured electronic operative reporting: comparison with dictation in kidney cancer surgery. Int. J. Med. Inform. 81 (3), 182–191.
Hovy, D., 2015. Demographic factors improve classification performance. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, Beijing, China, pp. 752–762.
Hovy, D., Prabhumoye, S., 2021. Five sources of bias in natural language processing. Lang.Ling.Compass 15 (8), e12432.
Humphreys, M., 2015. Reflections on the ethics of social experimentation. J.Glob.Dev. 6 (1), 87–112.
Hutchinson, B., Mitchell, M., 2019. 50 years of test (un) fairness: lessons for machine learning. In: Proceedings of the Conference on Fairness, Accountability, And Transparency, pp. 49–58.
Hutchinson, T., 2020. Natural language processing and machine learning as practical toolsets for archival processing. Rec. Manag. J. 30 (2), 155–174.
Jha, A.K., Agi, M.A., Ngai, E.W., 2020. A note on big data analytics capability development in supply chain. Decis. Support. Syst. 138, 113382.
Jick, T.D., 1979. Mixing qualitative and quantitative methods: triangulation in action. Adm. Sci. Q. 24 (4), 602–611.
Jing, N., Wu, Z., Wang, H., 2021. A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. Expert Syst. Appl. 178, 115019.
Johannsen, A., Hovy, D., Søgaard, A., 2015. Cross-lingual syntactic variation over age and gender. In: Proceedings of the Nineteenth Conference on Computational Natural Language Learning. Association for Computational Linguistics, Beijing, China, pp. 103–112. https://www.aclweb.org/anthology/K15-1011.
John-Mathews, J.M., 2021. Critical empirical study on black-box explanations in AI arXiv preprint arXiv:2109.15067.
John-Mathews, J.M., 2022. Some critical and ethical perspectives on the empirical turn of AI interpretability. Technol. Forecast. Soc. Chang. 174, 121209.
John-Mathews, J.M., Cardon, D., Balagué, C., 2022. From reality to world. A critical perspective on AI fairness. J. Bus. Ethics 1–15.
Kang, K., Wang, M., Luan, X., 2021. Decision-making and coordination with government subsidies and fairness concerns in the poverty alleviation supply chain. Comput. Ind. Eng. 152, 107058.
Kazim, E., Koshiyama, A.S., 2021. A high-level overview of AI ethics. Patterns 2 (9), 100314.
Khan, A.T., Cao, X., Li, S., Katsikis, V.N., Brajevic, I., Stanimirovic, P.S., 2022. Fraud detection in publicly traded US firms using Beetle Antennae Search: a machine learning approach. Expert Syst. Appl. 191, 116148.
Khosroshahi, H., Dimitrov, S., Hejazi, S.R., 2021. Pricing, greening, and transparency decisions considering the impact of government subsidies and CSR behavior in supply chain decisions. J. Retail. Consum. Serv. 60, 102485.
Kim, B.O., Lee, S.M., 1996. Logistics information's role within an IT systems architecture in a world-class organization. Logist. Inf. Manag. 9 (3), 19–26.
Kim, B., Park, J., Suh, J., 2020. Transparency and accountability in AI decision support: explaining and visualizing convolutional neural networks for text information. Decis. Support. Syst. 134, 113302.
Kiritchenko, S., Mohammad, S., 2018. Examining gender and race bias in two hundred sentiment analysis systems. In: Proceedings of the Seventh Joint Conference on Lexical And Computational Semantics, pp. 43–53.
Kleanthous, S., Kasinidou, M., Barlas, P., Otterbacher, J., 2022. Perception of fairness in algorithmic decisions: future developers' perspective. Patterns 3 (1), 100380.
Knechel, W.R., Thomas, E., Driskill, M., 2020. Understanding financial auditing from a service perspective. Acc. Organ. Soc. 81, 101080.
Kosack, S., Fung, A., 2014. Does transparency improve governance? Annu. Rev. Polit. Sci. 17.
Kotler, P., Zaltman, G., 1971. Social marketing: an approach to planned social change. J. Mark. 35 (3), 3–12.
Krupiy, T.T., 2020. A vulnerability analysis: theorising the impact of artificial intelligence decision-making processes on individuals, society and human diversity from a social justice perspective. Comput.LawSecur.Rev. 38, 105429.
