



Regulating generative AI: The limits of technology-neutral regulatory frameworks. Insights from Italy's intervention on ChatGPT

Antonio Cordella^a, Francesco Gualdi^{a,b,*}

^a Department of Management, The London School of Economics and Political Science, Houghton Street, WC2A 2AE London, UK

^b Department of International Business, Regent's University London, Inner Circle, NW1 4NS London, UK

ARTICLE INFO

Keywords:

Artificial intelligence
Generative AI
Regulation
Technology-neutral regulatory framework
ChatGPT

ABSTRACT

Existing literature has predominantly concentrated on the legal, ethical, governance, political, and socio-economic aspects of AI regulation, often relegating the technological dimension to the periphery, reflecting the design, use, and development of AI regulatory frameworks that are technology-neutral. The emergence and widespread use of generative AI models present new challenges for public regulators aiming at implementing effective regulatory interventions. Generative AI operates on distinctive technological properties that require a comprehensive understanding prior to the deployment of pertinent regulation. This paper focuses on the recent case of the suspension of ChatGPT in Italy to explore the impact the specific technological fabric of generative AI has on the effectiveness of technology-neutral regulation. By drawing on the findings of an exploratory case study, this paper contributes to the understanding of the tensions between the specific technological features of generative AI and the effectiveness of a technology-neutral regulatory framework. The paper offers relevant implications to practice arguing that until this tension is effectively addressed, public regulatory interventions are likely to underachieve their intended objectives.

1. Introduction

In recent years, the landscape of Artificial Intelligence (AI) regulation has been enriched by a burgeoning array of bills, acts, and norms. This surge reflects a growing consensus on the necessity of regulating AI development within a framework that ensures ethical use, fairness, and accountability (Erdélyi & Goldsmith, 2022; Veale, Matus, & Gorwa, 2023; Wirtz, Weyerer, & Kehl, 2022). As an example, in early 2024, the United Nations General Assembly approved a landmark resolution encouraging the promotion of safe, secure, and reliable AI to support sustainable development. This is the first instance in which the Assembly has formally addressed the regulation of AI. However, despite the relevance of the issue, governments and international public organizations, including the European Union, grapple with the complexities of crafting effective regulatory mechanisms (Hacker, Engel, & Mauer, 2023; Mökander, Axente, Casolari, & Floridi, 2022).

The challenges they face are manifold, encompassing the selection of appropriate regulatory instruments, defining clear regulatory objectives, and the overarching logic that should guide these efforts (Smuha, 2021b; Veale & Zuiderveen Borgesius, 2021).

The increasing and undisputed relevance of AI regulation has

triggered the attention of scholars who have focused on different dimensions of this phenomenon. This scholarly work spans the legal challenges and opportunities presented by AI (Hacker, Cordes, & Rochon, 2024; Robles Carrillo, 2020), governance issues, including the structure and enforcement of AI regulation (König, Wurster, & Siewert, 2023; Wirtz et al., 2022), and the ethical considerations that must inform regulatory frameworks (Calvi & Kotzinos, 2023; Taddeo & Floridi, 2018). Moreover, there is a growing discourse on the societal and political implications of AI regulation (Carlsson & Rönnblom, 2022; Helberger, 2024), highlighting the profound impact these technologies have on society at large.

Although relevant, these contributions overlook the specificity of the technological fabric at the core of generative AI, which presents the most challenging case for AI regulatory interventions (Ferrari, van Dijck, & van den Bosch, 2023; Hacker et al., 2023). Existing regulatory frameworks that apply to AI do not consider the specific characteristics of the technology they govern and hence are technology-neutral (Petit, 2017).

The extant literature, while invaluable, underscores the need for a more nuanced understanding of the technological underpinnings of AI and generative AI to better inform regulatory approaches that are specific to the technologies they seek to govern. To address this gap, the

* Corresponding author at: Department of Management, The London School of Economics and Political Science, Houghton Street, WC2A 2AE London, UK.

E-mail addresses: A.Cordella@lse.ac.uk (A. Cordella), F.Gualdi@lse.ac.uk (F. Gualdi).

paper investigates the challenges that generative AI poses to technology-neutral regulatory frameworks. It analyzes one of the world's first attempts to regulate generative AI, the Italian regulatory intervention on OpenAI, which relies on the technology-neutral regulatory framework of the GDPR.

Indeed, the GDPR is designed to protect “the fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data” (European Parliament, 2016), and in recital 15 it states that “the protection of natural persons should be technologically neutral and should not depend on the techniques used” (European Parliament, 2016).

This case offers a unique opportunity not only to study the effects of the most relevant intervention so far to regulate generative AI but also to analyze the impact that the specific technological characteristics of the AI in question have on the technology-neutral framework used to regulate it.¹

The paper unfolds in a structured manner: section two offers a critical appraisal of prevailing AI regulation discourses, pinpointing existing gaps. Section three delineates the selected research methodology. Section four unveils the case study focusing on Italy's regulatory stance on ChatGPT. Section five illustrates the findings from the case study. Section six analyses the impact of the regulatory effort against ChatGPT's technological particularities. Section seven delineates the key contribution of the paper. The concluding section eight offers implications for practice, addresses limitations, and envisions future opportunities for research.

2. Background

Investigating AI regulation first requires establishing a clear and comprehensive understanding of what regulation entails. Among the many definitions available, we adopt the one proposed by Black and Murray (2019), which is widely accepted in studies on the regulation of digital technologies (Eneman, Ljungberg, Raviola, & Rolandsson, 2022; Galaz et al., 2021; Koniakou, 2023). According to their definition, regulation refers to “sustained and focused attempts to change the behavior of others to address a collective problem or attain an identified end or ends, usually but not always through a combination of rules or norms and some means for their implementation and enforcement, which can be legal or non-legal” (Black & Murray, 2019).

This broad definition is particularly well-suited to address the diverse approaches used to regulate AI. It encompasses multiple forms, types, actors, and subjects of regulation, which are critical factors given the emerging and largely uncharted nature of AI governance.

Recognizing the growing significance of AI and the intensified academic attention on its regulation, we conducted a systematic literature review to map the research landscape on AI regulation across various disciplines. Given that AI regulation is a broad and multifaceted concept, our literature review categorizes the different contributions according to their specific research focus.

Leveraging established literature review approaches, such as the one on AI adoptions in the public sector by Zuidervijk, Chen, and Salem (2021), we conducted a comprehensive search across three main databases: Web of Science, SCOPUS, and The Digital Government Reference Library (DGRL). This was complemented by an additional search on Google Scholar to ensure a thorough exploration of the topic.

Since the purpose of the literature review was to identify works focusing on AI regulation, we utilized the search terms “AI regulation” OR “Regulating AI”, OR “Artificial intelligence regulation”, OR “Regulation of Artificial Intelligence” OR “Regulation of AI” in all the

databases. In Web of Science and Scopus, the terms were searched in the Title, Abstract, and Keywords. In DGRL, we searched in the Keywords. As per the identification of works, we started by applying the following inclusion criteria: (a) journal articles, conference proceedings, and book chapters; (b) English-written research; (c) timeframe of publication ranging from 2018 to 2024.

In Web of Science, we limited the search to the following research areas: Computer Science, Government Law, Science Technology other topics, Business Economics, Social Sciences other topics, Public Administration, Telecommunication. In SCOPUS, we limited the search to subject areas of Computer Science, Social Sciences, Business, Management, and Accounting. The initial search resulted in 141 papers: 73 from Web of Science; 60 from SCOPUS; 8 from DGRL. To complement our pool, we ran a search on Google Scholar through “Publish and Perish” software. We applied the same search terms and inclusion criteria. In line with Zuidervijk et al. (2021), we selected the 50 most cited works from Google Scholar. In total, we obtained a pool of 202 records, from which 11 duplicates were removed.

At this stage, we examined all 191 abstracts to exclude from the pool all those papers that (a) did not focus primarily on AI regulation and/or (b) focused on AI regulation in a specific field that was too distant from our focus or too narrow (such as Fintech, Healthcare, or Marketing). This process reduced the pool to 107 papers. We also excluded policy papers and documents not downloadable in any way. We obtained 94 papers, which constitute our final sample. Fig. 1 below illustrates the process we followed.

All 94 papers were read and examined in full length by the two authors independently. Accordingly, we clustered the papers along the main dimensions of AI regulation identified in the literature review. Six dimensions were identified: legal (28 papers); ethics (24); governance (18); politics (8); policies (5); socioeconomics (7); technology (4). The results of the review are exposed in Table 1.

The literature review reveals that scholars investigating AI regulation focus mainly on the legal dimension of the phenomenon. This literature primarily challenges the adequacy of the existing legal instruments to regulate AI (Veale & Zuiderveen Borgesius, 2021; Wan et al., 2022), introducing various proposals to enhance existing regulatory frameworks for AI with new legislative instruments. The suggested mechanism consists of soft law (Han et al., 2022), combination of different legal sources (Viljanen & Parviainen, 2022), right to contest AI post-facto (Kaminski & Urban, 2021), as well as innovative approaches such as sandbox solutions (Truby et al., 2022).

The scholarly discourse has also thoroughly examined the governance aspect, spotlighting the necessary evolution within governance frameworks for effective AI regulation. This includes a variety of transformative strategies such as amplifying human oversight (Lazcoz & de Hert, 2023); adopting a framework for control and accountability of AI that covers the entire lifecycle of the system (Restrepo-Amariles & Baquero, 2023); and replacing rigid regulations with more adaptable practices (Lucaj et al., 2023). Moreover, there is a growing consensus in the academic field that questions the traditional dominance of national public administrations in leading AI regulatory efforts (Wirtz et al., 2020). Scholars criticize this centralized approach proposing that a transnational phenomenon like AI necessitates a global governance model for regulation (Erdélyi & Goldsmith, 2022; Veale et al., 2023).

Research focusing on ethical dimension mainly discusses the necessity of integrating AI regulation with the protection of fundamental human rights and mechanisms to safeguard against discriminations and bias (Kiskis, 2023) such as appointing humans to oversee AI actions (Nwafor, 2021). This discourse extends to the “ethification” of AI regulation (Mesarčík et al., 2023) where regulation is discussed as the infusion of moral principles into the legal framework (Hoffmann & Hahn, 2020), and to proposals for the development of “detection mechanism” (Knott et al., 2023) to ensure pursuing specific ethical values such as fairness (Calvi & Kotzinos, 2023), trustworthiness (Díaz-Rodríguez et al., 2023), transparency (Reed, 2018).

¹ The EU AI Act came into force on August 1, 2024. However, many of its provisions will not take effect until August 2, 2026, allowing for a two-year implementation period. It is important to note that the AI Act, much like the GDPR, is designed to be technology-neutral by the choice of the legislators.

