June 2024

Unpopular Policies, Ineffective Bans: Lessons Learned from ChatGPT Prohibition in Italy

Francesco Bolici
University of Cassino and Southern Lazio, f.bolici@unicas.it

Alberto Varone
University of Cassino and Southern Lazio, alberto.varone@unicas.it

Gabriele Diana
University of Cassino and Southern Lazio, gabriele.diana@unicas.it

UNPOPULAR POLICIES, INEFFECTIVE BANS: LESSONS LEARNED FROM CHATGPT PROHIBITION IN ITALY

Completed Research Paper

Francesco Bolici, University of Cassino and Southern Lazio, Italy, f.bolici@unicas.it
Alberto Varone, University of Cassino and Southern Lazio, Italy, alberto.varone@unicas.it
Gabriele Diana, University of Cassino and Southern Lazio, Italy, gabriele.diana@unicas.it

Abstract

The rapid diffusion of disruptive technologies is generating a revolutionary and tangible impact over individuals, organizations and society. However, this rapid pace of development is not matched by up-to-date regulations, which makes the relationship between institutional policies and technological advancements complex and controversial. Taking as a reference generative AI, this work studies how individuals respond to public interventions banning disruptive technologies, exploring the arguments and sentiment they express towards it. By analysing approximately 15,000 X contributions on the suspension of ChatGPT in Italy, our work provide evidence that banning disruptive technologies is likely ineffective and unpopular. This was highlighted by the strong prevalence of individuals expressing a negative perception on the ban, by the presence of users actively and collaboratively searching solutions to bypass it, and a perceived institutional backwardness in terms of technology development.

Keywords: Innovation; AI; Technology; Policy.

1 Introduction

In the last decades, technology affordance has been growing at an exponential rate (Kurzweil, 2004; Marino, Cacciamani and Veneziano, 2021). This culminated in the development of disruptive technologies that demonstrated direct and significant impacts over individuals, organizations, and society (Tamkin et al., 2021). While the diffusion of such technologies is fast and extensive, public institutions have been slower in regulating their development and use (Calo, 2017; Erdélyi and Goldsmith, 2018; Rádi, 2023). This regulatory gap makes the reconciliation of institutional and technological dynamics controversial.

Generative Artificial Intelligence (AI) stands out as one of the most transformative technologies currently in the arena. Indeed, as emphasized by Eloundou et al. (2023), 80% of the US labour force could have at least 10% of their tasks affected by its introduction, with dramatic effects in terms of productivity and employment. The widespread integration of generative AI into both professional and personal spheres has underscored the risks associated with these technologies, challenging existing regulatory and normative frameworks. Due to its dependence on large data volumes and often opaque processing methods, this disruptive technology raises serious concerns regarding privacy, security, and transparency (Curzon et al., 2021; Rádi, 2023). While the development of comprehensive guidelines for the ethical implementation of AI systems is required (Erdélyi and Goldsmith, 2018; Curzon et al., 2021), this task remains complex. Thus, AI regulation is currently fragmented and, in some cases, not truly effective (Rádi, 2023).

ChatGPT, developed by OpenAI in November 2021, has rapidly become the most famous AI-tool in the market. Indeed, registering record-braking rates of adoption, it achieved 100 million users in only two months after its launch (Milmo, 2023). This language model, enhanced for conversation through Reinforcement Learning with Human Feedback (OpenAI, 2023), offers diverse capabilities, including assistance in creative writing, essay composition, prompt generation, coding, and answering queries (Taecharungroj, 2023). Despite the opportunities it generates, ChatGPT has raised concerns over user privacy protection (GPDP, 2023). Indeed, on March 30th, 2023, the Italian Data Protection Authority (GPDP) detected illegal data collection and processing practices operated by OpenAI through ChatGPT. This led to a temporary suspension of processing Italian users' personal data, rendering ChatGPT inaccessible in Italy from March 31st to April 28th, when OpenAI complied with the GPDP claims. Taking as a reference the ban of ChatGPT, this study has the purpose of answering the following research question: how does the public opinion respond to the ban of disruptive technologies in terms of arguments debated and sentiment expressed? Utilizing a dataset of 15,005 posts from X (formerly Twitter), this research analyses the content, sentiment, and arguments articulated by the Italian community in reaction to the ChatGPT ban, in order to advance our understanding of the perceived popularity and effectiveness of public interventions banning the adoption of disruptive technologies. Investigating this relationship is vital for designing successful regulatory interventions. In mature information societies, moral evaluations shape the public opinion, which in turn affects what becomes legally enforceable (Dimaggio and Powell, 2004; Floridi, 2018). Over time, as ethical norms shift, regulatory frameworks may become outdated, leading to discrepancies. These misalignments can spark resistance to restrictive public policies, potentially driving actions to evade or oppose them. Hence, understanding public sentiment is crucial for developing regulatory measures that are both effective and accepted.

The structure of this paper is as follows: section two will lay out the theoretical background of this work, concentrating on the diffusion of innovation, technology acceptance, and data privacy. Section three will detail the methodological approach employed. Sections four and five will present our findings and discuss their implications, respectively. Lastly, the conclusions and limitations of our study are presented in sections six and seven.

2 Theoretical background

2.1 Innovation diffusion and technology acceptance

According to the Diffusion of Innovation (DoI) theory, innovation can be defined as an idea, practice, or object perceived as new by its adopters (Rogers, 1983). Newness can manifest in knowledge, persuasion, or decision to adopt (Rogers, 1983). Building on this definition, innovation diffusion is conceptualised as the process through which an innovation is communicated, through some channels and over time, among the members of a social system. As such, This process is fundamentally a social one (Burt, 1987; Valente, 1996), with the adoption rate influenced by both the characteristics of the innovation and the attributes of the social system it permeates. Key innovation characteristics include: i. relative advantage, ii. compatibility, iii. complexity, iv. trialability, and v. observability (Rogers, 1983). Factors related to the social system encompass social contagion and structural equivalence (Burt, 1987), the role of opinion leaders and knowledge brokers (Rogers, 1983; Burt, 2004; Aula and Parviainen, 2012), as well as cultural similarity and societal pressure (Strang and Soule, 1998).

