Business Horizons (2022) 65, 765—776

Available online at www.sciencedirect.com

ScienceDirect

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY
www.journals.elsevier.com/business-horizons

All that glitters is not real affiliation: How to handle affiliate marketing programs in the era of falsity

Federico Mangiò [a], Giandomenico Di Domenico [b,*]

[a] University of Bergamo, Via dei Caniana 2 Bergamo, 24127 Italy
[b] Cardiff Business School, Cardiff University, Aberconway Building, Colum Road, Cardiff CF10 3EU, U.K.

KEYWORDS
Affiliate marketing;
Falsity;
Deceptive behaviors;
Social media;
Influencers

Abstract Affiliate or partnership marketing programs are a performance-based approach to online marketing whereby brands only pay when a sale occurs and is traced back to the affiliate who made it happen. Affiliate marketing programs conquered Web 2.0 and are now one of the most used channels for marketers and publishers online. Despite their success, affiliate marketing programs are exposed to different and problematic degrees of falsity, which finally threaten both consumers and brands. Acknowledging the lack of strategic and academic guidance about how to prevent and handle affiliate frauds, in this article we provide an original classification differentiating between noninfluencer and influencer falsity. We describe the direct and indirect costs that the various techniques belonging to each category cause and outline the best strategies that brands can implement to preserve their economic and reputational integrity. We then propose a two-stage affiliate listening protocol and show its applicability with an illustrative case on real influencer affiliate data. We offer several insights to marketers who need to manage their brands in an era of falsity, suggesting that continuous affiliate listening is needed to identify falsity in affiliate marketing programs and carefully select the affiliate influencers with whom to partner.

1. Affiliate marketing programs in the era of falsity

From Amazon to Instagram and Snapchat, from BuzzFeed to YouTube and Twitch, affiliate marketing programs have flooded Web 2.0, often even without us noticing. We often come across statements like "This content is sponsored by," or we hear our favorite content creators and social media influencers exclaiming: "Swipe up to take advantage of this incredible sale in my bio!" In these instances, chances are high that we are moving into the space of affiliate marketing

* Corresponding author
E-mail addresses: federico.mangio@unibg.it (F. Mangiò), didomenicog@cardiff.ac.uk (G. Di Domenico)

https://doi.org/10.1016/j.bushor.2022.07.001

programs. Amid the recent digital marketing revolution that has seen brands increasingly abandon owned media in favor of earned media, affiliate or partnership marketing programs represent one of the most dominant digital tools for online marketers; they have generated as much as 15% of global digital media revenues (CHEQ, 2021), and the great majority of marketing executives globally are eager to invest in this channel (Enberg, 2021).

These programs were first introduced and implemented with excitement by advertisers, who saw this tool as a safer means to implement online marketing (Edelman & Brandi, 2015). The initial adopters of affiliate marketing programs were small partners using their blogs or websites to earn money on commissions. Today, social media influencers have increasingly become an integral part of affiliate programs, raising the complexity of the affiliate marketing landscape and exposing brands to new, subtler perils. Affiliate marketing programs indeed show some structural flaws that mainly stem from the affordances of digital environments (Di Domenico et al., 2021), where fraudsters can develop and refine deceptive behaviors from digitally advanced techniques (e.g., cookie stuffing; Chachra et al., 2015) to more social media-sized frauds (e.g., engagement manipulation; Alba, 2019). Affiliate falsity threatens brands' image, reputation, and economic resources; in 2020, affiliate frauds cost brands $1.4 billion (CHEQ, 2021).

As the size of the affiliate marketing industry continues to grow—it was worth more than $15 billion in 2020 (CHEQ, 2021)—how can marketers protect their affiliate marketing programs from falsity? Affiliate frauds can take many forms. As such, there is no silver bullet for handling this problem, and the gray and academic literatures to date have failed to provide a meaningful characterization of affiliate frauds that would help brands to better understand all facets of this phenomenon and plan appropriate coping strategies. In this article, we provide an original classification of affiliate frauds based on the identity of the affiliate. In this sense, we distinguish between noninfluencer falsity and influencer falsity and describe how the various tactics belonging to each category impact brands. We also outline the appropriate strategies that brands can implement to identify affiliate fraud and preserve their economic and reputational integrity. Then, we propose a two-stage protocol that specifically helps brands to manage influencer affiliate falsity with the support of computer-aided textual analysis, or CATA (Brunzel, 2021). We conclude with an illustrative case in which this protocol is applied to real influencer affiliate data.

2. Affiliate falsity: What is it, and why does it matter?

An affiliate marketing program is a performance-based online marketing strategy whereby an actor (merchant) makes an agreement with another actor (affiliate or publisher) to feature a link from its websites on affiliated sites (Dwivedi et al., 2017). In particular, an affiliate earns a commission if a user browses an affiliate's site or social media account, clicks the affiliate's link to the merchant, and makes a purchase from the merchant (Edelman & Brandi, 2015). Initially, affiliate marketing programs were proclaimed "the holy grail of online advertising" (The Economist, 2005) as the pay-per-sale mechanism they are based on promised to liberate brands from blindly investing resources in advertising through the older pay-per-thousands mechanism. Previously, affiliates used to be small publishers who posted their affiliate links on websites, discussion forums, or blogs to redirect users to the merchants' websites (Enberg, 2021). However, social media influencers have increasingly become an integral part of digital marketing strategies, leveraging the influence they hold on their follower base to promote products and services (Leung et al., 2022). More importantly, they started to earn commissions within affiliate marketing programs, giving rise to the practice of influencer affiliate marketing (Bradley, 2021).