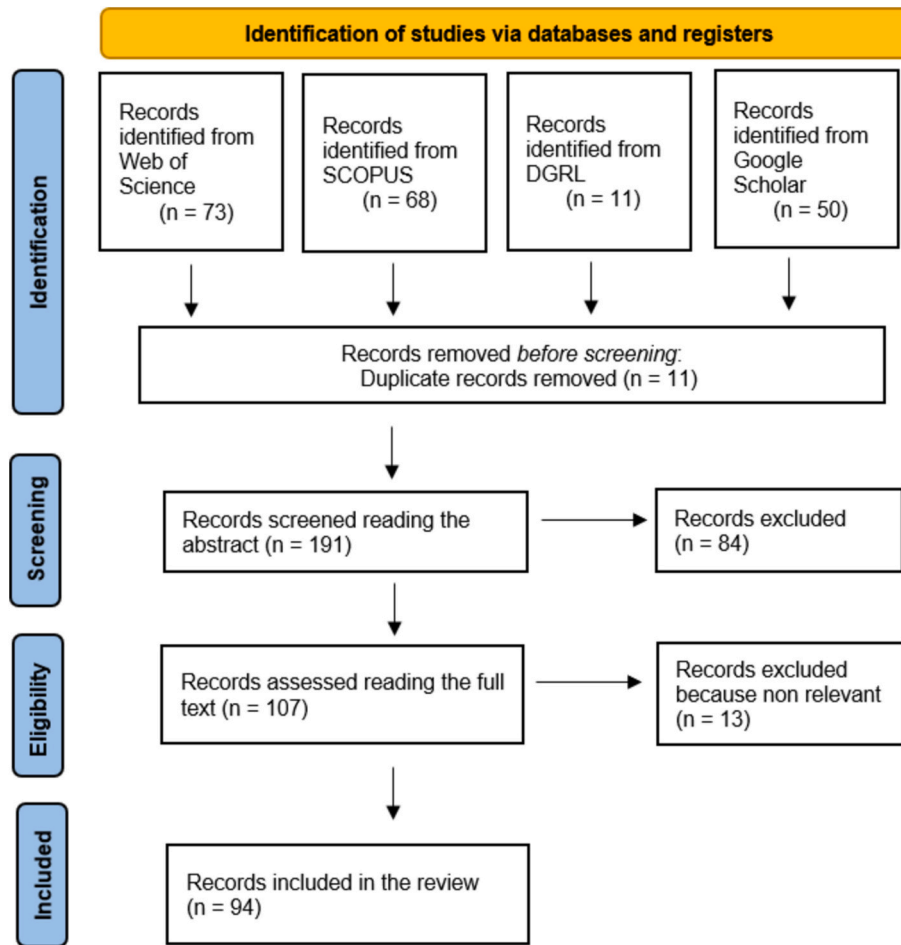


Fig. 1. PRISMA diagram illustrating the choices in the literature review process.

The literature also delves into the political and policy dimensions of AI regulation examining the strategic use of AI regulation approaches by various governments and international bodies. These strategies not only reflect political values (Justo-Hanani, 2022; Krarup & Horst, 2023) but also serve to craft public organizations’ international projection (Belli et al., 2023). On the policy front, the regulation of AI has spurred a panoply of related policies, requiring additional effort to align AI regulation to existing legal frameworks. This includes integration with labour laws (Jarota, 2023), Smart City policies (Zhou & Kankanhalli, 2021), and copyright laws (Lucchi, 2023).

Ultimately, AI regulation has been discussed vis-à-vis its wider socioeconomic impacts, including environmental change (Pagallo et al., 2022); economic competitiveness (Acemoglu, 2021; Mitchell et al., 2023); and the transformation of journalism and media (Helberger, 2024).

However, a notable gap identified through the literature review concerns the technological dimension of AI regulation. Within our sample of 94 papers, only a few works delve into AI regulation by examining the specific technological attributes of AI and generative AI (Anderljung et al., 2023; Ferrari et al., 2023; Hacker, 2023; Zanol et al., 2022). These studies stand out for their focus on how regulations can be aligned with, and possibly adapted to, the characteristics of AI systems, particularly generative AI. These studies identify the necessary conditions to regulate AI in inspecting and modifying the algorithm (Ferrari et al., 2023); the implementation of ex-ante and ex-post regulatory risk assessment tailored to the evolving nature of generative AI (Anderljung et al., 2023); and the efforts to ensure trustworthiness in the process of regulation (Hacker, 2023). However, they do not discuss how the technological characteristics of AI and generative AI affect its

regulation.

The unique characteristics of AI are inadequately considered in AI regulation research, leaving critical gaps in the understanding of how to effectively regulate these technologies. This lack of focus on AI’s specificities raises significant concerns regarding the efficacy of current regulatory efforts, which might fall short without a thorough comprehension of the vast spectrum of existing AI systems (Helberger & Diakopoulos, 2023). This challenge is particularly pronounced when it comes to generative AI (Hacker et al., 2023). In line with Black and Murray (2019), we argue that regulators must develop a deeper engagement with the technical underpinnings of AI technologies.

Rather than relying on technology-neutral frameworks, which can overlook the nuances of AI, regulatory approaches should be tailored to address the distinct characteristics of the specific AI systems they seek to govern. This will ensure more effective and adaptable regulatory mechanisms that better account for the rapidly evolving nature of AI technologies.

Against this background, our research sets out to bridge the identified gap by delving into the critical role of the technological dimension in AI regulation. By doing so, we aim to address the following research question:

How do the technological characteristics of generative AI constrain the effectiveness of technology-neutral regulation?

Focusing on the regulation of generative AI as a case study, we underscore the imperative for a robust understanding of AI’s technological underpinnings to achieve effective regulations. Governments and public organizations neglecting the technological dimension of AI risk forsaking opportunities when crafting and implementing vital legislative pieces focusing on AI.

Table 1
Overview of the dimensions of the AI regulation identified in the literature review.

Dimension	Topics explored	Sources
Ethics	Human rights and AI	Greiman (2020), Kiškis (2023), Nwafor (2021), Salgado-Criado and Fernández-Aller (2021).
	Tools and ways to increase ethical dimension in regulating AI	Bélisle-Pipon, Monteferrante, Roy, and Couture (2023), Buiten (2019), Calvi and Kotzinos (2023), de Laat (2021), Gaeta, Aulino, and Troisi (2023), Hoffmann and Hahn (2020), Knott et al. (2023), Mesarcík, Podroužek, and Gavorník (2023), Stahl, Rodrigues, Santiago, and Macnish (2022), Taddeo and Floridi (2018), White and Lidskog (2022).
	Focus on specific ethical issues (fairness, trustworthiness, transparency, arm)	Díaz-Rodríguez et al. (2023), Hacker and Passoth (2022), Knowles and Richards (2021), Middleton, Letouze, Hossaini, and Chapman (2022), Reed (2018), Smuha (2021a).
	Effects of ethical dimension of AI regulation	Cuéllar, Larsen, Lee, and Webb (2022), Papsyshev and Yarime (2024), Rességuier and Rodrigues (2020).
Legal	Adequacy of existing legal framework and tools	Atabekov and Yastrebov (2018), Chae (2020), Ebers (2019), Edwards (2022), Gervais (2023), Hacker et al. (2024), Jackson (2020), Kaminski (2023), Kashkin (2021), Khisamova, Begishev, and Gaifutdinov (2019), Smuha (2021b), Veale and Zuiderveen Borgesius (2021), Wan et al. (2022).
	Alternative legal solutions to regulate AI (i.e., soft law, multi-layered law, sandbox)	Bello y Villarino (2023), Chamberlain (2023), Clarke (2019), Han, Lenaerts, Santos, and Pereira (2022), Kaminski and Urban (2021), Lupo (2023), Robles Carrillo (2020), Stuurman and Lachaud (2022), Truby, Brown, Ibrahim, and Parellada (2022), Viljanen and Parviainen (2022).
	New concepts and perspectives to regulate AI	Black and Murray (2019), Hacker (2023), Hildebrandt (2018), Schuett (2023), Wulf and Seizov (2020).
Governance	Drivers and obstacles of a global governance on regulating AI	Erdélyi and Goldsmith (2022), Nitzberg and Zysman (2022), Veale et al. (2023).
	Identification and proposal of new governance frameworks	Bannister and Connolly (2020), Chauhan (2023), de Almeida, dos Santos, and Farias (2021), Mökander et al. (2022), Robles and Mallinson (2023), Wirtz et al. (2022), Wirtz, Weyerer, and Sturm (2020).
	Effects of AI governance on services and systems	Chatterjee, and N.S. (2023), Kokshagina, Reinecke, and Karanasios (2023), Paul (2023).
	Transformations in the AI regulation governance	Carter (2020), Charles, Rana, and Carter (2022), Lazcoz and de Hert (2023), Lucaj, van der Smagt, and Benbouzid (2023),

Table 1 (continued)

Dimension	Topics explored	Sources
Politics	Political and geopolitical drivers of the EU AI Act	Restrepo-Amariles and Baquero (2023), Dignam (2020), Finocchiaro (2024), Justo-Hanani (2022), Krarup and Horst (2023).
	Government strategies to regulate AI	Belli, Curzi, and Gaspar (2023), Cyman, Gromova, and Juchnevicius (2021), Papsyshev and Yarime (2023), Roberts et al. (2023).
Policy	Limits of policies related to AI regulation	Carlsson and Rönnblom (2022), Green (2022).
Socioeconomics	Assessment of the EU AI Act against Smart City policies and Labour laws	Jarota (2023), Ulnicane (2022), Zhou and Kankanhalli (2021).
	Impact of the EU AI Act on society, economics, and environment	Acemoglu (2021), Fernandes, Santos, and Lopes (2020), Helberger (2024), Mitchell, Let, and Tang (2023), Pagallo, Ciani Sciolla, and Durante (2022).
Technology	Citizens' support for the EU AI Act	König et al. (2023), Tallberg, Lundgren, and Geith (2024).
		Anderljung et al. (2023), Ferrari et al. (2023), Hacker et al. (2023), Zanol, Buchelt, Tjoa, and Kieseberg (2022)

As AI increasingly mediates business and social interactions, it is essential to assess the effectiveness of technology-neutral regulatory frameworks in governing these technologies. This paper addresses the limitations of frameworks like the GDPR in regulating generative AI, emphasizing the limits of technology-neutral regulatory approaches.

3. Research design

3.1. Research settings

To investigate the increasing intricate terrain of how generative AI impacts technology-neutral regulation, this paper adopts an exploratory research approach (Yin, 2018). This positions the paper in line with the most recent exploratory research on AI within the digital government field, see for instance the work by Fountain (2022). The choice to employ exploratory research emanates from its aptness in contexts where the phenomenon in question is either in its infancy or insufficiently explored (Stebbins, 2001). Furthermore, as posited by Eisenhardt and Graebner (2007), an exploratory case study holds a strategic position, especially for the generation of new theoretical contributions, diverging from explanatory or descriptive case studies which primarily serve the function of theory testing.

3.2. Case selection

This paper focuses on the Italian regulation of ChatGPT, marking the first instance in which a public entity – the Italian Data Protection Authority (*Garante per la protezione dei dati personali*, hereinafter “Garante”) – implemented a binding normative measure on the provision of generative AI services. This intervention marks a significant milestone in the regulation of AI technologies. In Italy, the Garante is responsible for overseeing data protection across all sectors, without targeting specific technologies. It intervenes whenever it suspects data protection laws violations. In this case, the Garante applied a technology-neutral framework (the GDPR) to regulate generative AI, in the absence of any technology-specific regulatory framework.

While numerous governments and institutions are in the process of debating or formulating AI legislations, none, including the recently approved European Union’s AI Act, have been applied yet. The Italian

case is particularly noteworthy for its focus on generative AI, which is not addressed by any other existing regulatory efforts, including the EU AI Act (Hacker, 2023). The significance of the Italian case is amplified by its broader influence, notably serving as a precedent for other countries, with Canada being a primary example (Vermees, 2023), and triggering the launch of a dedicated task force on ChatGPT by the European Data Protection Board. The case study is poised to provide valuable insights into how the specific technical aspects of generative AI affect the success of a regulatory framework that aims to remain neutral towards technology. Opting for the Italian regulation of ChatGPT as the case study fits with our exploratory research approach, especially considering that the main events took place between March and April of 2023. As the sole example of its kind, it represents the first instance to be explored. To better account for the uniqueness of the Italian case, we examined similar examples across other EU countries – each subject to the provision of the forthcoming EU AI Act. However, we found only limited evidence of regulatory initiatives, and no concrete examples of enacted regulation specifically addressing generative AI, aside from the Italian case. Table 2 provides a comparison of the regulatory initiatives, key actors, and outcomes across major EU countries.

3.3. Data collection and analysis

Building on Yin’s paradigm for exploratory research, the formulation of the research question followed an initial phase of data collection (Yin, 2018). During this phase, our systematic literature review uncovered that AI’s technological attributes are often overlooked in discussions within AI regulation literature. Noting the call in the literature to consider the importance of thoroughly investigating AI different configurations to understand their potential impact (Black & Murray, 2019), we developed an exploratory research question to analyze how the technological characteristics of generative AI affect a technology-neutral regulatory framework.

