Lies and fake news: a Kantian approach

Mentiras e fake news: uma interpretação kantiana

Karine Cristine de Souza Barboza

Universidade Federal do Paraná (UFPR)/Universität Vechta

karinesouzabarboza@gmail.com

Abstract: This article explores the contributions of Kantian philosophy to contemporary debates on information disorders, with a focus on fake news. It begins with Kant's definition of lying, emphasizing the conditions of declaration, untruthfulness, and intentionality as fundamental criteria for distinguishing misinformation, disinformation, and malinformation. The article argues that, unlike approaches centered on the medium of dissemination, these Kantian criteria offer a more precise delimitation of the concept of fake news. It then develops a critical analysis of Kant's definition of lying within the framework of Doctrine of Right, highlighting the distinction between private law, which protects individual rights, and public law, which safeguards the rights of humanity and social trust—the foundation of the civil pact. The article concludes that the integration of these criteria provides a robust theoretical framework for critically understanding fake news and related information disturbances, contributing to ethical and legal discussions on legitimacy and responsibility in the public sphere.

Keywords: declaration; fake news; information disorders; intentionality; lie; untruthfulness.

Resumo: Este artigo examina as contribuições da filosofia kantiana para o debate atual sobre os distúrbios informacionais, especialmente as fake news. Parte da definição kantiana de mentira, destacando as condições de declaração, inveracidade e intencionalidade como critérios fundamentais para distinguir misinformation, disinformation e malinformation. Defende que, ao contrário das abordagens centradas nos meios de veiculação, esses critérios kantianos oferecem uma delimitação mais precisa do conceito de fake news. Em seguida, desenvolve uma análise crítica da definição kantiana de mentira na Doutrina do Direito, ressaltando a distinção entre direito privado, focado na proteção dos direitos individuais, e direito público, que protege o direito da humanidade e a confiança social, base do pacto civil. Conclui que a articulação desses critérios fornece um referencial teórico sólido para compreender criticamente as fake news e demais distúrbios informacionais, contribuindo para debates éticos e jurídicos sobre legitimidade e responsabilidade na esfera pública.

Palavras-chave: declaração; distúrbios informacionais; fake news; intencionalidade; mentira; inveracidade.

1. Introduction

Although Kant does not explicitly offer a specific or extensive discussion of the concept of fake news, the centrality of the theme of lying in his philosophy is undeniable. In the *Groundwork of the Metaphysics of Morals* (GMS), Kant begins with the imperative "thou shalt not lie" from which he explores the scope of the moral duty (GMS, AA, 04: 389 | Kant, 1998a, p. 2-3)[1]. Throughout his writings, the immorality of lying is chiefly examined through the example of false promises. In the *Metaphysics of Morals* (MS), Kant revisits and expands his analysis of lying, characterizing it as "the greatest violation of a human being's duty to himself regarded merely as a moral being" (MS, AA 06: 429 | Kant, 1996, p. 182). There, he distinguishes between external and internal lying, emphasizing the precedence of the former. According to Kant,

> Such insincerity in his declarations [...] still deserves the strongest censure [...] that the ill of untruthfulness spreads into his relations with other human beings as well, once the highest principle of truthfulness has been violated (MS, AA 06: 430-31 | Kant, 1996, p. 183).

As in the *Metaphysics*, Kant also emphasizes the gravity of internal lying in *Religion within the Boundaries of Mere Reason* (RGV). In this context, he addresses it through the lens of untruthfulness in relation to one's own belief in God. Moreover, he defines lying as a form of moral perversity (RGV, AA 06: 38 | Kant, 1998b, p. 60). It is in *Toward Perpetual Peace* (ZeF), however, that we find a discussion of lying in the political context—the very context in which fake news is most commonly debated.

Moreover, despite the editorial and interpretative complexities that characterize the *Lectures on Ethics*, it is possible to identify, in a lecture specifically dedicated to the topic of lying, a passage in which Kant directly addresses the issue of spreading false reports. As the philosopher explains:

> If a man publishes a false report, he thereby does no wrong to anyone in particular, but offends against mankind, for if that were to become general, the human craving for knowledge would be thwarted; apart from speculation, I have only two ways of enlarging my store of information: by experience, and by testimony. But now since I cannot experience everything myself, if the reports of others were to be false tidings, the desire for knowledge could not be satisfied. A *mendacium* is thus a *falsiloquium in praejudicium humanitatis*, even when it is not also in violation of any particular *jus quaesitum* of another. In law a *mendacium* is a *falsiloquium in praejudicium alterius*, and cannot be anything else there, but from the moral viewpoint it is a *falsiloquium in praejudicium humanitatis* (V-Mo/Collins, 27: 447-448 | Kant, 1997, p. 203-204)[2].

The dissemination of false reports, constituting a breach of the fundamental *pactum sociale* described in this passage, finds substantial support in Kant's work, particularly in the *Metaphysics*. This correlation demonstrates that Kant not only acknowledged the far-reaching political and social implications of lying, but also emphasized the importance not merely of truthfulness, but equally of mutual trust among agents as an indispensable condition for both the construction of knowledge and the possibility of social life. Although Kant did not develop a systematic typology of the various forms of falsehood dissemination, this article argues that his concept of lying offers

---

[1] All references to Kant's works follow the Akademie-Ausgabe (AA) edition (Kant, 1900) accompanied by the corresponding translation.

