



Old crimes reported in new bottles: the disclosure of child sexual abuse on Twitter through the case #MeTooInceste

Jesús C. Aguerri¹ · Lorena Molnar² · Fernando Miró-Llinares¹

Received: 7 July 2022 / Revised: 16 January 2023 / Accepted: 18 January 2023 / Published online: 1 February 2023
© The Author(s) 2023

Abstract

Movements such as #MeToo have shown how an online trend can become the vehicle for collectively sharing personal experiences of sexual victimisation that often remains unreported to the criminal justice system. These social media trends offer new opportunities to social scientists who investigate complex phenomena that, despite existing since time immemorial, are still taboo and difficult to access. They also bring technical difficulties, as the challenge to identify reports of victimisation, and new questions about the characteristic of the events, the role that victimisation testimonies play and the capacity to detect them by analysing their characteristics. To address these issues, we collected 91,501 tweets under the hashtag #MeTooInceste, posted from the 20 to 27 January 2021. A model was fitted using Latent Dirichlet Allocation that detected 1688 tweets disclosing experiences of child sexual abuse, with an accuracy of 91.3% [$\pm 3\%$] and a recall of 93.1% [$\pm 5\%$]. We performed Conjunctive Analysis of Case Configurations on the tweets identified as disclosures of victimisation and found that long tweets posted by users with small accounts, without URL or picture, were more likely to be related to disclosure of child sexual abuse. We discuss the possibilities of these trends and techniques offer for research and practice.

Keywords Social media · Dark figure of crime · Cyber-reporting · Child sexual abuse · Victimisation

1 Introduction

In January 2021, a trend emerged in France on the social network Twitter under the hashtag #MeTooInceste. This trend occurred after the well-known French political scientist and TV commentator—Olivier Duhamel—was accused of sexually abusing his adolescent 14-year-old stepson for approximately 2 to 3 years. The accusation came from Kouchner (2021), the sister of the supposed victim, who published a book called “La familia grande”, in which she mentions these acts. From this moment, there was an avalanche of tweets with the hashtag #MeTooInceste, in which many victims of child sexual abuse disclosed online victimisation they had endured during childhood and adolescence (Méheut 2021). The hashtag was amongst the top trending topics on

twitter for 3 days, and during its 4 days of activity, it reached a total of 91,501 tweets. In the context of #MeTooInceste, it was argued that the reason where the victims were reporting their experiences was to inspire others to break their silence, to be recognised by society as victims, and to seek catharsis through the revelations. Nonetheless, the movement also received criticism, namely the risks of falsely accusing someone in the public arena, as well as the possible negative or traumatic consequences for other victims of child sexual abuse who have not yet reported and who might be in a vulnerable position (Radio Télévision Suisse 2021; Watson.ch 2021).

In recent decades, society has made a notable transition from the physical world to cyberspace (Lupton 2015; Twenge et al. 2019), as evidenced by these online social movements (Fileborn 2022). In this sense, the internet has become a new public sphere (Papacharissi 2002) and societal debate no longer occurs in the *agora* but on social media platforms such as Twitter. The online discussions about sexual victimisation which followed the revelations of Kouchner are not an isolated event but the continuation of a trend of online mass disclosures from survivors of sexual harassment, abuse or assault which was originally started by Tarana Burke in

✉ Jesús C. Aguerri
j.aguerri@crimina.es

¹ CRÍMINA Research Center for the Study and Prevention of Crime, Miguel Hernández University of Elche, Avda. de la Universidad s/n. Edif. Hélike, 03201 Elche, Spain

² École des Sciences Criminelles, Université de Lausanne, Lausanne, Switzerland

2007 with the popular #MeToo, relaunched in October 2017 by actress Alyssa Milano vis-à-vis the allegations of sexual assault against Hollywood producer Harvey Weinstein (The Guardian 2017). Other related events were, for instance, the Canadian #AggressionNonDénoncée in 2014, the French #BalanceTonPorc (the equivalent to the #MeToo movement) and the worldwide #MeTooGay (Huffington Post 2021). The hashtag #MeTooInceste has notable similarities with these trends, but its national character, its clear trigger and the particularity of dealing with a crime with such a dark figure make it a particularly relevant object of analysis. This is due to the vulnerability of the victims—underaged individuals in most cases—the social condemnation of such kinds of crimes (Jahnke et al. 2015), as well as the underreporting of child sexual abuse (Chandran et al. 2020).

Given this context, this article addresses this novel trend of child sexual abuse reporting on social media, and our aim is two-fold. First, to test whether it is possible to identify with automated techniques of the disclosures of experiences of child sexual abuse amongst the trend #MeTooInceste. Second, to identify patterns amongst the posts that disclose child sexual abuse. In our view, both the identification of tweets and of patterns are relevant to better inform victim detection policies that can increase victims' support as well as to decrease the dark figure of child sexual abuse.

In the following sections, we first review the existing studies that have addressed the topic; then, we present the methodology—based on *Latent Dirichlet Allocation* (LDA) and *Conjunctive Analysis of Case Configurations* (CACC)—that we used for detecting and analysing the tweets belonging to #MeTooInceste. Third, we illustrate our findings, namely the classification algorithm that we constructed and the combinations of metadata that are most frequently associated with tweets disclosing victimisation experiences, which will allow us to observe the role that these testimonies play within the trend towards This issue will be addressed using, as far as possible, the metadata of the tweets and without quoting any of the tweets, due to the ethical implications of this kind of data (Ayers et al. 2018; Fiesler and Proferes 2018), especially in research dealing with such a sensitive topic as child sexual abuse. We conclude by discussing the implications of our results regarding research on social media of child sexual abuse disclosure and other victimisation experiences, as well as the possibilities for intervention that these kinds of techniques offer.

2 Child sexual abuse and its traditionally high dark figure

In the context of #MeTooInceste, social media users discussed child sexual abuse as a synonym of incest. Nonetheless, nuances should be drawn on these two terms because

they do not possess the same meaning. On one hand, incest refers to both consensual and non-consensual sexual intercourse between close blood relatives (Kar and Swain 2020). Therefore, sexual intercourse between two siblings who are both adults is qualified as incest. On the other hand, child sexual abuse is defined as “the involvement of a child in sexual activity that he or she does not fully comprehend, is unable to give informed consent to, or for which the child is not developmentally prepared and cannot give consent, or that violates the laws or social taboos of society” (WHO 2003: 75). In this sense, child sexual abuse within the family would be qualified as non-voluntary incest with a child belonging to the family. Therefore, the movement *Me too incest* refers to the latter definition of incest when it is non-voluntary and committed with a child belonging to the perpetrator's family. Although they are different concepts, as illustrated above, we use herein *incest* when referring to the hashtag.