Beyond the general factors influencing innovation adoption, various theories have emerged to specifically address users' acceptance of new technologies and their willingness to use them. Prominently, the Technology Acceptance Model (TAM) (Davis, 1986; Venkatesh and Davis, 2000; Venkatesh and Bala, 2008) is recognized as a leading model in exploring factors affecting technology acceptance. Originating from Davis's doctoral research in Davis (1986), the TAM, based on the Theory of Reasoned Actions and the Theory of Planned Behaviour, was designed to understand the determinants of user acceptance and to offer practical evaluation methods for system designers prior to implementation. In its original formulation, the TAM posits that system adoption hinges on user attitude, which is influenced by i. perceived ease of use, and ii. perceived usefulness (Davis, 1986). These can be respectively defined as i. the extent to which an individual believes that using a system will be free of effort; and ii. the extent to which a person believes that the use of a system will helpful and enhance job performances (Davis, 1989). These factors are thought to be dependent on system design characteristics (Davis, 1989). Subsequent empirical validation led to the development of TAM2 (Venkatesh and Davis, 2000), which expanded the model by introducing additional factors like subjective norms, image, job relevance, output quality, and result demonstrability, which explain perceived usefulness and usage intentions. It also added moderating variables such as experience and voluntariness affecting the influence of subjective norms (Venkatesh and Davis, 2000). In parallel, the work of Venkatesh (2000) explored determinants of perceived ease of use, identifying four anchors (computer self-efficacy, perceptions of external control, computer anxiety, and computer playfulness) and two adjustments (perceived enjoyment and objective usability) that influence initial and subsequent perceptions, respectively. Building on these insights, TAM3 was developed as a comprehensive framework encompassing all identified determinants of IT adoption and use (Venkatesh and Bala, 2008). This iteration integrates prior factors into a unified model and proposes moderating effects of experience on the relationships between perceived ease of use and usefulness, computer anxiety and ease of use, and perceived ease of use and behavioural intention.

In recent years, the Technology Acceptance Model (TAM) and its modifications have become the principal framework for studying adoption decisions related to AI-powered conversational agents (Ling et al., 2021). Belda-Medina and Calvo-Ferrer (2022) utilised TAM2 to examine the integration of conversational AI tools in language learning among future educators. Their research indicated a positive correlation between perceived ease of use, attitudes towards adoption, and a moderate influence on behavioural intention. Another study by Richad et al. (2019) applied the TAM to explore factors influencing technology acceptance of chatbots in the Indonesian banking industry among millennials. The findings revealed that factors such as innovativeness, perceived usefulness, perceived ease of use, and attitudes towards using the conversational agent significantly impacted the behavioural intention to adopt. Furthermore, Ling et al. (2021) conducted a comprehensive review of various technology adoption models, including the TAM, proposing a 'collective model of acceptance and use of intelligent conversational agents'. According to this model, adoption and use are directly influenced by contextual factors and, through the mediation of usage benefits, by characteristics of both the agent and the user. Additionally, the model posits that the impact of usage benefits on adoption and use is exerted both directly and through evaluation factors.

In mature information societies (Floridi, 2018), limitations on the access of innovative technologies can be understood through the lenses of DoI and TAM as a matter of alignment among innovation's characteristics, subjective and system norms, and established regulatory frameworks (Rogers, 1983; Venkatesh, 2000; Dimaggio and Powell, 2004). TAM literature (e.g. Sohn and Kwon, 2020; Venkatesh and Bala, 2008) highlights the role of social influence in technology acceptance decisions through identification and internalisation (Venkatesh, 2000). Similarly, the DoI postulates that the diffusion of innovative technologies is influenced by the norms of the systems in which they proliferate (Rogers, 1983). The relationship between norms and regulatory frameworks is intricately linked through what Floridi (2018) terms as normative cascade, where the moral evaluation of what is deemed right, good, or necessary influences the public opinion. This, in turn, shapes what is socially acceptable or preferable, what is politically feasible, and consequently, what becomes legally enforceable. However, as public opinion shifts, regulatory frameworks may become outdated, ethically compelled to evolve. As further observed by Floridi (2018, p. 7) "such a normative cascade becomes obvious mainly when backlash happens, i.e., mostly in negative contexts, when the public rejects some solutions, even when they may be good solutions". This observation highlights the challenges faced by restrictive regulations on innovative technologies that do not align with subjective (TAM) and systemic norms (DoI). Such misalignments, by hindering the acceptance and diffusion of innovative technologies, may encounter resistance, as well as efforts to boycott or bypass the public intervention.

2.2 Data privacy

The concept of data privacy, as defined by Willems et al. (2022), relates to the extent of control an individual possesses over the distribution of their personal information. This concept encompasses the processes of acquiring, using, accessing, and correcting personal data. Smith et al. (1996) identified unauthorised data acquisition, secondary usage and access, and data inaccuracies as the four main concerns in data privacy. Further expanding on this, Solove (2006) developed a taxonomy of critical practices that pose risks to data protection, including the collection, processing, dissemination, and invasion.

Privacy is recognized as a societal good (Margulis, 2003) and a socially constructed need (Solove, 2006). However, studies by Kokolakis (2017) and Zafeiropoulou et al. (2013) highlight a paradoxical relationship between individual attitudes and behaviours regarding the disclosure of personal information. Termed the "privacy paradox" (Brown, 2001), this phenomenon describes how people, despite expressing concern over the collection and use of their personal data, often willingly exchange it for relatively minor rewards (Dienlin and Trepte, 2015). Kokolakis (2017) offers several explanations for this paradox, drawing from diverse perspectives: i. privacy calculus; ii. social theory; iii. cognitive biases and heuristics; iv. bounded rationality, incomplete information, and information asymmetries; and v. quantum theory. In addition, Acquisti et al. (2015) observed that people often lack awareness of the information they share and its potential uses. In contrast, Hargittai and Marwick (2016) found that, particularly among young adults, there is both awareness and concern about the risks of disclosing personal information online. They note, however, that these individuals feel control over their data diminishes once shared, due to unclear institutional practices, the nature of social media platforms, and the concept of networked privacy, where one's privacy can be compromised within social networks by others. They conclude that the coexistence of privacy concerns and extensive information disclosure is not contradictory but a pragmatic response to the modern networked landscape.

Data privacy has emerged as a pivotal ethical issue in the information age (Willems et al., 2022). The advent of AI-powered systems has intensified concerns regarding privacy protection (Brandtzaeg and Følstad, 2017; Calo, 2017; Cheng et al., 2022), as these systems increase users' vulnerability to privacy breaches, potentially leading to emotional distress, embarrassment, humiliation, and financial losses. The risk is particularly acute with highly interactive tools such as AI-powered chatbots (Følstad and Brandtzæg, 2017; Du and Xie, 2021). Kronemann et al. (2023) highlight that AI conversational agents, including chatbots, with anthropomorphic features (i.e., human-like characteristics) are more likely to encourage users to disclose personal data. Ischen et al. (2020), port this view, noting that traits like friendliness and socialness in these agents are linked to lower privacy concerns, increased comfort in disclosing information, and higher adherence to recommendations. While the need for a comprehensive regulatory framework governing AI development and use is recognized (Acquisti, Brandimarte and Loewenstein, 2015), such a system is still in development (Curzon et al., 2021; Rádi, 2023). In the interim, it is advised that companies clearly and transparently communicate their privacy policies, offer tangible benefits for users who provide and share personal information, and grant consumers greater control over data collection and management decisions (Du and Xie, 2021).