Lanfang, F., Peng, Z., 2014. Battle in the name of privacy and security: a comment on Tencent vs. Qihoo (Supreme People's Court of China 2014). Comput.LawSecur.Rev. 30 (5), 586–592.
Lee, T.W., Lee, T., 1999. Using Qualitative Methods in Organizational Research. Sage.
Li, J., Luo, G., Cheng, N., Yuan, Q., Wu, Z., Gao, S., Liu, Z., 2018. An end-to-end load balancer based on deep learning for vehicular network traffic control. IEEE Internet Things J. 6 (1), 953–966.

Li, Y., Thomas, M.A., Liu, D., 2020. From semantics to pragmatics: where IS can lead in Natural Language Processing (NLP) research. Eur. J. Inf. Syst. 1–22.
Liu, X., Shin, H., Burns, A.C., 2021. Examining the impact of luxury brand's social media marketing on customer engagement: using big data analytics and natural language processing. J. Bus. Res. 125, 815–826.
Liu, Z., Li, X., Zhu, X., 2022. Scenario modeling for government big data governance decision-making: Chinese experience with public safety services. Inf. Manag. 2022, 103622 https://doi.org/10.1016/j.im.2022.103622.
Lotz, S., Christandl, F., Fetchenhauer, D., 2013. What is fair is good: evidence of consumers' taste for fairness. Food Qual. Prefer. 30 (2), 139–144.
Lourenço, R.P., 2015. An analysis of open government portals: a perspective of transparency for accountability. Gov. Inf. Q. 32 (3), 323–332.
Lu, Y., Keung, P., Ladhak, F., Bhardwaj, V., Zhang, S., Sun, J., 2018. A neural interlingua for multilingual machine translation arXiv preprint arXiv:1804.08198.
Lubberink, R., Blok, V., Van Ophem, J., Omta, O., 2017. Lessons for responsible innovation in the business context: a systematic literature review of responsible, social and sustainable innovation practices. Sustainability 9 (5), 721.
Maas, J., 2022. A neo-republican critique of AI ethics. J.Responsib.Technol. 9, 100022.
Maass, W., Storey, V.C., 2021. Pairing conceptual modeling with machine learning. Data Knowl. Eng. 134, 101909.
Malhotra, D., Rishi, O., 2021. An intelligent approach to design of E-commerce metasearch and ranking system using next-generation big data analytics. J.King Saud Univ.Comput.Inf.Sci. 33 (2), 183–194.
Marulli, F., Verde, L., Campanile, L., 2021. Exploring data and model poisoning attacks to deep learning-based NLP systems. Procedia Comput.Sci. 192, 3570–3579.
Mathrani, A., Susnjak, T., Ramaswami, G., Barczak, A., 2021. Perspectives on the challenges of generalizability, transparency and ethics in predictive learning analytics. Comput.Educ.Open 2, 100060.
Messner, W., 2017. The role of gender in building organisational commitment in India's services sourcing industry. IIMB Manag. Rev. 29 (3), 188–202.
Miner, G., Elder IV, J., Fast, A., Hill, T., Nisbet, R., Delen, D., 2012. Practical Text Mining And Statistical Analysis for Non-structured Text Data Applications. Academic Press.
Mitri, M., 2020. Story analysis using natural language processing and interactive dashboards. J. Comput. Inf. Syst. 1–11.
Mittelstadt, B., 2019. Principles alone cannot guarantee ethical AI. Nat.Mach.Intell. 1 (11), 501–507.
Mourby, M., Cathaoir, K.Ó., Collin, C.B., 2021. Transparency of machine-learning in healthcare: the GDPR & European health law. Comput.LawSecur.Rev. 43, 105611.
Mulki, J., Lassk, F.G., 2019. Joint impact of ethical climate and external work locus of control on job meaningfulness. J. Bus. Res. 99, 46–56.
Nasaescu, E., Marin-Lopez, I., Llorent, V.J., Ortega-Ruiz, R., Zych, I., 2018. Abuse of technology in adolescence and its relation to social and emotional competencies, emotions in online communication, and bullying. Comput. Hum. Behav. 88, 114–120.
Nersessian, D., 2018. The law and ethics of big data analytics: a new role for international human rights in the search for global standards. Bus.Horiz. 61 (6), 845–854.
O'Neill, B., Stapleton, L., 2020. Social responsibility, human centred systems and engineering ethics: a new manifesto for systems engineering design praxis. IFAC-PapersOnLine 53 (2), 17451–17456.