In general, sexual crimes have a high dark figure, but child sexual abuse dark figure is especially high, being one of the offences with the highest dark figure due to the many barriers to reporting that victims face (Alaggia et al. 2019; Pereda et al. 2016, 2018). Consequently, disclosure often happens during adulthood, in many cases when the offence has already legally expired (Alaggia et al. 2019; Bennett and O'Donohue 2014; Kennedy and Prock 2018; Mojallal et al. 2021; Cyr et al. 2002; Ohlert et al. 2017; Giglio et al. 2011; Ullman 2007; Vollman 2021; WHO 2003; Zalberg 2017). Furthermore, when the perpetrator of sexual abuse is a family member, the victims report later, the abuse is likely to be more severe than that suffered by acquaintances or strangers, and it is likely that the incest begins much earlier but lasts longer (Ullman 2007). Moreover, during adulthood, survivors of incest face new challenges and dilemmas, such as to whom, when and how to report after so much time has passed (Tener and Murphy 2015). In general, the consequences for victims of child sexual abuse are extensive and they manifest higher risk of suffering mental health problems later in life (see reviews by Hillberg et al. 2011; Ohlert et al. 2017; Wolf and Pruitt 2019). Moreover, Pereda and Gallardo-Pujol (2011) found that victims of child sexual abuse even suffer neurobiological alterations which have a long-term negative impact both in childhood and in adulthood.

Despite the dark figure of the phenomenon, the known prevalence of child sexual abuse is not insignificant. Pereda et al. (2009) carried out a meta-analysis on the prevalence of child sexual abuse ($N=65$) and found that overall, in many continents, the mean prevalence of child sexual abuse in men was around 7% and 19% in women. The highest prevalence was reported in Africa (34%) and the lowest in Europe (9%). One contemporary French study contradicts the earlier finding and notes that 10% of French people have been

victims of child sexual abuse at least once in their lives; although this later study employs a broad definition of incest, including behaviours such as *inappropriate touching*, *exhibitionism* or *being exposed to sexual secrets*, as well as rape (IPSOS 2020). French police data on sexual child abuse indicate that in 2016, there were 13 suspected offenders per 100,000 inhabitants (Aebi et al. 2021)—the third-highest rate of offenders amongst the member countries of the Council of Europe—which remains relatively consistent at the conviction stage. Amongst the people convicted of sexual abuse of a minor, 31.5% were other underage individuals.

In brief, in spite of the efforts of public authorities to encourage victims to report child sexual abuse, there is a large dark figure and relatively high prevalence. These two characteristics seem inherent and unchanged over time, partly due to the highly sensitive nature of the offence, its family settings, the vulnerability of the victims, as well as the imbalance of power between victims and perpetrators.

3 Online disclosure of sexual victimisation

Research on the reporting of child sexual abuse through social media is rather scarce, with the noteworthy recent exceptions of Alaggia, et al. (2019), Lusky-Weisrose et al. (2022), Vollman (2021) and Idoiaga Mondragon et al. (2002). Alaggia et al. (2019) reviewed the research ($N=33$) on the facilitators and barriers of child sexual abuse disclosure, finding that barriers—shame, self-blame and fear—were more prevalently researched than facilitators. Interestingly, disclosure was found to be more likely to occur within a dialogical context activated by discussions about abuse or prevention forums providing information about sexual abuse. Lusky-Weisrose et al. (2022) studied disclosure of child sexual abuse that was committed by officials belonging to the Ultraorthodox Jewish Community in Israel. Through an analysis of Facebook posts, they were identified the factors that enabled survivors to overcome cultural barriers and disclose their experiences. In particular, their motives were related to seeking emotional and practical advice, revenging themselves by shaming their perpetrators, the desire for empowerment, preventing the perpetrators of committing further abuse, as well as the desire to promote change in their community by spreading awareness and knowledge. Conversely, the barriers that rendered difficult their disclosure were shame, uncertainty about the occurrence of the abuse, fear of social critiques, fear of triggering other survivors with their post, the taboo nature of sex ‘and prohibitions against reporting to secular authorities, speaking ill of a fellow Jew, degrading a rabbi, and the very use of social media’ (Lusky-Weisrose et al. 2022, p. 11). Vollman (2021) conducted a qualitative study about the testimonies of men that have been victims of both intra- and extra-familial

sexual abuse in childhood. The scholar collected the content of English-speaking websites from 2018 by hand and analysed them through *content analysis*. The author suggested that the victims posted their abuse on the internet in order to find peace, as well as to regain control of their lives in adulthood. This research also suggests that under the original, #MeToo posts have served a variety of purposes such as (1) disclosure of victimisation, (2) expressions of solidarity with the victims, (3) criticism of the #MeToo movement, (4) impersonal information and (5) expression of an opinion about the topic. In link to our study, Idoiaga Mondragon et al. (2002) used a lexical clustering system implemented in the Iramuteq software to extract the topics from more than 27,000 tweets with the hashtag #MeToo-Inceste in order to classify them and highlight the manner in which the hashtag offered to the victims a space to disclose their experiences and to find support. The researchers found that the different tweets could be classified into five classes, i.e.: (i) *Public incest scandals*—references to the scandal (see Sect. 1) that originated the hashtag—(22.78% of the tweets), (ii) *The horror of the victims*—disclosure of past experiences of child sexual abuse—(13.33%), (iii) *Call for change in child protection policies* (21.11%), (iv) *The #MeTooInceste tsunami on Twitter*—tweets that discussed the success of the hashtag—(22.22%) and (v) *Breaking the silence* (20.56%)—tweets about the importance of breaking the silence in regards to child sexual victimisation.

On the other hand, disclosure of sexual victimisation has been extensively studied on adult samples. These studies demonstrate that social networks are operating as a new context for movements that are based on the reporting of traditionally neglected victimisations and that provide victims with new opportunities to make their experiences public. It is nevertheless to note that child sexual abuse and adult sexual abuse are phenomena of a different nature, and the victim’s motivations for disclosure are not necessarily similar. Hence, their manifestation in social media may have notable differences. Nevertheless, the studies conducted on adult sexual victimisation can shed light on methods and patterns as well as the social responses of other Twitter users. Many studies focussed on the original *Me too movement*, framing their analyses on the first months of the reappearance of this trend, between 2017 and 2018 (Bogen et al. 2019; Brünker et al. 2020; Hosterman et al. 2018; Khatua et al. 2018; Li et al. 2020; Manikonda et al. 2018; Schneider and Carpenter 2019; Suk et al. 2021; Xiong et al. 2019). An interesting phenomenon highlighted by Clark-Gordon et al. (2019) is the disinhibition effect encouraged by the internet. In other words, the authors found that, when anonymous individuals communicate in a bolder manner on the internet, through their meta-analysis on the self-disclosure as a form of ‘benign disinhibition’, they found that anonymity was positively correlated with self-disclosure. Nevertheless, results must be interpreted cautiously since the effect was

heterogeneous, suggesting that moderating variables could influence the outcomes (Clark-Gordon et al. 2019). In addition, Bogen et al. (2019) studied the manner in which Twitter users mobilised the hashtag #MeToo when disclosing and responding to sexual violence ($N=1660$ tweets). Survivors of sexual violence disclosed the identity of who had assaulted them, what and how it happened, as well as details regarding the moment and place of the attack. Moreover, other tweets discussed the prevalence of violence in society, namely, violence against women. Bogen et al. (2019) argued that Twitter was a space where victims of trauma could connect with others—victims and non-victims—to find and offer support.