In this article, we adopt the distinction between influencer and noninfluencer affiliate marketing to illustrate how affiliates use falsity in both realms, instantiating affiliate frauds: activities which are explicitly forbidden under the terms and conditions of affiliate programs or by the law (Snyder & Kanich, 2016). Our distinction builds on the identity and modus operandi of the affiliate. Influencer affiliate marketing refers to instances in which the affiliate is a social media influencer: an individual with a considerable network of followers who creates and shares content on social media (Campbell & Grimm, 2019). Noninfluencer affiliate marketing programs are enforced by other actors who use different digital marketing tools such as websites, email marketing, or banner ads for their affiliate marketing activities. This distinction is relevant for two reasons:

1. Distinguishing between influencer and noninfluencer affiliate marketing helps us better

explain how affiliate frauds are perpetrated. While social media represents fertile ground for fraudster influencers, non-influencer affiliate frauds are realized by hidden fraudsters who exploit the shortcomings of the digital world outside of social media platforms.

2. Influencer and noninfluencer affiliate falsity have different impacts on brands and consumers. Noninfluencer affiliate falsity exerts a direct economic impact on brands due to the misattribution of sales and commissions to the deceptive affiliate; in these cases, consumers are usually unaware of the fraud being realized and are not impacted. Influencer affiliate falsity impacts brands directly and indirectly. The direct effect is due to deceptive influencers who buy fake followers and ask for higher compensation to promote the brand. The indirect effect instead passes through consumers, as a lack of transparency from the influencer can inhibit perceived trustworthiness and engagement with social media posts on the side of consumers (Karagür et al., 2022), ultimately hampering the performance of the campaign and potentially the brand's reputation.

Our classification helps to reconcile the knowledge about the different types of affiliate frauds, clarifying how they are carried out, their impact on brands and consumers, and the different solutions that brands can adopt to prevent them. Figure 1 summarizes our classification.

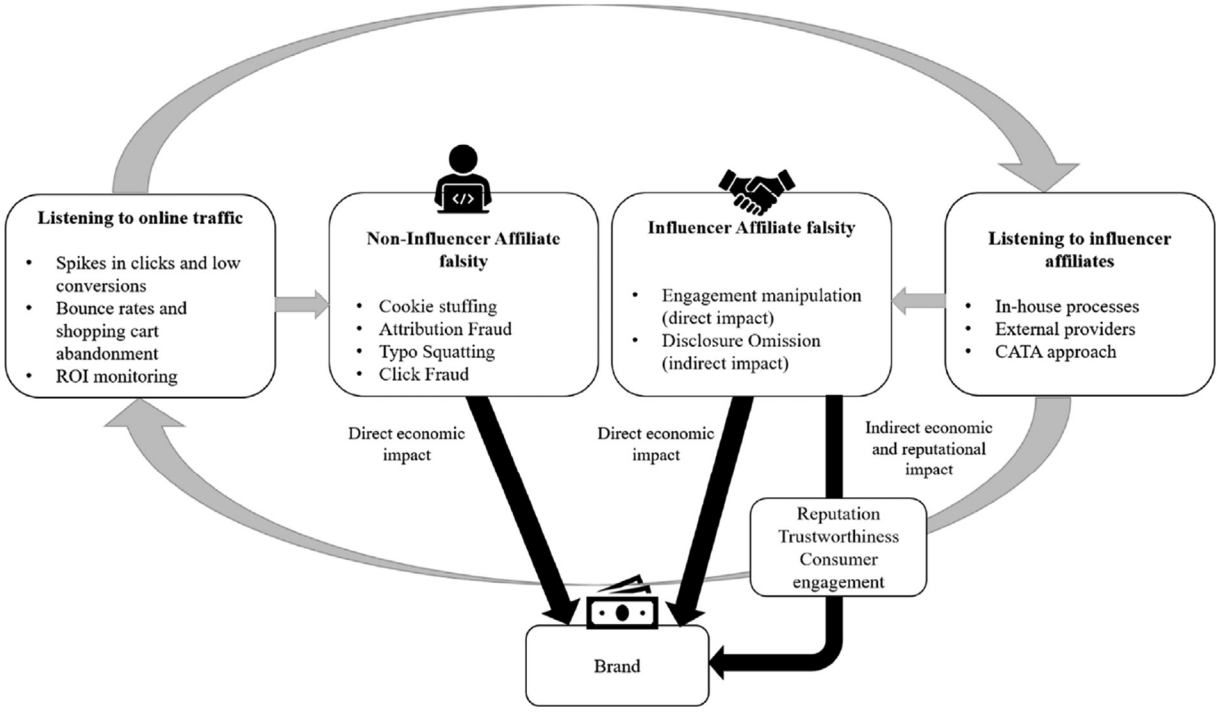
2.1. Noninfluencer affiliate falsity: Frauds and scams, from cookies to farms

Noninfluencer affiliate falsity is performed by fraudsters who, by exploiting or forcing technical shortcomings in the affiliate tracking and attribution systems, receive commissions they did not really earn. To date, these frauds have represented the main concern of marketers engaging in affiliate marketing programs as the misattribution of commissions results in a direct economic cost for the brand. Noninfluencer affiliate falsity involves many activities often undertaken by the same actor simultaneously and automatized through bots. The activities that most frequently affect brands are cookie stuffing, attribution frauds, typosquatting, and click frauds.

2.1.1. Cookie stuffing

Through cookie stuffing, fraudulent affiliates drop small HTTP files—called tracking cookies—from third-party advertisers onto the users' browsing history every time they visit their websites. In this way, if the user subsequently visits one advertiser

Figure 1. Affiliate falsity classification and brand coping strategies



and completes a purchase, the fraudster can claim a commission without actually having directed the user to the advertiser. Given their invisibility and ubiquity online, cookies can be dropped in multiple ways without getting noticed—for example, by including them in decoy pictures and redirect links. Usually placed on the websites of big retailers such as Walmart, Amazon, and eBay, cookie stuffing costs brands thousands of dollars every year. In 2014, an elaborated cookie stuffing scheme cost eBay $28 million in online marketing fees (Chachra et al., 2015). Despite the global digital marketing ecosystem eventually moving toward a cookieless future (Fou, 2021) with companies like Apple and Google planning a complete ban of third-party cookies, cookie stuffing is still going to represent a real threat to affiliate marketing—at least in the near future. Currently, not all global internet users navigate the web via cookie-free browsers like Safari and Firefox, and they are still easy targets for cookie-stuffing frauds.

