Chasing spammers: Using the Internet Protocol address for detection.

Abstract

The proliferation of reviews evaluating different services on social networks and online platforms and their importance in consumer decision-making has led some unscrupulous individuals to take advantage of the anonymity offered by the Internet to manipulate these reviews and influence customers' decisions. The main objectives of this study are: (1) to test whether spammers usually perform their misdemeanours from the same IP address; (2) to explore whether there are differences between stated sexes in this regard; (3) to detect the main motivations for posting fraudulent reviews; and (4) to determine the motivations for doing so from the same IP address. These objectives were achieved by means of a quasi-experiment with a sample of 7,192,487 users, and a qualitative investigation in which 37 users who had falsified information were interviewed. The results show that spammers who tend to fake their identity do so from the same IP address, and that they tend to be male. Four types of motivation are presented: revenge, entertainment, opportunity for profit, and self-esteem; as well as a further three to explain the use of the same IP: convenience, limited resources, and complacency.

Keywords: Reviews, Spammers, IP address, Fraudulent, Motivations

Type of article Research article

1 | INTRODUCTION

The popularity of the Internet and social media continues to grow, with an increasing number of people using them as a source of information in the purchasing process (Sáez-Ortuño et al., 2023c; Shu et al., 2017). In fact, the ease with which news and opinions about products or services can be searched, consumed, or shared on social media means online channels have surpassed traditional sources of information (Shu et al., 2017). This digitalised social environment has created a parallel marketplace, triggering the emergence of e-commerce and digital marketing (Kumar et al., 2016).

The digital marketplace, thanks to disintermediation, has shortened the distance between producers and customers, reduced time-to-market and increased the distribution of one-to-one communication (De Ruyter et al., 2018). At the same time, social networks enable consumers to connect with family, friends, acquaintances, and other consumers, for whom they are collecting, creating, and distributing increasingly visual, selfie-centric content, generating an unprecedentedly large flow of recorded information (Ludwig & de Ruyter 2016). When the content of the information being exchanged is related to the purchase, use, assessment of products or their sellers, this is called electronic word of mouth (eWOM) (Litvin et al., 2008). Although eWOM can consist of many types of comments about and ratings of products, the most important are online reviews (Kim et al., 2016).

Reviews are customers' assessments or opinions regarding their experiences with products (goods and services) and are published on online platforms as a testimony for future consumers (Chatterjee et al., 2021). These assessments are important for both consumers and digital managers alike (De Ruyter et al., 2018). Online consumers use them during the pre-purchase phase to form an idea about product features before deciding whether to buy. In fact, they are

the second most important source of information after recommendations from family and friends (Salehan & Kim, 2016; Septianto & Garg, 2021). Since reviews are testimonies of other customers' experiences, and are independent from marketing communication, they convey credibility and significantly influence new customers' decision-making (Wu et al., 2022). For example, there is evidence that 80% of consumers change their judgement and purchase decision after reading negative reviews, and 87% after reading positive ones (Zhang et al., 2016). Digital marketing managers, in awareness of the importance of reviews for changing consumer attitudes, need to know that they affect the evolution of sales (Moon et al., 2021, Saboo et al., 2016), brand and company image (Rambocas & Pacheco, 2018; Zaman et al., 2023) and even stock prices (Tirunillai & Tellis, 2012). In other words, they need to track the evolution of feedback, as a way of listening to their customers' voices, to thus modulate marketing policies (De Ruyter et al., 2018).

However, the fact that not all the published information is true calls this communication channel into question. Some reviewers, taking advantage of the anonymity provided by the online environment, engage in writing false comments or testimonies (Allcott & Gentzkow, 2017; Bonald et al., 2009; Vosoughi et al., 2018; Wu et al., 2022). Testimonies are considered false when they express an insincere assessment of products and companies, either because the reviewer has no real experience of them (Zhang et al., 2016) or because they deliberately seek to boost or tarnish the image of a company and its products (Daiv et al., 2020). In both cases, whether unintentional (misinformation) or malicious (disinformation), this distorts the veracity of the channel itself. The proliferation of fake news affects consumers, businesses, and the market in general. On the one hand, consumers may decide to choose the products of one company over another based on reviews by previous customers, and feel disappointed when their expectations are not met. On the other hand, online managers may make incorrect decisions based on false reviews. In short, false reviews distort the rules of the competitive

online market (Leonidou et al., 2021; Sáez-Ortuño et al., 2023b; Zaman et al., 2023). Therefore, it is important to develop mechanisms to facilitate their detection and neutralisation. However, it is also important to understand the motivations that lead these reviewers to engage in these inappropriate behaviours.

Echoing the concerns of economic agents, academia has sought to address this challenge by developing two main lines of action. On the one hand, studies focused on the detection of false reviews, for example by looking for patterns in the way they are written or in the type of sources they use (Di Domenico et al., 2020; Moon et al., 2021). On the other hand, studies that have focused on detecting suspected spammers (Mukherjee et al., 2013), either by examining demographic profiles (Sáez-Ortuño et al., 2023b), understanding their motivations (Zaman et al., 2023), or by tracing the sources from which they post (Waggoner et al., 2019). However, although it is easier to spot spammers than to detect fake comments, there is a proliferation of studies that analyse review content to detect fakes (Kim et al., 2016; Moon et al., 2021; Ott et al., 2013), while the literature on the detection of spammers is much scarcer (Sáez-Ortuño et al., 2023c). For example, in one study using a sample of more than seven million Spanish users (21.5% of Spanish internet users according to INE (2023)), spammers amount to just 5.86% of the total, in line with previous studies (Islam et al., 2020; Pennycook & Rand, 2019). Therefore, there are around 1.9 million potential spammers to be detected in Spain, while if Heydari et al. (2015) are correct in claiming that 70% of spammers write more than five reviews per day, then there are around 5.6 million fake reviews per day to be detected.

This study aims to extend the latter line of research by characterising spammers to facilitate their location. It also attempts to understand their motivations for what could be considered unethical behaviour. Although IP addresses have previously been used as a market segmentation factor (Louvieris & Driver, (2001), they have not been used to detect spammers in marketing despite their use in computer science (Rao et al., 2021). Specifically, this research

proposes to transfer the findings of Waggoner et al. (2019), who detected, in a study of users who write fake answers to online surveys, a major match between fake answers and duplicate IP addresses. In our case, we wish to test whether spammers behave in a similar manner and, consequently, that the detection of duplicate IP addresses could be used as a tool to locate fraudsters.

However, as Moon et al. (2021) argue, it is not easy to determine whether reviews are fake, as there is no method that can irrefutably discern their veracity and so they are classified on the basis of assumptions and criteria established by researchers and, consequently, the results are far from conclusive (Salminen et al., 2022). In this study, our estimator of the inclination to post false content is the introduction of misrepresentative identification data when registering on online platforms.

Based on the above arguments, the following research questions (RQ) are proposed:

RQ1. Can the fact that several users register their information or write several reviews from the same IP address within a short period of time be used as a criterion to suspect spammers?

RQ2. Is there a different predisposition between reported sexes to post false data and reviews?

In addition, the second objective of this study is to find out the motivations that lead spammers to post misleading reviews and to do so from the same address. Again, although there is some previous research, the literature on this topic is very scarce (Sáez-Ortuño et al., 2023b). Regarding the motives for publishing false information or reviews, the work by Zaman et al., 2022 and Saez-Ortuño et al., 2023 is noteworthy. As for publishing from the same IP address, there are also some exceptions such as the work by Keusch et al. (2019), which found that some users use pseudonyms either to hide their own identity or to impersonate someone else's, and tend to have jocular (joking) or criminal (harassment) motivations, although they use the same

IP address (Keusch et al., 2019). Hence, despite the risk of being identified, most people who use social networks for inappropriate behaviour tend to use the same IP address, rather than taking steps to mask their identity and location (Waggoner et al., 2019). An understanding of the motivations behind this seemingly absurd behaviour may provide useful information for policymakers, marketers, brand managers, retailers, and academics to help them address the growing threat of this malicious phenomenon. Based on the evidence gathered, this study poses the following research question:

RQ3. What are the main motivations for falsifying information or posting fake reviews?
RQ4. What are the main motivations of users who falsify information or post fake reviews in a short period of time to keep the same IP address?

The paper is organized into three main chapters, each serving a distinct purpose. In Chapter Two, we present the theoretical framework, providing a comprehensive definition of online deception and exploring the utilization of IPs for fraud detection. Motivation theory is also explained. Chapter Three outlines the various studies conducted to date in relation to the research questions posed earlier. The paper concludes with a summary of the findings in the form of conclusions.

2 | THEORETICAL FRAMEWORK

2.1. | Research Context

2.1.1 | Online deception: Categorizing False Content and Detecting Deceivers

The proliferation of consumer reviews and assessments of actual experiences around the world has become the workhorse of digital marketing (Kannan & Li, 2017; Wessel et al., 2016). Consumers, with just a few clicks, can access a record of opinions about past consumption experiences and based on that information, decide to place an order or switch to other products (Moon et al., 2021). In fact, evidence has been gathered that feedback on clothes purchase

experiences has a higher impact on new consumers than other traditional marketing tactics (Goh et al., 2013; Moon et al., 2021).

However, e-commerce managers are concerned about the proliferation of fake reviews that alter the validity of genuine ones (Daiv et al., 2020). According to TripAdvisor's Review Transparency Report (2021), in 2020, more than two million of the 26 million reviews on its platform were rejected and removed, including almost one million (3.6%) that were labelled as fake (TripAdvisor, 2021). These often-biased fake reviews can generate misleading expectations about the quality of products and services and lead to incorrect purchasing decisions (Moon et al., 2021). In some cases, false positive reviews are used to artificially inflate the belief among new consumers that previous buyers were satisfied, while in other cases negative testimonies are designed to do the opposite by undermining new customers' confidence and encouraging them to turn to competitors instead (Savage et al., 2015). They may even challenge the platforms on which the reviews are posted and, in the long run, this can cause widespread damage to e-commerce (Daiv et al., 2020), the market and society as a whole (Birim et al., 2022).

Given the pernicious effects of the proliferation of fake reviews, researchers and practitioners alike have gone to enormous efforts to detect them and the spammers who post them (Mukherjee et al., 2012; Mohawesh et al., 2021; Birim et al., 2022). However, given the large amount of online content and its rapid growth rate, it is impractical to monitor it manually, hence the need to devise automated detection mechanisms (Conroy et al., 2015), including the design of algorithms for the classification of information once it has been collected (Conroy et al., 2015; Parikh & Atrey, 2018; Shu et al., 2017) for the detection of sources and authors (Bondielli & Marcelloni, 2019), as well as more holistic approaches that consider both fake reviews and their creators (Zubiaga et al., 2018). All these methods have their advantages and

disadvantages, as well as their limitations in detecting online fraudsters (Sáez-Ortuño et al., 2023b). Table 1 presents a summary of the studies considered.

AI algorithms are often used for the analysis and detection of fake reviews. These can detect recurring patterns of behaviour that the e-commerce analyst, however experienced, cannot (Li et al., 2017; Mohawesh et al., 2021). Patterns extracted through machine learning include text length, imbalance between positive and negative reviews, and the relatively short time in which reviews are written. Regarding the first of these issues, while 75% of fake reviews have less than 136 words, 90% of genuine reviews have more than 200 (Noekhah et al., 2014). Second, 85% of fake reviews are positive (Crawford et al., 2015). And regarding the excessively short time taken to write fake reviews (Alsubari et al., 2022), it has been found that around 70% of spammers wrote more than five reviews per day, while 90% of normal users usually only write a single review when they purchase a product or service (Heydari et al., 2015).

While these approaches based on the analysis of reviews have been used successfully, their major drawback is that the comparison of such a large amount of information resulting from the proliferation of fake news and the need for so much processing time can quickly make them unfeasible (Akoglu et al., 2013; Heydari et al., 2015). To make the process more affordable and efficient, the scope of study needs to be reduced by selecting suspect candidates (Luca & Zervas, 2016) or focusing on more specific product domains (Akoglu et al. 2013; Moon et al., 2021).