4 Research questions

As noted above, scholars have discovered useful methods for the detection of victimisation disclosed on social networks and for their analysis. It seems nevertheless still necessary to extend these techniques to the disclosure of child sexual abuse victimisation and analyse the role these testimonies play within the trend #MeTooInceste. The research questions of our study are the following:

- Q. 1. What are the characteristics of the trend #MeTooInceste? I.e. timeline and concentration of retweets per tweet and tweets per user.
- Q. 2. Is LDA a useful technique to detect tweets disclosing experiences of child sexual victimisation within the trend #MeTooInceste?
- Q. 3. How many testimonies of victimisation there are into the trend?
- Q. 4. What are the characteristics, in terms of meta-data, of the tweets disclosing child sexual victimisation?

5 Data and methods

For this study, we collected 91,501 tweets containing the hashtag #MeTooInceste, published from the beginning of the trend, on 14 January 2021 until 20 January 2021. This section addresses the method followed for our data collection, processing and analysis, namely *topic modelling*, *model validation* and *conjunctive analysis of case configurations*.

5.1 Data collection and processing

We collected the tweets under the hashtag #MeTooInceste on 20 January 2021, using R software (v. 3.6.1) (R-Core-Team 2021) and the rtweet package (v. 0.7) (Kearney 2019). In total, 91,501 tweets were collected that were published between 16 and 20 January 2021 (Fig. 1), of which 10,011

are original tweets—content created by Twitter users—and the remaining 81,490 are retweets.

We generated a corpus of documents as a preliminary step to filtering—excluding stop words—and tokenise it (Silge and Robinson 2016). After this procedure, a document-term matrix (DTM) was generated. The values of the matrix used in this particular case are the frequency of each term within the document (Silge and Robinson 2017).

5.2 Topic modelling

In order to detect the tweets that disclose experiences of child sexual abuse, we applied topic modelling, a technique commonly used in text mining and natural language processing for unsupervised text classification (Kang et al. 2020; Watanabe and Zhou 2020), which has proven to be a useful methodology for the classification of posts published on social networks (Li et al. 2020). In essence, a topic model is a “generative model which provides a probabilistic framework for the term frequency occurrences in documents in a given corpus” (Gruen and Hornik 2011: 1).

Amongst the different approaches to topic modelling, the models based on Latent Dirichlet Allocation (LDA) stand out (Blei et al. 2010, 2003). LDA assumes that a topic is composed of an aggregate of terms and probabilistically relates certain terms with certain topics and, in turn, certain documents with certain topics. Consequently, this method allows algorithmic identification of topics within a collection of documents based on the terms that compose each document (Ghosh and Guha 2013). LDA calculates the probability that a term belongs to a topic—probability denoted as β —as well as the probability that a document belongs to a topic, and a value that is estimated by the proportion of terms in the document that belongs to each topic— γ —.

Following the recommendations of Gruen and Hornik (2011), we applied, on our previously constructed DTM, a fitted LDA model using Gibbs Sampling algorithms with 500 iterations (Griffiths and Steyvers 2004). It should be noted that the number of topics used by the model is a parameter that must be entered by the researcher and does not have a predefined value in the implementation performed by Gruen and Hornik (2011). Therefore, in our case, in order to determine the number of subjects (k) used, models with 2, 3, 4, 5 and 6 subjects were trained. The model finally selected used $k=4$ because, for the selection of the best parameter, the 15 terms to which the model assigned the highest β for each topic pointed towards $k=4$, which assigned to one of the topics—specifically topic 3—a series of terms (verbs formulated in the first person, possessive pronouns, etc.) that suggested that this topic was related to the disclosure of child sexual abuse cases. We are nonetheless aware that the usual method for selecting the value of k