[2] Although thematically related to this passage from the *Lectures on Ethics*, it is important to note that in the original German text Kant employs the term "*falsche Nachrichten*" (V-Mo/Collins, 27:447–448). In the Cambridge University Press English edition, this expression is rendered as "false report," preserving its broad original sense (Kant, 1997, p. 203). In the Brazilian Portuguese edition published by Unesp (Kant, 2018, p. 300), however, it appears as "notícias falsas" ("fake news"). It should be emphasized that in Kant's context, <Nachrichten> is not restricted to the modern journalistic meaning of "news," but rather denotes reports or information in a more general sense. Consequently, translations such as "false reports" or "false information" are more faithful to the original usage, avoiding the semantic narrowing that the contemporary term "fake news" may entail.

significant theoretical resources for the critical analysis and understanding of the contemporary phenomenon of fake news.

Accordingly, the structure of this article has been designed to ensure conceptual clarity and analytical coherence in addressing information disorders through the lens of Kant's practical philosophy. The article begins by proposing a precise delineation of information disorders, specifically disinformation, misinformation, and malinformation. Building on this foundation, it analyzes the threefold distinction found in the specialized literature in light of the conditions established in Kant's definition of a lie. This initial step is justified by the need to avoid terminological ambiguity and to ensure the appropriate applicability of Kantian concepts to the study of information disorders. Subsequently, the article turns to an examination of lying and fake news within the context of Kant's Doctrine of Right, highlighting that such practices constitute significant violations from a juridical perspective, particularly when they undermine public trust and the foundations of the social contract. By adopting this trajectory—from conceptual analysis to normative legal reflection—the study aims to articulate Kantian thought with the contemporary issue of information disorders, providing a robust and critical foundation for the ethical and legal understanding of fake news.

2. Fake news and information disorders

The dissemination and popularization of the term "fake news" gained significant prominence starting in 2016, particularly due to the U.S. presidential elections, when then-candidate Donald Trump accused certain media outlets of spreading false information and being untrustworthy[3]. Previously, the concept of fake news was predominantly associated with two main types of content: "The meaning of the term now ranges from fabricated news circulated via social media to a polemic umbrella term meant to discredit 'legacy' news media" (Quandt *et al.*, 2019, p. 1).

In contemporary discussions, however, the term has become highly ambiguous, due both to the lack of a precise definition and the challenges involved in identifying and characterizing the various modes of dissemination of false information. This concern is echoed by Kant, who holds that lies—particularly false reports—breach a fundamental social pact and engender distrust not only within information channels but also systemic distrust that permeates interpersonal relations, the relationship between citizens and the State, and relations among States, thereby undermining all contracts and, consequently, the very foundation of law.

Regarding the semantic ambiguity of the term fake news, there is a noticeable diversity of approaches and interpretations in the specialized literature. While some authors adopt a broad perspective, others advocate for more restrictive conceptual delimitations, highlighting the complexity and multiplicity of phenomena encompassed by this terminology. Nielsen and Graves (2017, p. 5) report that audiences tend to adopt a broad understanding of the term fake news, encompassing any form of dissemination of false information—including satire, fabricated news, poor journalism, propaganda, and more—regardless of the intent to deceive. By contrast, authors such as Alcott and Gentzkow (2017, p. 213), Lazer *et al.* (2018, p. 1094), Himma-Kadakas (2017), and Gelfert (2018, p. 84) advocate for a more specific use of the term, defining fake news as entirely fabricated content that mimics the form of traditional news, but does not adhere to journalistic processes for verifying the accuracy of information. This group of authors agrees that

---

[3] For a brief overview of the history of fake news, see Posetti and Matthews (2018).

fake news typifies news presented in the format of traditional media. Gelfert emphasizes that the creation of fake news carries the intention to deceive and is, as the author states, "misleading *by design*" (Gelfert, 2018, p. 84).

Due to the polysemy of the term, the European Commission (Lesher; Pawelec; Desar, 2022, p. 10) advocates abandoning the term fake news, whereas authors such as Lazer *et al.* defend its use "because of its value as a scientific construct, and because its political salience draws attention to an important subject" (Lazer *et al.*, 2018, p. 1094).

Despite the divergences surrounding the definition of the term, there is a growing consensus in the specialized literature regarding the importance of distinguishing different forms of false or distorted information, often grouped under the label of "information disorder". The Canadian Centre for Cyber Security (CCCS, 2024) and the European Commission (Lesher; Pawelec; Desar, 2022) all propose a classification into three categories: disinformation, misinformation, and malinformation.

Disinformation refers to false or misleading information that is deliberately created with the intent to manipulate public opinion, generate confusion, obtain political or economic advantages, or cause harm to individuals, institutions, or social groups. The central element of this category is the intention to deceive and produce negative consequences.

Misinformation, in turn, refers to the dissemination of false, inaccurate, or misleading information without deliberate intent to deceive. It often results from errors, misunderstandings, or naïve reliance on unverified sources, and is frequently shared by individuals who believe they are accurately informing others.

Finally, malinformation—or contextual misinformation—involves the use of truthful information that is presented out of context, distorted, or selectively manipulated with the intent to deceive, defame, or cause harm. Typical cases include the disclosure of private messages or genuine data obtained illicitly, as well as the use of accurate content to fuel hate speech or undermine someone's reputation.

In brief, both disinformation and malinformation involve a creator who is aware of the falsity of the information and intends to deceive. These are deliberate acts, unlike misinformation, which results from error or misunderstanding. In the case of misinformation, the spreader is unaware of the falsehood, is not the originator of the content, and lacks any intention to deceive.