3 Methodology

In our study, we utilised two primary data sources: the social media platform X (formerly known as Twitter) and Google Trends. Our objective was to evaluate the public response to the GPDP's intervention by analysing the sentiments and motivations expressed in the X contributions of Italian users, including tweets, retweets, replies, quoted tweets, and quoted replies. X allows users to publish messages, known as tweets, limited to 280 characters, facilitating user interactions through mentions, responses, and retweets (Schuster and Kolleck, 2020). X's global presence, vibrant community, and historically accessible data make it an invaluable source of relational data, particularly for studying diffusion processes and emerging behaviours (Steinert-Threlkeld, 2018). For instance, Bolici et al. (2020) leveraged X data to investigate the spread of blockchain technology in the tourism industry, demonstrating its effectiveness in analysing how technology diffusion is influenced by individual interactions within a social system. Other significant studies utilising X data include González et al. (2019), who examined public reactions to the Cambridge Analytica scandal, and Vogler and Meissner, (2020) who explored responses to the Ticketfly data breach. Their findings suggested that even after a significant cyber-attack, users in the U.S. did not view cybersecurity as a major concern. Thus, X data emerges as a potent tool for analysing social diffusion processes in terms of both innovation and public perceptions.

In this study, we utilised the X APIs to collect a dataset of public contributions. Our criteria for selection were: i. contributions posted between March 24th and May 4th; ii. written in Italian; and iii. containing the keyword ChatGPT. This process yielded a dataset comprising 31,260 contributions from X. The time frame of analysis spans the entire duration of the ChatGPT's suspension, also incorporating the week prior to its unavailability (March 24th to March 30th) and the week after its reinstatement (April 28th to May 4th). This approach is intended to offer a comprehensive view into the debate around the ChatGPT suspension before, during and after the GPDP's intervention, establishing a baseline for tracking changes in discourse throughout the whole period considered.

We then turned to Google Trends to analyse search queries related to ChatGPT during our study period (March 24th – May 4th, 2023). This analysis helped us identify the most central keywords associated with ChatGPT. The key terms we identified were: i. Privacy; ii. Dati personali (Personal data); iii. blocc* (Block*); iv. Ban*; v. Access*; vi. Intelligenza artificiale (Artificial intelligence); vii. Tecnologi* (Technolog*); viii. Innov* (Innovat*); ix. VPN; x. Altenativ* (Alternative*). Utilising these keywords, we refined our dataset to 15,005 contributions that contained the term ChatGPT and at least one of these ten keywords. This structured approach ensures a comprehensive and relevant collection of data, pivotal for analysing public sentiment and reactions regarding the ChatGPT suspension in Italy.

Queries registering the highest rising association with the term ChatGPT (March 24th – May 4th, 2023)				
chatgpt vpn	chatgpt italia vpn	chatgpt bloccato in italia	chatgpt italia blocco	opera browser
vpn gratis	pizzagpt	opera vpn	nova chatgpt	nova
chat gpt vpn	garante chatgpt	come usare chat gpt in italia	chatgpt banned in italy	my ip
free vpn	chatgpt proxy	nord vpn	free vpn chrome	chatgpt bloccata in italia
come usare chatgpt in italia	vpn chrome	proton vpn	vpn online	chat gpt bloccato

Table 1. Queries registering the highest rising association with the term ChatGPT in the study period.

Our analysis comprised several stages, each designed to provide a comprehensive understanding of the public debate surrounding the ChatGPT suspension.

1. Historical Analysis: we conducted a historical analysis to assess the evolution of the debate across three distinct periods: pre-suspension (March 24th – March 30th; 7 days), suspension (March 31st – April 27th; 28 days), and post-suspension (April 28th – May 4th; 7 days). During each phase, we examined both the volume of contributions and the sentiments expressed within them. Sentiment analysis was performed using the tool developed by Barbieri et al. (2021), a multilingual XLM-roBERTa-base model trained on approximately 198 million tweets and fine-tuned for sentiment analysis in multiple languages, including Italian.

2. Keyword Co-occurrence Analysis: We then proceeded with a keyword co-occurrence analysis to identify connections among the discussed topics. This step helped us understand how various themes and issues were interlinked in the public discourse.

3. Finally, we performed a qualitative content analysis on a sample of 750 contributions. Our objective was to identify the most prominent arguments both supporting and opposing the ban. This method, as outlined by White and Marsh (2006), employs an inductive and interpretative approach to discern concepts, patterns, and meanings from texts. Unlike quantitative content analysis, which seeks replicable

and valid inferences, qualitative content analysis aims to capture the essence, themes, and organisation of the messages. It involves developing coding schemes through iterative reading cycles, allowing for a deeper understanding of the underlying concepts and patterns. Compared to quantitative methods, however, this approach is subject to higher levels of subjectivity. A similar inductive approach in classifying X data was adopted by So et al. (2016), promoting the coherence between the classes identified and the phenomenon investigated. As in Chew and Eysenbach (2010) and So et al. (2016), in this study two of the authors implemented the coding process.

First, we refined our data by eliminating duplicated texts such as retweets. This process yielded 7,022 unique contributions. Second, based on a preliminary review, the authors inductively and independently identified a set of macro-categories of arguments related to the GPDP intervention. Categories that showed significant agreement between the assessments of two authors were validated; those with discrepancies were subject to further discussion to reach a consensus on their inclusion or exclusion. This consensus-based methodology facilitated the delineation of nine distinct categories of argumentation: three in support of and six against the intervention. Third, we randomly selected 750 contributions from the pool of 7,022 for classification based on the identified categories, with two authors working independently. Contributions that did not present an argument regarding the investigated phenomenon were excluded. Among these, approximately 48% were merely reporting the news or providing contextual information on the ban; 32% were more focused on the broader theme of AI development, opportunities and risks rather than the GPDP intervention per-se; 20% were commenting on the ban either expressing neutral comments or polarised views without motivating them (e.g. spreading information on how to overcome the ban). Contributions that both authors classified in the same class were accepted, while those with differing classifications were discussed further until a resolution was reached. This process led to the identification of 152 contributions expressing an argument on the GPDP intervention. Among these, 52 supported the ban, and 100 presented critical perspectives against it.

The authors utilised ChatGPT for proofreading. The contents so generated were checked and eventually edited by the authors.