In the second phase, we collected data on the case study of the Italian regulation of generative AI. Data collection encompasses the following documentary sources: (a) all the executive orders issued by the Garante, totalling two documents; (b) all the media releases by the Garante, numbering seven; (c) all General Data Protection Regulation (GDPR) legislative documents, amounting to five; and (d) all the documents disseminated by OpenAI on the issue, two in total. Table 3 provides an overview of the secondary sources consulted for the documentary analysis.

The second phase of data collection provided the researchers with empirical evidence that could be utilized for the development of initial propositions, as suggested by Eisenhardt and Graebner (2007). The collected documents were consulted independently by the members of

Table 3
List of documentary sources consulted.

Type of document	Source	Date	Code
Executive order 9870832	Garante	30 March 2023	A1
Executive order 9874702	Garante	11 April 2023	A2
Press release	Garante	31 March 2023	B1
Press release	Garante	4 April 2023	B2
Press release	Garante	6 April 2023	B3
Press release	Garante	8 April 2023	B4
Press release	Garante	12 April 2023	B5
Press release	Garante	13 April 2023	B6
Press release	Garante	28 April 2023	B7
Digital Article	OpenAI	5 April 2023	C1
Digital Article	OpenAI	28 June 2023	C2
Regulation 2016/679 (GDPR) Art. 5	The European Union		D1
Regulation 2016/679 (GDPR) Art. 6	The European Union		D2
Regulation 2016/679 (GDPR) Art. 8	The European Union		D3
Regulation 2016/679 (GDPR) Art. 13	The European Union		D4
Regulation 2016/679 (GDPR) Art. 25	The European Union		D5

the research team, who took notes and prepared memos with the most significant patterns identified in the documentary sources. Following this phase, the researchers met to collectively interpret the independently analysed data and agreed on a cohesive narrative for the case study. This collaborative phase was also crucial to identify inconsistencies and addressing outstanding questions.

Subsequently, a third phase of data collection was undertaken to evaluate the reliability of the documentary sources referenced. In this stage, the researchers gained exclusive and direct access to the Italian regulatory body overseeing OpenAI, known as the Garante. Given that the Garante operates as a collective entity with decisions made by a four-member panel, with one member tasked with leading each investigation, the researchers determined that interviewing all members would yield minimal additional insight. Therefore, the focus was directed towards a singular interview with the member tasked with leading the investigation against OpenAI.

The significance of the data gathered from the interview is underscored by the exceptional expertise and stature of the interviewee (Myers & Newman, 2007): we engaged with the foremost authority, the individual directly overseeing the procedure against OpenAI. Furthermore, methodology literature concurs that a single interview can be

Table 2
Overview of the regulatory initiatives on AI in the main EU countries. Source: authors’ elaboration on data available in the European Commission’s AI Watch repository.

	France	Germany	Spain	Italy	Netherlands
Regulatory initiatives focused on AI	Absent	Absent	Absent	Absent	Absent
Related regulatory initiatives	Legislative proposal to align the French Intellectual Property Code (IPC) to AI	Legislative amendment that references to AI in three provisions of the German Labor Law	Royal Decree “Sandbox” to set a test environment in compliance with the AI Act	Absent	Absent
Public bodies overseeing AI development	DPA launched the AI Service to steer AI and data protection	Parliament’s Study Commission on AI	Spanish Agency for the Supervision of AI tasked with implementing the AI Act	I3A – National Centre for AI	The Dutch DPA launched the Department for the Coordination of Algorithmic Oversight (DCA)
Regulatory initiatives on generative AI	Launch of the Generative AI Committee with informative purpose	Regional (Lander) DPAs started investigating compliance of generative AI with GDPR	Absent	DPA investigated compliance of generative AI with GDPR	The DCA will coordinate investigations on generative AI risks
Outcomes of the regulatory initiatives on generative AI	Absent	Absent	Absent	DPA imposed regulatory measure on OpenAI	Absent

adequate when the research is centred on a singular expertise (Baker, Edwards, & Doidge, 2012). The interview to the Garante member, whose expertise is the one which determined the intervention on OpenAI, occurred on December 4, 2023, strategically timed to allow a six-month interval between the second and third phases of data collection.

Analyzing the data gathered through the interview enabled the research team to reconcile inconsistencies and fill gaps in the case narrative that had emerged during the analysis of the secondary sources. The combination of documentary sources and the interview data allowed for effective triangulation of the collected data (Patton, 2014). Specifically, we adopted two of the four types of triangulations: “triangulation of sources” (comparing and cross-checking information from different data sources) and “analyst triangulation” (different investigators independently analyzing the same qualitative data and comparing the findings) (Patton, 2014, p. 556). Relying on multiple forms of triangulation strengthens case studies, particularly those involving “complex interventions” (Yin, 2013, p. 324), such as the Italian case of regulating OpenAI.

4. Case study

The study delves into the regulatory measures taken by the Italian Garante against OpenAI, culminating in a four-week suspension of its services in Italy and subsequent modifications to ChatGPT’s global services.

On March 30, 2023, the Garante launched an investigation into ChatGPT due to suspected violations of the GDPR. This investigation prompted OpenAI to temporarily suspend the processing of personal data from Italian users (Source A1). Following the Garante’s inquiry, OpenAI proactively implemented a geo-block on ChatGPT’s services in Italy (Source B3) from March 30 to April 28, 2023. The Garante’s investigation highlighted several non-compliances with the GDPR principles. The identified issues related to insufficient transparency about data collection, challenges in ensuring data fairness and accuracy, lack of legitimate basis for employing data in training algorithms, and insufficient measures for verifying users’ age, specifically articles 5, 6, 8, 13, and 25 (Source B1) of the GDPR. In the next section, we delve deeper into the role and responsibilities of the Garante, outlining the specific tasks it undertakes. We also explain the rationale behind its decision to act against OpenAI, providing a clear account of the facts and actions surrounding the intervention.

4.1. The Garante intervention

The Garante, an autonomous public regulatory entity, is charged with overseeing compliance with data protection norms and legislation regarding the processing of personal data. The Garante holds the legal competency to mandate modifications or impose immediate sanctions against entities that contravene data privacy standards. In particular, the Garante’s board members have the discretion to initiate investigations ex officio upon suspicion of privacy norm violations. This proactive approach was applied in the scrutiny of OpenAI. As the leading investigator explains, the Garante reached the unanimous decision to impose regulatory measures on OpenAI because of the perceived risk to personal data posed by ChatGPT: “We, the four members of the Garante, agreed that in ChatGPT, due to the high number of data subjects, personal data was at stake, which imposed the need to ensure the respect of the law” (Garante, interview). Furthermore, the Garante determined that the rapid adoption of ChatGPT services necessitated urgent action, particularly as OpenAI pioneered this service, thus establishing a precedent. The urgency was underscored by the notion that failing to enforce legal compliance could lead to technology dictating legal norms: “This was one of those cases where either you try to enforce the law, or the technology becomes the law” (Garante, interview). Commencing on March 30, 2023, the Garante’s exertion of its authority entailed an

investigation into ChatGPT for GDPR breaches, resulting in the temporary suspension of Italian citizens’ personal data processing by OpenAI (Source A1). This investigation motivated OpenAI to proactively implement a geo-block on ChatGPT services within Italy (Source B3). The initiation of this geo-block by OpenAI was an act of compliance and engagement in discussions to address the Garante’s findings about the several GDPR violations. Subsequently, OpenAI committed to modify its services and to implement changes to be compliant with the Garante’s concerns. Even if these corrections did not lead to a “complete reprogramming” (Garante, interview) of ChatGPT services, the Garante considered OpenAI’s efforts promising and decided to revoke the suspension order on April 11, 2023. ChatGPT services were fully restored in the Italian territory on April 28, 2023. Nevertheless, OpenAI’s practices continue to be the subject of a formal investigation and remain under the vigilant observation of the Garante into early 2024. The investigation concluded in January 2024, revealing OpenAI’s non-compliance with the GDPR. To date, both OpenAI and the Garante remain engaged in discussions to arrive at a mutually agreeable solution.

Figure 2 presents a detailed overview of the actions taken by the Garante, OpenAI’s responses, and the key events that define the regulatory intervention concerning ChatGPT during March and April 2023.

5. Findings of the case study

5.1. GDPR breaches: data collection process

The Garante has voiced concerns over the transparency practices of OpenAI in the data collection process for ChatGPT particularly in relation to the compliance with GDPR Article 13. This article requires that users are adequately informed about the collection of their personal data, demanding clarity on several fronts: who is collecting the data, the purposes of data collection and processing, the modalities of data sharing, and the justification of legitimate interests (Sources C1, D4). OpenAI has been criticized for not offering clear and comprehensible information on these critical aspects of its data collection methodology (Source B1). Moreover, the Garante points out the lack of publicly accessible documentation that delineates the methods OpenAI uses to accrue data for algorithm training. Contrary to OpenAI’s claim of relying on “public sources” for training its algorithms, the Garante asserts that these sources are not tailored specifically for algorithm training, highlighting that each public source is created with a unique intent (Garante, interview) which is not algorithmic training.

5.2. GDPR breaches: respect of fairness and accuracy of data

The Garante has invoked Article 5 of the GDPR to express concerns regarding the veracity of the outputs produced by ChatGPT, emphasizing the article’s mandate for private organizations to guarantee the accuracy of personal data and actively rectify any inaccuracies (Source D1). The Garante criticized ChatGPT for occasionally generating information that deviates from factual accuracy, thereby constituting a breach of GDPR’s Article 5. This infraction is underscored by the failure of OpenAI to consistently adhere to the data accuracy principle in its data processing activities (Source B3). Furthermore, such inaccuracies could precipitate discriminatory effects (Hacker et al., 2023), contravening not only the accuracy but also the fairness principle embedded within the GDPR. Additionally, the Garante highlights a significant concern regarding individuals’ inability to exercise their right of access, request data deletion, or correct inaccuracies, specifically referring to challenges in addressing algorithmic errors, termed as “algorithmic hallucination” (Garante, interview).

5.3. GDPR breaches: lack of appropriate legal justification for using data to train algorithms

The Garante highlighted a significant concern regarding the services

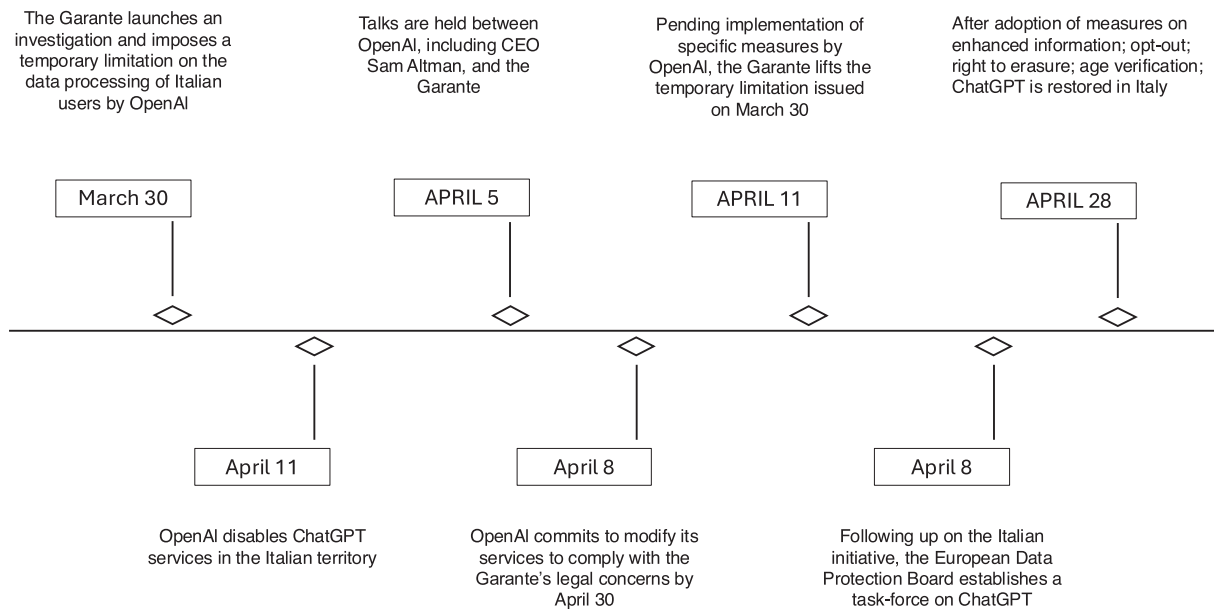


Fig. 2. Timeline of the regulatory initiative against OpenAI by the Italian Garante in March and April 2023.

offered by OpenAI via ChatGPT, notably the absence of “an appropriate legal basis” for the collection and processing of user data (Source A1). The investigation focused on the legitimacy of utilizing users’ personal data for the purpose of training ChatGPT’s algorithms. Article 6 of the GDPR stipulates the necessity for a legal basis prior to the collection and utilization of personal data. The Garante pointed out that OpenAI’s collection and usage of users’ data to train its algorithms lacked a satisfactory legal rationale (Source A1). The authority emphasized that, even when considering “legitimate interest” as a possible legal basis, such a justification requires adherence to principles of transparency, including the provision of information to users and acknowledgment of their right to opt-out. According to the Garante, both prerequisites were conspicuously absent, particularly in relation to the training processes, suggesting that the data of countless Italian citizens may have been utilized without prior notification or the provision of an opt-out option (Garante, interview).