2.1.2. Attribution frauds

Another way in which fraudsters manipulate the affiliate system is through the attribution fraud of fake app installs. This fraud allows deceptive affiliates to claim credits for app installs not generated by them performing sophisticated and subtle techniques. One of the most common is called click injection. Fraudsters develop a mobile app that, once installed by users on their smartphones, tracks the download of any other app. When fraudsters realize that an app has been downloaded, they generate new windows and force users to perform a series of clicks before the app installation is completed. In this way, the tracking system is deceived and the installation is attributed to the fraudulent source. Such activities infest ad networks with hundreds of thousands of malicious apps (Benes, 2018). For instance, Uber wasted more than $100 million in affiliate marketing investments due to attribution fraud (Silverman, 2018).

2.1.3. Typosquatting

The third type of noninfluencer affiliate falsity is called typosquatting. This illegal tactic involves bad actors registering online domains that show poor grammar or misspellings of an actual merchant's domain and tricking users into clicking on their link. Conversely, by clicking on hijacked URLs such as those depicted in Figure 2, the user is ultimately redirected to the merchant's website but the affiliate will collect a commission not rightfully earned. To avoid consumer backlash, some brands

Figure 2. Examples of common typosquatting techniques affecting affiliate programs

preventively register typo versions of their domains to anticipate fraudsters, but this is not always effective. For example, in 2006 Land's End proved in court that its affiliates registered a variety of domains misspelling the original Land's End website to earn commission from simply redirecting users to their website.

2.1.4. Click frauds

Lastly, we have click frauds. In the beginning, fraudsters would create computer programs specifically designed to generate fake clicks, also called click bots, to artificially inflate the revenues from affiliate marketing. Companies then started to protect themselves by applying CAPTCHA tools—systems intended to tell humans from machine inputs apart—on their websites to block such malicious click bots. However, fraudulent affiliates responded by creating even more sophisticated automated fraud schemes able to bypass CAPTCHA. Alongside this, fraudsters started to use humans to overcome the evolving antifraud systems as well. This is the case of human click farms, where real people click on ads, fill out forms, and even put items into online carts to trick marketers and merchants into thinking they are getting real leads. Usually, click farms are located in countries where the labor cost is minimal and outweighed by profits. Indeed, click farm workers in Bangladesh or India get paid as much as $120 per year (The Guardian, 2013), whereas the industry of click farms generates $152 billion yearly (CHEQ, 2021).

2.2. What can brands do about noninfluencer affiliate falsity?

As frauds can come in many forms, there is no one-size-fits-all solution for managers to protect their brands. However, monitoring the traffic quality represents a proper practice. There are three indicators to keep track of that might signal the

brand is under affiliate fraud attack. These are summarized as follows.

2.2.1. Spike in clicks and low conversions

If the number of clicks suddenly increases and it is not followed by a proportional increase in conversions, that very likely means that bots or click farms are in action. Managers should keep track of sources of traffic, as very unfamiliar sources or the same IP addresses can be evidence of bots or click farms.

2.2.2. High bounce rates and shopping cart abandonment

Brand managers might notice that many users immediately abandon the brand's website after a visit. The duration of the session provides insights to spot the action of malicious actors. If users stay on the website for zero seconds, then they are likely bots and the brand might be under the attack of click frauds. Sometimes bots are trained to disguise themselves as humans so that they fill a shopping cart, but since they cannot purchase anything, they bounce and abandon the cart. Thus, high shopping cart abandonment rates might signal bot activity.

2.2.3. Budget and falling ROI

Brand managers should always keep an eye on the budget and ROI of affiliate campaigns. The performance of affiliate campaigns could be affected by a variety of factors that managers should constantly monitor, including industry trends and other crises. If an unexpected bad performance cannot be attributed to any other factors, this should be a warning sign that the brand is being attacked via affiliate fraud.

Though it is not easy to accomplish, spotting and preventing noninfluencer falsity and fraud is possible through continuous and deep monitoring of traffic. Fortunately, brand managers have various third-party solutions available for them to better understand the effectiveness of their affiliate marketing campaigns. Among them, Anura.io (www.anura.io) and SEON.io (www.seon.io) offer brands the opportunity to develop tailored traffic monitoring solutions that aim to uncover hidden fraudsters and detect suspicious usage.

2.3. Influencer affiliate falsity: Mocking the system through engagement manipulation and disclosure omission

Influencer affiliate falsity occurs when a deceptive affiliate exploits the logics at the base of influencer marketing to earn undue affiliate commissions, brand promotions, and partnerships. These activities represent a direct and indirect cost for the brand as they may cause an erosion of reputation and consumer trust due to the association with deceptive and unlawful influencer affiliates (Leung et al., 2022). Different types of influencers populate the social media landscape, fulfilling different purposes (Bentley et al., 2021). Typically, the scale of influencers, ranging from nanoinfluencers (0—10,000 followers) to celebrity influencers (1+ million followers), affects their perceived level of authenticity and cultural impact, and it defines the relationship they have with their network (Campbell & Farrell, 2020). The influencer falsity tactics that we describe might be implemented by influencers at any level. However, in the domain of affiliate marketing frauds, several signs suggest brand managers should monitor smaller-scaled influencers. First, while the partnerships with celebrity influencers are regulated by well-established contracts, this is generally not the case with smaller influencers, who might escape the brand's control over how they operate—especially for brands owning large portfolios of influencers. Second, celebrity influencers already possess a large network of followers. Nanoinfluencers and microinfluencers might be more tempted to give an initial boost to their influencer activity—they may try, for example, purchasing fake followers—or try to preserve their perceived authenticity (Campbell & Farrell, 2020) by not disclosing the commercial nature of the post. Influencer affiliates perform falsity in two main ways: inflating the engagement metrics of social media (i.e., engagement manipulation) or concealing their commercial identity (i.e., disclosure omission).