In addition to these three distinctions, the European Commission also differentiates between propaganda and satire, thereby broadening the scope of the discussion on forms of misinformation. Regarding the specific placement of fake news within this taxonomy, authors such as Lazer *et al.* (2018) note a significant overlap with the category of disinformation, since fake news is generally fabricated with the deliberate intent to deceive. Distinctly, the European Commission (Lesher; Pawelec; Desar, 2022) following the interpretative framework proposed by Zimdars and McLeod (2020), explicitly classifies fake news as a subtype of disinformation.

Situated within this debate, Kant's definition of lying anticipates and grounds fundamental distinctions essential for understanding the phenomenon of fake news. I will argue below that the concept of fake news aligns closely with Kant's notion of lying, particularly with regard to the conditions of declaration, untruthfulness, and intentionality. These conditions are more pertinent

for delimiting the concept of fake news than the focus on journalistic design, as advocated by authors such as Gelfert (2018, p. 84). The relevant point is that the information is presented within a context of seriousness and veracity, rather than being limited to the format of traditional journalistic media. Although these broader criteria expand the range of cases considered as fake news, they highlight the crucial issue at stake in understanding the manipulation of false information, namely that such information is disseminated within a context that authorizes the receiver's trust both in the veracity of the message and in the sender's commitment to the truthfulness of what is communicated. Therefore, based on a Kantian philosophical framework, both fake news disseminated with deliberate intent to deceive by politicians and state actors, as well as those spread through social media and digital applications, fall within the concept of fake news, not being limited solely to information conveyed through traditional journalistic media formats.

Kant does not restrict false reports to the traditional journalistic format. Rather, he conceives them within a broader spectrum that includes cases such as defamation, false contracts and promises, as well as the employment of lying as a strategy in political practice. Thus, the Kantian definition of lying aligns closely with the broader concept of disinformation, understood as the intentional dissemination of factually false information with the deliberate purpose of deceiving. As will be detailed in the following section, the condition of declaration is fundamental: in order for a lie to occur, the false information must presuppose a declaration of honesty, either explicitly expressed by the communicator or implicitly inferred from the context. This declaration grants the receiver the legitimate right to take the information seriously and to assume that the communicator believes in what is being conveyed. For this reason, based solely on the condition of declaration, the Kantian concept of lying encompasses the tripartite distinction of contemporary information disorders—disinformation, malinformation, and misinformation—as all involve an explicit or implicit commitment by the communicator to the truthfulness of the information or the context in which it is conveyed. However, although the Kantian definition accommodates disinformation (and thus the specificity of fake news) and malinformation, misinformation does not fully fit, as it fails to satisfy the condition of intentionality: in this case, the spreader of misinformation does not have the deliberate intention to deceive but rather shares an error.

The distinction between disinformation and malinformation (or contextual deception) closely parallels the differentiation between lying and paltering. Unlike lying, in paltering the agent is truthful regarding the content of the declaration but deceives by manipulating the context in which the information is conveyed, thereby leading the interlocutor to a mistaken interpretation. This linguistic dynamic has been conceptualized by various authors under different terms: as paltering (Schauer & Zeckhauser, 2009 cited in Mahon, 2015, p. 6); as a deliberate attempt to deceive without resorting to outright falsehood (Saul, 2012); or as a form of linguistic fraud that, while misleading, does not technically constitute lying (Fried, 1978, p. 68). Paltering refers to situations in which the interlocutor is misled through truthful statements combined with the strategic omission of relevant information. It constitutes a form of evasion of sincerity, wherein the speaker, by selecting convenient true propositions, leads the listener to believe they have received a clear and direct answer to the question posed—even though this is not the case. This rhetorical device also frequently appears, for example, in cases of fake news, where real facts are presented out of context, causing the message recipient to infer meanings that were not present in the reported event. This type of malinformation can be illustrated by the case of Menno, presented

in the *Lectures on Ethics*: an example of deliberate ambiguity that simulates a direct answer while avoiding formal lying. As Kant states:

> When Menno, the founder of the sect, was due to be arrested, and escaped by mail coach, the arrestwarrant arrived first, at one of the stages, and the postmaster asked each of the passengers if Menno was on the coach. Instead of lying, that he was not on board, he asked his companion if it was being asked whether Menno was on board; but since the latter did not know Menno, he remained undetected (V-MS/Vigil, AA 27: 702 | Kant, 1997, p. 428).

Kant also points out that, in certain social situations, it is possible to avoid giving a direct answer to a question by means of an unexpected deviation in the conversation, thereby keeping uncertain what the speaker's true or apparent judgment is: "[...] by an unexpected turn of the conversation to divert the others in a direction where it remains doubtful what my true or ostensible judgement will be" (V-MS/Vigil, AA 27: 701 | Kant, 1997, p. 427). This rhetorical device, as described by Kant, constitutes a form of moral ambiguity employed with the intent to deceive. He explains that "*aequivocatio* is permitted, in order to reduce the other to silence and get rid of him, so that he shall no longer try to extract the truth from us, once he sees that we cannot give it to him, and do not wish to tell him a lie" (V-Mo/Collins, AA 27: 449 | Kant, 1997, p. 204–5). As Kant indicates, this kind of subterfuge, known as paltering, does not constitute a lie, as it does not involve the untruthfulness of the information itself: although there is contextual untruthfulness, the information is not factually false.