5.4. GDPR breaches: absence of mandatory checks to protect under aged users

The Garante observed a significant discrepancy between GDPR requirements and the operational reality of ChatGPT, particularly highlighting the platform’s failure to restrict access for users below the age of 13, in contradiction to both GDPR stipulations and OpenAI’s own declared policies (Sources D3, B6). This lack of age-appropriate protective measures for younger users engaging with ChatGPT’s interface was identified as a breach of Article 8 of the GDPR. This article mandates that entities collecting and processing personal data of minors must undertake “reasonable efforts” to verify that consent is provided by guardians or parents. The investigation by the Garante revealed the absence of effective age-verification mechanisms, such as age filters, potentially exposing minors to inappropriate interactions with ChatGPT. Given ChatGPT’s popularity among young audiences, the Garante underscored the imperative for more stringent adherence to age verification protocols.

5.5. Measures taken by OpenAI to rectify GDPR compliance issues

In response to the GDPR compliance issues raised by the Garante, OpenAI undertook a series of negotiations resulting in an interim consensus on remedial actions. An executive order issued by the Garante

detailed the commitments made by OpenAI, which were prerequisites for the suspension lift (Source A2). To address data collection and handling concerns, OpenAI pledged to release a comprehensive document elucidating its data access and usage policies, and to launch a communication campaign to increase the awareness about its services. To mitigate accuracy issues, the company introduced a mechanism allowing users to request data access, deletion, or modification via a link provided upon registration to ChatGPT. OpenAI also addressed the legal concerns regarding data usage for algorithm training by enabling a user opt-out feature for data utilization. Additionally, the implementation of an age verification tool was aimed at bolstering minors’ protection.

The corrections enabled by OpenAI addressed some of the concerns raised by the Garante in relation to the protection of users’ personal data, however they didn’t affect the technical way ChatGPT operates. Rather, OpenAI introduced corrections such as mechanisms and filters that improved and clarified the usability of its services. For example, to address the opt-out issue raised by the Garante, OpenAI introduced a tool that enables the user to opt-out from the utilization of their own personal data to train the algorithm. However, as noted by the Garante, this “does not technically mean having the possibility to see their data deleted by OpenAI or having the alternative option to request correction of misinformation or deletion of data concerning you” (Garante, interview). OpenAI corrections did not change the way the algorithm handled and reused data: it added tools such as filters to prevent the algorithm to release erroneous answers so that it could be more respectful of GDPR regulation: “Since they are unable to correct this information, what they do is prevent the chatbot from answering a question, in order to avoid providing incorrect responses” (Garante, interview). In the view of the Garante, OpenAI could not decouple the provision of ChatGPT services from the specific way it processes users’ data: “There hasn’t been any significant modification regarding the training of the algorithm with user data” (Garante, interview).

The Garante’s regulatory oversight primarily focuses on data privacy and usage, with less attention given to the intricate technological foundations of ChatGPT. However, a detailed exploration of these technological dimensions is paramount to understand the intricacies involved in regulating generative AI systems under a regulatory framework which is technology-neutral. In the next sub-section, we provide an in-depth analysis that examines the key technological functionalities at the core of ChatGPT’s operations.

5.6. ChatGPT algorithm: the technological architecture

ChatGPT operates through a sequential algorithmic process encompassing three critical stages: pre-training, fine-tuning, and response generation. The pre-training is a crucial phase in the development of the model: it involves assembling vast datasets of textual content, which are subsequently broken down into smaller units, called “tokens”. These tokens can be individual words, parts of words, or even characters. This process of tokenization enables the algorithm to learn contextual relationships and word associations, by breaking down text into smaller, more digestible pieces. This allows the model to analyse patterns and relationships between these pieces over time, facilitating its ability to predict interrelations between text segments. The fine-tuning stage further refines the model by introducing specialized datasets that closely align with the anticipated user inputs and their corresponding outputs. This process relies on supervised learning, where the algorithm is tuned to produce more accurate and contextual relevant responses. Human intervention plays a significant role in this stage, as experts guide the model’s optimization to produce specific outputs (Radford, Narasimhan, Salimans, & Sutskever, 2018). Additionally, user interactions during this phase can contribute to the model’s continuous improvements, with feedback and updates influencing future iterations to enhance its accuracy and responsiveness.

User interactions trigger the response generation phase, where the model processes and encodes the input text. At this point, the model predicts the most likely sequence of words to create a coherent and contextually appropriate response. It employs a probabilistic approach, drawing from patterns and structured information embedded in the training dataset, while also considering the specificities of the input’s context. The algorithm’s output is generated as a sequence of words that forms a meaningful response, aiming to reflect the input’s context as accurately as possible, guided by the logic encoded during the training (Radford et al., 2018). This phase represents the final stage in the interaction, where the generated response is presented to the users, acting as the model’s reply to their queries or prompts.

ChatGPT’s proficiency in generating contextually relevant text responses is attributed to its underlying algorithmic modelling logic (Breiman, 2001). This logic suggests that natural language generation reflects complex and partly uncharted cognitive processes, presenting challenges in forecasting the variables influencing output generation. ChatGPT distinguishes itself from other AI models by employing a stochastic approach to predict outputs, rather than relying solely on data modelling logic to establish deterministic relationships among variables (Breiman, 2001).

The essence of algorithmic modelling lies in identifying a range of variable combinations that can accurately but not precisely predict outputs, emphasizing the generation of probable predictions over definitive causality. This approach allows for the generation of diverse outputs from identical inputs, showcasing the algorithm’s generative capacity (Vapnik, 2000). Utilizing methods like bootstrapping, the algorithm isolates pertinent variables or their combinations for predictive analysis without relying on explicit programming or rule-based problem-solving (Helberger & Diakopoulos, 2023).

Generative AI, as exemplified by ChatGPT, harnesses extensive data compilations from varied online sources to produce outputs through probabilistic calculations (Bender, Gebru, McMillan-Major, & Shmitchell, 2021). These generated responses, in turn, serve as new inputs for the algorithm, fostering a continuous cycle of data processing and response formulation. This dynamic interplay between the evolving dataset and the algorithm’s variable-weighting mechanism underscores the inherent unpredictability of generative AI outputs. The mutable nature of probabilistic text distributions and the algorithm’s adaptive response generation contributes to the complexity and unpredictability characteristic of generative AI systems.

Therefore, understanding ChatGPT’s underlying technological mechanisms offers crucial insights into the challenges posed by existing

regulatory frameworks which are technology-neutral. It underscores the need for a nuanced approach to regulation that can accommodate the unique operational paradigms of generative AI systems.

6. Discussion

The Garante’s 2023 decision to apply the GDPR in addressing privacy concerns related to ChatGPT arose due to the absence of AI-specific legal frameworks, as the EU’s AI Act had been approved but was not yet enforced. This decision was based on the understanding that the GDPR is technology-neutral (Demetzou, 2019) and up until then has proven effective to regulate any technology dealing with data processing. By trusting the powerfulness of a technology-neutral legislation like the GDPR, the Garante intended to bring ChatGPT’s data collection practices and algorithmic training processes into alignment with the GDPR’s stringent legal and ethical requirements. Furthermore, it sought to manage ChatGPT access in a manner consistent with GDPR mandates.

The case in question presents an invaluable opportunity to assess the effectiveness of a technology-neutral regulatory framework when applied to highly specific and innovative generative AI like ChatGPT. It highlights the critical need of examining the impact of distinct technological traits on the regulatory measures intended to govern them.

6.1. GDPR principles and ChatGPT

The GDPR articulates seven foundational principles that govern the management of personal data: lawfulness, fairness, and transparency; purpose limitation; data minimization; accuracy; storage limitation; integrity and confidentiality; and accountability. These guiding principles address the collection, processing, storage, and sharing of personal data, ensuring its management is conducted in a controlled and observable manner. As the case study previously illustrated, OpenAI has undertaken measures to comply with GDPR principles, including the revelation of further details and the introduction of features designed to prevent violations of GDPR standards within its operations. However, this approach to data protection has inherent limitations especially in regulating generative AI models like ChatGPT because of the characteristics of the technology at the core of the functioning (Ferrari et al., 2023; Hacker et al., 2023).

Yet, these scholarly objections do not definitively resolve the question of whether the GDPR’s regulatory scaffold can effectively tackle privacy concerns arising from generative AI utilization. The functional mechanisms of ChatGPT pose formidable challenges for the implementation of GDPR principles. Embedded within the algorithm’s architecture is the capability to generate varied responses to a singular query, predicated on probabilistic analysis. These probabilities, in turn, reflect the data distribution across the dataset – namely, the organization of sentences. The dataset, however, is dynamic by nature. Each algorithmic output, informed by previous inputs, transforms into a novel input for the database. This regeneration of output thus continuously updates the input dataset, leading to modifications in the probabilistic distribution that the algorithm exploits to create new responses. This recursive process underpins the algorithm’s generative function and its adaptive learning mechanism, perpetually iterating in reaction to its outputs, thereby imposing significant complexities on adhering to GDPR guidelines.

6.1.1. ChatGPT and the principles of lawfulness, fairness, and transparency

The architecture of generative algorithms presents significant challenges in establishing a legal foundation for data processing, especially because generative AI continuously processes and transforms original inputs to create new data in an ongoing cycle. This raises important questions about the legality of using these derivative forms of data under the GDPR, which mandates fairness in data processing to prevent harm to individuals. Generative models function by identifying a wide array of variable combinations to make predictions that are accurate, though not

entirely definitive. These models rely on probabilities rather than direct causality, allowing them to generate different outputs from the same inputs. However, this probabilistic approach may inadvertently produce biased outcomes, creating challenges for compliance with the GDPR's fairness principle. Additionally, the GDPR requires data subjects to be informed about how their data are processed (Source A1, p. 2), a requirement that becomes difficult to meet due to the complex and evolving nature of generative algorithms. Every input shall be tokenized for the algorithm to produce an output, and this input data is often incorporated into the broader dataset for future use. This continuous incorporation of data complicates the task of ensuring transparency and maintaining informed consent, as the data subject may not fully understand or foresee how their data will be used. As a result, achieving full transparency and meeting the GDPR's requirements for informed consent and fairness can become nearly impossible, posing significant barriers to legal compliance.

6.1.2. *ChatGPT and the principle of purpose limitation*

The tokenization process essential to ChatGPT's algorithm, which is required both for training and for stochastically processing data to meet its generative needs, complicates adherence to the GDPR's principle of data collection and processing, particularly in maintaining alignment with the original purposes of data use. Through tokenization, the algorithm gains the ability to comprehend contextual connections and relationships between words, enhancing its capacity to predict how different segments of text relate to one another. It generates responses by leveraging a variety of relationships discovered within its dataset, identifying patterns consistent with its probabilistic models. In this process, the initial user query plays a central role in guiding the response generation.