2.3.1. Engagement manipulation

A brand's major concern in choosing which social media influencers to work with is the influencer's follower base and engagement rates (Leung et al., 2022). Fraudsters attempt to strategically manipulate these criteria to ask for higher compensation in partnership negotiations. As such, engagement manipulation represents a direct economic cost to the brand. Similar to click frauds, engagement metrics can be artificially inflated through both human and computer-based tactics.

Among the human tactics, one common fraud is sockpuppeting: the administration of plural, fake accounts by one single actual user. Hiding behind puppet profiles and pseudonyms, the scammers manage to interact at will with online content to amplify the metrics on which their income depends. Another mainstream human-based affiliate fraud involves lobbying activities executed by pod

communities: secret groups of online users who systematically endorse a mutual exchange of fictitious online engagement interactions during planned drops that exploit the affordances of specific social media platforms. For example, pod communities will share threads such as #likeforlikes or #followforfollows on Instagram, hacking the platform's ranking algorithm and placing them among the first results posts that record sky-high numbers of false likes overnight.

Computer-based fraud tactics involve the use of computer programs that grant the actual purchase of followers and the activity of bots that are specifically designed to artificially increase engagement metrics by creating false accounts. Despite attempts by social media platforms to curtail this engagement manipulation activity, it is still simple for users to purchase fake engagement. With as little as $330 it is possible to purchase over 3,500 comments, 25,000 likes, or 5,000 followers (Alba, 2019).

2.3.2. Disclosure omission

The second influencer affiliate falsity strategy, disclosure omission, is aimed at concealing the commercial nature behind the affiliate's online activity from their audience. To regulate the digital advertising environment, consumer protection authorities like the Federal Trade Commission (FTC) require content creators to clearly disclose to users their relationships with merchants every time an affiliate link is presented in online advertising content (Campbell & Grimm, 2019). All influencer affiliate content must include endorser-advertising disclosures. These can span from indirect affiliate link disclosures (i.e., merely specifying the merchant nature of the URLs embedded in the content) and channel support disclosures (i.e., promoting a financial contribution to the content creator from users to support their channel) to more explanatory disclosures in which the endorsers explicitly state they receive commissions upon clickthroughs (i.e., explanation disclosures). However, recent studies have found that less than 10% of affiliates on YouTube disclose the presence of affiliate links in their videos (Mathur et al., 2018). This not only violates international advertising regulations but also poses an indirect threat to the advertised brands that could suffer reputational damage for being associated with bad influencers. Recent evidence suggests that consumers today are more knowledgeable about the commercial nature of social media influencer posts (Statista, 2019) and, thus, expect the existence of a commercial/affiliate relationship between the brand and influencer even when not explicitly disclosed. Consequently, not disclosing a commercial partnership decreases the audience's perception of an influencer's trustworthiness and lowers engagement intentions (Karagür et al., 2022), which brings indirect negative consequences for the brand.

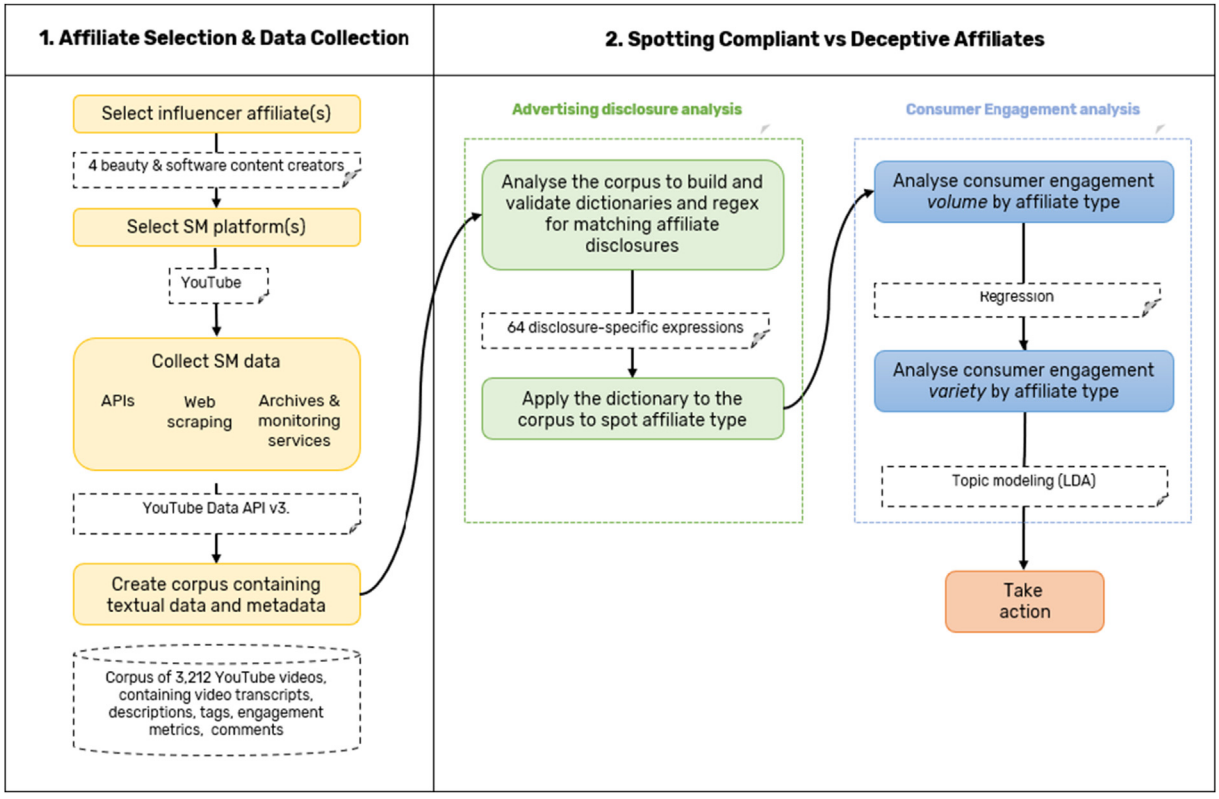
2.4. What can brands do about influencer affiliate falsity?