The second line of research aims to detect, study, and categorise spammers and understand the reasons for their behaviour. This is commonly done by comparing their behaviour with that of regular reviewers. For example, Savage et al. (2015), in a context of rating with 1 to 5 stars (i.e., without considering the written comment) proposed that spammers could be detected by analysing the deviation of their ratings from the majority opinion, as estimated by binomial regression. One of the first attempts to categorise spammers was the study by Luca and Zervas

(2016), who focused on restaurant ratings published on the Yelp platform. Even though the platform itself filters out approximately 16% of reviews, they found that the number of false favourable reviews was higher for restaurants that only had a few reviews or had recently received negative ones. They also ranked establishments and found that single restaurants were more likely to commit review fraud than chain restaurants (Luca & Zervas, 2016).

Another attempt to classify spammers was that by Akoglu et al. (2013), who established a correlation between fraudulent consumers, the type of product they evaluated and the type of ratings they always gave. They found that fraudsters are inclined to rate the same products (e.g., restaurants) and to rate good ones as bad or vice versa, suggesting a certain specialisation by product type and the systematic use of the same rating (Akoglu et al., 2013).

Spammers also tend to concentrate their fake reviews in a short period of time (Alsubari et al., 2022), which provides another criterion for their delimitation. However, categorisation of fake reviews by their writers' addresses is a less common practice (Waggoner et al., 2019) and could be another tool for their detection and delimitation.

Since the early 21st century, a digital ecosystem has been taking shape in which the exchange of data has become a generator of economic value (Sestino et al., 2023). This perception of value has led to the proliferation of misleading behaviour in the form of misrepresentation of opinions, news and advertisements on websites and social networks (Mintz, 2002; Wendling, 2018). Moreover, this phenomenon is transcending into all areas of business and society (Allcott & Gentzkow, 2017; Da Fonseca & Borges-Tiago, 2021).

Whereas in ancient societies, formed by small groups, there was a high risk associated to deceptive behaviour, because its discovery often led to expulsion or ostracism, in modern societies, particularly in large cities, fraudsters exploit their anonymity to go unpunished (Cosmides & Tooby, 2016). Similarly, online cheaters take advantage of the anonymity

provided by the digital environment to act improperly under the perception that they will not be caught (Allcott & Gentzkow, 2017).

Although writing and distributing false information is considered a form of deception (Lwin et al., 2016), the literature usually distinguishes between that which is generated unconsciously or unintentionally (misinformation) and that which is generated maliciously (disinformation) (Sáez-Ortuño et al., 2023b; Zubiaga et al. 2018). Undoubtedly, it is disinformation that has generated the greatest deal of academic attention (Shu et al., 2017). For example, Wu et al. (2021), in line with Bagozzi (1992), consider that consumers who generate disinformation do so because they are pursuing some kind of economic or psychological goal and, furthermore, if that goal is achieved, this will satisfy their psychological needs.

Table 1. Summary of studies considered

Reference	**Focus of Study**	**Key Findings/Contributions**
Bonald et al., 2009	Anonymity and dishonesty online	Discussed how anonymity on the internet can lead to dishonest actions by users.
Litvin et al., 2008	Nature of electronic word of mouth (eWOM)	Identified different types of eWOM, emphasizing the importance of online reviews.
Mukherjee et al., 2013	Detection of false reviewers (spammers)	Investigated methods for detecting spammers by analysing their profiles and behaviours.
Kim et al., 2016	Importance of online reviews	Highlighted the significance of online reviews in customers' assessments of products and services.

Reference	Focus of Study	Key Findings/Contributions
Kumar et al., 2016	E-commerce and digital marketing emergence	Explored the impact of digitalized social environments on marketplaces and marketing strategies.
Ludwig & de Ruyter 2016	Social networks and content generation	Analysed the massive flow of information generated by users on social networks.
Saboo et al., 2016	Influence of reviews on brand image	Explored how reviews contribute to brand and company image configuration.
Salehan & Kim, 2016	Source of product information	Noted that online reviews are a critical source of product information for consumers.
Shu et al., 2017	Online vs. traditional information sources	Found that online channels have surpassed traditional sources in influencing consumer decisions.
Allcott & Gentzkow, 2017	Nature of online misinformation and disinformation	Examined the dishonest behaviours of reviewers and the false information they spread.
Zhang et al., 2016	Effects of reviews on judgment and purchases	Provided evidence of the extent to which reviews can change consumer opinions and purchasing behaviour.
Chaffey & Ellis-Chadwick, 2019	Definition of digital marketing	Provided a comprehensive definition of digital marketing and its objectives.

Reference	Focus of Study	Key Findings/Contributions
De Ruyter et al., 2018	Marketing communication in the digital era	Examined how the digital marketplace has changed consumer-producer communication.
Vosoughi et al., 2018	Spread of false information	Studied the way false information is disseminated online.
Waggoner et al., 2019	Tracing the source of false reviews	Studied techniques for tracing the origin of spammers' activities.
Daiv et al., 2020	False reviews to manipulate company image	Analysed how false reviews are used to deliberately alter the perception of a company.
Di Domenico et al., 2020	Classifying true vs. false reviews	Focused on identifying patterns to distinguish between authentic and fake reviews.
Leonidou et al., 2021	Impact of fake reviews on market competition	Addressed how fake reviews can bias market competition.
Moon et al., 2021	Impact of reviews on product sales	Demonstrated the correlation between online reviews and product sales.
Septianto & Garg, 2021	Influence of reviews on decision-making	Confirmed that reviews significantly affect consumers' purchase decisions.
Chatterjee et al.,	Role of online reviews	Discussed the value of reviews for consumers

Reference	Focus of Study	Key Findings/Contributions
2021		and marketing managers.
Wu et al., 2022	Impact of reviews on consumer behaviour	Demonstrated the credibility and influence of online reviews with regard to consumer decision-making.
Sáez-Ortuño et al., 2023c	Influence of online information on purchasing	Discussed the increasing role of online channels in consumer decision-making processes.
Sáez-Ortuño et al., 2023d	Challenges in online marketing research	Discussed the difficulties in the design of instruments for analysing heterogeneous online information.
The present research	Detection of spamming activities	Found that spammers tend to conduct their activities from the same IP address and are predominantly male. Identified motivations for posting fraudulent reviews and using the same IP address for such activities.

2.1.2 | Using IP addresses to detect spammers

The use of address tracing to detect the sources of fraudulent traffic (e.g., fake clicks) comes primarily from the literature on online advertising fraud (Wu et al., 2021) and aims to classify posters as legitimate or fraudulent by observing the information flows that they generate (Hu et al., 2017; Oentaryo et al., 2014).

Cookies (Wang et al., 2017) or IP addresses (Zhang & Guan, 2008) are often used to locate and identify malicious users. For example, Wang et al. (2017) used cookies because of their relative stability and veracity, as they are specifically designed for web browsers to track, personalise, and store information about each user's session. Cookies mean that both benign and malicious users can be tracked on both private and public devices (Wang et al., 2017). In another study on advertising on mobile apps, Hu et al. (2017) proposed an algorithm based on bipartite graphs that matches users (identified by cookies) with app publishers, which can be used to detect people whose abnormal behaviour raises suspicions of click fraud.

There has been very little research on the use of IP addresses to identify users. Exceptions are the studies by Zhang and Guan (2008) and Wu et al. (2021). The former used IP addresses in an online advertising context to identify users and found that publishers of fraudulent content (numerous duplicate clicks) tend to focus their activity on a single ad for a short period of time from a single IP address. Meanwhile, Wu et al. (2021) used an algorithm and a bipartite graph that relates users (identified by IP address) to online advertising campaigns. These graphs detect IPs that generate fraudulent behaviour patterns, to which a random forest algorithm is applied to estimate their probability of defrauding. In our study, we use IP addresses to identify users suspected of defraud.

IP is a fundamental protocol in Internet communication as it provides the necessary addressing scheme and routing mechanisms for devices to connect and exchange information within the network (Postel, 1981). Each IP is unique to each device connected to a network and ensures that data is properly routed and delivered to its intended destination, regardless of the underlying network infrastructure (Kent & Seo, 2005).

However, there are still gaps in the address detection literature. Since the assignment of falsehood to an address is based on behavioural patterns (e.g., multiple reviews in a short period of time and from the same IP address), all fraudsters need to do to avoid detection is to replicate

the behaviour of legitimate users, which makes pattern-based methods increasingly inadequate (Wang et al., 2017) as fraudulent posters can go undetected simply by obtaining multiple IP addresses and spacing out their reviews.

2.2 | Motivation Theory

Motivation is one of the most important topics in psychology and the study of consumer behaviour. In psychology, the structural basis of motivation was established in the 1950s when Hull (1952) introduced the concept of drive and homeostasis to describe the most basic forms of instinctive or animal motivation. Drive is an internal animal reaction to a state of restlessness due to an imbalance in the satisfaction of psychological needs, while homeostasis is the state of equilibrium that the drive (the motivator) seeks to recover by reducing tension. Subsequently, Rotter (1954) highlighted the importance of expectations for shaping motivations, suggesting that the likelihood of an individual engaging in a certain behaviour depends on their expectations of achieving some specific goal, and the personal value that he or she assigns to that goal. Moreover, a discrepancy between expectations and reality may motivate an individual to react through corrective behaviours (Festinger, 1957). In turn, Heider (1958) proposed a classification of expectations based on the claim that they are related to motivational forces of different origins: internal or external. He called the internal ones dispositional forces (e.g., the individual makes little effort or has insufficient intelligence), and the external ones situational forces (e.g., the environment in which the activity is performed is unfair or biased). On the way in which both motivations act on behaviour, Atkinson (1964) proposes the additive formula. That is, both motivations must be added together to make up the total motivation. However, Deci (1971) challenged these assumptions and, through a series of experiments, showed that both motivations can act in both synergistic and delimiting ways depending on their type. For example, he showed that if an intrinsic motivation to perform a task is reinforced by extrinsic monetary compensation, and is then removed, this reduces intrinsic motivation, but if verbal

rewards (i.e. positive feedback) were used instead, intrinsic motivation was not undermined (Deci & Ryan, 2000). That is, the type of extrinsic motivation can help or hinder the intrinsic motivation to perform an act. All these lessons from the classics of psychology were adapted by Bagozzi (1992) and other consumer behaviour researchers to, for example, describe the drive towards a commercial objective. Thus, individuals often engage in activities (e.g., writing false reviews, false identities, etc.) with the aim of achieving a certain outcome, which he called the "outcome-desire unit" (Bagozzi, 1992). He also considered the role of expectations in commercial exchanges, whereby individuals who have completed the activity perceive that they have achieved their goal and hence desire-outcome fulfilment occurs, which is followed by an affective response (e.g., satisfaction) (Bagozzi, 1992). In line with Appraisal Theory, Lazarus (1991) proposed that appraisal (evaluation of the fulfilment of expectations) is followed by an emotional response that leads to coping. That is, emotional responses lead to modifications or persistence of behaviour with respect to the activity. The classification of motivations has also been adapted to the commercial world, such as when Davis et al. (1992) considered the different roles of intrinsic and extrinsic motivations in the inclination of workers to use computers in the workplace. However, it is also necessary to take into account the hypotheses proposed by self-determination theory. This considers that the drivers of behaviour are not only determined by the strength or quantity of motivation, but also by the type or quality thereof, which can act in a synergistic or undermining manner (Deci & Ryan, 2000).

However, when studying any activity or behaviour, and the motivators that explain it, a balance must be struck between explanatory value and the number of motivators or variables. For example, an exploratory study into the falsification of data and reviews could find an almost unlimited number of motivators. However, if an explanatory model is constructed to explain falsification behaviour and motivations are added indiscriminately, this model becomes less parsimonious and operable (Bagozzi, 1992). In other words, the researcher must select the

minimum number of motivators that generate the greatest explanatory capacity for the phenomenon.