However, this approach also leads to the alteration of the original data's intent and context. As the algorithm breaks down and reassembles data during processing, it can shift the purpose of the data away from its original collection goals (Source B1, p. 2).

For example, data collected for a specific purpose may evolve through multiple cycles of processing and learning, potentially diverging from the initial intent, which raises concerns about the ability to preserve the original purposes of data collection.

Moreover, ChatGPT's continuous learning and generative capabilities intensify these challenges, making it increasingly difficult to ensure that the foundational objectives of data collection and processing are maintained over time. This ongoing evolution poses a significant challenge to compliance with the GDPR's principle of purpose limitation, which requires that data be used only for the specific purposes for which they were collected. As the model's operations inherently alter the context and purpose of data, ensuring adherence to this principle becomes problematic. This situation highlights the broader difficulties that generative algorithms face in complying with the GDPR, particularly when it comes to maintaining compliance, specifically with purpose limitation.

6.1.3. *ChatGPT and the principle of data minimization*

The principle of data minimization, a cornerstone in privacy regulations such as the GDPR, mandates that personal data collection should be limited to what is strictly necessary for processing. Only the minimum amount of data required to achieve specific goals should be collected and used. However, the nature of generative algorithms, such as ChatGPT, creates significant inconsistencies in aligning with this principle.

Generative models rely on vast datasets to function effectively. Their ability to generate detailed, context-rich responses is directly tied to the size, variety, and comprehensiveness of the data they are trained on (Source B1, p. 2). These algorithms extract patterns and relationships from massive amounts of data, which enhances their predictive accuracy and flexibility. As a result, limiting the amount of data they process could significantly impair their effectiveness, making it difficult to strike

a balance between performance and compliance with data minimization requirements.

Furthermore, the generative process itself is continuously iterative. As these models interact with new inputs and generate new outputs, they contribute to an ongoing life cycle of data production that further refines and enhances their performance. This constant need for large datasets contrasts to the GDPR's data minimization ethos, which aims to restrict data usage to the absolute minimum necessary for specific tasks (Source A1, p.2).

This tension between the operational demands of generative AI and the principle of data minimization creates a fundamental conflict. The very design of these models, which thrives on expansive data to improve over time, is difficult to reconcile with a regulatory framework that seeks to limit data usage.

6.1.4. *ChatGPT and the principle of data accuracy*

ChatGPT generates outputs using a stochastic approach that may not always produce precise results. This raises concerns, especially when handling complex datasets whose relationships are constructed using probabilistic models, or when processing inputs that may contain inaccuracies (Source B5, p.3). The nature of generative AI complicates the task of ensuring data accuracy, as these algorithms rely on stochastic and dynamic constructed data relationships that continuously adapt and evolve by incorporating new information. As more data is processed, the relationships refine themselves, making it increasingly unachievable to verify and maintain the accuracy of the data.

This challenge is further exacerbated by the vast scale of datasets these models manage. Generative algorithms like ChatGPT require vast amounts of data to function effectively, and the sheer volume introduces more potential for inaccuracies to arise, persist, or propagate through the system. As a result, ensuring compliance with the GDPR's principle, which mandates that personal data must be kept accurate and up to date, becomes an increasingly complex task (Source B1, p. 2). The self-evolving nature of these algorithms further complicates the issue. Generative algorithms continuously refine their understanding and outputs based on new inputs, which can cause shifts in the data relationships over time. As the models evolve, it becomes more difficult to trace, validate, and correct inaccuracies that may emerge within this constantly adapting framework. Consequently, adhering to the GDPR's data accuracy principle is not only challenging but becomes increasingly arduous as the model processes larger and more complex datasets.

6.1.5. *ChatGPT and the storage limitation principle*

The storage limitation principle in GDPR mandates that personal data be kept only as long as necessary to fulfil its processing purposes, after which organizations must delete or review the data. Yet, this principle conflicts with the inner workings of generative algorithms, which tokenize data and use methods like bootstrapping to continually use, generate, and reintegrate data. These algorithms repeatedly isolate, combine, and process data for predictive analysis, making it difficult to comply with storage limitations requirements.

The self-sustaining nature of generative algorithms exacerbates this issue. These models constantly generate new data from existing inputs and reintegrate it into their system, creating a continuous cycle of data use and production. Setting and enforcing deletion or review timelines in such a system becomes highly challenging, if not nearly impossible. The ongoing generation and absorption of new data make it difficult to track when or if specific pieces of personal data should be deleted or reviewed.

Furthermore, the reintegration of newly generated data can inadvertently include personal identifiers or sensitive information, further complicating efforts to maintain GDPR compliance. Even when data are tokenized, the complexity of identifying and isolating personal data within vast and dynamic datasets becomes a significant barrier. This ongoing cycle, where data are reused and repurposed, blurs the lines between the original and newly generated data, making it difficult to

ensure that personal information is properly managed and eventually deleted. As a result, the storage limitation principle becomes particularly difficult to enforce within the framework of generative AI systems, highlighting a fundamental clash between GDPR requirements and the operational needs of these algorithms.

6.1.6. ChatGPT and the principle of integrity and confidentiality

The data security principle, as outlined in the GDPR, mandates that personal data must be processed securely to safeguard against unauthorized access, accidental loss, or data breaches. This requires the implementation of robust protective measures. However, ensuring data security becomes increasingly arduous with generative algorithms which process and assemble vast amount of data that are transformed through tokenization. Tokenization breaks data into smaller, more manageable tokens altering how data are analysed, stored, and utilized.

While tokenization is essential in generative models, it also introduces risks to data integrity and confidentiality. These risks originate from the token generation process, which segments data into smaller units that can later be reassembled in different combinations. When these tokenized datasets undergo probabilistic modelling, they are recombined to generate new outputs, such as sentences or other data constructs. This process can exacerbate the likelihood of security risks, as the recombination of data in unpredictable ways may expose sensitive information or compromise the integrity of the original data.

Probabilistic modelling which underpins the functionality of generative algorithms, operates by analysing and linking data through reflexive processes involving input and output generation. This dynamic approach to data handling can heighten the risk of compromising data security and confidentiality. The constant generation of new data and its reintegration into the model’s dataset creates crucial vulnerabilities, such as the potential for data alteration, unauthorized access, or accidental loss.

Furthermore, as these algorithms continually process and generate data, ensuring the integrity of the data – particularly that they remain accurate and unaltered – becomes a complex task. The iterative nature of these systems can make it difficult to track where data originate and whether they have been compromised during processing.

OpenAI has recognized these security challenges in the context of generative AI, emphasizing the complexity of maintaining data security in a statement on April 5, 2023 (Source C1).

Tokenization and probabilistic modelling are fundamental to the operation of generative AI, yet they complicate efforts to adhere to the GDPR’s data security requirements. The inherent risks of reassembling data in new ways, combined with the constant reintegration of newly generated data, create significant flaws in safeguarding data integrity and confidentiality in these systems.

6.1.7. ChatGPT and the principle of accountability

The dynamic nature of generative algorithms, such as those underpinning ChatGPT, presents significant challenges in complying with the GDPR’s requirements for detailed documentation and explanation of data processing activities (Source A2, item 5, p.6). The stochastic nature of these algorithms and the ever-evolving data relationships that continuously redefine the dataset hinder the ability to track, document, and explain each step in the data processing chain. This unpredictability hinders adherence to the GDPR’s accountability principle, which is a key concern for regulators like the Garante. The accountability principle requires organizations not only to follow data protection rules but also to be able to demonstrate their compliance effectively. Moreover, the capacity of these algorithms to affect individuals autonomously intensifies the necessity for human oversight, as mandated by the GDPR to prevent significant automated decisions without human intervention.

In addition to these complexities, organizations that employ generative algorithms must go beyond mere compliance with GDPR principles. They are required to actively demonstrate compliance through comprehensive documentation of their data processing activities. This

involves thoroughly mapping out and reviewing data protection policies, ensuring that data processing steps, policies, and safeguards are aligned with GDPR standards. Organizations must establish a clear audit trail that explains how data are processed, why it is necessary, and what safeguards are in place to protect data subjects.

6.2. The technological dimension: implications for generative AI regulation

To concisely present the interplay between ChatGPT’s technological characteristics and GDPR principles, as well as the challenges these features pose to achieving GDPR’s regulatory goals, a detailed table is constructed. Table 4 aligns specific GDPR principles with the inherent features of the ChatGPT algorithm and outlines the challenges that arise in ensuring compliance with these principles.

This table systematically highlights why ChatGPT’s core technological features are not compatible with the key principles of the GDPR technology-neutral regulatory framework.

OpenAI has implemented various solutions designed to ensure that ChatGPT adheres to the minimum level of compliance with data protection principles. OpenAI’s solutions range from introducing filters to access and use the service, to including a more detailed privacy policy and description of the relevant legal basis for processing. Additionally, they involve implementing opt-out mechanisms from processing of personal data for AI system training purposes, granting an effective right to erasure, and adopting an age gate.

Although these measures provide a degree of protection for GDPR principles, they fall short of fully addressing the challenges posed by the datasets and probabilistic algorithms that ChatGPT uses. ChatGPT falls short in fully upholding the principles of lawfulness, fairness, and transparency. These principles are tested by the continuous generation and reprocessing of data, which complicate maintaining a lawful basis

Table 4
Impact of generative AI on fundamental principles of GDPR.

GDPR Principle	ChatGPT technology features	Challenges in Data Set Dimensions
Lawfulness, Fairness, and Transparency	Generative algorithms continuously process and generate new data	Difficulty in maintaining lawful basis Risk of biased outcomes Challenges in providing clear and comprehensive explanations to data subjects
Purpose Limitation	Generative algorithms process data cyclically and data use is altered.	Maintaining the integrity of the data’s original purpose Potential divergence from initial data use intentions
Data Minimization	Generative algorithms rely on large datasets and continuous data production	Balancing the need for extensive data with the principle of using minimal data
Data Accuracy	Generative algorithms utilize probabilistic analysis to process data, with the generated outputs being continuously integrated into the evolving dataset	Ensuring accuracy and up-to-date information in a dynamic and expansive data environment
Storage Limitation	Generative algorithms continuously generate data, leading to self-perpetuating processes that evolve autonomously	Difficulty in enforcing data deletion or review time limits Tracking persistence of personal data
Integrity and Confidentiality	Generative algorithms tokenize and recombine data properties into new datasets	Maintaining data security Protecting against unauthorized access or data loss
Accountability	Generative algorithms evolve with minimal or no human oversight	Challenges in documenting and explaining data processing activities

for data processing, increase the risk of biased outcomes, and present obstacles in explaining processes to data subjects comprehensively. The principle of purpose limitation faces threats due to the cyclical data processing nature of these systems, which may lead to a divergence from the initial intentions of data use. Data minimization becomes challenging with the reliance on extensive datasets and the production of new data. Ensuring data accuracy is problematic in the face of probabilistic analysis and constantly evolving datasets. Storage limitation is difficult to maintain due to the self-perpetuating nature of continuous data utilization and generation, raising issues with enforcing data deletion or review timelines. The principles of integrity and confidentiality are pressured by the continuous and dynamic processing of data, requiring robust measures to protect against unauthorized access or data loss.

As discussed above, the essence of these challenges is rooted in the inherent characteristics of ChatGPT and generative AI technologies at large, marked by the uncontrollable, constantly evolving nature of the datasets and algorithms that drive these systems.

Implementing filters to manage queries has not succeeded in preventing the spread of incorrect, misleading, or deceptive information by the generative AI model. The constantly changing datasets used to train and refine ChatGPT, coupled with the probabilistic nature of generating outputs, perpetually risk inherent inaccuracies. In the same way, despite GDPR's restrictions, the extensive collection of personal data essential for training the algorithm continues unchanged. The regulatory framework of the GDPR was crafted without foreseeing such advancements. The Garante's leading investigator pointedly remarked: "No one ever envisioned that trawling billions of personal data points for commercial purposes could become a reality, despite its potential positive impacts on society" (Garante, interview).