To face the challenge of managing the affiliate-influencer landscape, brands can rely on two options (Edelman & Brandi, 2015). The first involves implementing in-house processes aimed specifically at selecting, verifying, and monitoring everything that is said in the name of the brand by all the publishers, brand ambassadors, and influencers. For example, global consumer goods leader Unilever recently devised a multilayered internal procedure specifically designed to enhance its long-term relationship with the influencers of its many brands, ensuring that virtuous influencers are rewarded whilst inauthentic ones are staved off. Otherwise, influencer management tasks can be outsourced to external service providers. The market of influencer marketing platforms is fertile and expanding as many specialized providers like Upfeat (www.upfeat.com), SEON.io (www.seon.io), and Feedzai (www.feedzai.com) offer AI-driven solutions aimed at helping their clients to minimize the cost of influencer frauds with a relatively low effort from their clients. However, although useful, these solutions can be problematic for brands for two reasons. First, with social media analytics generally (Lee, 2018), not every brand can devote part of their marketing budget to outsourcing influencer fraud management processes. Second, even when managers rely on external providers, they would need to understand the underpinnings of the statistical and AI tools used by influencer marketing platforms to better evaluate their value proposition and avoid being deceived by the hype surrounding these buzzwordy technologies. To account for these issues, CATA offers an effective and affordable way for brands of all sizes to navigate the complex intersection between influencer marketing and affiliate marketing programs.

3. How to use CATA to prevent influencer affiliate falsity

Among the methods stemming from the intersection of computer science, linguistics, and the AI domain, CATA has recently gained academic and

Figure 3. Stages of influencer affiliate CATA-based listening protocol



practitioner attention (Brunzel, 2021). CATA comprises a wide array of techniques and tools, ranging from sentiment analysis and text categorization to information extraction that merges statistics, rule-based, and AI approaches for making replicable and valid inferences from textual data on a large scale. Bridging quantitative and qualitative methods, CATA not only outperforms traditional qualitative business research but also allows brands to extract additional insight from the ocean of textual data that in the past has been left untouched.[1] Today, brands successfully introduce CATA protocols into their marketing intelligence routines and use them for several scopes and purposes. Social media analytics, under which influencer analysis also falls, is precisely one of them (Delbaere et al., 2021; Lee, 2018). We suggest that brands can monitor their affiliate space independently and systematically through the two-stage CATA-based protocol shown in Figure 3. Stage 1 involves the selection of the influencer marketing campaigns, the social media platform where it takes place, and the collection of online textual data. Stage 2 focuses on how to spot compliant and deceptive influencers based on the digital traces they leave on social media platforms. Even though CATA can include very sophisticated techniques, we posit that the steps included in this protocol can be understood and performed by any social media manager with a basic background in data management and analysis, given that user-friendly and scalable visual-programming CATA software is increasingly available (Ordenes & Silipo, 2021).

3.1. Stage 1: Affiliate selection and online textual data collection

The first stage regards the identification and selection of the proper social media platform and influencer affiliates to focus on. Influencer affiliate marketing programs have flooded social media platforms that foster influencer-follower interactions through interactive content like TikTok, YouTube, or Instagram. The analyst can focus either on a single platform or, more simultaneously, as affiliate marketing campaigns are often performed on multiple platforms. Similarly, the analyst can decide whether to focus on the content related to one specific campaign and influencer, or multiple ones at the same time. The

[1] For a primer on CATA, see Brunzel (2021).

selection of a specific platform also involves some technical considerations regarding how to collect large volumes of brand-related textual data online. Indeed, analysts can do so in three ways. First, an analyst can engage an application programming interface (API), which is an interface conceived to let the analyst's software communicate with the target source where data is displayed. Today, all the main websites publicly offer access to their APIs with some limitations regarding the volume and type of data that can be accessed. When an API is not enough, it is possible to opt for web scraping: the automatic extraction of structured information from unstructured web sources. Though extremely flexible, web scraping is not always considered a legitimate activity from an ethical and legal perspective, as it can breach the data privacy and security policies of the target online source. One last convenient option is to resort to the online archival and monitoring services provided by external content marketing platforms such as CrowdTangle[2] and BuzzSumo,[3] which provide multiplatform interfaces able to return the most relevant influencers' data for target domains or keywords. Once the analyst identified the target influencers and the social media platform and gathered enough textual data through APIs, scraping, or archival sources, it is time to examine the influencer-follower interactions.

3.2. Stage 2: Spotting compliant and deceptive influencers

The second stage concerns spotting and determining who is a compliant influencer and who instead might be a fraudster. To do so, brands can follow the digital traces that fraudsters leave behind online. Some of these red flags can be checked manually, for example, by examining the influencers' accounts to check for missing information in the bios, strange or misspelled usernames, and geographical locations far away from the actual market served. However, in a big data environment, a sounder strategy is to implement automated CATA approaches to detect fraud. We specifically focus on two: advertising disclosure and consumer social media engagement analysis.