In this context, Sandel (2012) explores the implications of Kant's definition of lying through his analysis of Bill Clinton's statements during the sex scandal that undermined Clinton's popularity and influenced the impeachment proceedings. After Clinton's romantic involvement with a White House staffer was confirmed, his lawyer argued that the president had misled the public. However, he maintained that Clinton had not technically lied when he claimed not to have had sexual relations with the staffer, since their interaction did not fall under the dictionary definition of "sex". In his provocatively titled essay *Would Kant have defended Bill Clinton?*, Sandel suggests that Kant's conception of lying is significantly narrower than commonly assumed, such that evasive statements—even if technically true, as in the case of paltering—might, under this definition, be considered morally permissible (Sandel, 2012, p. 134).

It is important to note, however, that Kant treats paltering as a strategy employed in situations of social inconvenience, aimed at avoiding embarrassing disclosures. As we will see later in the Doctrine of Right, Kant does not limit legal responsibility to cases of outright lying, but extends it to any form of deception that undermines the freedom of another's choice. In this sense, even within legal contexts, there would be no room for the legitimacy of malinformation, even if it relies on the isolated truthfulness of certain facts, insofar as it distorts the context of their presentation and thereby misleads the recipient.

In what follows, I present the Kantian definition of a lie, elaborating on the conditions of declaration, untruthfulness, and intentionality that constitute it, and examining how these elements relate to the tripartite distinction among types of information disorder.

3. A conceptual framework for fake news grounded in Kant's definition of a lie

Although Kant does not present an exhaustive analysis of the concept of lying, §9 of the Doctrine of Virtue constitutes the primary reference for its definition[4]. Kant defines lying as the "[...] communication of one's thoughts to someone through words that yet (intentionally) contain the contrary of what the speaker thinks" (MS, AA 06: 430 | Kant, 1996, p. 182)[5]. Kant's stylistic choice to place the term "intentionally" <absichtlich> in parentheses within the definition of lying, beyond a mere rhetorical device, indicates that intentionality, although not syntactically integrated into the main structure, is nonetheless presupposed as a necessary condition of lying. In this regard, the adverb *absichtlich* performs a dual semantic function: on the one hand, it indicates that lying is a deliberate choice by the agent, rather than a mere mistake; on the other hand, it points to the duplicity of the liar, whose intention to deceive is concealed beneath a pretense of truthfulness. As Kant proceeds, lying is the "[...] renunciation by the speaker of his personality, and such a speaker is a mere deceptive appearance of a human being, not a human being himself". Thus, beyond the intentional expression of the opposite of what one thinks—that is, a conscious propositional untruth—Kant includes in the definition of lying the condition of intentionality, understood as the duplicity of the liar who conceals their true intention.

Kant's definition of lying, grounded in the condition of intent to deceive, encompasses the concepts of disinformation and contextual deception (or malinformation). However, it is important to note that, although the aforementioned tripartite distinction constitutes the prevailing conceptual framework in the study of information manipulation in digital environments, Kant's concept of lying is not limited to the medium through which falsehoods are disseminated. Kant does not confine the concept of lying to oral or written communication, nor does he restrict it to a deceptive interaction between two individuals (one-to-one). Moreover, the act of lying is not limited to speech and can be conveyed through other means such as letters, magazines, television commercials, among others (Mahon, 2015, p. 8).

Consequently, Kant's definition of lying establishes untruthfulness as a fundamental criterion, without necessarily requiring falsehood. As Mahon (2015, p. 5) clarifies, untruthfulness presupposes that the agent utters a declaration he believes to be false, implying the subjective awareness that his proposition does not correspond to reality. In contrast, the condition of falsehood entails that the statement is objectively false. Although contemporary distinctions regarding forms of misinformation predominantly use falsehood as a criterion, it is important to emphasize that for Kant it suffices that the agent intends to deceive through untruthfulness. As will be demonstrated below, when specifically addressing false reports, Kant clearly applies the criterion of falsehood to evaluate the content of the declaration, while employing the criterion of untruthfulness to judge the moral agent's intention.

Related to this distinction between untruthfulness and falsehood, the criterion of intentionality in the definition of lying serves as a basis for differentiating between lying and error, much like the distinction between disinformation and misinformation. According to the philosopher,

---

[4] For a detailed study of the conditions of the concept of lying in Kant, see Barboza and Klein (2024), upon which this section draws.

[5] In the original text: "[...] die doch das Gegenteil von dem (absichtlich) enthalten, was der Sprechende dabei denkt [...]" (MS, AA 06: 429).

> One cannot always stand by the truth of what one says to oneself or to another (for one can be mistaken); however, one can and must stand by the truthfulness of one's declaration or confession, because one has immediate consciousness of this. For in the first instance we compare what we say with the object in a logical judgment (through the understanding), whereas in the second instance, where we declare what we hold as true, we compare what we say with the subject (before conscience). Were we to make our declaration with respect to the former without being conscious of the latter, then we lie, since we pretend something else than what we are conscious of (MpVT, AA 08: 267-268 | Kant, 1998b, p. 27).

In this passage, Kant emphasizes that the decisive moral standard for evaluating declarations is not their conformity with the facts, but rather their truthfulness—understood as the correspondence between what is declared and what is believed. This stems from the fact that, while error concerning objects external to consciousness is always possible, the same does not apply to internal content—that is, to what we are subjectively aware of. Conceptually, therefore, truthfulness and untruthfulness refer to the relationship between a declaration and the agent's belief, whereas truth and falsehood pertain to the correspondence between the declaration and cognitive content.