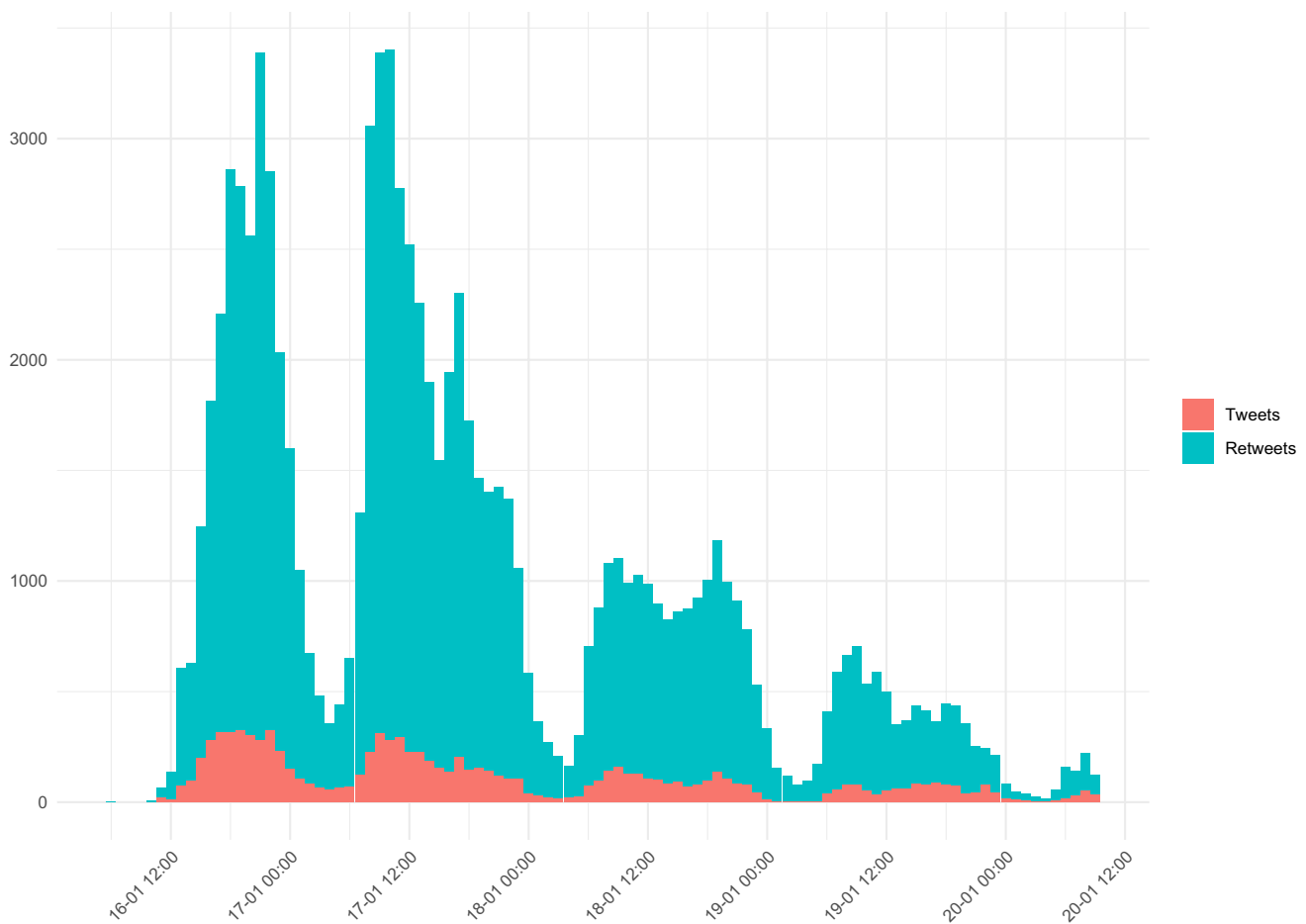


Fig. 1 Temporal distribution of tweets

that best fits the corpus typically uses *the perplexity* of each model as a criterion (Ghosh and Guha 2013); however, we have opted for this selection method based on expert criteria since, in our view, it was the most suitable for our aim of detecting the disclosure of child sexual abuse.

5.3 Model validation and exploration of the differences between categories

The model determined *topic 3* as the largest—in terms of higher γ —with regard to reports of cases of child sexual abuse, specifically in 2350 tweets. However, in some of these documents, the value of γ assigned to the different topics was very close, so a value of γ equal to or greater than 0.3 in topic 3 was set to assign the tweets related to this topic. This new adjustment was made in order to improve the model’s ability to detect tweets reporting cases of child sexual abuse endured by the user or by people close to him or her.

As a result, we identified 1688 tweets referring to disclosure of child sexual abuse. In order to determine the validity of the prediction made by the model, two samples

were constructed by simple random sampling, one drawn from those tweets that the model had identified as child sexual abuse disclosure and the other from the tweets that had not been identified as such. The first sample aimed to estimate the accuracy of the model—proportion of “true positives”—for which a sample of $n=655$ was drawn (confidence level = 95%, margin of error = $\pm 3\%$ and $p=0.5$). The second sample was drawn to determine recall—the proportion of true negatives—and its size was $n=364$ (confidence level = 95%, margin of error = $\pm 5\%$ and $p=0.5$). Both samples were manually reviewed by two of the researchers to determine how many of the documents identified by the model as disclosure actually contained disclosure of child sexual abuse—suffered in the first person or by a family member—as well as how many documents not identified as disclosure did in fact disclose experiences of child sexual abuse.¹

¹ There were no discrepancies between both reviewers.

Table 1 Variables encoding

Favourites		Retweet		Followers		Photo	URL	Verified	Days	Text size	
Low	[0, 421]	Low	0	Very small	[0, 33]	No	No	No	1	Small	[13, 102]
Medium	(421, 3373]	Low	0	Small	(33, 246]	Yes	Yes	Yes	2	Medium	(102, 179]
High	(3373, 14491]	Medium	(0, 2]	Medium	(246, 1091]				3	Long	(179, 255]
Very high	(14491, 731763]	High	(2, 3774]	Big	(1091, 9128715]				4	Very long	(255, 280]

5.4 Conjunctive analysis of case configurations

Once the model had been validated, we explored the metadata associated with tweets classified as disclosure with the aim of finding the characteristics and role of these. For this purpose, we analysed the data related to the tweets themselves, which previous research highlighted as relevant for the analysis of posts published on social networks, these being (1) number of retweets, (2) number of favourites, (3) time of publication and (4) whether or not the tweet has images or URLs (Miró-Llinares, et al. 2018; Nesi et al. 2018; Peng et al. 2011; Pezzoni et al. 2013; Zaman et al. 2014), as well as metadata related to the account that posted the tweet—(1) verified account status and (2) number of followers.

For this purpose and given the heterogeneity of the variables, we used *Conjunctive Analysis of Case Configurations* (CACC), an exploratory data analysis technique that allows the identification of complex relationships between the attributes of variables that are related to a specific outcome (Moneva et al. 2020). Since its inception (Miethe et al. 2008), the CACC has been successfully used on multiple occasions within the field of criminology, allowing researchers to unravel complex causal relationships in diverse domains, such as fear of crime (Hart 2017) or criminogenic microenvironments (Hart and Miethe 2015).

The CACC, as a case-oriented technique based on *configurational thinking* (Hart 2020), allows us to analyse the combinations between attributes of the variables that make up the most common profiles and that are most frequently associated with the dependent variable (Hart and Moneva 2018), the latter being, in our case, the disclosure of child sexual abuse. To perform the CACC, we transformed the different variables, where necessary, into ordinal variables, and divided the continuous variables into quartiles (as recommended by Miethe et al. 2008). Table 1 illustrates the encoding of the variables retained in our model. Once the variables had been adapted to the CACC requirements, we transformed the database into a *truth table*, in which the columns represent the different predictor variables, as well as the number of times a particular case configuration is observed in the database, and the probability of each configuration resulting in the outcome that is being analysed (Moneva et al. 2020). This table allows assessing which

specific case configurations are most frequently associated with, in our case, the presence of disclosure of child sexual abuse in the tweet and which case configurations are dominant. Per recommendation of Hart (2020), we consider dominant case configurations of those that are observed 10 or more times.

6 Results

6.1 #MeTooInceste's evolution

Figure 1 shows the evolution of the trend #MeTooInceste from 16 January 2021. The hashtag started to be active on the morning of 16 January and remained active for five days (16, 17, 18, 19 and 2 January), with scarce activity on the 20 January. Each day, the wave of tweets grew rapidly in the morning and, after remaining active throughout the day, decreased at night until it reached its lowest point of activity at 3:00 a.m. Central European Time (CET). After this moment, the number of users participating in the trend and the number of tweets that refer to the hashtag started to grow again, and a new wave began the following morning.