Advertising disclosure analysis aims to tell regulation-compliant content creators and fraudsters apart by mapping the presence of affiliate disclosures in endorsed content through processes that count the presence or absence of disclosure statements. The analysis proceeds in two steps. First, the analyst performs a preliminary qualitative inspection of the textual data gathered, looking for textual patterns through which affiliates disclose the nature of their partnership (e.g., "I can receive commission if you click on this link"). In this way, the analyst proceeds to build what is called a custom dictionary: a textual list containing recurring affiliate links, channel support, explanation disclosure statements, or a regular expression ("regex") able to match them. Second, once validated, these dictionaries and rules are processed by wordcount software—such as LIWC[4], Provalis Researcher's Wordstat[5], or Gate[6]—to check the presence of disclosure statements in the areas of the entire influencer-generated textual data supposed to contain affiliate links (e.g., YouTube video description sections, Instagram or TikTok content captions). In this way, the analyst can determine which affiliate is compliant with online advertising regulations.

The second fraud detection analysis involves systematic assessment of the consumer engagement generated on the social media platform. Consumer engagement is a multidimensional phenomenon of particular interest to brands and marketers due to its predictive power on consumer and firm outcomes (de Oliveira Santini et al., 2020). In social media contexts, it is commonly operationalized and tracked through the accumulated volume of likes, comments, and shares that specific brand-related content records. Different types of influencers are characterized by different follower bases and different engagement relationships (Britt et al., 2020; Campbell & Farrell, 2020). Thus, analysts can control whether consumer engagement is aligned with expectations in terms of two dimensions: volume and variety.

3.2.1. Volume

Given that influencer affiliate fraudsters inflate their engagement metrics—through, among others, sockpuppets, bots, and pods—they hardly create sustained engagement interactions in terms of volume of likes, favorites, and comments with their audiences. Therefore, spikes in follower counts reached overnight should represent an alarm signaling the presence of a fraudster, along with distribution of engagement metrics that differ too much from the ordinary, or followers-to-engagement ratios too large given the actual size of the influencer's network.

[2] www.crowdtangle.com
[3] www.buzzsumo.com
[4] http://liwc.wpengine.com
[5] www.provalisresearch.com
[6] www.gate.ac.uk

3.2.2. Variety

Analysts should also consider the affective and semantic variety of consumer engagement interactions (i.e., the actual content of the user-generated comments). Microinfluencers tend to build and maintain more intimate connections with their followers, engaging in considerably more two-way and personalized interactions than their mega counterparts (Britt et al., 2020). For example, the comments generated by pod groups tend to be very generic and decontextualized, using low-informative emojis and generic comments (e.g., "love this"). Such textual patterns can be identified automatically via topic discovery algorithms (e.g., topic modeling) and traced back to the creators they are associated with. Luckily, to perform these analyses, the analyst does not necessarily need an in-house data science function. Today, both commercial and noncommercial software-as-a-service providers[7] offer point-and-click, visual programming platforms to perform CATA without any kind of coding requirement, making it easier to include these tools in business intelligence operations.

4. An illustration: Influencer affiliate analysis on YouTube

The next section shows how the protocol illustrated above can be performed with real influencer affiliate data. For analysis, we selected eight popular content creators operating in two industries—beauty/cosmetics and consumer software—in which affiliate marketing is a predominant advertising strategy. Then, we proceeded with the two-stage protocol. First, we decided to focus on YouTube. To collect the textual data, we called YouTube Data API v3 using the R software. This enabled us to create a dataset in which each row corresponds to one of the 3,212 videos posted over time by the content creators, and each column contains textual data about the videos and the related engagement metrics (Table 1). Then, we applied advertising disclosure and consumer social media engagement analyses to identify compliant and fraudster influencers.

As for the first analysis, we created a dictionary containing words and sentences expressing affiliate disclosures and applied it to the description section of the YouTube videos. This allowed us to identify the videos from compliant (1,848) and deceptive (1,465) influencers. For the consumer engagement analysis, we firstly analyzed the volume of consumer engagement generated by the two types of affiliates with regression analysis. A dummy indicating the affiliate influencer type (compliant vs. deceptive) based on whether the video included an affiliate disclosure or not served as an independent variable, along with other relevant control variables—namely, the complexity, length, and type of content, measured respectively via words and prepositions count, duration of the videos, and YouTube tags and video category associated. Consumer engagement—operationalized as the sum of views, net likes, and comments count—served as the dependent variable. A negative binomial regression was run on weekly aggregated data. Figure 4 shows the effects of these variables on the log count of consumer engagement as well as their significance, indicated by the error bars. As the coefficient of compliant influencers is significantly higher than that of their deceptive counterparts, keeping the effect of control variables constant, the regression analysis confirms that compliant content creators can generate more volume of consumer engagement than deceptive ones.

Finally, we further analyzed the variety of consumer engagement triggered by the two types of content creators by performing an automated detection of the topics discussed by their followers in about 300,000 unique comments left below the videos. We applied an extension of the same topic modeling algorithm that can be quickly mobilized from programs like Knime or Leximancer. We then statistically tested whether the identified topics are more or less strongly associated with the type of content creators. We identified 11 unique topics discussed by users in reaction to the videos. The results of this analysis allow the analyst to disentangle with more granularity the variety of consumer engagement the content creators can elicit on the platform. Figure 5 shows how prevalent each of the 11 topics identified is among the comments left below the videos of the two types of influencer affiliates, which lay at the two sides of a continuum. Compliant content creators generate more engaged reactions than their counterparts. For example, followers are likely to express gratitude, thanking the influencers for their contents or sharing engaged suggestions about the products being advertised. Conversely, prevalent among the comments to the videos of deceptive content creators are very cold topics

[7] Commercial examples include Provalis Research's Wordstat (www.provalisresearch.com), MeaningCloud (www.meaningcloud.com), and Leximancer (https://www.leximancer.com); Noncommercial examples include Knime Analytics (www.knime.com) and RapidMiner (www.rapidminer.com) text mining extensions.