Another important point, clarified by Wood (Wood, 2008, p. 240), is that lying must be understood as a declaration (*Aussage, Deklaration, declaratio*), that is, a form of expression that grants the listener the right to presume that what is said corresponds to what is believed. A declaration, therefore, entails a commitment to truthfulness, as it occurs in contexts in which the interlocutor is entitled to rely on the conformity between speech and the speaker's internal conviction—which makes the speaker morally responsible if they declare something they know to be false (Wood, 2008, p. 241).

The commitment to truthfulness may be explicitly stated by the speaker or, as is often the case, implied by the context itself. In normative settings, such as ethical or legal domains, the requirement of honesty is presumed, since the context authorizes interlocutors to expect that what is said corresponds to what is believed. However, there are discursive situations in which this requirement is suspended, either explicitly or contextually—as in the case of satire or fictional works. In such cases, the utterance is not subject to the implication of truthfulness, and the listener is not expected to take what is said at face value. As Kant observes, when a speaker signals that their words are not to be taken seriously—whether by stating so directly or through the circumstances in which the statement is made—one cannot say that a lie has been told, even if someone is misled as a result. In such a case, the deception does not stem from a deliberate intention to deceive (MS, AA 06: 238 | Kant, 1996, p. 30). Although such expressions may, in fact, lead to misunderstanding, the absence of an intention to deceive, combined with the suspension of declarative commitment, exempts the speaker from moral responsibility for lying—even if they may, in certain cases, still be held accountable for the harm caused by the misunderstanding, regardless of whether it was grounded in a lie.

The European Commission (Lesher; Pawelec; Desar, 2022, p. 9) defines satire as a form of social and political critique that uses humor and exaggeration to address relevant societal issues. However, a disinformation problem arises when the original satirical context is lost, often due to mass sharing. This loss of context may lead the reader to interpret the content as a literal assertion. Such a situation is particularly problematic for the declarative condition in the Kantian definition of lying, as the recipient is left uncertain about how to interpret the information. However, even when the liar takes advantage of such ambiguity to deceive, the Kantian definition of lying does not exempt them from the harm resulting from their action. The Kantian definition of lying

includes the condition of untruthfulness, which is grounded in the agent's intention rather than the falsehood of the information. Kant states that the liar "makes himself contemptible in his own eyes and violates the dignity of humanity in his own person" (MS, AA 06: 429 | Kant, 1996, p. 182). In the legal sphere, although the ambiguity of the context may hinder the proof of intent to deceive, the agent can still be held liable for the damages caused by their deception. Thus, according to the tripartite distinction of information disorders, the loss of the satirical context and contextual ambiguity can be understood as a use of contextual deception or malinformation by the spreader, which may generate misinformation in the receiver—these points will be further developed in the following section, dedicated to lying in the Doctrine of Right.

As initially argued, Kant's concept of lying provides central conceptual elements for understanding the contemporary phenomenon of fake news. Thus, even the emphasis on the condition of declaration—fundamental in Kant's definition—indicates that the study of misinformation should not be confined to the traditional journalistic context. Restricting the analysis to this specific domain risks overlooking other equally relevant communicative contexts. The criterion of declaration, as employed by Kant, is more appropriately used to delineate cases of information disorders: what matters is not the medium used, but whether the interlocutor is entitled to take the received information seriously, presuming that the communicator is committed to the truthfulness of what is conveyed. As highlighted by various contemporary studies, this type of communicative commitment manifests across multiple mediation forms that go beyond magazines and newspapers, including social networks, messaging applications, and other digital formats of information circulation.

4. Fake news and lying in the Doctrine of Right

In the passage where Kant asserts that "if a man publishes a false report, he thereby does no wrong to anyone in particular, but offends against mankind", the philosopher revisits the distinction between the public and private spheres of law in the Doctrine of Right (V-Mo/Collins, AA 27: 448 | Kant, 1997, p. 203-204). In private law, lying is narrowly defined as a declaration contrary to what one thinks, uttered with the intention to deceive and resulting in harm to another's free will. In this sense, Kant restricts the legal use of the term "lie" to falsifications that directly infringe upon the rights of another person, as exemplified by the simulation of a contract (MS, AA 06: 238 | Kant, 1996, p. 30). He clarifies that the jurist recognizes and applies this definition only when there is a violation of legal duties toward others (*officii juridicorum*), understanding lying in this case as a *falsiloquium dolosum in praejudicium alterius*—that is, a deceitful falsehood to the detriment of another (V-MS/Vigil, AA 27: 604–605 | Kant, 1997, p. 350–351).

Consequently, public law understands *mendacium* more broadly, recognizing it as a violation not only of legal duties between individuals but also as an infringement of the right of humanity. Kant observes that even when a lie does not cause direct harm to anyone, it still constitutes a serious breach of duty in general, as it undermines the credibility of declarations. As he states,

> I do wrong to duty in general in a most essential point. That is, as far as in me lies I bring it about that statements (declarations) in general find no credence, and hence also that all rights based on contracts become void and lose their force, and this is a wrong done to mankind in general (VRML, AA 08: 426 | Kant, 1993, p. 64).

Thus, Kant concludes that lying inevitably causes harm: if not to an individual in particular, then at least to humanity as a whole, since it corrupts the very source of law. In this context, trust in the truthfulness of declarations is a structural and indispensable element for legal stability

and legitimacy. Therefore, whether lying in court, in contracts, or disseminating fake news, the offender does not only harm the directly affected party but also weakens the entire legal system based on the credibility of declarations (Wood, 2008, p. 243). For this reason, the conditions of declaration, untruthfulness, and intentionality are crucial for the configuration of criminal or civil infractions, as well as for the proper delimitation of corresponding legal responsibility.