Ocone, R., 2020. Ethics in engineering and the role of responsible technology. Energy AI, 100019. https://doi.org/10.1016/j.egyai.2020.100019.
O'Neill, O., 2002. Autonomy And Trust in Bioethics. Cambridge University Press.
Owen, R., Macnaghten, P., Stilgoe, J., 2012. Responsible research and innovation: from science in society to science for society, with society. Sci. Public Policy 39 (6), 751–760.
Owen, R., Macnaghten, P., Stilgoe, J., 2020. Responsible research and innovation: from science in society to science for society, with society. In: Emerging Technologies: Ethics, Law And Governance. Routledge, pp. 117–126.
Pagliaro, S., Lo Presti, A., Barattucci, M., Giannella, V.A., Barreto, M., 2018. On the effects of ethical climate (s) on employees' behavior: A social identity approach. Front. Psychol. 9, 960.
Pandey, S., Pandey, S.K., 2019. Applying natural language processing capabilities in computerized textual analysis to measure organizational culture. Organ. Res. Methods 22 (3), 765–797.
Perazzoli, S., de Santana Neto, J.P., de Menezes, M.J.M.B., 2022. Systematic analysis of constellation-based techniques by using Natural Language Processing. Technol. Forecast. Soc. Chang. 179, 121674.
Perboli, G., Gajetti, M., Fedorov, S., Giudice, S.L., 2021. Natural language processing for the identification of human factors in aviation accidents causes: an application to the SHEL methodology. Expert Syst. Appl. 186, 115694.
Perez, C.C., 2019. Invisible Women: Exposing Data Bias in a World Designed for Men. Random House.
Ploug, T., Holm, S., 2020. The four dimensions of contestable AI diagnostics-a patient-centric approach to explainable AI. Artif. Intell. Med. 107, 101901.
Prabhu, O.S., Gupta, P.K., Shashank, P., Chandrasekaran, K., Usha, D., 2021. Towards a federated learning approach for NLP applications. In: Applications of Artificial Intelligence And Machine Learning. Springer, Singapore, pp. 157–167.
Prabhumoye, S., Boldt, B., Salakhutdinov, R., Black, A.W., 2020. Case study: deontological ethics in NLP arXiv preprint arXiv:2010.04658.
Raab, C.D., 2020. Information privacy, impact assessment, and the place of ethics. Comput.LawSecur.Rev. 37, 105404.
Radkov, R., 2018. An approach to choosing an optimal IT infrastructure in accordance with an assignor's requirements. In: 2018 20th International Symposium on Electrical Apparatus And Technologies. IEEE, pp. 1–4.
Ratten, V., 2015. Factors influencing consumer purchase intention of cloud computing in the United States and Turkey: the role of performance expectancy, ethical awareness and consumer innovation. EuroMed J. Bus. 10 (1), 80–97.
Remes, J., MlEttinen, H., Reunanen, A., Pyörälä, K., 1991. Validity of clinical diagnosis of heart failure in primary health care. Eur. Heart J. 12 (3), 315–321.
Ryan, M., 2022. The ethics of dietary apps: technology, health, and the capability approach. Technol. Soc. 101873.
Schrecker, T., 2013. Interrogating scarcity: how to think about 'resource-scarce settings'. Health Policy Plan. 28 (4), 400–409.
Schwepker Jr., C.H., Hartline, M.D., 2005. Managing the ethical climate of customer-contact service employees. J. Serv. Res. 7 (4), 377–397.
Schwepker, C.H., Ferrell, O.C., Ingram, T.N., 1997. The influence of ethical climate and ethical conflict on role stress in the sales force. J. Acad. Mark. Sci. 25 (2), 99–108.
Scuotto, V., Magni, D., Palladino, R., Nicotra, M., 2022. Triggering disruptive technology absorptive capacity by CIOs. Explorative research on a micro-foundation lens. Technol. Forecast. Soc. Chang. 174, 121234.
Shahriari, K., Shahriari, M., 2017. IEEE standard review—ethically aligned design: a vision for prioritizing human wellbeing with artificial intelligence and autonomous systems. In: 2017 IEEE Canada International Humanitarian Technology Conference. IEEE, pp. 197–201.
Shearer, T., 2002. Ethics and accountability: from the for-itself to the for-the-other. Acc. Organ. Soc. 27 (6), 541–573.