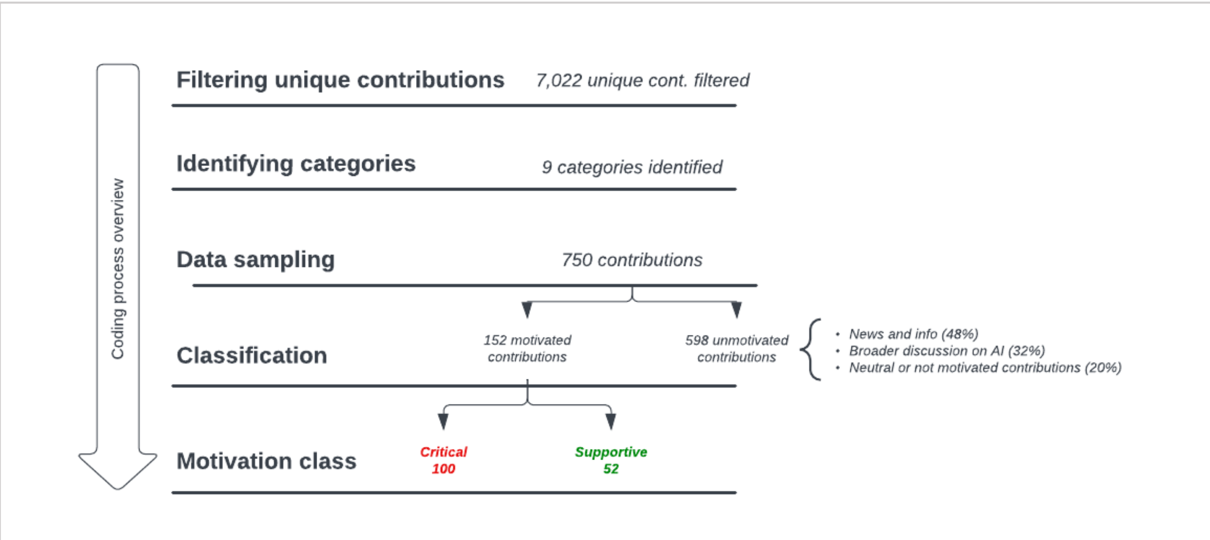

Figure 1. *Overview of the coding process.*

4 Results

4.1 Historical analysis

The GPDP intervention generated a vocal reaction of the public opinion. The data were divided for convenience in three periods: pre-suspension (March 24th – March 30th; 7 days); suspension (March 31st

– April 27th, 28 days); post-suspension (April 28th - May 4th, 7 days). The contributions collected are respectively 643, 12,942 and 1,420. Other than the longest, the suspension period registered the highest number of contributions per day (462), followed by the post-suspension and pre-suspension periods (respectively 203 and 92). The major discussion windows occurred in the period within March 31st and April 10th, which account for approximately 71% of all the contributions analysed. The suspension period also registered the highest rates of responsiveness and engagement, respectively measured in terms of replies per tweet (0.32) and retweets per tweet (1.54) collected.

In terms of sentiment, the suspension determined a significant increase in the portion of contributions negatively polarised. This shifted from 34% (pre-suspension period) to 61% (post-suspension period). After the ban, this approximately returned to its original level (36%). Similarly, the share of contributions expressing a positive sentiment decreased during the ban and reached, in the post-suspension period, a share approximately 20% higher than in the pre-ban. While in the pre-suspension period the main sentiment expressed was clearly neutral, after the ban the distribution appears more levelled, with approximately 6 percentage points dividing the prevailing and least observed sentiment classes.

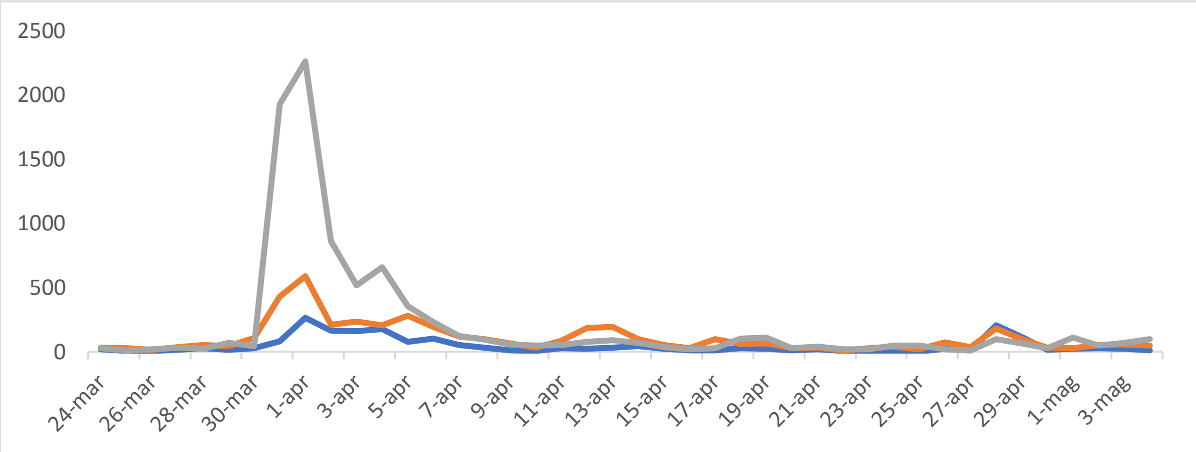

Figure 2. Historical analysis of the X contributions (negative sentiment in grey; neutral sentiment in orange; positive sentiment in blue).

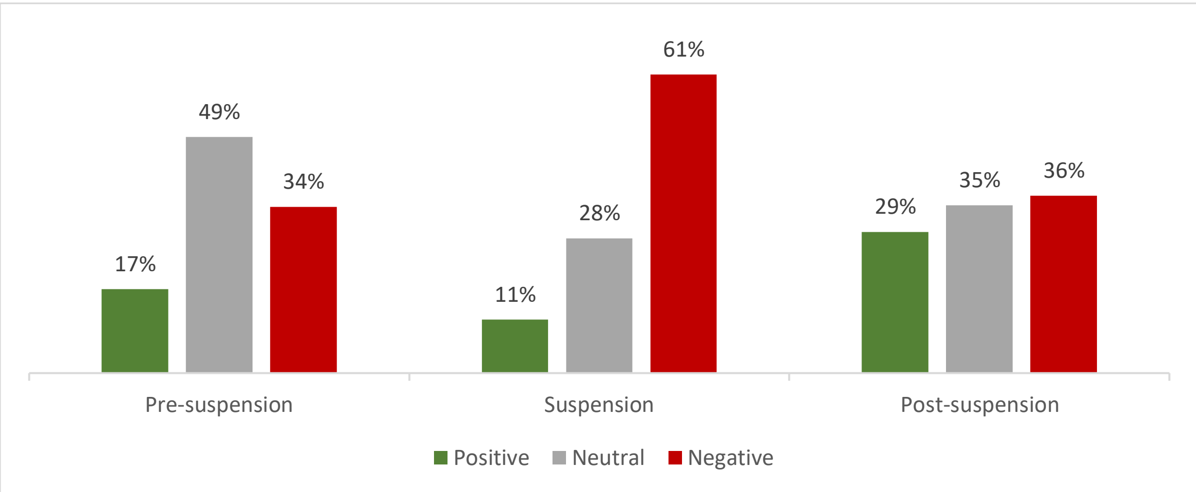

Figure 3. Sentiment frequencies by period.