Based on Green's analysis of perjury, for example, it is possible to highlight the importance of distinctions grounded in the criteria present in Kant's conception of lying. As Green emphasizes, the legal treatment of deception varies depending on the social role of the agent—whether an ordinary citizen or a public authority—and on the specific context in which the deception occurs, such as judicial hearings, commercial transactions, police stations, or situations of sexual intimacy (Green, 2019, p. 483). The social role of the individual who deceives, as well as the context in which the deception occurs, reinforce the significance of the declaration condition. For example, a public authority, in the exercise of their duties, is committed to adhering to the procedures established by the norms regulating their civil competence. Moreover, the context in which the declaration is made supports or delineates its declarative character, determining the specific requirements applicable to each situation.

The legal relevance of the untruthfulness condition in lying is particularly evident in cases of perjury. According to Green, statutes concerning perjury have traditionally been interpreted as requiring that the declaration be objectively verifiable with respect to its untruthfulness or falsehood. Declarations based on beliefs or opinions do not constitute perjury, except in exceptional cases where the witness claims to hold a belief or opinion that they, in fact, do not possess—this existence or nonexistence of the belief being a fact susceptible to verification.

The author further highlights the complexity of the evidentiary requirement regarding whether it is necessary to demonstrate that the statement was literally false, or if a "merely misleading" declaration, even if literally true, would suffice to support a conviction. As Green explains, the rule currently followed in most common law jurisdictions requires that, for a witness to be convicted of perjury, they must actually lie under oath, and not merely mislead (Green, 2019, p. 485). In other words, just as in Kant's definition of lying, the distinction between the untruthfulness of the declaration and factual falsity, along with the intention to deceive, are fundamental elements for delimiting the legal offense. Thus, it becomes essential to determine whether there was a declaration contrary to what the speaker believes, characterizing a lie, or whether the deception occurred through other modalities, such as in the case of paltering.

Green illustrates this distinction by revisiting the case of Bill Clinton, cited in Sandel's critique of Kant's definition of lying (2012, p. 134). Despite being charged with perjury by the Independent Counsel, Green contends that Clinton's response about being alone in the Oval Office with the White House intern does not amount to perjury. Green would likely agree with Sandel on this point, stating that "Clinton offered an evasive, non-responsive, and factually true reply to the question posed; but he did not actually lie" (Green, 2019, p. 486). Although Clinton intended to deceive, Green argues that his conviction could not be grounded in the legal definition of perjury: the legal context of Clinton's testimony certainly imposed the declarative condition, and there was intent to deceive; however, by providing an evasive yet factually true response, no falsification of the declaration occurred. It should be noted, however, that neither Green nor Sandel deny Clinton's intent to mislead the Court.

The intentionality condition inherent in Kant's concept of lying is fundamental for analyzing the responsibility for untruthfulness. Kant reaffirms the formal injustice of lying in his treatment of false reports, stating that "every lie is objectionable and deserving of contempt, for once we declare that we are telling the other our thoughts, and fail to do it, we have broken the pactum, and acted contrary to the right of humanity" (V-Mo/Collins, 27: 448 | Kant, 1997, p. 203–204). The broken pact is described by Kant as the second condition of sociability: "but the liar destroys this fellowship, and hence we despise a liar, since the lie makes it impossible for people to derive any benefit from what he has to say" (V-Mo/Collins, 27: 444 | Kant, 1997, p. 201). Truthfulness, as a condition of sociability, requires correspondence between the intention declared in discourse and the agent's actual intention. Hence, there is an intrinsic tension between lying and truthfulness: the former presupposes the intention to deceive, whereas the latter explicitly rejects it. Although Kant repeatedly emphasizes the immorality of lying in various texts, the Doctrine of Right also provides sanctions for other forms of deception, since any falsification of the declared intention violates the legal duty of truthfulness. What makes lying—and, in legal terms, perjury—particularly serious is the deliberate declaration of the opposite of what one believes, directly concealing one's true intention in contexts where there is an explicit duty of truthfulness.

Kant characterizes lying as a formal injustice against the right of humanity also in the *Metaphysics*, denying that false reports constitute a right even in the case of war between states, since such behavior "would make its subjects unfit to be citizens [...] in a word, using such underhanded means as would destroy the trust requisite to establishing a lasting peace in the future" (MS, AA 06: 348 | Kant, 1996, p. 117).

The first relevant point regarding this passage concerns the dissemination of false reports as a harm to the free will of others, such that the falsity of these declarations results in damage both epistemologically and politically. Kant emphasizes the epistemological aspect in the *Lectures on Ethics* when he states that "[...] since I cannot experience everything myself, if the reports of others were to be false tidings, the desire for knowledge could not be satisfied" (V-Mo/Collins, 27: 448 | Kant, 1997, p. 203–204). For Kant, although it is not essential that the ideas expressed be true (since error is possible), the lack of truthfulness would undermine the credibility of discourse, reducing it to a mere collection of assertions without commitment on the part of their authors. This situation would open the door to the unrestricted propagation of unfounded information, thereby weakening standards of responsibility, including within the legal sphere. Therefore, Kant maintains that assuming responsibility for one's own declarations is fundamental for the legitimate exercise of public reason.