Table 1. Number of collected videos and mean engagement metrics[a]

Category	Videos	Views	Net likes	Comments	Duration (sec)
Beauty	1,577	477,863 (893,199)	19,091 (30,517)	1,496 (2,282)	737 (382)
deceptive	*625*	*818,534 (1,213,799)*	*31,964 (37,593)*	*2,422 (2,636)*	*749 (361)*
compliant	*952*	*254,208 (478,609)*	*10,569 (20,825)*	*790 (1,652)*	*730 (395)*
Software	1,735	181,271 (442,802)	5,214 (13,083)	258 (555)	735 (469)
deceptive	*811*	*245359 (532,603)*	*8,797 (16,997)*	*302 (455)*	*803 (584)*
compliant	*924*	*125021 (335,620)*	*2,070 (6,912)*	*219 (627)*	*676 (327)*

[a] Standard deviation

that share deal-oriented tones (e.g., "How do contributors to free sites get rewarded?"; "Use my links, please!" [topic "links"]) or very generic ones, typical of pods communities ("love it," "OMG" [topic "generic"]).

The illustrative case shows that CATA can be used to build straightforward but predictive protocols that can help brand managers to identify and prioritize affiliate influencers by analyzing in a big-data-friendly way the relationship between their contents and consumer engagement. Although more complex protocols and models are available, relevant signals and trends that should sound an alarm regarding the presence of deceptive behaviors in online brand-related contexts can be quickly and automatically grasped already through these CATA. In particular, by using this protocol, we discovered that compliant content creators stimulate more engaged reactions in their followers, who not only show appreciation and gratitude for the influencer's activity but are also engaged more in meaningful brand- or product-related discussions. On the other hand, the engagement stimulated by deceptive content creators likely comes from pod communities, is less authentic, and is more oriented toward exploiting community-based mechanisms in the affiliate environment. Even if in this illustrative case we focused on the comparison of two relevant groups of affiliate influencers, the same approach can be easily adapted for other affiliate management analyses illustrated in this article, such as the detection of overnight spikes in followers count by including time variables in the CATA. In the same vein, the proposed techniques can spot fraudulent activities enforced by any scale influencers.

5. Listen, act, and repeat

In an ever-more content-driven digital economy, affiliate marketing programs present real and florid opportunities for brands to reach consumers in new, meaningful ways. As the size of ecommerce is steadily growing, so are revenues from and investments in affiliate marketing (CHEQ, 2021). Relatedly, affiliate partnerships represent a key monetization source for social media influencers of all sizes (Enberg, 2021). Just recently, the leading social media platform Instagram (2021)

Figure 4. Effects of affiliate influencer type and controls on consumer engagement

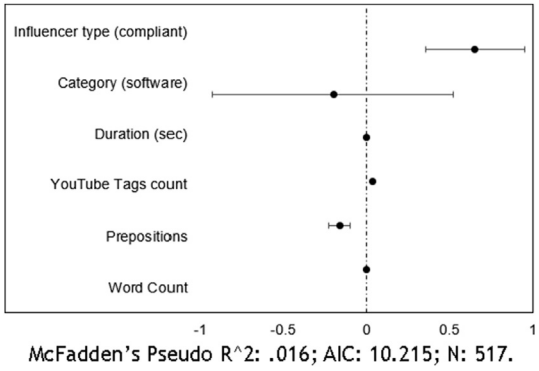
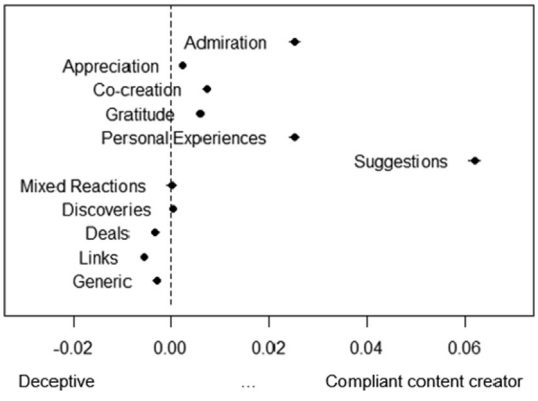


Figure 5. Prevalence of consumer discussion topics, by affiliate influencer type



launched a native affiliation tool that allows users to include affiliate links into their stories and sell their merchandise. Therefore, brand managers should be proactive in their approaches to these digital marketing tools, seizing the vast opportunities they offer. This, first and foremost, entails being able to protect the brand from the different types of deceptive behaviors that both influencer and noninfluencer affiliates put into practice online. To effectively manage their brands in an era of increasing falsity, we suggest that brand managers should follow the listen, act, and repeat guideline. By listening to specific indicators relative to online traffic (e.g., bounce rates, shopping cart abandonment, spikes in clicks, low conversions), as well as to the affiliates' online activities (e.g., volume and variety of consumer engagement generated), managers will not only get a more meaningful understanding of their digital marketing dynamics and performance but also disentangle potential frauds from deceptive affiliates in a short time. Not all frauds are created equal, and thus recognizing in which domain of affiliate falsity a brand is eventually trapped allows fine tuning of the most proper copying strategies. Also, systematically monitoring the affiliates' activity is pivotal to keeping the brand under control and preventing cost and reputational damages from escalating. This should be done, though, without hampering the partners' ability to create content freely and creatively. Among the different tools and techniques brands have at their disposal to analyze the influencer affiliate ecosystem, CATA approaches precisely allow monitoring in almost real-time, but without being intrusive. By adopting the protocol we propose in this article, marketers can also develop a profound knowledge of their influencer affiliates and be able to select the compliant ones, amplifying the reach of the brand, opening more market opportunities, and protecting themselves from being associated with deceptive affiliates. Finally, and most importantly, this whole process should not be intended as exclusive and one-shot, as the various monitoring solutions should be applied in parallel and constantly to protect brands more efficiently.