4.2 Keywords analysis

The analysis of Google trends data showed a strong connection between queries mentioning ChatGPT. Although less marked, this connection also emerges in the analysis of X contributions.

The keyword analysis reveals high levels of interconnectedness between the topics analysed, both at the network and individual pairs level. The co-citation network is remarkably dense, indicating that 44 of the 45 possible connections among all pairs of keywords considered were actually realised. The network diameter, expressed as the longest distance between the two furthest nodes in the network, equals two. This confirms that most keywords were directly connected, and so mentioned together within the same contribution. The 44 edges emerged revealed a total of 9,544 keywords co-citations. The 10 keywords were instead individually mentioned (with the search term "ChatGPT") 9,265 times. On average, each keyword was directly connected to 8.8 others (average degree), and co-mentioned approximately 1,909 times with each (average weighted degree). At the micro level, the pair of keywords most frequently mentioned together were privacy & blocc* (1,713 times), blocc* & intelligenza artificiale (1106 times) and privacy & dati personali (758 times). Only in combination with the search term "ChatGPT", privacy, blocc* and tecnologi* were the most mentioned. VPN, which was found in strong connection with ChatGPT in Google Trends, occupies the fifth position among the most quoted keywords in combination with the search term (558 times) and approximately 9% of the 9,544 co-citations pairs identified above.

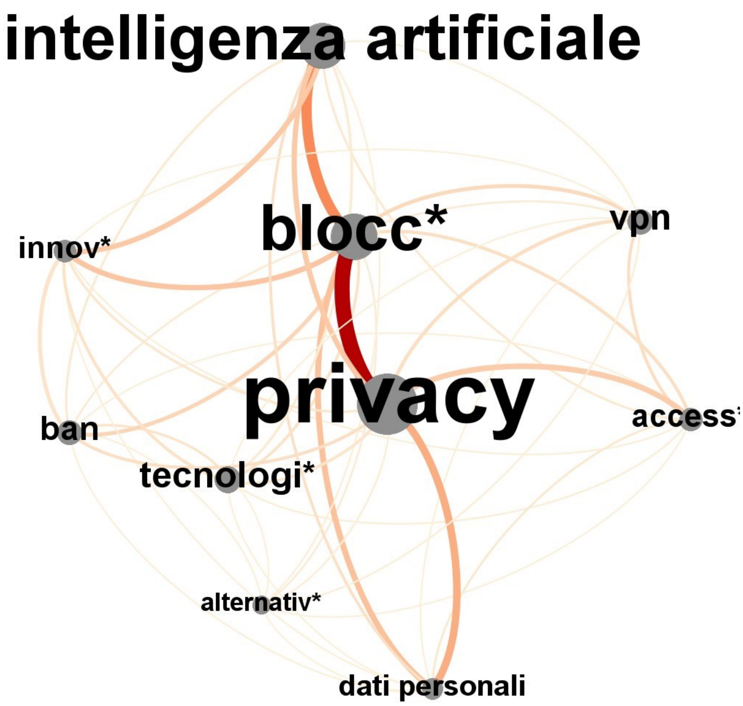

Figure 3. Keywords co-occurrence network.

4.3 Content analysis

The content analysis identified a total of 152 arguments expressed in relation to the ban. Among these, 100 expressed a negative opinion, while the remaining 52 expressed a positive one. These arguments were categorised into nine motivation classes, reflecting 6 critical perspectives and 3 supportive ones. This varied landscape of viewpoints ranges from concerns over technological stagnation and democratic freedom to emphasis on legal adherence and data security.

Among the critical perspectives, the most emphasised were “Anti-democracy”, “Technological Lag” and “Data control disillusion”, with 25, 22 and 22 instances observed respectively. The "Anti-democracy" perspective criticises the ban as overly authoritarian or potentially infringing on personal freedoms. Some opinions within this class draw parallels between the GPDP's actions and those taken by authoritarian regimes globally. The "Technological Lag" group views the ban as hindering technological progress, and perceive Italy as lagging behind other nations due to conservative technology policies. Those in the "Data Control Disillusion" category express frustration over the perceived futility of the ban, arguing that user data and privacy are already compromised by other entities, and see the ban as unfairly discriminatory for ChatGPT. Other critical views include: i. the "Anti-system" stance, which opposes the ban on political grounds, emphasising distrust in public institutions rather than practical concerns about ChatGPT ii. concerns over the ban's detrimental effects on innovation within businesses and the broader economy (Hampering product, process, and business model innovation); iii. the belief that the ban was unnecessary, based on observations of minimal changes post-ban or the perception that ChatGPT was already in substantial compliance with the law prior to the GPDP's intervention.

Among the supportive arguments for the ban, the focus on "Data Privacy and Security" emerged as the most cited (28 instances). Individuals in this class express concern over the misuse of personal data by OpenAI and the unintended consequence of increasing the use of insecure VPN services further compromising user privacy. Following this, the Institutional alignment class was identified (19 instances). This perspective stresses the importance of adhering to national regulations, legal foundations, and the protection of rights as institutional pillars. The least emphasised class of supportive motivations is "Compatibility between Ban and Innovation," observed in 5 instances. This viewpoint suggests the ban acts as a targeted measure rather than a general hindrance to technological advancement, advocating for a balance where innovation, ethical standards, and privacy rights coexist within sustainable technological progress.

Other than identifying distinctive motivations for supporting and criticising the ban, the analysis highlighted 17 contributions providing instructions on how to overcome the ban utilising VPN services, 14 contributions in which users explicitly stated of being using VPN during the suspension period and 11 contributions advertising the existence of alternative chatbots.

Class	Frequency	Description	Stand
Data privacy and security	28	Opinions concerned with the misuse of personal information, and on the implications the ban may cause in the same respects. Among the latter, the proliferation of insecure VPN services, further endangering the privacy of users overcoming the GPDP intervention	Supportive
Anti-democracy	25	Opinions describing the ban as a too extreme, overly coercive and authoritarian measure, possibly in violation of personal freedom. Some of these opinions relate the GPDP intervention to those operated by authoritarian regimes around the world	Critical
Technological lag	22	Opinions suggesting that the ban hindered innovation, or that Italy is overly conservative on policies connected to technology development, especially in comparison to other countries	Critical
Data control disillusion	22	Opinions expressing disappointment with the ban, as user data and privacy are inevitably and by long compromised by other companies, tools, and websites. Thus, the ban of ChatGPT is perceived discriminatory and not useful if seen under the lens of a digital world	Critical

Institutional alignment	19	Opinions emphasising the importance of adhering to national regulations, legal foundations, and the protection of rights. These opinions identify the centrality of supporting legal adherence as an institutional pillar	Supportive
Anti-system	16	Opinions with that present a rather political view in opposition with public institutions, rather than practical concerns regarding ChatGPT's usage. A notable comparison is drawn between the ban and the privacy-restricting implemented by the state during the COVID-19 pandemic (e.g. the implementation of green pass). This inconsistency in behaviour is seen as irksome	Critical
Hampering product, process, and business model innovation	12	Opinions concerned with the negative impact of the ban on businesses and the economy. Some examples include concern for those who had integrated ChatGPT into their workflows or business models, or that were developing products, services or applications based on it	Critical
Compatibility between ban and innovation	5	Opinions suggesting that the ban is a specific, targeted action that doesn't broadly inhibit technological progress. Thus, innovation, ethical standards and privacy rights can coexist, favouring sustainable technological development	Supportive
ChatGPT was already compliant with the law	3	Perceived the ban as unnecessary. This opinion derives from perceiving only minimal changes in ChatGPT after the elimination of the ban; thus, it is assumed that it had to be fundamentally compliant with the law even before the suspension	Critical

Table 2. Classes of arguments, frequencies, description and stand.