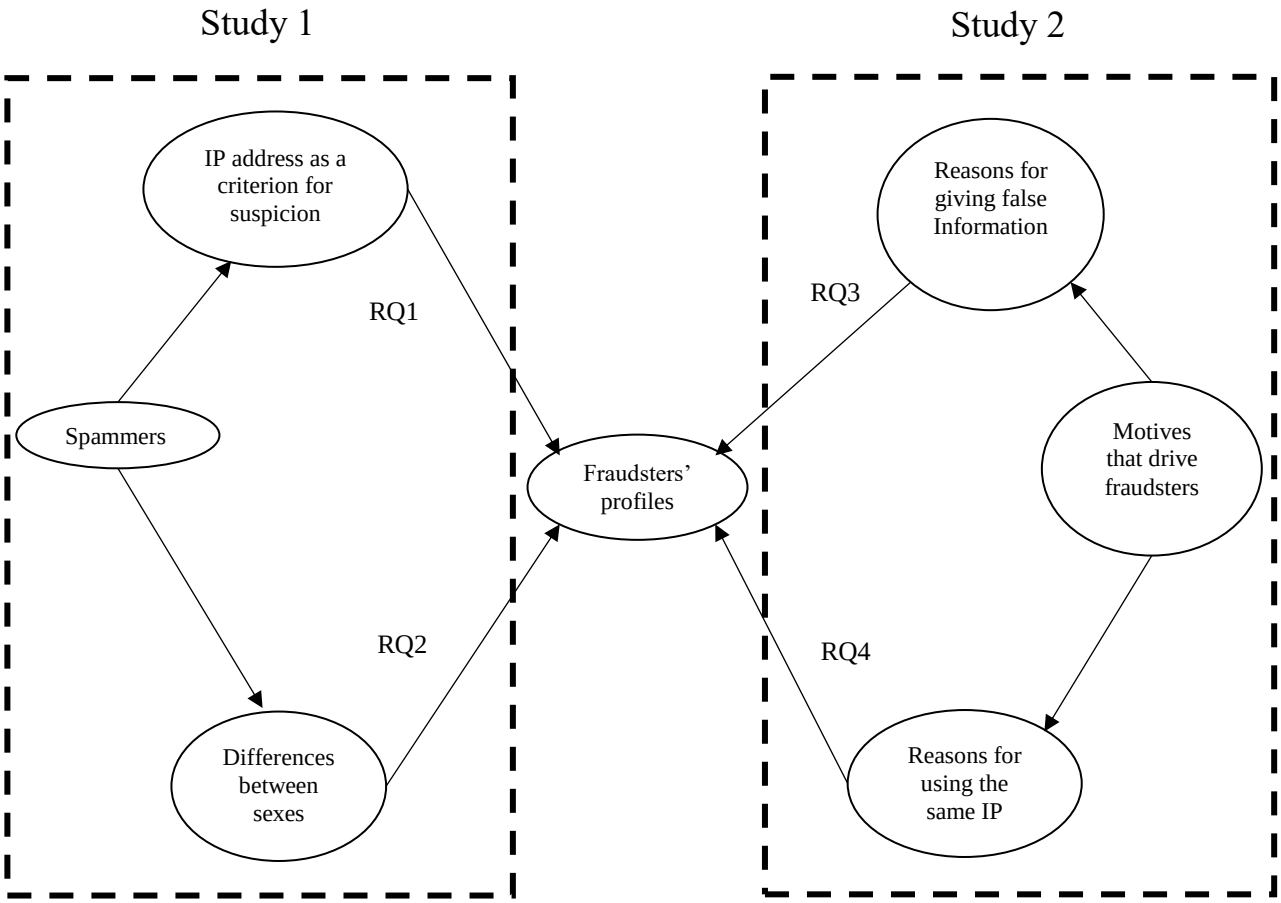
Authors like Zaman et al. (2023) consider that the sources that motivate disinformation should be classified as either intrinsic or extrinsic. Financial incentives have been identified as the main extrinsic motivator. For example, restaurant owners with little customer traffic or that have received negative ratings often post false positive reviews to attract more customers (Luca & Zervas, 2016). In another study, Wang et al. (2017) analyse the tactic of rewarding consumers who post five-star ratings with vouchers that can be redeemed for cash. In contrast, it has also been reported that some guests make illegitimate complaints and then request financial compensation (e.g., free dinners), for example by threatening to post a bad review if their demands are not met (Gössling et al., 2018).

The literature has highlighted several intrinsic motivations, including feeling upset, self-appointed brand managers, and social status (Wu et al., 2021). If clients are upset by a brand or company, this is often the result of a bad experience, disappointment, or perceived betrayal, which causes them to act vindictively by posting false reviews not only about the product that disappointed them but about the brand's products in general (Anderson & Simester, 2014; Wu et al., 2021). The self-appointed brand manager is a loyal customer who strongly identifies with the brand’s image and systematically posts positive reviews of any of its products, thus enhancing its prestige, while often also publishing negative reviews of rival brands' products (Thakur et al., 2017). Moon et al. (2021) add that positive reviews are more likely to be misrepresentative than negative ones due to the perception that they are "white lies". Regarding motivations linked to boosting social status, several pieces of evidence have been collected (Anderson & Simester, 2014; Wu et al., 2021). In a recent study, Moon et al. (2021) find that false reviewers view themselves as opinion leaders. In the same vein, some customers believe

that by posting negative comments about products, even if they have not bought them, they will be perceived as experts (Anderson & Simester, 2014).

In all these studies, data and comment forgers carry out their activities under the perceived anonymity of social networks and the internet, and they would probably act very differently if they knew that they could be fully identified. In this study, we explore the motivations when forgers are aware that they are disclosing their identity whenever they post comments.

Figure 1. Conceptual framework



3. | OVERVIEW OF THE STUDIES

The research questions were addressed by taking a triangulation approach, which consists of analysing the same phenomenon from different perspectives and methods (Denzin, 1978). Two

studies were carried out with three different methodologies. Study 1 is a quasi-experiment to test whether multiple users of the same IP address are more likely to be spammers than when each user has their own IP address. It also explores differences between stated sex. Study 2 is qualitative and explores the motives that drive fraudsters to do this and to do so from the same IP address.

3.1 | **Study 1**

The aim of this study is to show that multiple users using the same IP address are significantly more likely to be spammers than users who only use a single IP address, and explores their composition by sex. It achieves this goal by means of a quasi-experiment. Unlike experiments that require the random assignment of subjects, quasi-experiments are empirical studies that investigate the causal impact of a specific phenomenon on the population, but without random assignment (Kirk, 2013). The latter factor raises concerns about the internal validity of the results, particularly in small samples or when there are confounding variables that cannot be controlled or accounted for (Rossi et al., 2018). However, our quasi-experiment is conducted on a sample of 7,192,487 Spanish users (approximately 15% of the population), using a selection procedure based on AI algorithms. We therefore believe that the sample size and recruitment method compensate for any internal validity concerns.

It is not easy to detect fake reviews and spammers, as it is so difficult to verify that their posts are false. Some studies, such as Moon et al. (2021), check whether reviews are fake by asking one group to post fake reviews and another to post genuine ones in order to compare their characteristics. However, most studies do not know whether the reviews are genuine and rely on estimates or the use of indirect cues. For example, Ott et al. (2011) use subjective criteria to classify reviews collected from TripAdvisor as misleading or truthful, without knowing anything about their actual veracity. However, the assumption that the credibility of 'truthful' reviews makes them genuine has been criticised (Luca & Zervas, 2016). Indirect cues to

recognise the veracity of reviews have included corroborating that the guest really did stay in a hotel (Moon et al., 2021) or checking the IP address from which the review was written to discern whether it came from a server farm (Waggoner et al., 2019).

This study also uses indirect cues, as the detection of users who falsify identification data is considered to be an estimate of spammers. We build on previous work that has already established the relationship between the falsification of identification data and the intention to post false ratings (Wang et al., 2017; Hu et al., 2017; Thakur et al. 2017).

The process depicted in Fig. 2 was followed to detect users who falsify their data on the internet and to corroborate their IP addresses:

First, a landing page data collection company granted permission to use its database of 7,192,487 Spaniards who provided information by participating in online sweepstakes and contests between 2010 and 2023. The lead-gathering company issues advertisements inviting internet users of legal age to participate in sweepstakes to win an Alexa Echo Dot in exchange for providing personal data (generating 97% of the data), as well as to play quizzes on history, geography, cooking, etc. (generating 3% of the data) (See an example in Figure 2). Further examples of landing pages are available on the company websites (https://www.sorteopremios.com, https://www.mitest.de). Before we continued with the analysis, the company was asked to verify that it observes the European Data Protection Regulation and the corresponding Spanish law LOPD-RGPD, "Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de Los Derechos Digitales" (AEPD, 2018). This means that all users who accessed the landing page were notified in advance, ticked the consent box to accept the various data collection purposes, declared themselves over 18 years of age, confirmed that they had read and accepted the terms of participation and privacy policy, and agreed to receive promotional information from the sponsors.

Figure 2. Example of a sample data collection form from www.sorteopremios.com

As discussed above, the database contains 7,192,487 registered users with information on their names, e-mail addresses, telephone numbers, stated sex and age, postcodes, IP addresses, time of registration and registration start page. Based on IP addresses, postcodes and first and last names, the sample was divided into two: Same IP and Unique Users, and Same IP and Different Users. The result of this division revealed that 94.8% were unique users with the same IP and 5.1% were different users with the same IP.

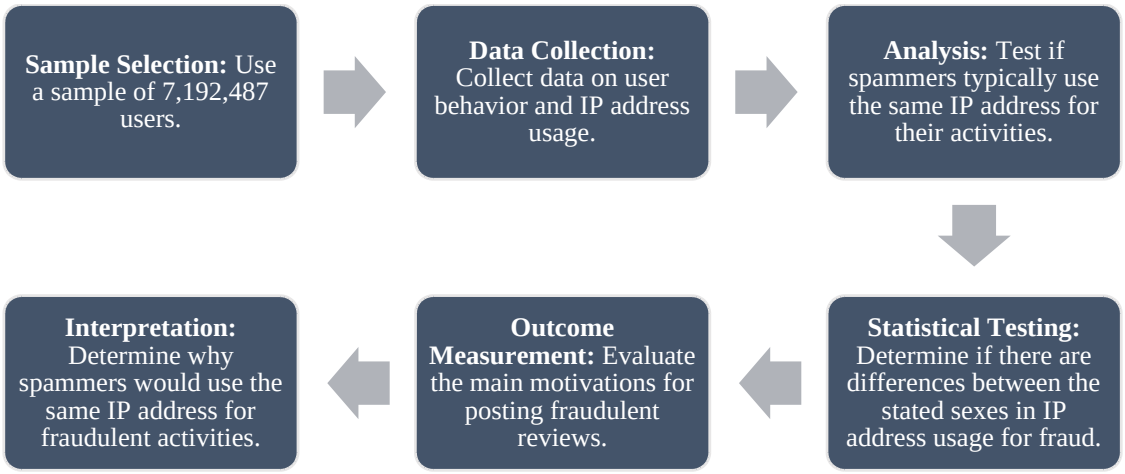
Second, to estimate potentially fraudulent users, we followed the procedure used by Sáez-Ortuño et al. (2023b) of sequentially filtering the pieces of information registered by users in the personal data form. In the first step, three data vectors (columns) containing IP address, postal code and name were selected. Starting with the name records, the Node.js® algorithm was applied, which uses JavaScript to create command lines (Escobar-Jeria et al., 2007). The algorithm, which cannot be published for copyright reasons, compares all the names registered by the participants with the data in the repository of the National Institute of Statistics and the

IDA-Padrón. After comparison, names that did not match the official data or that were infrequent (less than 20 times in Spain or 5 per province) were highlighted (Sáez-Ortuño et al., 2023b). Next, the postcodes recorded on the form were checked against those assigned by geographical location, and those that did not match were marked (Sáez-Ortuño et al., 2023a). The second step was a verification exercise consisting of additional checks of telephone number, email address and age (Sáez-Ortuño et al., 2023b). Application of all these filters generated a list of suspicious participants that had one or more ticks against the information they had registered in their account. At this point, as in the study by Ott et al. (2011), the researchers, after a reflective process, determined the basic criteria for identifying a participant as fraudulent. In the case of Ott et al. (2011), the researchers used three criteria to determine whether the wording of a hotel review was false: (1) use of the first-person singular, (2) non-use of spatial references (e.g. bathroom, far from the centre) and (3) increased use of negative emotion terms. In this study, given that the authors were dealing with personal data and, based on previous results (Sáez-Ortuño et al., 2023b), it was considered that one error could be unintentional, but that from two errors onwards there is a clear intention to mislead. The total number of observations is 422,193 users (249,640 stated male and 172,553 female) estimated as fraudsters, representing 5.86 % of the sample.