7. Contribution of the paper

Italy's approach to regulating ChatGPT and generative AI at large has revealed a substantial divide between the existing AI regulation literature, regulatory actions, and the sophisticated demands of generative AI technological landscapes. This discrepancy signals the need for an in-depth analysis and possible overhaul of current regulatory strategies and practices concerning the changeable and uncertain nature of generative AI's datasets and algorithms.

The Garante's pioneering steps to regulate ChatGPT mark a world-first attempt to navigate the complexities inherent in generative AI, shedding light on a global challenge for regulators.

The unique and noteworthy characteristics of generative AI technologies (Moses, 2013), which warrant academic scrutiny, starkly contrast with the technology-neutral regulatory frameworks, creating a trajectory divergence that complicates the regulation of technology through legal means alone (Buiten, 2019).

Although the GDPR principles are ideally suited for technologies that do not change post-deployment and lack generative capabilities, the evolving datasets and generative nature of systems like ChatGPT reveal the shortcomings of this technology-neutral framework in the context of generative AI.

In response to the Garante's concerns regarding personal data processing, OpenAI implemented several measures, including content filters, an opt-out option for data inclusion, and mechanisms for data deletion or correction. Despite these efforts, the Garante recognized that such measures do not fully address the intrinsic challenges of ChatGPT's technology (Garante, interview), noting the impracticality of ceasing all personal data processing due to the system's reliance on diverse datasets for training and its potential to generate incorrect outcomes and log personal data during interactions (Garante, interview).

The findings of the paper contribute to the existing literature by emphasizing the critical importance of integrating a deeper understanding of technology into the regulation of generative AI. The paper underscores the challenges of effectively addressing the far-reaching

impact of generative AI on public sectors and societal outcomes without a thorough grasp of the underlying technological fundamentals. The unique and distinctive characteristics of generative AI technology warrant focused scholarly investigation (Moses, 2013) to enhance our understanding of how regulatory intervention can be made more effective and adaptive to the evolving nature of AI systems.

8. Conclusions and outlook

The challenge of regulating generative AI models, such as ChatGPT, is notably highlighted by the Italian regulatory efforts. These efforts illuminate the struggle of traditional, technology-neutral regulations to adapt to the specific technological features of generative AI, which relies on algorithmic rather than data modelling logic. This regulatory misalignment, demonstrated in the Italian context, serves as a cautionary tale for policymakers focused on developing AI-specific norms or guidance (such as Canada, the EU, the US, and the UK).

8.1. Implications for practice

The paper highlights the critical need for practitioners to develop a deep understanding of the foundational technology behind generative AI to enact impactful legislation. A nuanced regulatory approach, tailored to the specific contexts and applications of generative AI, is essential to ensure that regulations are both practical and forward-thinking. By adjusting or regulating the use of generative AI, regulators can better strike a balance between ensuring compliance with the regulatory requirements and unlocking the vast potential of generative AI.

The analysis of the specific findings from the case study on the application of technology-neutral regulations, as exemplified by the GDPR, offers several key recommendations for shaping regulatory approaches to generative AI. However, it is crucial to recognize that these recommendations stem from a single case study and may therefore have limitations in terms of broader applicability. The unique conditions and scope of this case might not adequately address the full spectrum of regulatory challenges that generative AI presents across different contexts, industries, and jurisdictions. As such, these insights should be considered as preliminary guidelines, subject to further validation and adaptation in more varied regulatory environments.

Enhanced regulations frameworks could focus on several key areas:

- Enhancing transparency through data provenance tools. Service providers and businesses relying on generative AI should implement data provenance tools that monitor the origin, history, and modifications of data throughout their lifecycle. These tools enable regulators and users to trace data as they flow through the AI system, helping to maintain data integrity and foster trust.
- Maintaining purpose alignment and ethical data usage. Generative AI systems must adhere to the original purposes of data collection, preventing the repurposing of information in ways that deviate from initial intentions. The development of ethical guidelines around data usage is essential to ensure that AI systems operate with integrity.
- Strengthening accountability mechanisms for AI-driven decisions. Responsibility for AI-generated decisions should be shared between service providers and organizations using the technology. Regulations could require AI providers to implement mechanisms to rectify or revise harmful decisions and organizations to disclose information on the role of AI in decision-making.
- Balancing data requirements with minimization. While generative AI models rely on large, diverse datasets, it is important to develop methods that balance the need for extensive data with the principle of data minimization. This requires ensuring that only the necessary amount of data is used, without compromising the functionality of the AI, thereby meeting both regulatory requirements and operational demands.

- Maintaining data accuracy and security. Regulators should seek the implementation of mechanisms for continuous data accuracy in dynamic AI environments, ensuring that data remain accurate and up to date to maintaining reliable outputs. Further, regulation could focus on strengthening data security protocols to protect the integrity and confidentiality of data in environments where AI processes data continuously and dynamically.
- Addressing data retention and storage limitation challenges. Given the evolving nature of generative AI, developing solutions to effectively manage data retention and enforce storage limitations is critical. Credible regulations should mandate that organizations must create mechanisms allowing for proper data monitoring, deletion, and review, ensuring that personal data are managed in compliance with regulatory standards.

8.2. Limitations and future research

This research highlights two important boundary conditions. First, the exploratory nature of the case study may affect its generalizability. The case study focuses on the adoption of one specific technology-neutral regulation, the GDPR. It is possible that the findings of this work might not be extended to other technology-neutral AI regulations. However, to address this issue, further research could be conducted across different regulatory frameworks and diverse AI technologies to assess the broader applicability of the findings. By exploring multiple case studies or comparative analysis involving various technology-neutral regulations, researchers can determine whether the insights drawn from this study hold true in other contexts and regulations beyond the GDPR (Yin, 2013). This approach would enhance the generalizability of the research and provide a more comprehensive understanding of the effectiveness and challenges of technology-neutral AI regulation.

Second, we acknowledge that data collection is limited, as it relies on the documentary sources publicly available and one single interview with the key Italian regulator. Nevertheless, the dataset is exhaustive of the information that was accessible, given that the conversations between the Garante and OpenAI were not disclosed. To address this limitation, we believe that the direct access to the highest-level Italian regulator could compensate the limitations in the sample (Buchanan, Boddy, & McCalman, 2013), further aligning with Myers and Newman's (2007, p. 4) argument that "the level at which the researcher enters the organization is crucial".

Lastly, considering these limitations, the paper intentionally focused on exploring the challenges of applying technology-neutral regulations to the specific and unique characteristics of generative AI. The scope was purposefully confined to analyse these limitations rather than proposing alternative regulatory approaches, as doing so would fall outside the feasible boundaries of this study.

However, we do provide recommendations on what an alternative framework might consider, such as incorporating stricter consent protocols, enhanced transparency through data provenance tools, and clearer accountability mechanisms, all tailored to the evolving nature and complexity of generative AI technologies.

Our research seeks to catalyse a scholarly debate on the intricacies of generative AI regulation, advocating for in-depth studies into the characteristics and challenges of generative AI. By encouraging a theoretical engagement with these emerging phenomena, we aim to bolster academic contributions that can help policymakers and practitioners to better design and deploy effective regulatory approaches towards generative AI. Specifically, we propose two research streams for scholars to explore.

The first stream aligns with the contribution of this paper: we encourage scholars to investigate the challenges that arise from applying technology-neutral regulatory frameworks to generative AI, replicating or expanding upon the approach taken in this research. In this context, the forthcoming application of the EU AI Act offers a timely and valuable

opportunity. Further research is needed to assess whether the AI Act adheres to the principle of technology neutrality, as intended by the European legislators (European Parliament, 2023), or if it deviates from this principle, as some critics have suggested (Grady, 2023). Given that the AI Act has already captured significant scholarly attention along its approval and deliberation (Veale & Zuiderveen Borgesius, 2021), it will be essential to closely monitor its implementation and effectiveness, particularly in the application to generative AI systems.

The second stream of research approaches the problem from the opposite angle, investigating whether a regulatory approach that is not technology-neutral may be better suited to addressing the challenges posed by generative AI. As governments are contemplating various strategies to regulate generative AI, research is needed to determine whether the principle of technology neutrality will continue to be upheld, or if a departure from this principle may be warranted to ensure more robust protection for citizens and society.

Investigating these alternative, context-specific regulatory frameworks could offer insights into whether targeted regulations may prove more effective in managing the complexities of generative AI.

Credit authorship contribution statement

Antonio Cordella: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Francesco Gualdi:** Writing – review & editing, Writing – original draft, Resources, Methodology, Data curation, Conceptualization.

The two authors are listed in alphabetical order only and have contributed equally to the paper.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Academic Assistant Pro to check for the correctness of the language. After using this system, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. This research received no grants or funds from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Acemoglu, D. (2021). *Harms of AI*. National Bureau of Economic Research Working Paper Series (vol. 29247). <https://doi.org/10.3386/w29247>
- de Almeida, P., dos Santos, C., & Farias, J. (2021). Artificial intelligence regulation: A framework for governance. *Ethics and Information Technology*, 23(3), 505–525. <https://doi.org/10.1007/s10676-021-09593-z>
- Anderljung, M., Barnhart, J., Korinek, A., Leung, J., O'Keefe, C., Whittlestone, J., Avin, S., Brundage, M., Bullock, J., & Cass-Beggs, D. (2023). Frontier AI regulation: Managing emerging risks to public safety. arXiv preprint: 2307.03718. Doi: [10.48550/arXiv.2307.03718](https://arxiv.org/abs/2307.03718).
- Atabekov, A., & Yastrebov, O. (2018). Legal status of artificial intelligence across countries: Legislation on the move. *European Research Studies Journal*, 21(4), 773–782. <https://doi.org/10.35808/ersj/1245>
- Baker, S. E., Edwards, R., & Doidge, M. (2012). How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research. <https://research.brighton.ac.uk/en/publications/how-many-qualitative-interviews-is-enough-expert-voices-and-early>.
- Bannister, F., & Connolly, R. (2020). Administration by algorithm: A risk management framework. *Information Polity*, 25, 471–490. <https://doi.org/10.3233/IP-200249>
- Bélisle-Pipon, J.-C., Monteferrante, E., Roy, M.-C., & Couture, V. (2023). Artificial intelligence ethics has a black box problem. *AI & SOCIETY*, 38(4), 1507–1522. <https://doi.org/10.1007/s00146-021-01380-0>
- Belli, L., Curzi, Y., & Gaspar, W. B. (2023). AI regulation in Brazil: Advancements, flows, and need to learn from the data protection experience. *Computer Law and Security Review*, 48, Article 105767. <https://doi.org/10.1016/j.clsr.2022.105767>