Table 2 shows the distribution of tweets during the observation period. On the first day, 12,890 users participated in the hashtag with 27,008 tweets, of which 14% were original tweets—content contributions made by a user—and the rest retweets. On the second day, participation in the hashtag increased regarding both the number of users and tweets under the hashtag #MeTooInceste; however, the percentage of original tweets out of the total number of tweets dropped to 10% when compared to the first day. In addition, during the second day, the number of verified accounts that participated in the trend also grew and the accounts that posted original tweets had more followers than those participating on the first day. These two characteristics—the increment in verified accounts and the participation of users with more followers—suggest that during the second day the hashtag managed to attract more attention from influential personalities and organisations. On days 3 and 4, the hashtag lost strength both in number of users and volume of tweets, with only 9055 new tweets being posted on the fourth day, only 18% of them original tweets. This percentage, which is much higher than the previous days, indicates that there was still

Table 2 Descriptive statistics across the days

	Nb users	Nb tweets	Nb verified accounts	Followers mean	Followers sd	Fol-lowers median	Followers iqr
<i>Day 1</i>							
Tweet original	2397	3234	52	7843.9	117,991	212	819.25
Retweet	10,493	23,774	115	1421.09	12,537.1	332	860
<i>Day 2</i>							
Tweet original	2585	3481	81	19,493	275,706	272	1107
Retweet	17,492	34,236	169	1380	16,922.3	290	738
<i>Day 3</i>							
Tweet original	1421	1920	92	30,500	294,856	284	1492.5
Retweet	10,959	15,801	112	1871	28,214.1	289	827
<i>Day 4</i>							
Tweet original	1035	1376	72	28,550	227,334	220	1217.5
Retweet	5565	7679	87	2589	40,399.5	246	755

Table 3 Retweet per tweet distribution

Days	Number of original tweets with retweets															
	=0		≥1		>5		>10		>50		>100		>200		>1000	
1	1777	54.95%	1457	45.05%	470	14.53%	341	10.54%	128	3.96%	83	2.57%	53	1.64%	11	0.34%
2	2007	57.66%	1474	42.34%	391	11.23%	226	6.49%	56	1.61%	33	0.95%	18	0.52%	2	0.06%
3	1083	56.41%	837	43.59%	217	11.30%	139	7.24%	33	1.72%	18	0.94%	6	0.31%	1	0.05%
4	873	63.44%	503	36.56%	102	7.41%	54	3.92%	8	0.58%	3	0.22%	3	0.22%	0	0.00%

a relevant number of users associating their tweets with the hashtag; however, the diffusion of the hashtag is already much lower as the number of retweets associated with the hashtag decreased significantly.

Additionally, this evolution in the ratio of original tweets to retweets can also be observed in the different metrics relating to the number of retweets which each tweet receives. For instance, on day 4, 63% of the tweets published and received no retweets, a significantly higher percentage than on previous days—see Table 3. In fact, if we look at the distribution of the number of retweets received by each tweet, we can see that on the first day, there is greater *horizontality*, with the number of retweets per tweet being less concentrated around the mode, and that this decreases as the days go by—see Fig. 2.

6.2 LDA results

The LDA model showed that theme 3, which was identified as corresponding to child sexual abuse cases, was the main theme in 2350 tweets, of which 1688 had a gamma value assigned to theme 3 that was greater than 0.3. The sample (c.l. = 95%, $e = \pm 3%$) constructed to test the reliability of

the model showed that only 57 tweets had been misclassified,² which allows us to infer that the model—with 95% confidence—has an accuracy of 91.3% [$\pm 3%$]. On the other hand, the sample checking (n.c. = 95%, $e = \pm 5%$) we used to identify the number of false and true negatives shows that the model has a recall of 93.1% [$\pm 5%$].

The precision and recall values suggest that the model is useful for the detection of tweets that disclose cases of incest. In fact, it allows us to state, with a degree of caution, that there were around 1600 tweets reporting cases of child sexual abuse during the period of activity of the hashtag. Furthermore, these tweets were posted by 1350 different users. Of course, we do not have the capacity to verify the truthfulness of these testimonies, but this is not the purpose of this research.

6.3 Conjunctive analysis of case configurations

Given the variables we used and their coding, 6144 theoretical possible case configurations can be generated with our data; however, our whole dataset is based on 1772 case

² The errors come mainly (73.21%) from the presence in the tweets of first-person testimonies referring to the subject’s childhood, or from first-person descriptions of the emotional reaction to the other testimonies of the hashtag.