Third, the fraudulent users were divided into single users using one IP versus multiple users sharing the same IP. The result was 72,647 unique users (37,117 stated male and 35,530 female) with the same IP address and 349,546 users sharing the same IP address (212,523 stated male and 137,023 female). In other words, almost five times as many fraudsters were operating from a shared network location as those that did so from a unique address. Meanwhile, the distribution of unique users with a single IP is fairly similar for both sexes, but far more of the users of a shared IP were male than were female. We were also able to verify different

registration times from the same location, i.e., the creation of multiple accounts from the same address.

Figure 3. Sequence of steps observed in Study 1



3.1.1 | Results

To verify the results noted in the descriptive analysis and given that the data collected features frequencies, the most appropriate test was the Chi-square test (Bearden et al., 1982). The test of differences in frequency distribution using chi-square is shown in Table II. Estimated fraudulent users were analysed in terms of being Same IP and Single Users (72,647 observed users), Same IP and Different Users (349,546 observed users) and by stated sex. When comparing the frequency distribution of the sample with respect to fraudulent estimates using the Same IP and Unique Users or Same IP and Different Users, the results present statistically significant differences ($\chi 2(2) = 5{,}239{,}056$, $p = 0.000$). Specifically, the observed numbers of unique users of the same IP (72,647 vs. 400,560 expected) and different users of the same IP (349,546 vs. 21,632 expected) show the strong correspondence between multiple users using the same IP and being fraudulent. Thus, RQ1 is supported, as several users registering their information or writing several reviews from the same IP address can be used as an indication of fraud.

The same pattern is repeated by declared sex. Male users with the Same IP and Single Users or Same IP and Different Users present significant differences ($\chi 2(2) = 3{,}287{,}194$, $p = 0.000$), specifically the observed numbers of unique male users of the same IP (37,117 versus the expected 236,848) and different users of the same IP (212,523 versus the expected 12,791). Females with Same IP and Unique Users or Same IP and Different Users also present significant differences ($\chi 2(2) = 1{,}958{,}732$, $p = 0.000$), specifically, unique users of the same IP (35,530 versus the expected 163,711) and different users of the same IP (137,023 versus the expected 8,841). The frequency distribution of reported male and female fraudsters was also estimated, and the results indicate significant differences ($\chi 2(2) = 4{,}464$, $p = 0.000$). In other words, male respondents are more likely than female respondents to misreport information in the case of different users of the same IP. However, there are no significant differences when comparing stated sex according to Same IP and Unique Users ($\chi 2(2) = 0.24$, p = n.s.), as these are mostly found in Same IP and different users ($\chi 2(2) = 14084$, $p = 0.000$). The answer to RQ2 is thus positive but qualified, as there is a greater predisposition of male than female respondents to publish false data and reviews, but only when different users do so from the same IP.

Table II. The distribution of identical IP addresses, unique users, and different users sharing IP addresses among fraudulent registrants relative to the total sample.

	Total	**Fraudulent users**	**Male**	**Female**
Same IP, Unique Users	6,823,954	72,647	37,117	35,530
		(400,560)	*(236,848)*	*(163,711)*
%	94.87%	1.01%	0.51%	0.49%
Same IP, Different Users	368,533	349,546	212,523	137,023
		(21,632)	*(12,791)*	*(8,841)*
%	5.12%	4.85%	2.95%	1.90%
Total sample	7,192,487	422,193	249,640	172,553
%		5.86%	3.47%	2.39%

Notes: In each cell: top figure, absolute frequencies; in brackets, expected values; and percentage of total sample. Fraudulent users Same IP, Unique Users and Same IP, Different Users ($\chi 2$ (2) = 5239056, p = .000); Male Same IP, Different Users and Same IP, Different Users ($\chi 2$ (2) = 3287194, p = .000); Female Same IP, Different Users and Same IP, Different Users ($\chi 2$ (2) = 1958732, p = .000); Male vs Female ($\chi 2$ (2) = 4464, p = .000).

3.1.2 | Discussion

Users who misrepresent information in online reviews are not easy to detect as it is not possible to verify whether the information is false, and so we can only work with behavioural estimates (Wu et al., 2021). In this study, building on previous work that established the relationship between the falsification of one's identification data and the intention to post fake reviews (Wang et al., 2017; Hu et al., 2017; Thakur et al. 2017), the observation that the user has falsified their personal data is proposed as an estimator of their inclination to post fake reviews. In addition, an indirect process is proposed that relates the estimation of fraudulent behaviour to the provenance of that information, i.e. the IP address. In other words, it establishes implicit relationships between two types of information, IP user characteristics and fraudulent behaviour.

The findings show, in line with previous studies (Islam et al., 2020; Pennycook & Rand, 2019), that the percentage of users who post fraudulent information on social networks is small, in this study 5.86%. However, there is a disproportionately high rate of fraudulent behaviour among registrants who share an IP address, among both males and females. In other words, the fact that several users use the same IP raises suspicion that their aim is to generate multiple entries on product or service rating platforms (e.g., TripAdvisor) and to post fake reviews that compromise the integrity of the data. The practice of quickly producing different profiles from a single public connection also aligns with the use of automated bots (server farms) rather than separate individuals using the same IP.

With respect to stated sex, statistical tests revealed a greater predisposition among male users to intentionally provide inaccurate information, particularly in registrations from overlapping IPs.

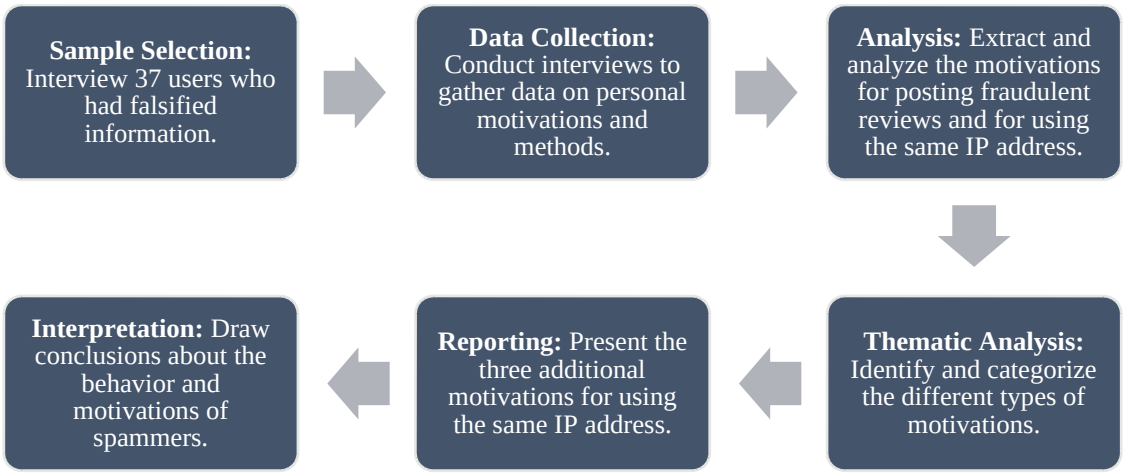
3.2 | Study 2

Bagozzi's (1992) proposal that every activity pursues a goal, and that this configures the drivers that motivate that behaviour, implies the consideration of different motivators of a subjective nature (Rotter, 1954). However, as postulated by Heider (1958), the motivations that drive an individual's behaviour are explained by intrinsic drivers (generated by the subject him/herself) and extrinsic drivers (generated by the environment). In sum, the ontological nature of motivation appears to be multidimensional and, moreover, may differ in terms of quantity as well as type (Zaman et al., 2022).

The multidimensional and subjective nature of the motivations that lead users to keep the same IP address when providing false information or engaging in deception was researched qualitatively. Qualitative research generates an in-depth understanding of the drivers that incite consumers to engage in a particular behaviour, in this study to post false information while keeping the same IP address (Vo-Thanh et al., 2021).

Given the origin of the database, from information captured by leads for online contests and sweepstakes, participants in these contests were deliberately chosen to attend semi-structured face-to-face interviews. As some authors point out, the nature and type of service considered for posting fake reviews affects consumers' online decisions differently (Zaman et al., 2022; Xu, 2020). For example, Park and Lee (2009) point out that experience goods (those that can only be accurately evaluated after the experience) are highly sensitive to negative eWOM comments. This is true for the service offered by lead platforms. Figure 4 shows the sequence of steps observed in Study 2.

Figure 4. Sequence of steps observed in Study 2



3.2.1 | Data collection

Participants in online sweepstakes and contests residing in the city of Barcelona were invited to attend semi-structured interviews. Since the objective was to explore the motivations for posting fake reviews by pretending to be someone else from the same IP, only users who met these requirements were recruited. Candidates were selected by asking the landing page data collection company to provide us with a sample of about 850 users.

Participants were invited by phone to attend the interview in the centre of Barcelona and were offered a 50 euro cheque as an incentive. The research assistant, who recruited the participants, was asked to select only people living in Barcelona and who are fluent in Spanish, and to try to achieve a balance in terms of declared sex and cohort. However, the final selection was based on the semantic saturation criterion as recommended in the literature (e.g., Vo-Thanh et al., 2021; Zaman et al., 2022; Sáez-Ortuño et al., 2023b), which meant the resulting sample was somewhat skewed towards digital natives. In total, 834 telephone calls were made, of which 327 were answered, and 221 expressed their intention to participate in the study. Finally, 37

individuals (19 males (M) and 18 females (F)) between 18 and 65 years of age (including 9 aged 18 to 25 years and 7 aged 58 to 65 years) were selected. The interviews were conducted in Spanish and lasted between 20 and 35 minutes. Table III shows the distribution by declared sex and age.

Table III. Distribution of interviewees by declared sex and age.

Age/Stated sex	Total	Male	Female
18-24 years	9	4	5
25-34 years	12	5	7
35-44 years	4	1	3
45-54 years	5	4	1
More than 54	7	5	2
Total	37	19	18

3.2.2 | Field work process

Interviewees were summoned one by one to a central location in the city where they were given a brief introduction to the objectives of the study. They were also informed of their rights when participating in a research study and asked to sign the consent form. Since the object of the study could be considered unethical conduct and the interviewees might have been reluctant, the recommendations of Sannon et al. (2018) were followed before starting the interviews. These basically consist of the interviewer downplaying socially reprehensible behaviour and saying that the research does not qualify the use of lies or falsification of data as inherently good or bad, but simply aims to analyse aspects of human communication.

To clarify the purpose of the study and avoid possible confusion, the interviewer presented several examples of false information posted on online contest and sweepstake platforms by participants masquerading as other people from the same and different IP addresses. For example, regarding false information, the name "Pepe Pelopincho" was registered, and with

regard to impersonation, some participants registered using different versions of the name of the famous footballer Lionel Messi, such as "Leo Messi" and "Lio Mesi".

The structured interview itself was based on previous studies (Zaman et al., 2022; Saez-Ortunño et al., 2023b):

1) How many times have you posted false information on online contest and sweepstake platforms?

2) In that information, was the data completely fictitious or did it impersonate someone else?

3) What were your motives for posting it?

4) Did you do it from the same IP or from different IPs? If from the same IP, why?

The audio of the interviews was recorded.

3.2.3 | Data analysis and results

The participants were spontaneously drawn from the population raised in the digital environment. As shown in Table III, more than 56% of the participants are from the younger generation (18-34 years old) despite the sample delivery requirements issued to the lead capturing company.

In line with the results obtained by Zaman et al. (2022), the participants who answered the relevant question (16 respondents) reported that they had posted false information between two and four times. Regarding the second question, 27 respondents answered that the information that they had posted was totally false and 14 answered that they had at some point impersonated a known person.