References

Alba, D. (2019, December 6). Fake 'likes' remain just a few dollars away, researchers say. *The New York Times*. Available at https://www.nytimes.com/2019/12/06/technology/fake-social-media-manipulation.html

Benes, R. (2018, November 19). Is app-fraud install on the rise? *eMarketer*. Available at https://www.emarketer.com/content/is-app-install-fraud-on-the-rise

Bentley, K., Chu, C., Nistor, C., Pehlivan, E., & Yalcin, T. (2021). Social media engagement for global influencers. *Journal of Global Marketing, 34*(3), 205–219.

Bradley, S. (2021, December 28). How influencers use affiliate marketing programs to make money and how much they earn. *Business Insider*. Available at https://www.businessinsider.com/how-influencers-make-money-using-affiliate-marketing-programs-examples-2021-7?r=US&IR=T

Britt, R. K., Hayes, J. L., Britt, B. C., & Park, H. (2020). Too big to sell? A computational analysis of network and content characteristics among mega and micro beauty and fashion social media influencers. *Journal of Interactive Advertising, 20*(2), 111–118.

Brunzel, J. (2021). Making use of quantitative content analysis: Insights from academia and business practice. *Business Horizons, 64*(4), 453–464.

Campbell, C., & Farrell, J. R. (2020). More than meets the eye: The functional components underlying influencer marketing. *Business Horizons, 63*(4), 469–479.

Campbell, C., & Grimm, P. E. (2019). The challenges native advertising poses: Exploring potential Federal Trade Commission responses and identifying research needs. *Journal of Public Policy and Marketing, 38*(1), 110–123.

Chachra, N., Savage, S., & Voelker, G. M. (2015). Affiliate crookies: Characterizing affiliate marketing abuse. In K. Cho & K. Fukuda (Eds.), *Proceedings of the 2015 Internet Measurement Conference* (pp. 41–47). New York, NY: Association for Computing Machinery.

CHEQ. (2021). *The economic cost of bad actors on the internet*. Available at https://cheq.ai/wp-content/uploads/2021/12/Economic-cost-of-affiliate-fraud-2020-Report-12.pdf

de Oliveira Santini, F., Ladeira, W. J., Pinto, D. C., Herter, M. M., Sampaio, C. H., & Babin, B. J. (2020). Customer engagement in social media: A framework and meta-analysis. *Journal of the Academy of Marketing Science, 48*, 1211–1228.

Delbaere, M., Michael, B., & Phillips, B. J. (2021). Social media influencers: A route to brand engagement for their followers. *Psychology and Marketing, 38*(1), 101–112.

Di Domenico, G., Sit, J., Ishizaka, A., & Nunan, D. (2021). Fake news, social media, and marketing: A systematic review. *Journal of Business Research, 124*, 329–341.

Dwivedi, Y. K., Rana, N. P., & Alryalat, M. A. A. (2017). Affiliate marketing: An overview and analysis of emerging literature. *The Marketing Review, 17*(1), 33–50.

The Economist. (2005, September 29). *Pay per sale*. Available at http://www.economist.com/node/ 4462811

Edelman, B., & Brandi, W. (2015). Risk, information, and incentives in online affiliate marketing. *Journal of Marketing Research, 52*(1), 1–12.

Enberg, J. (2021, March 12). Influencer monetization 2021. *eMarketer*. Available at https://www.emarketer.com/content/influencer-monetization-2021

Fou, A. (2021, March 27). The 'cookieless future' and what it means for marketers. *Forbes*. Available at https://www.forbes.com/sites/augustinefou/2021/03/27/the-cookieless-future-and-what-it-means-for-marketers/?sh=527bbe676808

The Guardian. (2013, August 2). *How low-paid workers at 'click farms' create appearance of online popularity*. Available at https://www.theguardian.com/technology/2013/aug/02/click-farms-appearance-online-popularity

Instagram. (2021). *New ways for creators to make a living*. Available at https://about.instagram.com/blog/announcements/creator-week-2021-new-ways-for-creators-to-make-a-living

Karagür, Z., Becker, J. M., Klein, K., & Edeling, A. (2022). How, why, and when disclosure type matters for influencer marketing. *International Journal of Research in Marketing, 39*(2), 313—335.

Lee, I. (2018). Social media analytics for enterprises: Typology, methods, and processes. *Business Horizons, 61*(2), 199—210.

Leung, F. F., Gu, F. F., & Palmatier, R. W. (2022). Online influencer marketing. *Journal of the Academy of Marketing Science, 50*, 226—251.

Mathur, A., Narayanan, A., & Chetty, M. (2018). An empirical study of affiliate marketing disclosures on YouTube and Pinterest. *arXiv*. Available at https://doi.org/10.48550/arXiv.1803.08488

Ordenes, F. V., & Silipo, R. (2021). Machine learning for marketing on the KNIME Hub: The development of a live repository for marketing applications. *Journal of Business Research, 137*, 393—410.

Silverman, C. (2018, November 26). These hugely popular Android apps have been committing ad fraud behind users' backs. *Buzzfeed News*. Available at https://www.buzzfeednews.com/article/craigsilverman/android-apps-cheetah-mobile-kika-kochava-ad-fraud

Snyder, P., & Kanich, C. (2016). Characterizing fraud and its ramifications in affiliate marketing networks. *Journal of Cybersecurity, 2*(1), 71—81.

Statista. (2019). *Number of brand sponsored influencer posts on Instagram from 2016 to 2020*. Available at https://www.statista.com/statistics/693775/instagram-sponsored-influencer-content/