Kant further develops the political harms of false reports in *Toward Perpetual Peace*. In addition to establishing falsehood as a moral limit to political prudence (ZeF, AA 08: 370 | Kant, 1917, p. 162), Kant rejects all peace treaties concluded with *reservatio mentalis*, that is, with mental reservations (ZeF, AA 08: 344 | Kant, 2006, p. 68). This prohibition is grounded in the moral imperative to avoid any form of deception in declarations of intent, especially those that concern the end of war. In the philosophical tradition, *reservatio mentalis* denotes a form of dissimulation in which a speaker limits the meaning of their words internally while outwardly expressing them without any restriction. It is a deceptive ambiguity: the speaker utters a declaration that appears complete and truthful, yet secretly adds an internal condition that alters its actual meaning.

For Kant, this kind of dissimulation is morally impermissible in juridical contexts such as treaties because it violates the condition of public truthfulness. Even if it does not involve an explicit lie—since the speaker does not affirm something they believe to be false—it still aims to mislead the other party. As such, *reservatio mentalis* constitutes a species of falsehood that undermines the trust necessary for rightful relations between states. In the context of peace treaties, it conceals the true intention not to maintain peace, thereby nullifying the validity of the agreement and transforming it into a covert instrument of war. Thus, Kant considers treaties made with mental reservation not only illegitimate but also morally equivalent to a falsification, because the communicative act pretends to bind the parties while inwardly intending the opposite.

As Wood argues, Kant was attentive to the political implications of permitting lying: "the issue that appears to have really concerned both Kant and Constant is the duty of politicians and statesmen to be truthful in their official declarations" (Wood, 2008, p. 249)[6].

Kant synthesizes in his response to Constant the arguments against political legitimacy for lying that he had previously developed in *Toward Perpetual Peace*. In this work, Kant presents the principles "Act and justify yourself" (*Fac et excusa*), "If you did it, deny it" (*Si fecisti nega*), and finally "Divide and rule" (*Divide et impera*) (ZeF, AA 08: 374–375 | Kant, 2006, p. 98). Kant describes that under the principle "Act and justify yourself", committed violence is falsely justified after the fact, generating lies regarding the legitimacy of the state's possession. Consequently, Kant describes that under the principle "If you did it, deny it", the politician lies about the mistakes committed, fabricating falsehoods to attribute the errors and illegalities to his subjects. Lastly, Kant revisits the famous principle "Divide and rule", portraying it as a political simulation aimed at fomenting discord among the people in order to deceive them with promises of greater freedom. Kant concludes this passage by stating that "it is true that nowadays nobody is taken in by these political maxims, for they are all familiar to everyone" (ZeF, AA 08: 376 | Kant, 2006, p. 172). In commenting on this passage, Kant criticizes the normalization of political cynicism among great powers, emphasizing that these unjust maxims—such as lying and manipulation—are no longer a source of scandal, since their public acknowledgment causes no shame. What discredits rulers is not the immorality of these principles, but their failure in practice, given that political prestige is measured by the expansion of power rather than by moral legitimacy. This position directly contrasts with Kant's own view that politics must remain subordinate to morality. As he argues in Toward Perpetual Peace, the success of political action can never justify the adoption of immoral maxims, for political conduct must be guided by universal principles of justice and respect for human dignity.

5. Conclusion

This article set out to investigate the complex conceptual issue of fake news through Kant's philosophical framework, aiming to overcome the ambiguities and disputes that permeate contemporary debates on information disorders. Reaffirming the central thesis, the criteria for defining lying in Kant—namely, declaration, untruthfulness, and intentionality—prove essential to deepen the contemporary discussion on information disorders.

---

[6] For a reconstruction of the dispute between Kant and Constant, as well as a proposed resolution of the moral casuistry, see Klein (2018).

The analytical function of the declaration condition is to define what is formally subject to verification of truth or veracity. That is, it delimits cases in which there is a clear and morally relevant propositional content. The condition of untruthfulness, in turn, allows distinguishing between factual falsity and the untruthfulness of the declaration—that is, between objectively false information and the absence of commitment to the truth. The condition of intentionality, like untruthfulness, enables the differentiation of lying proper from other types of false declarations that do not involve intent, such as errors or unintentional misinformation. This condition is fundamental, for example, in distinguishing disinformation (false and intentional) from misinformation (false, but not necessarily intentional).

In their articulation, these criteria provide a robust and precise conceptual framework that enables a rigorous analysis of the multiple facets of information disorders, distinguishing them clearly and offering solid grounds for ethical and legal analyses. Kant's work presents crucial criteria not only for the definition of information disorders but also addresses these issues within the theory of law, the legal system, and political contexts. Thus, Kant's definition of lying decisively contributes to a critical understanding of fake news, broadening and deepening reflections on the moral and social challenges inherent in information disorders.

The articulation between the criteria for defining fake news and the specificity of information disorders for its analysis within the Doctrine of Right is crucial. This demonstrates that lying, as a violation of the right of humanity and the social pact, goes beyond mere isolated falsehoods or traditional media formats. It constitutes a transgression of the epistemological and political foundations of social coexistence and the legal order. Kant distinguishes between lying that violates individual rights—typical of private law—and lying that affects the right of humanity, undermining social trust and the foundation of the civil pact in public law. Thus, Kantian conditions for the definition of lying allow for capturing decisive elements in the contemporary dynamics of fake news: the relevance of context and the imposed demands, delineated by the condition of declaration; as well as the untruthfulness and intentionality of the liar in declaring the opposite of what they intend. This approach broadens the analytical scope to encompass the multiple channels through which fake news spreads—from social media and digital applications to traditional media—and incorporates the diversity of political and interpersonal environments in which it manifests. In this regard, Kantian reflection not only offers conceptual clarity but also highlights the profoundly harmful nature of fake news. Consequently, Kantian thought proves to be an indispensable epistemological and normative tool for the critical understanding and confrontation of information disorders.