Responses to questions three and four were transcribed and analysed using thematic analysis, which consists of identifying and coding phrases or expressions that refer to the object of study, in this case the motives that drive or justify the behaviour (Braun & Clarke, 2019). Coding followed a three-step process (Tuomi et al., 2021; Sáez-Ortuño et al., 2023b). First, the recorded

responses were transcribed and read, and sections or expressions related to motives were highlighted. Second, the highlighted sections and expressions were assigned open codes, resulting in 18 initial codes. Third, the codes were structured in a pyramidal form with three levels. To improve internal validity, and following Zaman et al. (2022), two of the researchers read, highlighted, and coded the first three interviews separately, and the degree of convergence was checked by Cohen's Kappa estimation and, in both cases, indicated good agreement (value above 0.80) (Landis & Koch, 1977). However, points of divergence were examined until consensus was reached. As a result of this procedure and following Bagozzi's (1992) principle of maintaining parsimony and operationality, the responses to question three on the reasons for publishing false information were grouped into four themes: revenge against unpunished companies (14 respondents), fun (10 respondents), increasing the likelihood of reward (8 respondents) and increasing self-esteem (5 respondents). According to Heider's (1958) proposed classification, there are two intrinsic motivations: the search for entertainment and improvement of self-esteem; and two extrinsic motivations: revenge against an unfair environment or to achieve economic gain (Deci, 1971). However, in our study, the motives are independent of each other and no synergistic or undermining effects are observed between them (Deci & Ryan, 2000).

A brief description of the motivators follows, and Table IV presents some examples to illustrate the assignment of labels.

1) Revenge. Some respondents try to justify their behaviour as a kind of crusade against companies acting on the Internet. In other words, they claim to act in retaliation against illegal actions such as cheating, abuse of information provided, etc.
2) Entertainment. Some interviewees consider this behaviour fun and a way to play pranks on friends or relatives, which could be considered a kind of gamification (Zaman et al., 2022). In this case, impersonation is more frequent, although as one of the interviewees

says, "there is no bad intention behind it" (F, 19, Undergraduate Business student). In other words, these behaviours are viewed as innocent, inconsequential fun.

3) Opportunity for profit. Several interviewees say their motivation for faking data is to increase the probability of winning the prize. In other words, it is an argument linked to the possibility of achieving an economic return.
4) Self-esteem. A few interviewees suggested their self-esteem as a reason for falsifying information. That is, deceiving the machine with false information made them feel clever.

Table IV. Interview results from question three and categorisation from the open coding

Themes	Description	Sample Quotations	Listing Key Phrases
Revenge	It is alleged that users act in retaliation against companies that use the internet to make money from customers and go unpunished.	"I've written false information because two years ago I gave my support to a charity, and since then they have not stopped hammering me with messages in my mailbox" (F, 60, Daycare in nursery school). "These companies take advantage of the good faith of users, as they make false promises of prizes" (F, 23, Undergraduate Business student). "I like to take part in quizzes, and I do it a lot. But I have never received a prize, so I falsify my information" (F, 32, Civil servant).	Bad experience, mistrust, unethical agents, impunity
Entertainment	It is considered fun, entertaining and a way to play pranks by impersonating someone they know or giving false information.	"I find it amusing to enter my nephew's name and if he gets the prize, it will be a surprise for him" (F, 61, Cook at a daycare centre). "Sometimes, I put my boyfriend's name just as a joke, but there is no bad intention behind it" (F, 19, Undergraduate Business student).	As a game, entertainment, something fun, spontaneous.

		"It is true that I have sometimes entered my work colleague's name, but I didn't even think about it much, just for fun" (M, 47, Salesman).	
Opportunity for profit	The reason for falsifying data is to have a better chance of winning the prize.	"I often falsify my data so I can play more times and have a better chance of winning the prize. It's a very normal thing to do, all my colleagues do that too" (M, 18, Undergraduate Business student). "I meet up with my friends and we play the games together, so swap names to have a better chance to win" (F, 21, Undergraduate Arts student). "If I provide false data, I can enter several times and my chances of success are higher" (M 28, Shop assistant)	Higher chance of winning prize, higher probability, and expectations
Self-esteem	It seems to be a mechanism to boost the user's self-esteem, because they feel clever for outwitting the machine by giving false information.	"I love doing the intelligence tests posted on my browser and give false information as if it were a challenge" (M 26, Undergraduate Business student). "I'm a gaming geek, and sometimes I've faked my data and when I see that the platform doesn't detect it, it makes me feel good" (M, 22, Undergraduate in Computer Science student) "I challenge my colleague to see who can come up with the weirdest name, and I like to be the wittiest" (M, 25, Master Maths student).	Raise self-esteem, feel better by outdoing the machine, boost ego, think you are smarter than the machine.

Regarding the answers to question four, the main novelty of this study, on the reasons given by data forgers for carrying out their misdeeds from an identifiable IP address, some respondents

were not very aware that their IP address could be identified so relatively easily. Among the most aware, three reasons emerged: convenience, limited resources, and complacency about keeping the same IP address rather than frequently switching to new ones. Overall, it appears that the effort and disadvantages may outweigh the perceived benefits for many users. In this case, according to Heider's (1958) classification, there would be two sources of intrinsic motivation: convenience and limited resources; and one source of extrinsic motivation: the environment is incompetent and tolerates these activities. Again, we did not observe synergistic or undermining effects between the motivations considered (Deci & Ryan, 2000). A brief description of the motivators is given below, and Table V presents some examples to illustrate the assignment of labels.

1) Convenience. Some respondents claim that obtaining and rotating new IP addresses can be tedious and time-consuming. So, it is easier to stay on the same IP address.
2) Limited resources. Others mention their lack of skills, competence, time, and resources to acquire and manage multiple IP addresses. While some of the interviewees, because of their background, may have technical knowledge, they also claimed that they did not have the time or financial resources to acquire new IPs on a regular basis.
3) Complacency. Several interviewees argued that if they have been able to falsify information from the same IP address so far and had suffered no consequences, why should they change their ways? In other words, the "if it ain't broke, don't fix it" mentality comes into play.

In summary, both RQ3, on the main motivations for falsifying information or posting false reviews, and RQ4, on the motivations for doing so from the same IP address, are successfully supported.

Table V. Interview results from question four and categorisation from the open coding

Themes	Description	Sample Quotations	Listing Key Phrases
Convenience	Obtaining and rotating new IP addresses can be tedious and time-consuming for some users. It's often easier to just stay on the same IP address. (Saha et al.,2003)	"It is too much work to keep changing my IP" M, 21, Social Science undergrad student. "It's easier to just use the same one." F, 46, Nurse "I don't have the time to figure out how to change my IP address. I have other priorities." M, 42, English Teacher	Tedious, time-consuming, easier, simplicity
Limited Resources	Acquiring and managing multiple IP addresses may require additional technical skills, tools, or financial expenditure. Some users may lack the capabilities or funds to regularly get new IPs. (Saha et al.,2003)	"I don't know how to change my IP. It seems complicated and very expensive" F, 39, Librarian "I don't have the technical skills to change my IP address. I'm not sure where to start." M, 62, Statistics Professor	Technical skills, tools, financial cost, capabilities
Complacency	If a static IP address has worked without issues previously, some users may see no need to put in the extra effort to change it. The mentality of "if it ain't broke, don't fix it" comes into play. (Saha et al.,2003)	"My IP has always worked fine, so I don't see any reason to change it." M, 33, Accountant "I like having a static IP because it's easier to remember and use on my devices." F, 26, real estate agent	Effort, no issues, no perceived benefit from changing

3.2.4 | Discussion

The findings that could be compared with results from previous studies show a high degree of consistency. Thus, the number of times participants falsified their data ranges between two and

four, which is very similar to the results obtained by Zaman et al. (2022). Meanwhile, 14 of the 37 respondents stated that they had at some point impersonated an acquaintance. However, given the exploratory nature of the research, the results are merely indicative.

Four main motivations are alleged to justify the reprehensible behaviour of falsifying data and reviews: revenge, entertainment, opportunity for profit, and self-esteem. In general, similar reasons for all the motives obtained in the qualitative research were found in different previous studies.

The most common motive was revenge, as was also highlighted earlier by Thakur et al. (2017) in their study of perceived betrayal and the desire for revenge in relation to cybershilling. Furthermore, Zaman et al. (2022) proposed “retaliation” when discussing upset or betrayed customers who seek vengeance against brands with which they had bad experiences. However, some studies on cyberbullying suggest that revenge is more of an excuse to justify bad behaviour, as it is usually more present in aggressors than in victims (Fluck, 2017). In our study, the label “revenge” is used to refer to a supposed reaction against abuses committed by online companies. In other words, the participants, in full awareness of their misconduct, justified their actions by alluding to a kind of preventive punishment, as noted by Fluck (2017). In a way, these arguments reflect distrust of lead-gathering companies (Sáez-Ortuño et al., 2023b) and of the bad practices of some companies (Zaman et al., 2022).

Another motive which also appears in the literature is entertainment (Moon et al., 2021; Zaman et al., 2022; Sáez-Ortuño et al., 2023b). Although from different perspectives, several previous studies have observed similar motives. Moon et al. (2021) call it hedonic behaviour, while Zaman et al. (2022) introduced the link with gamification, whereby there is fun to be gained from faking reviews, and Sáez-Ortuño et al. (2023b) also discuss the concept of fun. To explain this behaviour, it has been suggested that "digital natives" find it difficult to distinguish between jokes and deliberate acts intended to cause harm, and that the lack of face-to-face

communication sometimes leads to biases in the interpretation of the meaning of the message (Talwar et al., 2014). Friends, relatives, or acquaintances are simply impersonated for amusement, as a joke, and users write fake reviews and enter false data simply for recreation, with no intention to do harm (Crosslin & Golman, 2014).

The third motive refers to the quest for financial gain. Humans have long made a profit from lying, as has already been considered from evolutionary psychology (Cosmides & Tooby, 2016). It is also one of the most defended motives in the literature. For example, some researchers describe businesses that get their employees to pose as their rivals' customers and post bad reviews (Litvin et al., 2008). Thakur et al. (2017) pointed out that the use of financial rewards is a key incentive in customers' willingness to engage in cyber-shilling. Also, in a context of posting fake reviews on cosmetic products, both South Korean and French customers confirmed that financial compensation was a strong motivator (Zaman et al., 2022). Therefore, the motivation to falsify data proposed in this study of standing a better chance of winning the prize, since participants can only play once, helps to validate the economic nature and relevance of this motivation.

The falsification of information to boost self-esteem has also been confirmed in previous studies (e.g., Zaman et al., 2022; Kapoor, 2021; Moon et al., 2021). In one of the first studies on motivations for e-WOM, Moe and Schweidel (2012) divide posters into "experts" and "less active", noting that the former tend to post very negative comments to attract attention and showcase their position as opinion leaders. Moon et al. (2021), in a quantitative investigation, point to indirect cues, such as the propensity to bargain, as well as prosocial and individualistic consumer behaviour when referring to self-esteem. However, the most direct evidence comes from Zaman et al. (2022), who consider self-esteem to be boosted by appearing to be an expert or opinion leader. In our case, self-esteem is linked to the feeling of superiority gained by cheating or outwitting the machine and its creators.

Regarding question four, on the reasons for spoofing data from an identifiable IP address, to the best of our knowledge there is no precedent with which to compare the results obtained and therefore the discussion is based on indirect elements. However, the arguments put forward by people who are most aware of the trail they leave behind when committing their misdeeds are quite reasonable.