- Bello, & Villarino, J.-M. (2023). Global standard-setting for artificial intelligence: Para-regulating international law for AI? *The Australian Year Book of International Law Online*, 41(1), 157–181. <https://doi.org/10.1163/26660229-04101018>
- Bender, E. M., Geburu, T., McMillan-Major, A., & Shmittchell, S. (2021). In *On the dangers of stochastic parrots: Can language models be too big?*. <https://doi.org/10.1145/3442188.3445922>
- Black, J., & Murray, A. (2019). Regulating AI and machine learning: Setting the regulatory agenda. *European Journal of Law and Technology*, 10(3).
- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16. <https://doi.org/10.1214/ss/1009213726>
- Buchanan, D., Boddy, D., & McCalman, J. (2013). Getting in, getting on, getting out, and getting back. In A. Bryman (Ed.), *Doing research in organizations* (1st ed., pp. 53–67). Routledge. <https://doi.org/10.4324/9780203385449>
- Buiten, M. (2019). Towards intelligent regulation of artificial intelligence. *European Journal of Risk Regulation*, 10(1), 41–59. <https://doi.org/10.1017/err.2019.8>
- Calvi, A., & Kotzinos, D. (2023). In *Enhancing AI fairness through impact assessment in the European Union: A legal and computer science perspective*. <https://doi.org/10.1145/3593013.3594076>
- Carlsson, V., & Rönnblom, M. (2022). From politics to ethics: Transformations in EU policies on digital technology. *Technology in Society*, 71, Article 102145. <https://doi.org/10.1016/j.techsoc.2022.102145>
- Carter, D. (2020). Regulation and ethics in artificial intelligence and machine learning technologies: Where are we now? Who is responsible? Can the information professional play a role? *Business Information Review*, 37(2), 60–68. <https://doi.org/10.1177/0266382120923962>
- Chae, Y. (2020). U.S. AI regulation guide: Legislative overview and practical considerations. *Journal of Robotics, Artificial Intelligence and Law*, 3(1), 17–40. doi: agispt.20230202082753.
- Chamberlain, J. (2023). The risk-based approach of the European Union's proposed artificial intelligence regulation: Some comments from a tort law perspective. *European Journal of Risk Regulation*, 14(1), 1–13. <https://doi.org/10.1017/err.2022.38>
- Charles, V., Rana, N. P., & Carter, L. (2022). Artificial intelligence for data-driven decision-making and governance in public affairs. *Government Information Quarterly*, 39(4), Article 101742. <https://doi.org/10.1016/j.giq.2022.101742>
- Chatterjee, S., & N.S., S. (2023). Impact of AI regulation and governance on online personal data sharing: From sociological, technology and policy perspective. *Journal of Science and Technology Policy Management*, 14(1), 157–180. <https://doi.org/10.1108/JSTPM-07-2020-0103>
- Chauhan, K. (2023). From 'what' and 'why' to 'how': An imperative driven approach to mechanics of AI regulation. *Global Jurist*, 23(2), 99–124. <https://doi.org/10.1515/gj-2022-0053>
- Clarke, R. (2019). Regulatory alternatives for AI. *Computer Law and Security Review*, 35(4), 398–409. <https://doi.org/10.1016/j.clsr.2019.04.008>
- Cuéllar, M.-F., Larsen, B., Lee, Y. S., & Webb, M. (2022). Does information about AI regulation change manager evaluation of ethical concerns and intent to adopt AI? *The Journal of Law, Economics, and Organization*, 40(1), 34–75. <https://doi.org/10.1093/jleo/ewac004>
- Cyman, D., Gromova, E., & Juchnevicius, E. (2021). Regulation of artificial intelligence in BRICS and the European Union. *Brics Law Journal*, 8(1), 86–115. <https://doi.org/10.21684/2412-2343-2021-8-1-86-115>
- Demetzou, K. (2019). Data protection impact assessment: A tool for accountability and the unclarified concept of 'high risk' in the general data protection regulation. *Computer Law and Security Review*, 35(6), Article 105342. <https://doi.org/10.1016/j.clsr.2019.105342>
- Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy artificial intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, 99, Article 101896. <https://doi.org/10.1016/j.inffus.2023.101896>
- Dignam, A. (2020). Artificial intelligence, tech corporate governance and the public interest regulatory response. *Cambridge Journal of Regions, Economy and Society*, 13(1), 37–54. <https://doi.org/10.1093/cjres/rtsaa002>
- Ebers, M. (2019). Regulating AI and robotics: Ethical and legal challenges. In M. Ebers, & S. Navas Navarro (Eds.), *Algorithms and law*. Cambridge University Press. <https://doi.org/10.2139/ssrn.3392379>
- Edwards, L. (2022). Regulating AI in Europe: Four problems and four solutions. <https://www.adalovelaceinstitute.org/report/regulating-ai-in-europe/>.
- Eisenhardt, K., & Graebner, M. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, 50(1), 25–32. <https://doi.org/10.5465/amj.2007.24160888>
- Eneman, M., Ljungberg, J., Raviola, E., & Rolandsson, B. (2022). The sensitive nature of facial recognition: Tensions between the Swedish police and regulatory authorities. *Information Policy*, 27, 219–232. <https://doi.org/10.3233/IP-211538>
- Erdélyi, O., & Goldsmith, J. (2022). Regulating artificial intelligence: Proposal for a global solution. *Government Information Quarterly*, 39(4), Article 101748. <https://doi.org/10.1016/j.giq.2022.101748>
- European Parliament. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). In , 119. *Official Journal of the European Union* (pp. 1–88). <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02016R0679-20160504>.
- European Parliament (2023, 18-06-2024). EU AI Act: first regulation on artificial intelligence. Retrieved 17 September from <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>.
- Fernandes, P. M., Santos, F. C., & Lopes, M. (2020). Norms for beneficial A.I.: A computational analysis of the societal value alignment problem. *AI Communications*, 33, 155–171. <https://doi.org/10.3233/AIC-201502>
- Ferrari, F., van Dijk, J., & van den Bosch, A. (2023). Observe, inspect, modify: Three conditions for generative AI governance. *New Media & Society*. <https://doi.org/10.1177/14614448231214811>
- Finocchiaro, G. (2024). The regulation of artificial intelligence. *AI & SOCIETY*, 39(4), 1961–1968. <https://doi.org/10.1007/s00146-023-01650-z>
- Fountain, J. E. (2022). The moon, the ghetto and artificial intelligence: Reducing systemic racism in computational algorithms. *Government Information Quarterly*, 39(2), 101645. <https://doi.org/10.1016/j.giq.2021.101645>
- Gaeta, M. C., Aulino, L., & Troisi, E. (2023). The possible relationships between law and ethics in the context of artificial intelligence regulation. *Humana Mente*, 16(44). <https://www.humanamente.eu/index.php/HM/article/view/435>.
- Galaz, V., Centeno, M. A., Callahan, P. W., Causevic, A., Patterson, T., Brass, I., ... Levy, K. (2021). Artificial intelligence, systemic risks, and sustainability. *Technology in Society*, 67, Article 101741. <https://doi.org/10.1016/j.techsoc.2021.101741>
- Gervais, D. (2023). Towards an effective transnational regulation of AI. *AI & SOCIETY*, 38(1), 391–410. <https://doi.org/10.1007/s00146-021-01310-0>
- Grady, P. (2023). The AI act should be technology-neutral. <https://itif.org/publications/2023/02/01/the-ai-act-should-be-technology-neutral/>.
- Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law and Security Review*, 45, Article 105681. <https://doi.org/10.1016/j.clsr.2022.105681>
- Greiman, V. (2020). *Artificially intelligent systems and human rights: A global perspective*. Reading, UK: International Conference on Cyber Warfare and Security.
- Hacker, P. (2023). The European AI liability directives – Critique of a half-hearted approach and lessons for the future. *Computer Law and Security Review*, 51, Article 105871. <https://doi.org/10.1016/j.clsr.2023.105871>
- Hacker, P., Cordes, J., & Rochon, J. (2024). Regulating gatekeeper artificial intelligence and data: Transparency, access and fairness under the digital markets act, the general data protection regulation and beyond. *European Journal of Risk Regulation*, 15(1), 49–86. <https://doi.org/10.1017/err.2023.81>
- Hacker, P., Engel, A., & Mauer, M. (2023). Regulating ChatGPT and other large generative AI models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, Chicago, IL, USA*.
- Hacker, P., & Passoth, J.-H. (2022). Varieties of AI explanations under the law. From the GDPR to the AIA, and beyond. In A. Holzinger, R. Goebel, R. Fong, T. Moon, K.-R. Müller, & W. Samek (Eds.), *xxAI - Beyond Explainable AI: International Workshop, Held in Conjunction with ICML 2020, July 18, 2020, Vienna, Austria, Revised and Extended Papers* (pp. 343–373). Springer International Publishing. https://doi.org/10.1007/978-3-031-04083-2_17.
- Han, T. A., Lenaerts, T., Santos, F. C., & Pereira, L. M. (2022). Voluntary safety commitments provide an escape from over-regulation in AI development. *Technology in Society*, 68, Article 101843. <https://doi.org/10.1016/j.techsoc.2021.101843>
- Helberger, N. (2024). FutureNewsCorp, or how the AI act changed the future of news. *Computer Law and Security Review*, 52, Article 105915. <https://doi.org/10.1016/j.clsr.2023.105915>
- Helberger, N., & Diakopoulos, N. (2023). ChatGPT and the AI act. *Internet Policy Review*, 12(1). <https://doi.org/10.14763/2023.1.1682>
- Hildebrandt, M. (2018). Algorithmic regulation and the rule of law. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2128), Article 20170355. <https://doi.org/10.1098/rsta.2017.0355>
- Hoffmann, C., & Hahn, B. (2020). Decentered ethics in the machine era and guidance for AI regulation. *AI & SOCIETY*, 35(3), 635–644. <https://doi.org/10.1007/s00146-019-00920-z>
- Jackson, M. (2020). Regulating AI. In M. Jackson, & M. Shelly (Eds.), *Legal regulations, implications, and issues surrounding digital data* (pp. 159–181). IGI Global. <https://doi.org/10.4018/978-1-7998-3130-3.ch009>.
- Jarota, M. (2023). Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an occupational health and safety perspective. *Computer Law and Security Review*, 49, Article 105825. <https://doi.org/10.1016/j.clsr.2023.105825>
- Justo-Hanani, R. (2022). The politics of artificial intelligence regulation and governance reform in the European Union. *Policy Sciences*, 55(1), 137–159. <https://doi.org/10.1007/s11077-022-09452-8>
- Kaminski, M. (2023). *Regulating the risks of AI*. 103. Boston University Law Review. <https://doi.org/10.2139/ssrn.4195066>, 22–21.
- Kaminski, M., & Urban, J. (2021). The right to contest AI. *Columbia Law Review*, 121(7), 1957–2048. <https://www.jstor.org/stable/27083420>.
- Kashkin, S. (2021). Robotics for modernization of legal regulation of artificial intelligence and robotics Technologies in Russia with platform legal models. *Kutafin Law Review*, 8(3), 370–389. <https://doi.org/10.17803/2313-5395.2021.3.17.370-389>
- Khisamova, Z., Begishev, I., & Gaifutdinov, R. (2019). On methods to legal regulation of artificial intelligence in the world. *International Journal of Innovative Technology and Exploring Engineering*, 9(1), 5159–5162. <https://doi.org/10.35940/ijitee.A9220.119119>
- Kiskis, M. (2023). Legal framework for the coexistence of humans and conscious AI. *Frontiers in Artificial Intelligence*, 6, Article 1205465. <https://doi.org/10.3389/frai.2023.1205465>
- Knott, A., Pedreschi, D., Chatila, R., Chakraborti, T., Leavy, S., Baeza-Yates, R., ... Bengio, Y. (2023). Generative AI models should include detection mechanisms as a condition for public release. *Ethics and Information Technology*, 25(4), 55. <https://doi.org/10.1007/s10676-023-09728-4>