Bibliographic References

ALLCOTT, H.; GENTZKOW, M. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, v. 31, n. 2, p. 211-236. DOI: 10.1257/jep.31.2.211

BARBOZA, K. C. S; KLEIN, J. T. 2024. O conceito de mentira segundo Kant. *Revista Instante*, v. 6, n. 3, p. 42–174. DOI: 10.29327/2194248.6.3-7

CANADIAN CENTRE FOR CYBER SECURITY. 2024. *How to identify misinformation, disinformation, and malinformation*. Awareness series. Available at: https://www.cyber.gc.ca/en/guidance/how-identify-misinformation-disinformation-and-malinformation-itsap00300#wb-tphp. Accessed in July 2025.

FRIED, C. 1978. *Right and wrong*. Oxford: Harvard University Press.

GELFERT, A. 2018. Fake news: a definition. *Informal Logic*, v. 38, n. 1, p. 84-117. DOI: 10.22329/il.v38i1.5068

GREEN, S. P. 2019. Lying and the law. In: MEIBAUER, J. (Ed.). *The Oxford handbook of lying*. Oxford: Oxford University Press, p. 483-493.

HIMMA-KADAKAS, M. 2017. Alternative facts and fake news entering journalistic content production cycle. *Cosmopolitan Civil Societies*, p. 25-40. DOI: 10.5130/ccs.v9i2.5469

KANT, I. 1900. *Gesammelte Schriften v. 1-24*. Berlin: Preussische Akademie der Wissenschaften; Deutsche Akademie der Wissenschaften zu Berlin (v. 23); Akademie der Wissenschaften zu Göttingen (v. 24).

KANT, I. 1993. *Grounding for the metaphysics of morals with on a supposed right to lie because of philanthropic concern*. Translated by James Ellington. Indianapolis: Hackett Publishing Company.

KANT, I. 1998a. *Groundwork of the metaphysics of morals*. Translated by Mary Gregor. Cambridge: Cambridge University Press.

KANT, I. 1997. *Lectures on ethics*. Translated by Peter Heath. Cambridge: Cambridge University Press.

KANT, I. 2018. *Lições de ética*. Traslated by Bruno Leonardo Cunha and Charles Feldhaus. São Paulo: Editora Unesp Digital.

KANT, I. 1917. *Perpetual peace: a philosophical essay*. Translated by M. Campbell Smith. London: George Allen & Unwin Ltd.

KANT, I. 1998b. *Religion within the boundaries of mere reason and other writings*. Translated by Allen Wood and George Di Giovanni. Cambridge: Cambridge University Press.

KANT, I. 1996. *The metaphysics of morals*. Translated by Mary Gregor. Cambridge: Cambridge University Press.

KANT, I. 2006. *Toward perpetual peace: a philosophical sketch*. Translated by David Colclasure. New Haven: Yale University Press.

KLEIN, J. T. 2018. Kant versus Constant: sobre um suposto direito de mentir. *Studia Kantiana*, v. 16, p. 95-126. DOI: 10.5380/sk.v16i3.89814

LAZER, D. M. J. *et al*. 2018. The science of fake news: addressing fake news requires a multidisciplinary effort. *Science*, v. 359, n. 6380, p. 1094-1096. DOI: 10.1126/science.aao2998

LESHER, M.; PAWELEC, H.; DESAR, A. 2022. Disentangling untruths online: creators, spreaders and how to stop them. *OECD Going Digital Toolkit Notes*, n. 23, p. 1-37.

MAHON, J. 2015. *The definition of lying and deception*. Stanford Encyclopedia of Philosophy. Available at: https://plato.stanford.edu/archives/fall2015/entries/lying-definition/. Accessed in July 2025.

NIELSEN, R. K.; GRAVES, L. 2017. *"News you don't believe"*: audience perspectives on fake news. Oxford: Reuters Institute for the Study of Journalism. Available at: https://reutersinstitute.politics.ox.ac.uk/our-research/news-you-dont-believe-audience-perspectives-fake-news. Accessed in July 2025.

POSETTI, J.; MATTHEWS, A. 2018. *A short guide to the history of 'fake news' and disinformation*: a learning module for journalists and journalism educators. ICFJ. Available at: https://www.icfj.org/news/short-guide-history-fake-news-and-disinformation-new-icfj-learning-module. Accessed in July 2025.

QUANDT, T. *et al*. 2019. Fake news. In: VOS, T. P. *et al*. (ed.). *The international encyclopedia of journalism studies*. Hoboken: John Wiley & Sons.

SANDEL, M. 2012. *Justice*: what's the right thing to do? New York: Farrar, Straus and Giroux.

SAUL, J. M. 2000. Did Clinton say something false? *Analysis*, v. 60, n. 3, p. 255-257. DOI: 10.1111/1467-8284.00235

SCHAUER, F.; ZECKHAUSER, R. 2009. Paltering. In: HARRINGTON, B. (Ed.). *Deception*: from ancient empires to internet dating. Stanford: Stanford University Press, p. 38-54.

WOOD, A. 2008. *Kantian ethics*. Cambridge: Cambridge University Press.

ZIMDARS, M.; McLEOD, K. 2020. *Fake news*: understanding media and misinformation in the digital age. Cambridge: MIT Press.