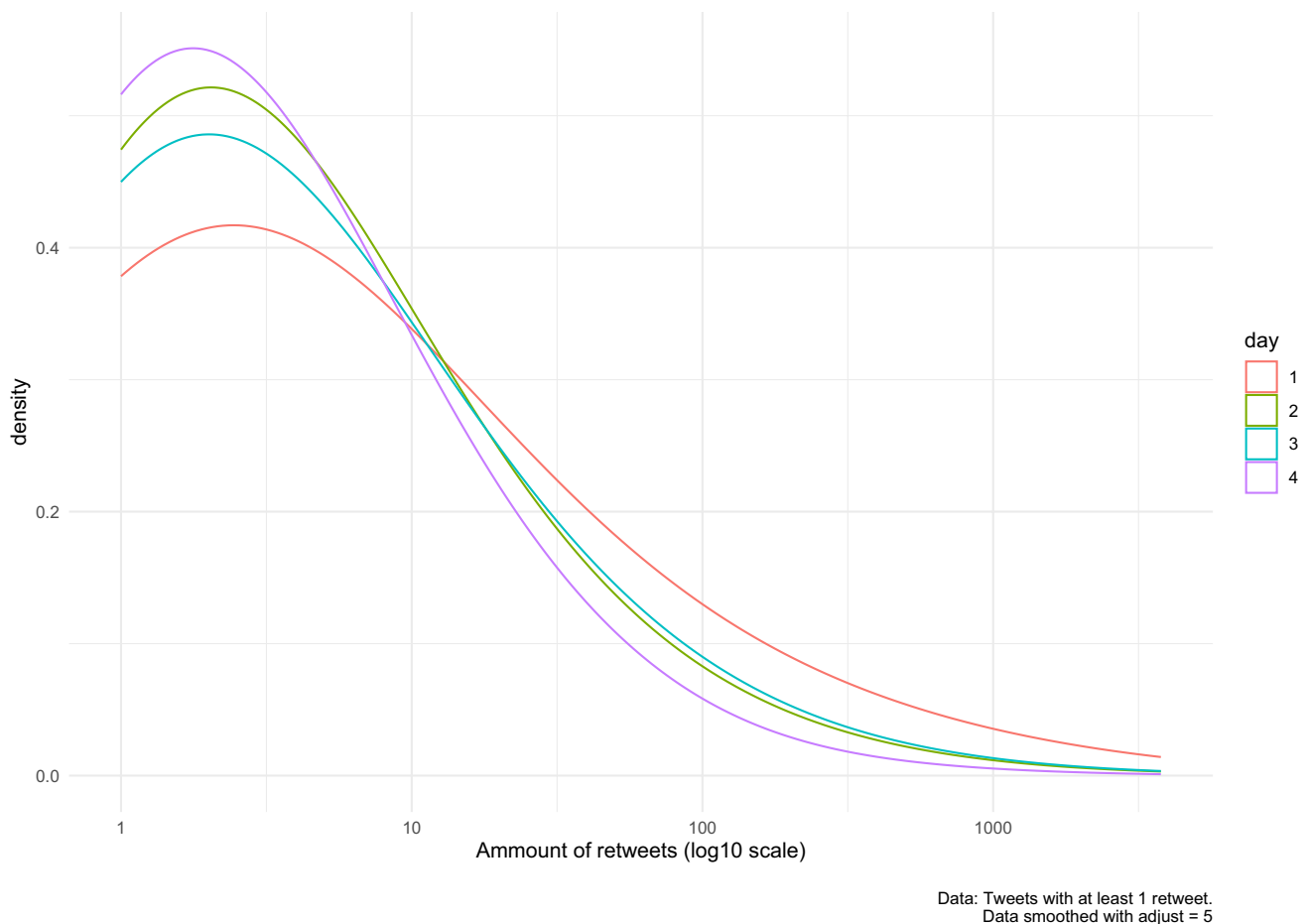


Fig. 2 Retweets per tweet frequency

configurations, roughly 28% of the theoretical configurations. Moreover, by transforming our data into a truth table, we have been able to detect 271 dominant case configurations (15.3% of the observed case configurations), which account for 57.6% of our data. Consequently, we can affirm that our variables have the capacity to "cluster" the data ($X^2 = 2531.19$, $df = 270$, $p < 2.2204e-16$). However, according to the *Situational Clustering Index* (Hart 2020), the magnitude of clustering generated around the dominant profiles ($SCI = 0.32$) is moderate. Each of the dominant case configurations has a probability that relates it to a certain value of the dependent variable, in our case the probability that the case configuration is associated with tweets containing disclosure of child sexual abuse— $P(D)$. Following Hart and Moneva (2018), we can take as "typical situational profiles" those case configurations that are two standard deviations above the mean of the probability that the tweet contains disclosure (therefore, we will consider typical situational profiles those with a $P(D)$ greater than 56.1%).

As Table 4 shows, there are 25 case configurations strongly associated with disclosure of child sexual abuse.

The first notable characteristic of these case configurations is that they all come from: unverified accounts and tweets without photos nor URLs. In 92% of the cases, the tweets are qualified as either "long"—(179:255]—or "very long"—(255:304]. Furthermore, it should be noted that 68% of the profiles correspond to very small accounts (less than 34 followers), which, added to the small accounts—{34:246}—represent 92% of the cases. The frequency of profiles in which we find small or very small account sizes is particularly relevant, considering that in 23 of the 25 profiles, the variable number of retweets is medium—(0:2)—or high (2:3774).

The two profiles that are most strongly associated with the disclosure of child sexual abuse correspond to long tweets, without URLs, without photos, posted on day 1, coming from unverified, small or medium-sized accounts that get a high number of retweets and a medium or large number of favourites. In particular, the profile in which the accounts are small gets a medium number of retweets and the profile in which the accounts are medium-sized gets a large number of favourites.

Table 4 Dominant case configurations

Favourites	Retweet	Followers	Photo	URLs	Verified	Days	Text size	N_Break	<i>p</i>
High	High	Small	No	No	No	1	Very long	17	0.94
Medium	High	Very small	No	No	No	1	Very long	11	0.91
Low	High	Very small	No	No	No	3	Very long	14	0.86
Low	Medium	Very small	No	No	No	2	Very long	35	0.83
Low	High	Very small	No	No	No	2	Long	22	0.82
Low	Medium	Very small	No	No	No	1	Very long	20	0.80
Low	Medium	Very small	No	No	No	3	Very long	62	0.77
Low	Medium	Very small	No	No	No	4	Very long	34	0.76
Low	High	Very small	No	No	No	1	Long	20	0.75
Low	Medium	Very small	No	No	No	2	Long	34	0.74
Low	High	Very small	No	No	No	2	Very long	25	0.72
Medium	High	Small	No	No	No	1	Very long	10	0.70
Low	High	Very small	No	No	No	1	Medium	10	0.70
Medium	High	Very small	No	No	No	1	Long	12	0.67
Low	High	Very small	No	No	No	1	Very long	32	0.66
Low	Medium	Very small	No	No	No	3	Long	31	0.65
Very high	High	Medium	No	No	No	1	Very long	14	0.64
Low	Medium	Very small	No	No	No	3	Medium	11	0.64
Very high	Medium	Small	No	No	No	1	Very long	13	0.62
High	Medium	Small	No	No	No	2	Very long	18	0.61
High	High	Big	No	No	No	1	Very long	10	0.60
Low	Medium	Small	No	No	No	2	Long	10	0.60
Low	Low	Very small	No	No	No	3	Very long	72	0.60
High	Medium	Small	No	No	No	2	Long	17	0.59
Low	Low	Very small	No	No	No	4	Very long	65	0.58

Despite what these two profiles suggest, the number of favourites tends to be low amongst the different profiles and seems to depend primarily on the size of the account. For instance, profiles 17 and 21 are identical to profile 1 in almost all variables except account size and the number of favourites the tweet gets, with larger accounts getting more favourites. Thus, the most common configuration amongst these profiles is that they (1) are small accounts which regardless of the day, (2) get a high number of retweets and (3) a low number of favourites, (4) publish a long or very long tweet, (5) do not contain a photo and (6) do not contain URLs.

This trend is also corroborated by checking the data from the opposite perspective, i.e. looking at the dominant case configurations that are less likely to contain tweets reporting child sexual abuse, as in Table 5. There are 86 dominant case configurations in which the tweet does not contain a disclosure of child sexual abuse. There is still a relationship between the size of the account and the number of favourites, although the most frequent value of this variable is very high, but the cases in which the number of retweets is low are the most frequent, and many of them are associated with large accounts and tweets with many favourites. In addition, we now find that photos or URLs are present in the tweets.