- Knowles, B., & Richards, J. (2021). In *The sanction of authority: Promoting public trust in AI*. <https://doi.org/10.1145/3442188.3445890>
- Kokshagina, O., Reinecke, P., & Karanasios, S. (2023). To regulate or not to regulate: Unravelling institutional tussles around the regulation of algorithmic control of digital platforms. *Journal of Information Technology*, 38(2), 160–179. <https://doi.org/10.1177/02683962221114408>
- Koniakou, V. (2023). From the “rush to ethics” to the “race for governance” in artificial intelligence. *Information Systems Frontiers*, 25(1), 71–102. <https://doi.org/10.1007/s10796-022-10300-6>
- König, P. D., Wurster, S., & Siewert, M. B. (2023). Sustainability challenges of artificial intelligence and Citizens’ regulatory preferences. *Government Information Quarterly*, 40(4), Article 101863. <https://doi.org/10.1016/j.giq.2023.101863>
- Krarup, T., & Horst, M. (2023). European artificial intelligence policy as digital single market making. *Big Data & Society*, 10(1). <https://doi.org/10.1177/20539517231153811>
- de Laat, P. B. (2021). Companies committed to responsible AI: From principles towards implementation and regulation? *Philosophy and Technology*, 34(4), 1135–1193. <https://doi.org/10.1007/s13347-021-00474-3>
- Lazcoz, G., & de Hert, P. (2023). Humans in the GDPR and AIA governance of automated and algorithmic systems. Essential pre-requisites against abdicating responsibilities. *Computer Law and Security Review*, 50, Article 105833. <https://doi.org/10.1016/j.clsr.2023.105833>
- Lucaj, L., van der Smagt, P., & Benbouzd, D. (2023). In *AI Regulation Is (not) All You Need*. <https://doi.org/10.1145/3593013.3594079>
- Lucchi, N. (2023). ChatGPT: A case study on copyright challenges for generative artificial intelligence systems. *European Journal of Risk Regulation*, 1–23. <https://doi.org/10.1017/err.2023.59>
- Lupo, G. (2023). Risky artificial intelligence: The role of incidents in the path to AI regulation. *Law, Technology and Humans*, 5(1), 133–152. <https://doi.org/10.5204/ltjh.2682>
- Mesarčík, M., Podroužek, J., & Gavorník, A. (2023). On defense of “ethification” of law: How ethics may improve compliance with the EU digital laws. *Computer Law and Security Review*, 50, Article 105852. <https://doi.org/10.1016/j.clsr.2023.105852>
- Middleton, S. E., Letouze, E., Hossaini, A., & Chapman, A. (2022). Trust, regulation, and human-in-the-loop AI: Within the European region. *Communications of the ACM*, 65(4), 64–68. <https://doi.org/10.1145/3511597>
- Mitchell, A., Let, D., & Tang, L. (2023). AI regulation and the protection of source code. *International Journal of Law and Information Technology*, 31(4), 283–301. <https://doi.org/10.1093/ijlit/eaad026>
- Mökander, J., Axente, M., Casolari, F., & Floridi, L. (2022). Conformity assessments and post-market monitoring: A guide to the role of auditing in the proposed European AI regulation. *Minds and Machines*, 32(2), 241–268. <https://doi.org/10.1007/s11023-021-09577-4>
- Moses, L. B. (2013). How to think about law, regulation and technology: Problems with ‘technology’ as a regulatory target. *Law, Innovation and Technology*, 5(1), 1–20. <https://doi.org/10.5235/17579961.5.1.1>
- Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and Organization*, 17(1), 2–26. <https://doi.org/10.1016/j.infoandorg.2006.11.001>
- Nitzberg, M., & Zysman, J. (2022). Algorithms, data, and platforms: The diverse challenges of governing AI. *Journal of European Public Policy*, 29(11), 1753–1778. <https://doi.org/10.1080/13501763.2022.2096668>
- Nwafor, I. (2021). AI ethical bias: A case for AI vigilantism (Allantism) in shaping the regulation of AI. *International Journal of Law and Information Technology*, 29(3), 225–240. <https://doi.org/10.1093/ijlit/eaab008>
- Pagallo, U., Ciani Sciolla, J., & Durante, M. (2022). The environmental challenges of AI in EU law: Lessons learned from the artificial intelligence act (AIA) with its drawbacks. *Transforming Government: People, Process and Policy*, 16(3), 359–376. <https://doi.org/10.1108/TG-07-2021-0121>
- Papyshev, G., & Yarime, M. (2023). The state’s role in governing artificial intelligence: Development, control, and promotion through national strategies. *Policy Design and Practice*, 6(1), 79–102. <https://doi.org/10.1080/25741292.2022.2162252>
- Papyshev, G., & Yarime, M. (2024). The limitation of ethics-based approaches to regulating artificial intelligence: Regulatory gifting in the context of Russia. *AI & SOCIETY*, 39(3), 1381–1396. <https://doi.org/10.1007/s00146-022-01611-y>
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage Publications.
- Paul, R. (2023). European artificial intelligence “trusted throughout the world”: Risk-based regulation and the fashioning of a competitive common AI market. *Regulation & Governance*. <https://doi.org/10.1111/rego.12563>
- Petit, N. (2017). *Law and regulation of artificial intelligence and robots-conceptual framework and normative implications*. Available at SSRN 2931339. <https://doi.org/10.2139/ssrn.2931339>
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. <https://openai.com/index/language-unsupervised/>.
- Reed, C. (2018). How should we regulate artificial intelligence? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2128), Article 20170360. <https://doi.org/10.1098/rsta.2017.0360>
- Rességuier, A., & Rodrigues, R. (2020). AI ethics should not remain toothless! A call to bring back the teeth of ethics. *Big Data & Society*, 7(2). <https://doi.org/10.1177/2053951720942541>, 2053951720942541.
- Restrepo-Amariles, D., & Baquero, P. M. (2023). Promises and limits of law for a human-centric artificial intelligence. *Computer Law and Security Review*, 48, Article 105795. <https://doi.org/10.1016/j.clsr.2023.105795>
- Roberts, H., Babuta, A., Morley, J., Thomas, C., Taddeo, M., & Floridi, L. (2023). Artificial intelligence regulation in the United Kingdom: A path to good governance and global leadership? *Internet Policy Review*, 12(2), 1–31. <https://doi.org/10.14763/2023.2.1709>
- Robles Carrillo, M. (2020). Artificial intelligence: From ethics to law. *Telecommunications Policy*, 44(6), Article 101937. <https://doi.org/10.1016/j.telpol.2020.101937>
- Robles, P., & Mallinson, D. J. (2023). Catching up with AI: Pushing toward a cohesive governance framework. *Politics & Policy*, 51(3), 355–372. <https://doi.org/10.1111/polp.12529>
- Salgado-Criado, J., & Fernández-Aller, C. (2021). A wide human-rights approach to artificial intelligence regulation in Europe. *IEEE Technology and Society Magazine*, 40(2), 55–65. <https://doi.org/10.1109/MTS.2021.3056284>
- Schuett, J. (2023). Defining the scope of AI regulations. *Law, Innovation and Technology*, 15(1), 60–82. <https://doi.org/10.1080/17579961.2023.2184135>
- Smuha, N. (2021a). Beyond the individual: Governing AI’s societal harm. *Internet Policy Review*, 10(3). <https://doi.org/10.14763/2021.3.1574>
- Smuha, N. (2021b). From a ‘race to AI’ to a ‘race to AI regulation’: Regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57–84. <https://doi.org/10.1080/17579961.2021.1898300>
- Stahl, B. C., Rodrigues, R., Santiago, N., & Macnish, K. (2022). A European Agency for Artificial Intelligence: Protecting fundamental rights and ethical values. *Computer Law and Security Review*, 45, Article 105661. <https://doi.org/10.1016/j.clsr.2022.105661>
- Stebbins, R. A. (2001). *Exploratory research in the social sciences* (Vol. 48). Sage Publications. <https://doi.org/10.4135/9781412984249>
- Stuurman, K., & Lachaud, E. (2022). Regulating AI. A label to complete the proposed Act on Artificial Intelligence. *Computer Law and Security Review*, 44, Article 105657. <https://doi.org/10.1016/j.clsr.2022.105657>
- Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science*, 361(6404), 751–752. <https://doi.org/10.1126/science.aat5991>
- Tallberg, J., Lundgren, M., & Geith, J. (2024). AI regulation in the European Union: Examining non-state actor preferences. *Business & Politics*, 26(2), 218–239. <https://doi.org/10.1017/bap.2023.36>
- Truby, J., Brown, R. D., Ibrahim, I. A., & Parellada, O. C. (2022). A sandbox approach to regulating high-risk artificial intelligence applications. *European Journal of Risk Regulation*, 13(2), 270–294. <https://doi.org/10.1017/err.2021.52>
- Ulinicane, I. (2022). Artificial intelligence in the European Union: Policy, ethics and regulation. In T. Hoerber, G. Weber, & I. Cabras (Eds.), *The Routledge handbook of European integrations*. Taylor & Francis. <https://doi.org/10.4324/9780429262081-19>
- Vapnik, V. N. (2000). *The nature of statistical learning theory* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-4757-3264-1>
- Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, 19, 255–275. <https://doi.org/10.1146/annurev-lawsocsci-020223-040749>
- Veale, M., & Zuiderveen Borgesius, F. J. (2021). Demystifying the draft EU artificial intelligence act — Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, 22(4), 97–112. <https://doi.org/10.9785/crl-2021-220402>
- Vermes, J. (2023). *Why regulators in Canada and Italy are digging into ChatGPT’s use of personal information*. CBC. Retrieved 01/09/2023, from <https://www.cbc.ca/news/world/openai-chatgpt-data-privacy-investigations-1.6804205>.
- Viljanen, M., & Parviainen, H. (2022). AI applications and regulation: Mapping the regulatory strata. *Frontiers in Computer Science*, 3. <https://doi.org/10.3389/fcomp.2021.779957>
- Wan, W. Y., Tsimplis, M., Siau, K. L., Yue, W. T., Nah, F. F.-H., & Yu, G. M. (2022). Legal and regulatory issues on artificial intelligence, machine learning, data science, and big data. In J. Y. C. Chen, G. Fragomeni, H. Degen, & S. Ntoa (Eds.), *HCI international 2022 – Late Breaking Papers*. Cham: Interacting with eXtended Reality and Artificial Intelligence.
- White, J. M., & Lidskog, R. (2022). Ignorance and the regulation of artificial intelligence. *Journal of Risk Research*, 25(4), 488–500. <https://doi.org/10.1080/13669877.2021.1957985>
- Wirtz, B., Weyerer, J., & Sturm, B. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, 43(9), 818–829. <https://doi.org/10.1080/01900692.2020.1749851>
- Wirtz, B. W., Weyerer, J., & Kehl, I. (2022). Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government Information Quarterly*, 39(4), Article 101685. <https://doi.org/10.1016/j.giq.2022.101685>
- Wulf, A. J., & Seizov, O. (2020). Artificial intelligence and transparency: A blueprint for improving the regulation of AI applications in the EU. *European Business Law Review*, 611–640. <https://doi.org/10.54648/eubl2020024>
- Yin, R. K. (2013). Validity and generalization in future case study evaluations. *Evaluation*, 19(3), 321–332. <https://doi.org/10.1177/1356389013497081>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.
- Zanol, J., Buchel, A., Tjoa, S., & Kieseberg, P. (2022). What is AI? Exploring the scope of the “artificial intelligence act”. In E. Weblaw (Ed.), *Proceedings of the 25th International Legal Informatics Symposium IRIS*.
- Zhou, Y., & Kankanahalli, A. (2021). AI regulation for smart cities: Challenges and principles. In E. Estevez, T. A. Pardo, & H. J. Scholl (Eds.), *Smart cities and smart governance: Towards the 22nd century Sustainable City* (pp. 101–118). Springer International Publishing. https://doi.org/10.1007/978-3-030-61033-3_5
- Zuidervijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research

agenda. *Government Information Quarterly*, 38(3), Article 101577. <https://doi.org/10.1016/j.giq.2021.101577>

Antonio Cordella is an Associate Professor in Management at the London School of Economics and Political Science (LSE), and a visiting professor at the Maastricht Graduate School of Governance, UNMERIT, The Netherlands. His research interests cover the areas of E-government, Economic Theories of Information Systems, and the Social Studies of Information Systems. His publications address problems related to the adoption of ICTs in the public sector and discuss the institutional and organisational implications of such adoptions. An Italian national, he holds a PhD in Information Systems from Gothenburg University, Sweden.

Francesco Gualdi is a Lecturer of International Business and Technology at Regent's University London. He is also affiliated to the Department of Management of the London School of Economics and Political Science (LSE), and he is a Research Fellow at the Center for Technology in Government (Albany, USA). He holds a PhD in Enterprise Engineering from Tor Vergata University in Rome. His main research interests cover the areas of AI adoption in public sector, digital government, impact of technology on policymaking activity. Prior to the academic career, he served as strategic adviser to the Minister of European Affairs within the Italian Government (2014–2018).