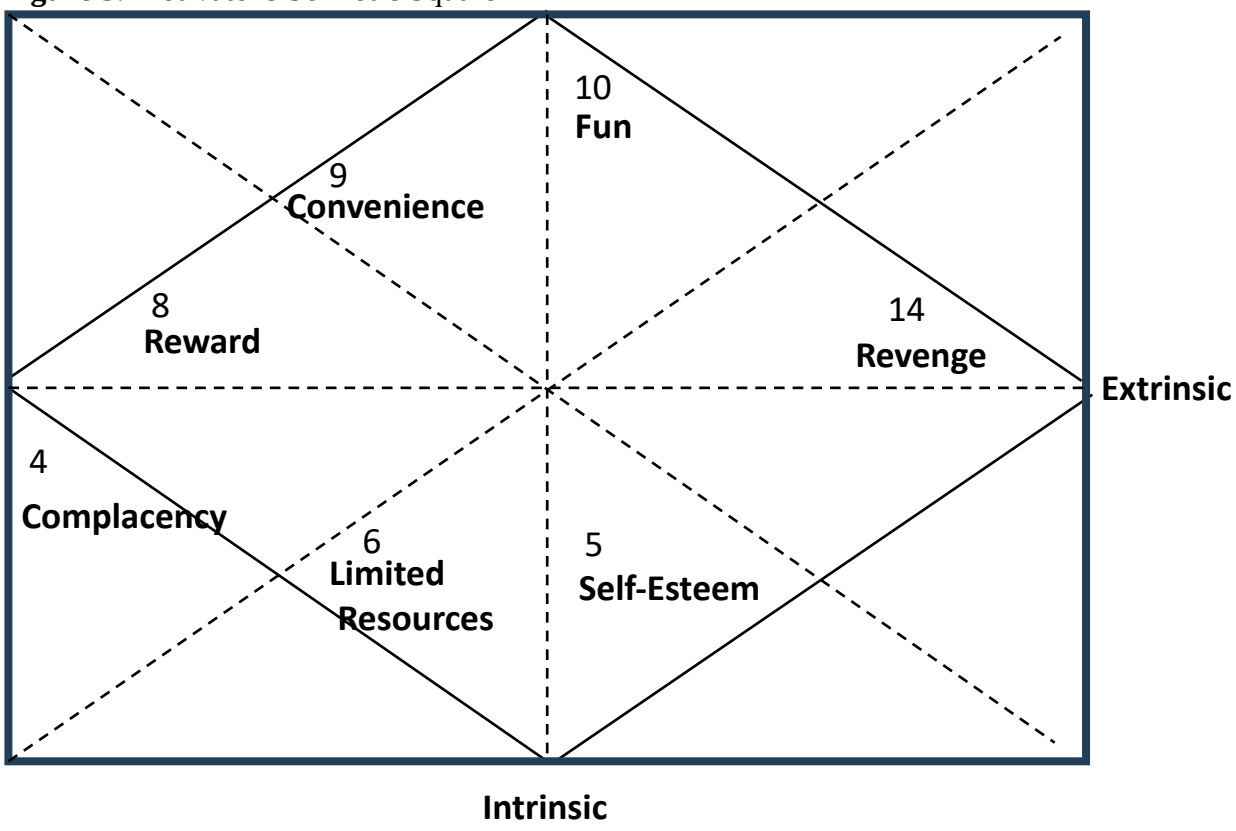
Convenience is labelled as an important motivation, the argument being that obtaining and rotating IP addresses can be tedious and time-consuming (Jaillant & Caputo, 2022). Motivation linked to saving time and effort has been one of the most considered since the first studies on distance shopping (Eastlick & Feinberg, 1999), and this factor has become even more prominent in online environments where location is an irrelevance (Swaminathan et al., 1999). Limited resources are another cited motivation, which includes both users' knowledge constraints and their financial ones. Studies of online consumers often differentiate between experienced consumers, who demonstrate relative mastery of the technologies involved, and more novice consumers who express unease when handling them (Bernard & Makienko, 2011). Indeed, a certain mastery of IT tools facilitates access, manipulation and dissemination of fake reviews and hinders their identification (Casas, Del Rey, & Ortega, 2013). As noted above, some respondents with little computer expertise were unaware that they were leaving a trail, and those who were argued that they simply did not have the technical skills to manage multiple IPs (Li, 2022).

The third motivation is complacency, which groups together arguments linked to the fact that as users have not had to deal with any problems so far with a static IP, they see no reason to make the effort to get a new one (Maaß et al., 2021). In fact, studies on online consumer behaviour have already established the relationship between the degree of satisfaction and the absence of problems with the desire to stay longer (Kim et al. 2007). This is undoubtedly a statement of perceived impunity, i.e., there is no punishment for posting false reviews or

information. Furthermore, studies on cyberbullying have argued that anonymity and the perception of impunity may encourage users to engage in ethically reprehensible behaviour, including the perpetration of assault and other types of cybercrime (Compton, Campbell, & Mergler, 2014).

Therefore, convenience, resource constraints and complacency may motivate users to stick to the same IP address rather than constantly switching to new ones (Jaillant & Caputo, 2022; Li, 2022; Maaß et al., 2021).

Figure 5. Motivatons Semiotic Square



4 | CONCLUSIONS

Increasing activity on the Internet and social media has led to the creation of an online ecosystem with its own marketplace, commerce, and marketing (Kumar et al., 2016; Sestino et al., 2023), where there is an abundant exchange of information and eWOM, and where comments about the shopping experience have an enormous influence on consumers' decisions (Wu et al., 2020). However, this power to influence is exploited by unscrupulous individuals

who post false testimonies, either rating products they have not used (Zhang et al., 2016) or rating them incorrectly to boost or tarnish the image of a company and its brands (Daiv et al., 2020).

The proliferation of fake reviews affects consumers, companies, digital platforms, and the market in general. In awareness of their pernicious effect, scholars and practitioners have gone to enormous efforts to detect both fake testimonies and their authors (Mukherjee et al., 2012; Mohawesh et al., 2021; Birim et al., 2022), the results regarding the latter being clearer than the former. This study enriches the literature on spammer detection by proposing the analysis of duplicate IP addresses to estimate their total proportion and by declared sex. The motivations for this incorrect behaviour are also explored and, in many cases, validated. Reasons for posting false comments from the same IP address are also examined.

The study was carried out with Spaniards in the context of entertainment services, specifically among participants in online competitions and sweepstakes. This is the kind of service where previous users' opinions are very important references on which to base decisions since the experience can only be evaluated after having lived it. Moreover, as noted by Akoglu et al. (2013), spammers tend to present a certain degree of specialisation by product type and systematically give the same rating.

4.1 | Theoretical implications

Although the number of users who provide fraudulent information is relatively small (Islam et al., 2020; Pennycook & Rand, 2019), in this study being estimated at 5.86%, their capacity to harm is very high. Indeed, 70% of spammers wrote more than five reviews per day (Heydari et al., 2015). Therefore, it seems that it would be more efficient to detect and neutralise spammers than to control the huge proliferation of fake reviews. To do this better, more needs to be known about their behavioural patterns and motivations (Cosmides & Tooby, 2016; Sáez-Ortuño et al.,

2023b). This study expands on the findings of Waggoner et al. (2019) to corroborate how, in the context of entertainment services (sweepstakes and online quizzes), users using different names from the same IP address are more likely to falsify their data (Song et al., 2021).

Waggoner et al. (2019) address reports that up to 25% of respondents in some MTurk studies provide false answers from server farms outside the United States (Ahler et al., 2018). These authors propose their detection by cross-checking respondents' IP addresses against up to three verification services (Waggoner et al., 2019). In our study, three items of information are also used to corroborate IP addresses, namely time of registration, postcode, and homepage. By cross-checking this information, shared IP addresses can be detected, where it is noted that registrations were virtually simultaneous, and discrepancies in terms of postcodes or landing pages provide valuable further incriminating evidence. However, it is important for the algorithms to be flexible enough for the researcher to adapt the criteria for either inclusion or exclusion and to deal with false positives/false negatives. Therefore, work needs to be done on the development of AI algorithms that enable such cross-checking of information in a streamlined manner that will facilitate the detection of suspects.

Since the sample collects information on users who have participated in entertainment activities, it is skewed towards females, who represent approximately 65% of the total. Since the emergence of online gambling, the demographic profile of participants has shifted towards a predominance of female users (Hing et al., 2016). According to McCormack et al. (2014), online gambling generates a greater sense of security for women compared to physical locations. However, despite the greater inclination for women to participate in online entertainment activities, no significant differences are estimated with respect to the willingness to provide false information. The sole exception is the case where several users use the same IP, when a greater inclination is detected in men than in women.

Finally, our study elucidates the motives associated with the posting of false information and reviews. These are many and diverse, including instinctive or visceral motivations based on revenge (Hull, 1952), expectations, whether for financial gain or entertainment (Rotter, 1954), or on internal dispositional forces and self-esteem (Heider, 1958). Unlike the study by Zaman et al. (2022), which generated fourteen motives for posting fake reviews, our study followed Bagozzi's (1992) principle of selecting the minimum number of motivators has concentrated them into four. It is encouraging to note that the motives characterised in this study match those reported in several other publications on fake reviews (Moon et al., 2021; Thakur et al., 2017; Zaman et al., 2022). Our data helps to corroborate and validate these previous results. However, the list is certainly not exhaustive as the type of product and the researcher's criteria may influence its classification (Moon et al., 2021). With respect to the motives for posting fraudulently from the same IP, two internal dispositional forces (the convenience of requiring very little effort, and limited resources or knowledge), and one external situational force (complacency, the environment in which the activity is carried out is very tolerant, as it does not punish those who act improperly). Self-determination theory posits that extrinsic motivations can act in a promotive or undermining manner with respect to intrinsic motivations (Deci, 1971), whereby the perception of anonymity and impunity encourages unethical behaviour and only an external event such as a threat of punishment could mitigate such behaviour (Deci & Ryan, 2000).

4.2 | Implications for management

Given that the digital ecosystem creates an environment where anonymity and the perception of impunity encourage dishonest behaviour, companies and organisations should prioritise the detection and neutralisation of wrongdoers. It has been proposed that the infringers themselves are more efficient to combat than the fake reviews and information that they create. Spammers

publish numerous false reviews that are not easy to identify upfront (Mukherjee et al., 2013) and can lead both consumers and managers to make incorrect decisions grounded on dishonest information (Moon et al., 2021). Therefore, the only option is to develop user verification measures to ensure both the integrity of systems and the accuracy of recorded data and opinions (Shu et al., 2020).

This study, in line with previous findings, proposes the use of IP addresses to track potential fraudsters. In other words, a relationship has been established between the use of the same IP by several users and their inclination to falsify information. To estimate whether different user profiles send messages from the same IP as if they were from different people, a filtering system consisting of at least three steps is proposed. These consist of matching IPs against registration time, postcodes, and landing pages. In addition, a quasi-experiment has corroborated that the main reason for duplicating IPs is to post false data. These results are in line with previous findings (Waggoner et al., 2019) and with the fact that a small number of sources tend to generate large amounts of fake news (Axelrod & Hamilton, 1981; Allcott & Gentzkow, 2017). This is particularly relevant for businesses operating in online environments, such as social networks, e-commerce platforms, and online service providers, as the identification of spammers' profiles can enable the implementation of specific security protocols and anti-fraud algorithms that target certain groups (Pattee et al., 2022). That is, algorithms for identifying IP duplication could provide a list of IP addresses suspected of harbouring spammers. However, no procedure is perfect. In line with what Pattee et al. (2022) point out regarding malware detection, most software can easily be detected by means of simple checks, but some are very similar to benign code and need to be handled more thoroughly.

One way to gain deeper insight is to understand the motivations behind these behaviours. In this study, four motives for falsifying data and three for doing so from the same IP were collected. Given that this is reprehensible behaviour, and although attempts have been made to

play down the issue, the motivations are more reminiscent of excuses than real causes. However, one of the underlying motives is the perception of immunity. As Compton et al. (2014) point out, anonymity and perceived impunity with regard to the digital ecosystem may encourage users to engage in ethically reprehensible behaviour. Therefore, if users are made more aware of the fact that they are easy to trace, even when they impersonate someone else, this will not only help to deter potential offenders, but should also raise the trust of honest consumers (Kumar et al., 2016; Sáez-Ortuño et al., 2023b; Sáez-Ortuño et al., 2023c).

4.3 | Limitations and future studies

Like any research, this study has some limitations. The results were based solely on the information contained in the database provided by the lead generation company in reference to online entertainment services, competitions, and sweepstakes. Given that there appears to be a certain degree of specialisation of fraudsters in different types of products and responses (Akoglu et al., 2013), future research could consider other products and services, or other types of sales, such as pre-order (Jha et al., 2019) to corroborate whether these motivations are cross-sector. Also, due to the restrictions imposed by the company on the availability of the data, fraudsters could only be adjudged on an aggregate level, i.e., no distinction was made between those captured by contests or sweepstakes, nor by user origin, social networks, etc. Such an extension is quite plausible, and would greatly enrich the results (Parekh et al., 2018).