5 Discussion

The findings from our research indicate a marked engagement and impassioned reaction from the Italian community to the ban of ChatGPT. This is evidenced by a significant increase in the number of contributions from the pre-suspension to the suspension period, coupled with a notable prevalence of contributions expressing negative sentiment during the latter. These observations are in line with existing literature, which points to AI-related systems as having not only a high level of public exposure but also a contentious relationship with privacy issues (Ischen et al., 2020; Raab, 2020; Saura, Ribeiro-Soriano and Palacios-Marqués, 2022; Willems et al., 2022).

Italian users demonstrated a tendency to circumvent the GPDP intervention by using VPN services, a trend observed both in Google searches and X discussions. On X, users explicitly mentioned using VPNs to bypass the ban and shared tutorials to assist others in doing the same. Most of these contributions lacked explicit reasons behind this choice. The relationship between ChatGPT and VPN in Google search trends further suggests that a considerable number of Italian users not only discussed but actively sought ways to evade the ban. This behaviour, combined with the prevalence of users expressing and justifying negative views on the ban, raises questions about the effectiveness of the GPDP's intervention. It suggests a possible disconnection between the legal measures taken and societal preferences regarding the intersection of AI and privacy. This observation points to the complexities and challenges in aligning regulatory actions with public sentiment in the rapidly evolving domain of AI and privacy.

The content analysis revealed six motivations criticising the ban and three supporting it. One significant concern among the critics was the fear of a technological lag — a relative disadvantage compared to more progressive countries in the exploration, adoption, and exploitation of AI systems. This apprehension aligns with the concept of social pressure, which arises from observing the behaviour of others. The principle of structural equivalence (Burt, 1987) suggests that agents in a social system, sharing similar connections and occupying equivalent positions, influence each other’s behaviours. Since Italy was the only country to restrict access to ChatGPT, Italian users may have found themselves at a coercive disadvantage within the broader context of developed countries actively engaging with this technology. This situation could justify the concerns expressed in the X discussions. Furthermore, this analysis provides insight into why some opinions viewed the ban as an overly coercive and authoritarian measure, particularly when contrasted with strategies implemented in countries with similar cultural and economic backgrounds. This highlights the importance of homogeneous AI regulation at the international level and underscores the necessity for common guidelines on the responsible development and use of AI-powered tools. Such a coordinated approach could help align national policies with global trends and mitigate concerns about technological disparities

Prior to the ban, qualities like accessibility and usefulness of ChatGPT may have facilitated its rapid diffusion in work-related contexts. This point is supported by studies that have shown efficiency improvements in various sectors, such as consultancy, through the adoption of this tool (Dell’Acqua et al., 2023). Our analysis resonates with this observation, suggesting that the ban might have adversely affected businesses and employees who had integrated ChatGPT into their workflows, or that were developing innovative products and business models centred around it.

Our results highlighted a nuanced view positing that the ban and innovation are not inherently incompatible. It was argued that although the ban temporarily restricted access to a specific application (ChatGPT), it did not preclude access to the underlying technology. This situation was perceived as an opportunity to encourage more responsible innovation, aligning with recent research emphasising the need of balancing technology affordances with ethical and legal considerations (Cheng et al., 2022; Rádi, 2023)

The classes Data privacy and security and Data control disillusion present contrasting viewpoints. The first class expresses concerns about OpenAI's misconduct and its implications for personal privacy and data security. In contrast, the Data control disillusion class reflects a sense of resignation regarding the ability to maintain control over personal information in the digital age, where data is constantly created and transmitted. According to the relevant literature, this disillusionment may be justified by the presence of cognitive biases influencing user persecutions, that favour the disclosure of personal information often in exchange for modest rewards (Dienlin and Trepte, 2015). Sensibilization campaigns focused on educating users about data collection and processing, as implemented during the ChatGPT suspension, may encourage more proactive management of personal data by users, helping them distinguish between potentially harmful and appropriate data management practices. Lastly, the classes Institutional alignment and Anti-system present divergent opinions and motivations regarding the GPDP intervention. Institutional alignment reflects satisfaction with the realignment of business practices to conform with established legal frameworks. In contrast, the Anti-system class criticises these frameworks, highlighting perceived inconsistencies and flaws in national policies over time.

6 Conclusions

The exponential growth in technological affordances experienced in the last decades has contributed to the development of disruptive technologies. Despite their massive potential, their development and use are currently under-regulated by public institutions, which have so-far struggled in developing shared and effective policies. Implementing two analytical techniques, namely sentiment and qualitative content analysis, this work studied the ban of ChatGPT operated by the Italian Data Protection Authority to answer the following research question: how does the public opinion respond to the ban of disruptive technologies in terms of arguments debated and sentiment expressed?

Based on the study of a specific case, the suspension of ChatGPT in Italy, our analysis suggests that banning disruptive technologies may be unpopular and ineffective.

The unpopularity of the ban was emphasised by the strong prevalence of negative sentiment in correspondence of the restrictive intervention (approximately 60% in the suspension period, versus 34% and 36% respectively before and after it). Second, 6 out of 9 of the classes of arguments proposed by the public opinion expressed a critical perspective, accounting for 100 critical motivations against 52 in support. Critical positions emphasised the fear of technological lag, perception of anti-democratic public conducts, data control disillusion, fear of business innovation hampering, the exhibition of anti-system attitudes and the perceived aligned compliance of the banned technology with the law even before the intervention. Although minoritarian, supportive arguments praised the elements of data privacy and security, institutional alignment and compatibility between ban and innovation.

The potential ineffectiveness of the ban was emphasised by an extensive and collaborative search for solutions to bypass it. This aspect further questions the implications of such policies, as the behaviours exhibited in the Italian case showed not only that the public engaged with the banned technology, but has potentially engaged in behaviours further endangering its privacy (i.e., using potentially insecure VPN services).