7 Discussion

Borrowing Grabosky's (2001) metaphor "Old wine in new bottles", referring to the use of the internet for the perpetration of traditional crimes—e.g. fraud—we state that, in the context of the #MeTooInceste movement, old crimes are reported in new bottles. Through the manuscript, we demonstrated, complementing previous research (Idoiaga Mondragon et al. 2002) on the #MeTooInceste phenomenon, the possibility of identifying the disclosure of childhood sexual abuse victimisation with content analysis techniques, specifically topic modelling through LDA, achieving results that are consistent. It is nevertheless to note that our study detected a slightly higher number of tweets containing disclosure of victimisation (16% vs 13% found by Idoiaga Mondragon et al. (2002)), but this is inside our margin of error of 3%. Furthermore, it is to note that the studies take slightly different time frames (the present study takes the first 2 days and the following 3 days). The emergence of the online trend #MeTooInceste seemed to have been a facilitator for disclosing the child sexual abuse experiences, which is consistent with former findings that disclosure is activated by discussions about abuse (Alaggia et al. 2019). From a criminology perspective, the internet is a prolific space not only for

Table 5 Top 10 bigger dominant case configurations with $p=0$

Favourites	Retweet	Followers	Photo	Urls	Verified	Days	Text size	N_Break	p
High	High	Small	No	No	No	1	Very long	17	0.94
Medium	High	Very small	No	No	No	1	Very long	11	0.91
Low	High	Very small	No	No	No	3	Very long	14	0.86
Low	Medium	Very small	No	No	No	2	Very long	35	0.83
Low	High	Very small	No	No	No	2	Long	22	0.82
Low	Medium	Very small	No	No	No	1	Very long	20	0.80
Low	Medium	Very small	No	No	No	3	Very long	62	0.77
Low	Medium	Very small	No	No	No	4	Very long	34	0.76
Low	High	Very small	No	No	No	1	Long	20	0.75
Low	Medium	Very small	No	No	No	2	Long	34	0.74
Low	High	Very small	No	No	No	2	Very long	25	0.72
Medium	High	Small	No	No	No	1	Very long	10	0.70
Low	High	Very small	No	No	No	1	Medium	10	0.70
Medium	High	Very small	No	No	No	1	Long	12	0.67
Low	High	Very small	No	No	No	1	Very long	32	0.66
Low	Medium	Very small	No	No	No	3	Long	31	0.65
Very high	High	Medium	No	No	No	1	Very long	14	0.64
Low	Medium	Very small	No	No	No	3	Medium	11	0.64
Very high	Medium	Small	No	No	No	1	Very long	13	0.62
High	Medium	Small	No	No	No	2	Very long	18	0.61
High	High	Big	No	No	No	1	Very long	10	0.60
Low	Medium	Small	No	No	No	2	Long	10	0.60
Low	Low	Very small	No	No	No	3	Very long	72	0.60
High	Medium	Small	No	No	No	2	Long	17	0.59
Low	Low	Very small	No	No	No	4	Very long	65	0.58

crime commission, but also for building a community that supports victims who disclose traumatic experiences such as child sexual abuse.

As a short-lived event—limited to its own nature as a hashtag—#MeTooInceste concentrates in an ephemeral place on the internet, especially during its first 3 days, a debate on child sexual abuse that is not evenly distributed throughout the social network or over time. Findings suggest that the hashtag was mainly linked to smaller accounts during the first day of activity. During the second day, the hashtag gained in activity, mainly due to the increase in accounts that started participating through retweets. On the other hand, the accounts that added content to the hashtag by publishing original tweets were larger and there were in larger proportion verified accounts, which might indicate that on the second day the hashtag attracted the attention of celebrities, organisations or media. On days 3 and 4 the activity associated with hashtag decreased, and though it still interested very large accounts, each new tweet got lower numbers of retweets. Thus, the hashtag lost the popularity that defined it on the first day, with the number of retweets per tweet now more concentrated around the mode of the distribution. It is to note that the characteristics mentioned above refer to the entire hashtag, where different types of

content are mixed, so they cannot be generalised as characteristics of disclosure of child sexual abuse victimisation on social networks. Instead, they characterise this hashtag as an event.

Furthermore, we identified, with a relatively small margin of error, 1688 tweets that disclosed experiences of child sexual victimisation, corresponding to 1350 accounts, which seems like a considerable role in the trend. In addition, the CACC analysis highlighted that tweets containing disclosure of victimisation are associated with a relatively high number of retweets, despite coming from small accounts and lacking URLs or images. This is, in our view, a remarkable finding. It is plausible that, due to the social shame and guilt that these experiences of victimisation trigger (Dorahy & Clearwater 2012), the users prefer to hide their identity (as also pointed out by Clark-Gordon et al. 2019). In addition, the length of the tweets can be understood as a necessity for the survivors to contextualise their experiences by providing details of them, as already found also in the tweets regarding the original #MeToo (Bogen et al. 2019).

The reader should bear in mind that the CACC is an exploratory technique and that our study must be replicated in other settings due to its lack of external validity. Nevertheless, we hope to have laid a foundation for further research

on the identification and study of crime reported on social networks by using metadata associated with the publication. This is a research line that is already open, for example, in the detection of fake news (see Ruchansky et al. 2017; Wang et al. 2018; Shin et al. 2018), hate crime (Miró-Llinares et al. 2018) or fear of crime (Castro-Toledo et al. 2020).

8 Implications for research and practice

The digital space is not only a new environment for the perpetration of crimes, but could also affect victims and the way they disclose their victimisation experiences, offering them an online community that often is seldom present in the offline world. This may be especially relevant regarding crimes such as child sexual abuse, of which the reporting rate is especially low (Alaggia et al. 2019; Bennett and O'Donohue 2014; WHO 2003; Zalcborg 2017). In our view, our study opens opportunities for both research and practice.

First, the emergence of these online events allows researchers to increase their knowledge on child sexual abuse. In this sense, beside the analysis of the metadata associated with the tweet, further research should carry a detailed analysis of the content disclosed—e.g. modus operandi, perpetrators, victim assistance experiences—and enlarge the current knowledge. The method presented in this article allows ‘to find needles in a haystack’, rendering easier to distinguish the tweets disclosing victimisation from others: messages of solidarity, support, critiques, *trolling* and false reports, etc.

The study also shows that social networks could be a new place where interventions should focus. Victim support organisations and police institutions could implement automated systems to detect this kind of content in order to offer support to users that disclosed victimisation or creating protocols for action for law enforcement agencies when these posts arise.