Moreover, in this research, the quantitative data exclusively relates to customers residing in Spain, and the qualitative data was only obtained from residents of Barcelona, although these were people who had posted false information or reviews in the past. However, previous work, such as the study by Barnes et al. (2007), has shown that the factors explaining market segmentation in Germany, the USA and France are the same, but that these differ in terms of the weight of each segment in each country. More recent studies, such as Borges-Tiago et al.

(2020) suggest that user motivations differ by country and region. Therefore, it would also be interesting to explore motivations through a cross-cultural analysis (Altman & Bland, 1998), as this could help to establish universal criteria to help prevent the prevalence of these behaviours among spammers.

Moreover, even though the sample available to the respondents' telephone captors was widely dispersed in age and that recruitment criteria were set to preserve it, the participants who were most willing to participate in the study were those under 35 years, i.e., the ones who are most educated in the field of new technologies (Purani et al., 2019). Future research could compare younger people's motivations with those of older adults, who are more likely to make unintentional errors, both when recording their data and when evaluating products or establishments (Saez-Ortuño et al., 2023b). Exploration of this phenomenon and of mechanisms to avoid unfairly penalising users who falsify data due to factors beyond their control is an interesting avenue with strong ethical implications.

REFERENCES

AEPD (Agencia Española de Protección de Datos). (2018). Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPD-GDD). Retrieved November 5, 2023, from https://www.aepd.es/

Ahler, D. J., Roush, C. E., & Sood, G. (2018). The micro-task market for "Lemons": Collecting data on Amazon's Mechanical Turk. *Working Paper*. Epub. Available at http://www.gsood.com/research/papers/turk.pdf.

Akoglu, L., Chandy, R., & Faloutsos, C. (2013). Opinion fraud detection in online reviews by network effects. In Proceedings of the International AAAI Conference on Webblogs and Social Media (Vol. 7, No. 1, pp. 2-11). https://ojs.aaai.org/index.php/ICWSM/article/view/14380/14229

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211-236.

Alsubari, S. N., Deshmukh, S. N., Alqarni, A. A., Alsharif, N., Aldhyani, T. H., Alsaade, F. W., & Khalaf, O. I. (2022). Data analytics for the identification of fake reviews using supervised learning. *Computers, Materials & Continua*, *70*(2), 3189-3204.

Altman, D. G., & Bland, J. M. (1998). Generalisation and extrapolation. *British Medical Journal*, *317*(7155), 409-410.

Anderson, E. T., & Simester, D. I. (2014). Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research*, *51*(3), 249-269.

Atkinson, J. W. (1964). *An introduction to motivation*. NJ: Van Nostrand.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390-1396.

Bagozzi, R. P. (1992). The self-regulation of attitudes, intentions, and behavior. Social *Psychology Quarterly, 55*(2), 178-204.

Barnes, S. J., Bauer, H. H., Neumann, M. M., & Huber, F. (2007). Segmenting cyberspace: a customer typology for the internet. *European Journal of Marketing, 41*(1/2), 71-93.

Bearden, W. O., Sharma, S., & Teel, J. E. (1982). Sample size effects on chi square and other statistics used in evaluating causal models. *Journal of Marketing Research, 19*(4), 425-430.

Bernard, E. K., & Makienko, I. (2011). The effects of information privacy and online shopping experience in e-commerce. *Academy of Marketing Studies Journal, 15*, 97-112.

Birim, Ş. Ö., Kazancoglu, I., Mangla, S. K., Kahraman, A., Kumar, S., & Kazancoglu, Y. (2022). Detecting fake reviews through topic modelling. *Journal of Business Research, 149*, 884-900.

Bonald, T., Feuillet, M., & Proutiere, A. (2009). Is the "Law of the Jungle" Sustainable for the Internet?. In *IEEE INFOCOM 2009*, pp. 28-36, Rio de Janeiro, Brazil. https://ieeexplore.ieee.org/document/5061903

Bondielli, A., & Marcelloni, F. (2019). A survey on fake news and rumour detection techniques. *Information Sciences*, *497*, 38–55.

Borges-Tiago, T., Tiago, F., Silva, O., Guaita Martinez, J. M., & Botella-Carrubi, D. (2020). Online users' attitudes toward fake news: Implications for brand management. *Psychology & Marketing, 37*(9), 1171-1184.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise, and health, 11*(4), 589-597.

Casas, J. A., Del Rey, R., & Ortega-Ruiz, R. (2013). Bullying and cyberbullying: Convergent and divergent predictor variables. *Computers in Human Behavior, 29*(3), 580-587.

Chaffey, D., & Ellis-Chadwick, F. (2019). Digital marketing (7th edition). Pearson Eds, Harlow, UK.

Chatterjee, S., Goyal, D., Prakash, A., & Sharma, J. (2021). Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application. *Journal of Business Research, 131*, 815-825.

Compton, L., Campbell, M. A., & Mergler, A. (2014). Teacher, parent and student perceptions of the motives of cyberbullies. *Social Psychology of Education, 17*, 383-400.

Conroy, N.J., Rubin, V.L., & Chen, Y. (2015). Automatic deception detection: methods for finding fake news. *Proceedings of the Association for Information Science and Technology, 52*(1), 1–4.

Cosmides, L., & Tooby, J. (2016). Adaptations for reasoning about social exchange. In D.M. Buss (Eds.), *The Handbook of Evolutionary Psychology* (pp. 625-668). New Jersey: John Wiley & Sons, Inc.

Crawford, M., Khoshgoftaar, T. M., Prusa, J. D., Richter, A. N., & Al Najada, H. (2015). Survey of review spam detection using machine learning techniques. *Journal of Big Data, 2*(1), 1-24.

Crosslin, K., & Golman, M. (2014). "Maybe you don't want to face it"–College students' perspectives on cyberbullying. *Computers in Human Behavior, 41*, 14-20.

Da Fonseca, J. M. R. R., & Borges-Tiago, M. T. (2021). Cyberbullying From a Research Viewpoint: A Bibliometric Approach. In M. Cruz-Cunha, & N. Mateus-Coelho (Eds.). *Handbook of Research on Cyber Crime and Information Privacy* (pp. 182-200). IGI Global.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace 1. *Journal of Applied Social Psychology, 22*(14), 1111-1132.

Daiv, K., Lachake, M., Jagtap, P., Dhariwal, S., & Gutte, V. (2020). An approach to detect fake reviews based on logistic regression using review-centric features. *International Research Journal of Engineering and Technology (IRJET), 7*(06), 2107-2112.

De Ruyter, K., Isobel Keeling, D., & Ngo, L. V. (2018). When nothing is what it seems: A digital marketing research agenda. Australasian marketing journal, 26(3), 199-203.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*(1), 105-115.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268.

Denzin, N.K. (1978). Sociological methods: A sourcebook. New York, NY: McGraw-Hill.

Di Domenico, G., Nunan, D., Sit, J., & Pitardi, V. (2021). Free but fake speech: When giving primacy to the source decreases misinformation sharing on social media. *Psychology & Marketing, 38*(10), 1700-1711.

Eastlick, M. A., & Feinberg, R. A. (1999). Shopping motives for mail catalog shopping. *Journal of Business Research, 45*(3), 281-290.

Escobar-Jeria, V. H., Martín-Bautista, M. J., Sánchez, D., & Vila, M.-A. (2007). Analysis of Log Files Applying Mining Techniques and Fuzzy Logic. H.G. Okuno, & M. Ali (Eds.). *BT - New Trends in Applied Artificial Intelligence* (pp. 483-492), Springer Berlin Heidelberg.

Festinger, L. (1957). A theory of cognitive dissonance. Evanston, IL: Row, Peterson, & Co.

Fluck, J. (2017). Why do students bully? An analysis of motives behind violence in schools. *Youth & Society, 49*, 567-587.

Goh, K. Y., Heng, C. S., & Lin, Z. (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user-and marketer-generated content. *Information systems research, 24*(1), 88-107.

Gössling, S., Hall, C. M., & Andersson, A. C. (2018). The manager's dilemma: a conceptualization of online review manipulation strategies. *Current Issues in Tourism, 21*(5), 484-503.

Heider, F. (1958). The psychology of interpersonal relations. New York: Wiley.

Heydari, A., ali Tavakoli, M., Salim, N., & Heydari, Z. (2015). Detection of review spam: A survey. *Expert Systems with Applications, 42*(7), 3634-3642.

Hing, N., Russell, A., Tolchard, B., & Nower, L. (2016). Risk factors for gambling problems: An analysis by gender. *Journal of Gambling Studies, 32*, 511-534.

Hu, J., Liang, J., & Dong, S. (2017). iBGP: A Bipartite Graph Propagation Approach for Mobile Advertising Fraud Detection. *Mobile Information Systems*, Article ID 6412521. https://doi.org/10.1155/2017/6412521.

Hull, C. L. (1952). A behavior system; an introduction to behavior theory concerning the individual organism. Yale University Press.

INE. (2020). Survey on Equipment and Use of Information and Communication Technologies in Households 2019, INEbase, National Institute of Statistics. Retrieved July 6, 2022 from https://www.ine.es/dynt3/inebase/en/index.htm?padre=6898&capsel=6933

INE (2023). Población que usa Internet (en los últimos tres meses). Tipo de actividades realizadas por Internet. Retrieved January 21, 2024 from https://ine.es/ss/Satellite?c=INESeccion_C&cid=1259925528782&p=1254735110672&pagename=ProductosYServicios%2FPYSLayout

Islam, M. R., Liu, S., Wang, X., & Xu, G. (2020). Deep learning for misinformation detection on online social networks: a survey and new perspectives. *Social Network Analysis and Mining, 10*(1), 1-20.

Jaillant, L., & Caputo, A. (2022). Unlocking digital archives: cross-disciplinary perspectives on AI and born-digital data. *AI & society, 37*(3), 823-835.

Jha, S., Deitz, G. D., Hart, P., & Royne Stafford, M. B. (2019). Sales promotions for preorder products: The role of time-of-release. *Psychology & Marketing, 36*(9), 875-890.

Kannan, P. K. & Li, H. "Alice" (2017). Digital marketing: A framework, review and research agenda. *International journal of research in marketing, 34*(1), 22-45.

Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in social media research: Past, present and future. *Information Systems Frontiers, 20*(3), 531-558.

Kent, S. & Seo, K. (2005). Security architecture for the internet protocol. RFC 4301 (Proposed Standard), Internet Engineering Task Force. https://www.ietf.org/rfc/rfc4301.txt

Keusch, F., Struminskaya, B., Antoun, C., Couper, M., & Kreuter, F. (2019). Willingness to participate in passive mobile data collection. *Public Opinion Quarterly, 83*, 210-235.

Kim, J., Fiore, A. M., & Lee, H. H. (2007). Influences of online store perception, shopping enjoyment, and shopping involvement on consumer patronage behavior towards an online retailer. *Journal of Retailing and Consumer Services, 14*(2), 95-107.

Kim, B., Kim, S., & Heo, C. Y. (2016). Analysis of satisfiers and dissatisfiers in online hotel reviews on social media. *International journal of contemporary hospitality management, 28*(9), 1915-1936.

Kirk, R. E. (2013). Experimental Design. Procedures for the Behavioral Sciences. Sage Editors.

Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. K. (2016). From social to sales: The Effects of Firm-Generated Content in Social Media on Customer Behavior. *Journal of Marketing, 80*(1), 7-25.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lazarus, R. S. (1991). Cognition and motivation in emotion. *American Psychologist, 46*(4), 352-367.

Leonidou, L. C., Aykol, B., Hadjimarcou, J., & Palihawadana, D. (2021). Unveiling the infidelity problem in exclusive manufacturer–distributor relationships: A dyadic perspective. *Psychology & Marketing, 38*(11), 2122-2141.

Li, L. (2022). Reskilling and upskilling the future-ready workforce for industry 4.0 and beyond. *Information Systems Frontiers*, 1-16.

Lin, S.-Y., Chen, Y.-W., Kang, H.-C., Wu, Y.-J., Chen, P.-Z., Wu, C.-W., Lin, C.-S., Wu, F.-L. L., Shen, L.-J., Huang, Y.-M., & Huang, C.-F. (2021). Effects of a pharmacist-managed anticoagulation outpatient clinic in Taiwan: evaluation of patient knowledge, satisfaction, and clinical outcomes. *Postgraduate Medicine, 133*(8), 964-973.