Simintiras, A.C., Dwivedi, Y.K., Kaushik, G., Rana, N.P., 2015. Should consumers request cost transparency? Eur. J. Mark. 49 (11/12), 1961–1979.
Skaug, S.H., 2022. The ethics of trading privacy for security: the multifaceted effects of privacy on liberty and security. Technol. Soc. 101854.
Sohal, A., Nand, A.A., Goyal, P., Bhattacharya, A., 2022. Developing a circular economy: an examination of SME's role in India. J. Bus. Res. 142, 435–447.
Srinivasan, R., González, B.S.M., 2022. The role of empathy for artificial intelligence accountability. J.Responsib.Technol. 9, 100021.
Stocking, J.C., Utter, G.H., Drake, C., Aldrich, J.M., Ong, M.K., Amin, A., Marmor, R.A., Godat, L., Cannesson, M., Gropper, M.A., Romano, P.S., 2020. Postoperative respiratory failure: an update on the validity of the Agency for Healthcare Research and Quality Patient Safety Indicator 11 in an era of clinical documentation improvement programs. Am. J. Surg. 220 (1), 222–228.
Sugumaran, R., Degroote, J., 2010. Spatial Decision Support Systems: Principles And Practices. CRC Press.
Sylvestre, P., Gaucher, N.O., Perez, T., Drouin, O., 2021. Rethinking informed consent in the age of behavioural sciences and relational autonomy. EthicsMed.Public Health 19, 100708.
Tan, S.Y., Taeihagh, A., Tripathi, A., 2021. Tensions and antagonistic interactions of risks and ethics of using robotics and autonomous systems in long-term care. Technol. Forecast. Soc. Chang. 167, 120686.
Tatman, R., 2017. Gender and dialect bias in YouTube's automatic captions. In: Proceedings of the First ACL Workshop on Ethics in Natural Language Processing. Association for Computational Linguistics, Valencia, Spain, pp. 53–59.
Treviño, L.K., Weaver, G.R., Gibson, D.G., Toffler, B.L., 1999. Managing ethics and legal compliance: what works and what hurts. Calif. Manag. Rev. 41 (2), 131–151.
Tripathy, J.K., Sethuraman, S.C., Cruz, M.V., Namburu, A., Mangalraj, P., Vijayakumar, V., 2021. Comprehensive analysis of embeddings and pre-training in NLP. Comput.Sci.Rev. 42, 100433.
Valentine, S., Godkin, L., 2019. Moral intensity, ethical decision making, and whistleblowing intention. J. Bus. Res. 98, 277–288.
Verde, L., Marulli, F., Marrone, S., 2021. Exploring the impact of data poisoning attacks on machine learning model reliability. Procedia Comput.Sci. 192, 2624–2632.
Vermanen, M., Rantanen, M.M., Harkke, V., 2021. Ethical framework for IoT deployment in SMEs: individual perspective. Internet Res. https://doi.org/10.1108/INTR-08-2019-0361.
Victor, B., Cullen, J., 1987. A theory and measure of ethical climate in organizations. In: Research in Corporate Social Performance And Policy, 9, pp. 51–71.
Victor, B., Cullen, J.B., 1988. The organizational bases of ethical work climates. Adm. Sci. Q. 101–125.
Vošner, H.B., Bobek, S., Kokol, P., Krečič, M.J., 2016. Attitudes of active older Internet users towards online social networking. Comput. Hum. Behav. 55, 230–241.
Wagner, B., 2020. Accountability by design in technology research. Comput.LawSecur. Rev. 37, 105398.
Wang, J., Li, S., Jiang, M., Wu, H., Zhou, G., 2018. Cross-media user profiling with joint textual and social user embedding. In: Proceedings of the 27th International Conference on Computational Linguistics. Association for Computational Linguistics, Santa Fe, New Mexico, USA, pp. 1410–1420.
Westbrook, R.A., 1981. Sources of consumer satisfaction with retail outlets. J. Retail. 57 (3), 68–85.
Yadav, R., Giri, A., Chatterjee, S., 2022. Understanding the users' motivation and barriers in adopting healthcare apps: a mixed-method approach using behavioral reasoning theory. Technol. Forecast. Soc. Chang. 183, 121932.
Yang, J., Tang, Z., Yuan, Q., Xu, B., 2021. The economic and social benefits of the government-backed credit guarantee fund under the condition of an economic downturn. Technol. Forecast. Soc. Chang. 166, 120632.