7 Limitations

Limitations of this work include the limited generalizability of our findings, the limited sample size and the potential for subjectivity biases in the qualitative analysis. First, as our findings are drawn from a specific case study involving ChatGPT, their applicability and generalizability to other contexts should be approached with caution. Factors such as the particular technology and tool in question, cultural dynamics, and the platform used for data collection (X) may have influenced our conclusions. Secondly, our content analysis is based on 750 randomly selected contributions, which represent only a fraction of the available data. Future development of this study includes integrating the analysis with a broader array of contributions to provide a more complete overview of the topic. Lastly, this study adopted a qualitative approach to identify and classify the supportive and critical motivations related to the ban of ChatGPT. Although this permits a deeper exploration of the landscape under investigation, this approach inherently introduces a risk of subjective bias. Despite efforts to minimise this through a double-blinded classification process, some level of subjectivity in interpreting motivations still influence our classification.

References

Acquisti, A., Brandimarte, L. and Loewenstein, G. (2015). “Privacy and human behavior in the age of information,” *Science* 347(6221), 509–514. Available at: https://doi.org/10.1126/science.aaa1465.

Aula, P. and Parviainen, O. (2012). “Communicating Connections: Social Networks and Innovation Diffusion,” in H. Melkas and V. Harmaakorpi (eds) *Practice-Based Innovation: Insights, Applications and Policy Implications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 49–63. Available at: https://doi.org/10.1007/978-3-642-21723-4_4.

Barbieri, F., Anke, L.E. and Camacho-Collados, J. (2021). “XLM-T: Multilingual Language Models in Twitter for Sentiment Analysis and Beyond”. Available at: https://doi.org/10.48550/ARXIV.2104.12250.

Belda-Medina, J. and Calvo-Ferrer, J.R. (2022). “Using Chatbots as AI Conversational Partners in Language Learning,” *Applied Sciences* 12(17), 8427. Available at: https://doi.org/10.3390/app12178427.

Bolici, F. *et al.* (2020). “Innovation diffusion in tourism: how information about blockchain is exchanged and characterized on twitter,” *The TQM Journal* [Preprint]. Available at: https://doi.org/10.1108/TQM-01-2020-0016.

Brandtzaeg, P.B. and Følstad, A. (2017). “Why People Use Chatbots,” in I. Kompatsiaris et al. (eds) *Internet Science*. Cham: Springer International Publishing (Lecture Notes in Computer Science), 377–392. Available at: https://doi.org/10.1007/978-3-319-70284-1_30.

Brown, B. (2001). "Studying the internet experience," *Publishing Systems and Solutions Laboratory* [Preprint].

Burt, R.S. (1987). "Social Contagion and Innovation: Cohesion versus Structural Equivalence," *American Journal of Sociology* 92(6), 1287–1335. Available at: https://doi.org/10.1086/228667.

Burt, R.S. (2004). "Structural Holes and Good Ideas," *American Journal of Sociology* 110(2), 349–399. Available at: https://doi.org/10.1086/421787.

Calo, R. (2017). "Artificial Intelligence Policy: A Roadmap," *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.3015350.

Cheng, X. *et al.* (2022). "The good, the bad, and the ugly: impact of analytics and artificial intelligence-enabled personal information collection on privacy and participation in ridesharing," *European Journal of Information Systems* 31(3), 339–363. Available at: https://doi.org/10.1080/0960085X.2020.1869508.

Chew, C. and Eysenbach, G. (2010). "Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak," *PLoS ONE*. Edited by M. Sampson, 5(11), p. e14118. Available at: https://doi.org/10.1371/journal.pone.0014118.

Curzon, J. *et al.* (2021). "Privacy and Artificial Intelligence," *IEEE Transactions on Artificial Intelligence* 2(2), 96–108. Available at: https://doi.org/10.1109/TAI.2021.3088084.

Davis, F.D. (1986). *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. Diss. Massachusetts Institute of Technology.

Davis, F.D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* 13(3), p. 319. Available at: https://doi.org/10.2307/249008.

Dell'Acqua, F. *et al.* (2023). "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality," *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.4573321.

Dienlin, T. and Trepte, S. (2015). "Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors: The relation between privacy attitudes and privacy behaviors," *European Journal of Social Psychology* 45(3), 285–297. Available at: https://doi.org/10.1002/ejsp.2049.

Dimaggio, P.J. and Powell, W.W. (2004). "Chapter 4 THE IRON CAGE REVISITED: INSTITUTIONAL ISOMORPHISM AND COLLECTIVE RATIONALITY IN ORGANIZATIONAL FIELDS," in F. Dobbin (ed.) *The New Economic Sociology*. Princeton University Press, 111–134. Available at: https://doi.org/10.1515/9780691229270-005.

Du, S. and Xie, C. (2021). "Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities," *Journal of Business Research* 129, 961–974. Available at: https://doi.org/10.1016/j.jbusres.2020.08.024.

Eloundou, T. *et al.* (2023). "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models". arXiv. Available at: http://arxiv.org/abs/2303.10130 (Accessed: 9 May 2023).

Erdélyi, O.J. and Goldsmith, J. (2018). "Regulating Artificial Intelligence: Proposal for a Global Solution," in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. *AIES '18: AAAI/ACM Conference on AI, Ethics, and Society*, New Orleans LA USA: ACM, 95–101. Available at: https://doi.org/10.1145/3278721.3278731.

Floridi, L. (2018). "Soft Ethics and the Governance of the Digital," *Philosophy & Technology*, 31(1) 1–8. Available at: https://doi.org/10.1007/s13347-018-0303-9.

Følstad, A. and Brandtzæg, P.B. (2017). "Chatbots and the new world of HCI," *Interactions*, 24(4), 38–42. Available at: https://doi.org/10.1145/3085558.

González, F. *et al.* (2019). "Global Reactions to the Cambridge Analytica Scandal: A Cross-Language Social Media Study," in *Companion Proceedings of The 2019 World Wide Web Conference*. *WWW '19: The Web Conference*, San Francisco USA: ACM, 799–806. Available at: https://doi.org/10.1145/3308560.3316456.

GPDP (2023). *Provvedimento del 30 marzo 2023 [9870832]*. Available at: https://www.garanteprivacy.it:443/home/docweb/-/docweb-display/docweb/9870832 (Accessed: 12 April 2023).

Hargittai, E. and Marwick, A. (2016). "“What Can I Really Do?” : Explaining the Privacy Paradox with Online Apathy”. Available at: https://doi.org/10.5167/UZH-148157.

Ischen, C. *et al.* (2020). “Privacy Concerns in Chatbot Interactions,” in A. Følstad et al. (eds) *Chatbot Research and Design*. Cham: Springer International Publishing (Lecture Notes in Computer Science), 34–48. Available at: https://doi.org/10.1007/978-3-030-39540-7_3.

Kokolakis, S. (2017). “Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon,” *Computers & Security* 64, 122–134. Available at: https://doi.org/10.1016/j.cose.2015.07.002.