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism management, 29*(3), 458-468.

Louvieris, P., & Driver, J. (2001). New frontiers in cybersegmentation: marketing success in cyberspace depends on IP address. *Qualitative Market Research: An International Journal, 4*(3), 169-181.

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science, 62*(12), 3412-3427.

Ludwig, S., & de Ruyter, K. (2016). Decoding social media speak: developing a speech act theory research agenda. *Journal of Consumer Marketing, 33*(2), 124-134.

Lwin, M.O., Wirtz, J., & Stanaland, A.J.S. (2016). The privacy dyad: Antecedents of promotion- and prevention-focused online privacy behaviors and the mediating role of trust and privacy concern. *Internet Research, 26*(4), 919-941.

Maaß, M., Clement, M.-P., & Hollick, M. (2021). Snail Mail Beats Email Any Day: On Effective Operator Security Notifications in the Internet. In *The 16th International Conference on Availability, Reliability and Security*, 11, pp.1-13. Vienna, Austria.

https://dl.acm.org/doi/10.1145/3465481.3465743
McCormack, A., Shorter, G. W., & Griffiths, M. D. (2014). An empirical study of gender differences in online gambling. *Journal of Gambling Studies, 30*, 71-88.
Mintz A.P. (2002). Web of deception: Misinformation on the Internet. New Jersey, USA: Information Today, Inc.,
Mohawesh, R., Xu, S., Tran, S. N., Ollington, R., Springer, M., Jararweh, Y., & Maqsood, S. (2021). Fake reviews detection: A survey. *IEEE Access, 9*, 65771-65802.
Moe, W. W., & Schweidel, D. A. (2012). Online product opinions: Incidence, evaluation, and evolution. *Marketing Science, 31*(3), 372-386.
Moon, S., Kim, M. Y., & Iacobucci, D. (2021). Content analysis of fake consumer reviews by survey-based text categorization. *International Journal of Research in Marketing, 38*(2), 343-364.
Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013). Fake review detection: Classification and analysis of real and pseudo reviews. UIC-CS-03-2013. Technical Report.
Noekhah, S., Fouladfar, E., Salim, N., Ghorashi, S. H., & Hozhabri, A. A. (2014, January). A novel approach for opinion spam detection in e-commerce. In Proceedings of the 8th *IEEE international conference on E-commerce with focus on E-trust* (Vol. 33). https://www.academia.edu/14458375/A_Novel_Approach_for_Opinion_Spam_Detection_in_E_Commerce
Oentaryo, R., Lim, E. P., Finegold, M., Lo, D., Zhu, F., Phua, C., ... & Berrar, D. (2014). Detecting click fraud in online advertising: a data mining approach. *The Journal of Machine Learning Research, 15*(1), 99-140.
Ott, M., Cardie, C., & Hancock, J. T. (2013, June). Negative deceptive opinion spam. In Proceedings of the 2013 *conference of the north American chapter of the association for computational linguistics: human language technologie*s (pp. 497-501). https://aclanthology.org/N13-1053.pdf
Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. In: *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies* vol 1. pp 309–319. https://aclanthology.org/P11-1032.pdf
Parekh, D., Amarasingam, A., Dawson, L., & Ruths, D. (2018). Studying jihadists on social media: A critique of data collection methodologies. *Perspectives on Terrorism*, *12*(3), 5-23.
Parikh S.B, & Atrey P.K. (2018). Media-rich fake news detection: a survey. In *2018 IEEE conference on multimedia information processing and retrieval (MIPR)*, Miami, FL, USA. https://www.researchgate.net/publication/325722738_Media-Rich_Fake_News_Detection_A_Survey
Park, C., & Lee, T. M. (2009). Information direction, website reputation and eWOM effect: A moderating role of product type. *Journal of Business Research, 62*(1), 61–67.
Pattee, J., Anik, S. M., & Lee, B. K. (2022). Performance Monitoring Counter Based Intelligent Malware Detection and Design Alternatives. *IEEE Access, 10*, 28685-28692.
Patton, M. Q. (1990). Qualitative evaluation and research methods. SAGE Publications, inc.
Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition, 188*, 39-50,
Postel, J. (1981). Internet Protocol. RFC 791 (Proposed Standard), Internet Engineering Task Force. https://www.ietf.org/rfc/rfc791.txt
Purani, K., Kumar, D. S., & Sahadev, S. (2019). e-Loyalty among millennials: Personal characteristics and social influences. *Journal of Retailing and Consumer Services, 48*, 215-223.

Rambocas, M., & Pacheco, B. G. (2018). Online sentiment analysis in marketing research: a review. *Journal of Research in Interactive Marketing, 12*(2), 146-163.

Rao, S., Verma, A. K., & Bhatia, T. (2021). A review on social spam detection: challenges, open issues, and future directions. *Expert Systems with Applications, 186*, 115742.

Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). Evaluation: A systematic approach. Sage publications.

Rotter, J. B. (1954). Social learning and clinical psychology. New York: Prentice-Hall.

Saboo, A. R., Kumar, V., & Ramani, G. (2016). Evaluating the impact of social media activities on human brand sales. *International Journal of Research in Marketing, 33*(3), 524-541.

Sáez-Ortuño, L., Huertas-Garcia, R., Forgas-Coll, S., & Puertas-Prats, E. (2023a). How can entrepreneurs improve digital market segmentation? A comparative analysis of supervised and unsupervised learning algorithms. *International Entrepreneurship and Management Journal*, 1-28.

Sáez-Ortuño, L., Forgas-Coll, S., Huertas-Garcia, R., & Sánchez-García, J. (2023b). Online cheaters: Profiles and motivations of internet users who falsify their data online. *Journal of Innovation & Knowledge, 8*(2), 100349.

Sáez-Ortuño, L., Forgas-Coll, S., Huertas-Garcia, R., & Sánchez-García, J. (2023c). What's on the horizon? A bibliometric analysis of personal data collection methods on social networks. *Journal of Business Research, 158*, 113702.

Sáez-Ortuño, L., Sanchez-Garcia, J., Forgas-Coll, S., Huertas-García, R., & Puertas-Prat, E. (2023d). Impact of Artificial Intelligence on Marketing Research: Challenges and Ethical Considerations. In L. Moutinho, L. Cavique, & E. Bigné (Eds), *Philosophy of Artificial Intelligence and Its Place in Society* (pp. 18-42). IGI Global.

Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems, 81*, 30-40.

Salminen, J., Kandpal, C., Kamel, A. M., Jung, S. G., & Jansen, B. J. (2022). Creating and detecting fake reviews of online products. *Journal of Retailing and Consumer Services, 64*, 102771.

Sannon, S., Bazarova, N. N., & Cosley, D. (2018). Privacy lies: Understanding how, when, and why people lie to protect their privacy in multiple online contexts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-13), Montreal, Canada. https://dl.acm.org/doi/pdf/10.1145/3173574.3173626

Savage, D., Zhang, X., Yu, X., Chou, P., & Wang, Q. (2015). Detection of opinion spam based on anomalous rating deviation. *Expert Systems with Applications, 42*(22), 8650-8657.

Septianto, F., & Garg, N. (2021). Grateful compliance: Gratitude effects on willingness to comply with responsible drinking messages. *Psychology & Marketing, 38*(9), 1460-1474.

Sestino, A., Kahlawi, A., & De Mauro, A. (2023). Decoding the data economy: a literature review of its impact on business, society and digital transformation. European Journal of Innovation Management. Vol. ahead-of-print No. ahead-of-print https://doi.org/10.1108/EJIM-01-2023-0078

Shu, K., Bhattacharjee, A., Alatawi, F., Nazer, T. H., Ding, K., Karami, M., & Liu, H. (2020). Combating disinformation in a social media age. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(6), e1385.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: a data mining perspective. *ACM SIGKDD Explorations Newsletter 19*(1), 22–36.

Shu, K., Mahudeswaran, D., Wang, S. H., Lee, D., & Liu, H. (2020). FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media. *Big Data, 8*(3), 171-188.

Song, C., Ning, N., Zhang, Y., & Wu, B. (2021). A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks.

Information Processing and Management, 58(1), 102437.
Swaminathan, V., Lepkowska-White, E., & Rao, B. P. (1999). Browsers or buyers in cyberspace? An investigation of factors influencing electronic exchange. *Journal of Computer-Mediated Communication, 5*(2), JCMC523.
Talwar, V., Gomez-Garibello, C., & Shariff, S. (2014). Adolescents' moral evaluations and ratings of cyberbullying: The effect of veracity and intentionality behind the event. *Computers in Human Behavior, 36*, 122-128.
Thakur, S., Meenakshi, Er., & Priya, A. (2017). Detection of malicious URLs in big data using ripper algorithm, In *2017 2nd IEEE international conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT)*(pp. 1296-1301). https://ieeexplore.ieee.org/document/8256808
Tirunillai, S., & Tellis, G. J. (2012). Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science, 31*(2), 198-215.
TripAdvisor. (2021). TripAdvisor Review Transparency Report 2021. Retrieved September 24, 2023, from https://www.tripadvisor.co.uk/TransparencyReport2021.
Tripathy, R. M., Bagchi, A., & Mehta, S. (2013). Towards combating rumors in social networks: Models and metrics. *Intelligent Data Analysis*, *17*(1), 149-175.
Tuomi, A., Tussyadiah, I. P., & Hanna, P. (2021). Spicing up hospitality service encounters: the case of Pepper™. *International Journal of Contemporary Hospitality Management*, *33*(11), 3906-3925.
Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146-1151.
Vo-Thanh, T., Zaman, M., Hasan, R., Rather, R. A., Lombardi, R., & Secundo, G. (2021). How a mobile app can become a catalyst for sustainable social business: The case of Too Good To Go. *Technological Forecasting and Social Change*, *171*, 120962.
Waggoner, P. D., Kennedy, R., & Clifford, S. (2019). Detecting fraud in online surveys by tracing, scoring, and visualizing IP addresses. *Journal of Open Source Software, 4*(37), 1285.
Wang, K., Xu, G., Wang, C., & He, X. (2017, August). A Hybrid Abnormal Advertising Traffic Detection Method. In 2017 *IEEE International Conference on Big Knowledge (ICBK)* (pp. 236-241). IEEE. https://ieeexplore.ieee.org/document/8023423
Wendling, M. (2018). The (almost) complete history of "fake news". Retrieved November 30, 2022 from https://www.bbc.com/news/blogs-trending-42724320
Wessel, M., Thies, F., & Benlian, A. (2016). The emergence and effects of fake social information: Evidence from crowdfunding. *Decision Support Systems, 90*, 75-85.
Wu, Y., Xu, Y., & Li, J. (2021). Fraudulent traffic detection in online advertising with bipartite graph propagation algorithm. *Expert Systems with Applications, 185*, 115573.
Wu, Y., Ngai, E.W.T., Wu, P., & Wu, C. (2022). Fake news on the internet: a literature review, synthesis and directions for future research. *Internet Research*, *32*(5), 1662-1699.
Xu, X. (2020). Examining an asymmetric effect between online customer reviews emphasis and overall satisfaction determinants. *Journal of Business Research, 106*, 196-210.
Zaman, M., Vo-Thanh, T., Nguyen, C. T., Hasan, R., Akter, S., Mariani, M., & Hikkerova, L. (2023). Motives for posting fake reviews: Evidence from a cross-cultural comparison. *Journal of Business Research, 154*, 113359.
Zhang, H., Alim, M. A., Li, X., Thai, M. T., & Nguyen, H. T. (2016). Misinformation in Online Social Networks: Detect Them All with a Limited Budget. *Acm Transactions On Information Systems*, *34*(3), 1-24.
Zhang, L. F., & Guan, Y. (2008). Detecting Click Fraud in Pay-Per-Click Streams of Online Advertising Networks. In 28th I*nternational Conference on Distributed Computing Systems*, Vols 1 and 2, Proceedings (pp. 77-84). Los Alamitos: Ieee Computer Soc.

https://ieeexplore.ieee.org/document/4595871
Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and resolution of rumours in social media: a survey. *ACM Computing Surveys, 51*(2), 1-36