Zhang, B., Huang, H., Pan, X., Ji, H., Knight, K., Wen, Z., Sun, Y., Han, J., Yener, B., 2014. Be appropriate and funny: automatic entity morph encoding. In: Proceedings of the

52nd Annual Meeting of the Association for Computational Linguistics, Vol. 2, pp. 706–711. Short Papers.
Zhang, Y., Wu, M., Tian, G.Y., Zhang, G., Lu, J., 2021. Ethics and privacy of artificial intelligence: understandings from bibliometrics. Knowl.-Based Syst. 222, 106994.
Zhdanov, D., Bhattacharjee, S., Bragin, M., 2021. Incorporating FAT and privacy aware AI modeling approaches into business decision making frameworks. Decis. Support. Syst. 113715.

Rajat Kumar Behera is the Associate Professor in the school of computer engineering at Kalinga Institute of Industrial Technology (KIIT). His research area includes Data Science & Business Analytics, Technology Adoption, Ethics & Information Technology and Software Engineering. He received B.E from VSSUT (formerly UCE Burla), M.Tech from IIT Delhi and PhD from IIM Ranchi. He has published research articles in leading journals like Information Systems Frontiers, Journal of Retailing and Consumer Services, Journal of Business Research, Industrial Marketing Management, Information Technology & People, International Journal of Medical Informatics etc. and holds professional designations including PMP, ITIL and Six Sigma Green Belt.

Pradip Kumar Bala is professor in the area of Information Systems & Business Analytics at Indian Institute of Management (IIM) Ranchi. He received his B.Tech., M.Tech. and Ph.D. from Indian Institute of Technology (IIT) Kharagpur in 1993, 1999 and 2009 respectively. He worked in Tata Steel before joining academics. He also worked as associate professor at Xavier Institute of Management Bhubaneswar and as assistant professor at IIT Roorkee before joining IIM Ranchi in 2012. His teaching and research areas include text mining & NLP, recommender systems, data mining applications, data mining and NLP algorithms, social media analytics and marketing analytics. He has conducted many training programmes in business analytics & business intelligence. He has published >100 research papers in reputed international journals, conference proceedings and book chapters. He is also a member of the International Association of Engineers (IAENG). He has served as Director In-charge, Dean (Academics), Chairperson, Post-Graduate Programmes, Chairperson, Doctoral Programme & Research, and Member of Board of Governors of IIM Ranchi.

Nripendra P. Rana is a Professor in Marketing at the College of Business and Economics at Qatar University, Doha, Qatar. His current research interests focus primarily on adoption and diffusion of emerging ICTs, e-commerce, m-commerce, e-government and digital and social media marketing. He has published >200 papers in a range of leading academic journals, conference proceedings, books etc. He has co-edited five books on digital and social media marketing, emerging markets and supply and operations management. He has also co-edited special issues, organised tracks, mini-tracks and panels in leading conferences. He is a Chief Editor of International Journal of Electronic Government Research and an Associate Editor of International Journal of Information Management. He is a Senior Fellow of the Higher Education Academy (SFHEA) in the UK. He is also a Distinguished Visiting Professor at University of Johannesburg, South Africa.

Zahir Irani studied Manufacturing Engineering at the University of Salford before working as a Project Engineer in industry. He then secured a scholarship from Brunel University London, where he completed a PhD in Investment Decision Making. His professional development has seen him complete his leadership and financial management training at Harvard Business School (HBS) and University of Cambridge, respectively. Zahir's research interests are multidisciplinary in nature, and developed from early work on the area of evaluating investments in Manufacturing Information Systems through to more recent works in Transformational Government. He has received significant levels of funding from across the world as Principal Investigator, including from the UK Research Councils (EPSRC, ESRC), European Commission, Qatar Foundation, Australian Research Council and QinetiQ. He also publishes in leading scholarly journals and has a H-index of almost 60. Zahir manages to find time to write press and thought leadership pieces on higher education and graduate employability that have appeared in The Guardian, The Independent, The Wall Street Journal, Financial Times (FT), Thomson Reuters, University Business, Research Foresight and Times Higher Education (THE). Zahir was named by the Sunday Times as a leading Bradford influencer and currently is the Deputy Vice-Chancellor (Academic, Innovation and Quality) at the University of Bradford.