Kronemann, B. *et al.* (2023). “How AI encourages consumers to share their secrets? The role of anthropomorphism, personalisation, and privacy concerns and avenues for future research,” *Spanish Journal of Marketing - ESIC* 27(1), 2–19. Available at: https://doi.org/10.1108/SJME-10-2022-0213.

Kurzweil, R. (2004). “The Law of Accelerating Returns,” in C. Teuscher (ed.) *Alan Turing: Life and Legacy of a Great Thinker*. Berlin, Heidelberg: Springer Berlin Heidelberg, 381–416. Available at: https://doi.org/10.1007/978-3-662-05642-4_16.

Ling, E.C. *et al.* (2021). “Factors influencing users’ adoption and use of conversational agents: A systematic review,” *Psychology & Marketing* 38(7), 1031–1051. Available at: https://doi.org/10.1002/mar.21491.

Margulis, S.T. (2003). “Privacy as a Social Issue and Behavioral Concept: Privacy as a Social Issue,” *Journal of Social Issues* 59(2), 243–261. Available at: https://doi.org/10.1111/1540-4560.00063.

Marino, N., Cacciamani, G. and Veneziano, D. (2021). “Exponential Technologies and Future Scenarios,” in D. Veneziano and E. Huri (eds) *Urologic Surgery in the Digital Era*. Cham: Springer International Publishing, 249–258. Available at: https://doi.org/10.1007/978-3-030-63948-8_16.

Milmo, D. (2023). “ChatGPT reaches 100 million users two months after launch,” *The Guardian*, 2 February. Available at: https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app (Accessed: 17 November 2023).

OpenAI (2023). *What is ChatGPT? | OpenAI Help Center*. Available at: https://help.openai.com/en/articles/6783457-what-is-chatgpt (Accessed: 17 November 2023).

Raab, C.D. (2020). “Information privacy, impact assessment, and the place of ethics,” *Computer Law & Security Review* 37, p. 105404. Available at: https://doi.org/10.1016/j.clsr.2020.105404.

Rádi, V.G. (2023). “Comparative Analysis of the AI Regulation of the EU, US and China from a Privacy Perspective,” in *2023 46th MIPRO ICT and Electronics Convention (MIPRO)*. *2023 46th MIPRO ICT and Electronics Convention (MIPRO)*, Opatija, Croatia: IEEE, 1446–1451. Available at: https://doi.org/10.23919/MIPRO57284.2023.10159864.

Richad, R. *et al.* (2019). “Analysis of Factors Influencing Millennial’s Technology Acceptance of Chatbot in the Banking Industry in Indonesia,” *International Journal of Management* 10(3). Available at: https://doi.org/10.34218/IJM.10.3.2019.011.

Rogers, E.M. (1983). *Diffusion of innovations*. 3rd ed. New York: London: Free Press; Collier Macmillan.

Saura, J.R., Ribeiro-Soriano, D. and Palacios-Marqués, D. (2022). “Assessing behavioral data science privacy issues in government artificial intelligence deployment,” *Government Information Quarterly* 39(4), p. 101679. Available at: https://doi.org/10.1016/j.giq.2022.101679.

Schuster, J. and Kolleck, N. (2020). “The Global Diffusion of Social Innovations – An Analysis of Twitter Communication Networks Related to Inclusive Education,” *Frontiers in Education* 5, p. 492010. Available at: https://doi.org/10.3389/feduc.2020.492010.

Smith, H.J., Milberg, S.J. and Burke, S.J. (1996). “Information Privacy: Measuring Individuals’ Concerns about Organizational Practices,” *MIS Quarterly*, 20(2), 167. Available at: https://doi.org/10.2307/249477.

So, J. *et al.* (2016). “What Do People Like to “Share” About Obesity? A Content Analysis of Frequent Retweets About Obesity on Twitter,” *Health Communication* 31(2), 193–206. Available at: https://doi.org/10.1080/10410236.2014.940675.

Sohn, K. and Kwon, O. (2020). “Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products,” *Telematics and Informatics* 47, p. 101324. Available at: https://doi.org/10.1016/j.tele.2019.101324.

Solove, D.J. (2006). “A Taxonomy of Privacy,” *University of Pennsylvania Law Review* 154(3), 477. Available at: https://doi.org/10.2307/40041279.

Steinert-Threlkeld, Z.C. (2018). *Twitter as Data*. 1st edn. Cambridge University Press. Available at: https://doi.org/10.1017/9781108529327.

Strang, D. and Soule, S.A. (1998). “Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills,” *Annual Review of Sociology* 24(1), 265–290. Available at: https://doi.org/10.1146/annurev.soc.24.1.265.

Taecharungroj, V. (2023). ““What Can ChatGPT Do?” Analyzing Early Reactions to the Innovative AI Chatbot on Twitter,” *Big Data and Cognitive Computing* 7(1), p. 35. Available at: https://doi.org/10.3390/bdcc7010035.

Tamkin, A. *et al.* (2021). “Understanding the Capabilities, Limitations, and Societal Impact of Large Language Models”. Available at: https://doi.org/10.48550/ARXIV.2102.02503.

Valente, T.W. (1996). “Social network thresholds in the diffusion of innovations,” *Social Networks*, 18(1), 69–89. Available at: https://doi.org/10.1016/0378-8733(95)00256-1.

Venkatesh, V. (2000). “Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model,” *Information Systems Research* 11(4), 342–365. Available at: https://doi.org/10.1287/isre.11.4.342.11872.

Venkatesh, V. and Bala, H. (2008). “Technology Acceptance Model 3 and a Research Agenda on Interventions,” *Decision Sciences* 39(2), 273–315. Available at: https://doi.org/10.1111/j.1540-5915.2008.00192.x.

Venkatesh, V. and Davis, F.D. (2000). “A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies,” *Management Science* 46(2), 186–204. Available at: https://doi.org/10.1287/mnsc.46.2.186.11926.

Vogler, D. and Meissner, F. (2020). “How users tweet about a cyber attack: An explorative study using machine learning and social network analysis,” *Journal of Digital Media & Policy* 11(2), 195–214. Available at: https://doi.org/10.1386/jdmp_00016_1.

White, M.D. and Marsh, E.E. (2006). “Content Analysis: A Flexible Methodology,” *Library Trends* 55(1), 22–45. Available at: https://doi.org/10.1353/lib.2006.0053.

Willems, J. *et al.* (2022). “AI-driven public services and the privacy paradox: do citizens really care about their privacy?,” *Public Management Review* 1–19. Available at: https://doi.org/10.1080/14719037.2022.2063934.

Zafeiropoulou, A.M. *et al.* (2013). “Unpicking the privacy paradox: can structuration theory help to explain location-based privacy decisions?,” in *Proceedings of the 5th Annual ACM Web Science Conference*. *WebSci ’13: Web Science 2013*, Paris France: ACM, 463–472. Available at: https://doi.org/10.1145/2464464.2464503.