9 Conclusion

This paper studied the distribution of tweets ($N=91,501$) published during the first week of the hashtag #MeTooInceste, which referred to cases of child sexual abuse within the family. The digital space has brought with it new ways of perpetrating crimes but also new ways of reporting them and a new mode of social communication with regard to the disclosure of victimisation experiences. Using LDA, we identified 1688 that disclosed experiences of child sexual abuse victimisation. These tweets contributed significantly to the dissemination of the hashtag #MeTooInceste, achieving

significant support from other users via retweets, especially during the first day of the hashtag's activity. As the days went by, the hashtag lost popularity but this did not prevent new experiences of victimisation from appearing. The tweets disclosing child sexual abuse were in more prevalence long, belonging to small amounts, with no URL nor link. Victim assistance services and law enforcement agencies should implement automated detection techniques in order to offer support to the users that disclosed these types of experiences. Moreover, research should enlarge the current knowledge by analysing the specific content of the tweets.

Acknowledgements This publication was made possible by Grant FJC2020-042961-I funded by Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/<https://doi.org/10.13039/501100011033> and by the European Union "NextGenerationEU"/PRTR"; and it is part of the project #FakePenal (PID2021-125730OB-I00) funded MCIN/AEI/<https://doi.org/10.13039/501100011033/FEDER>, UE

Author contributions J.A., L.M. and F.M wrote the main manuscript text. J.A. prepared the figures and carry out the data analysis, and F.M. reviewed the manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declarations

Conflicts of interest The authors have no conflicts of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aebi MF, Caneppele S, Hashimoto YZ, Jehle JM, Khan TS, Kühn O, Lewis C, Molnar L, Þórisdóttir R, Smit P, National Correspondents (2021) European Sourcebook of Crime and Criminal Justice Statistics 2021, 6th edn. Series UNILCRIM, (1)2021. Retrieved from: <https://wp.unil.ch/europeansourcebook/printed-editions-2/>
- Alaggia R, Collin-Vézina D, Lateef R (2019) Facilitators and barriers to Child Sexual Abuse (CSA) disclosures: a research update (2000–2016). *Trauma Violence Abuse* 20(2):260–283. <https://doi.org/10.1177/1524838017697312>
- Ayers JW, Caputi TL, Nebeker C, Dredze M (2018) Don't quote me: reverse identification of research participants in social

- media studies. *Nat Digit Med*. <https://doi.org/10.1038/s41746-018-0036-2>
- Bennett N, O'Donohue W (2014) The construct of grooming in child sexual abuse: conceptual and measurement issues. *J Child Sex Abus* 23(8):957–976. <https://doi.org/10.1080/10538712.2014.960632gu>
- Blei D, Nag AY, Jordan M (2003) Latent Dirichlet allocation. *J Mach Learn Res* 3:993–1022
- Blei D, Carin L, Dunson D (2010) Probabilistic topic models. *IEEE Signal Process Mag* 27(6):55–65. <https://doi.org/10.1109/msp.2010.938079>
- Bogen KW, Bleiweiss K, Leach N, Orchowski L (2019) #MeToo: disclosure and response to sexual victimization on Twitter. *J Interpers Violence*. <https://doi.org/10.1177/0886260519851211>
- Brünker F, Wischniewski M, Mirbabaie M, Meinert J (2020, January 7) The role of social media during social movements—observations from the #metoo debate on Twitter. <https://doi.org/10.24251/HICSS.2020.288>
- Castro-Toledo F, J, Gretenkort T, Esteve M, Miró-Llinares F, Ceccato V, Mahesh K (2020) “Fear in 280 characters”: a new approach for evaluation of fear over time in cyberspace. In: *Crime and fear in public places. Towards safe inclusive and sustainable cities*, pp 326–343. Routledge. <https://doi.org/10.4324/9780429352775-22>
- Chandran S, Bhargava S, Kishor M (2020) Under reporting of child sexual abuse—the barriers guarding the silence. *Telangana J Psychiatry* 4(2):57–60. <https://doi.org/10.18231/2455-8559.2018.0003>
- Clark-Gordon CV, Bowman ND, Goodboy AK, Wright A (2019) Anonymity and online self-disclosure: a meta-analysis. *Commun Rep* 32(2):98–111. <https://doi.org/10.1080/08934215.2019.1607516>
- Cyr M, Wright J, McDuff P, Perron A (2002) Intrafamilial sexual abuse: brother–sister incest does not differ from father–daughter and stepfather–stepdaughter incest. *Child Abuse Negl* 26(9):957–973. [https://doi.org/10.1016/s0145-2134\(02\)00365-4](https://doi.org/10.1016/s0145-2134(02)00365-4)
- Dorahy MJ, Clearwater K (2012) Shame and guilt in men exposed to childhood sexual abuse: a qualitative investigation. *J Child Sex Abus* 21(2):155–175. <https://doi.org/10.1080/10538712.2012.659803>
- Fiesler C, Proferes N (2018) “Participant” perceptions of twitter research ethics. *Soc Media + Soc*. <https://doi.org/10.1177/2056305118763366>
- Fileborn B (2022) Mapping activist responses and policy advocacy for street harassment: current practice and future directions. *Eur J Crim Policy Res* 28:97–116. <https://doi.org/10.1007/s10610-021-09479-2>
- Ghosh D, Guha R (2013) What are we “tweeting” about obesity? Mapping tweets with topic modeling and geographic information system. *Cartogr Geogr Inf Sci* 40(2):90–102. <https://doi.org/10.1080/15230406.2013.776210>
- Giglio JJ, Wolfeich PM, Gabrenya WK, Sohn ML (2011) Differences in perceptions of child sexual abuse based on perpetrator age and respondent gender. *J Child Sex Abus* 20(4):396–412. <https://doi.org/10.1080/10538712.2011.593255>
- Grabosky PN (2001) Virtual criminality: old wine in new bottles? *Soc Leg Stud* 10(2):243–249. <https://doi.org/10.1177/a017405>
- Griffiths TL, Steyvers M (2004) Finding scientific topics. *Proc Natl Acad Sci USA* 101:5228–5235. <https://doi.org/10.1073/pnas.0307752101>
- Gruen B, Hornik K (2011) Topicmodels: an R package for fitting topic models. *J Stat Softw* 40(13):1–30
- Hart TC (2017) Using typologies of victimisation worry to create strategies for reducing fear of crime. *Police Pract Res* 18(4):407–419. <https://doi.org/10.1080/15614263.2016.1230854>
- Hart TC (2020) Identifying situational clustering and quantifying its magnitude in dominant case configurations: new methods for conjunctive analysis. *Crime Delinq* 66(1):143–159. <https://doi.org/10.1177/0011128719866123>
- Hart TC, Miethe TD (2015) Configural behavior settings of crime event locations: toward an alternative conceptualization of criminogenic microenvironments. *J Res Crime Delinq* 52(3):373–402. <https://doi.org/10.1177/0022427814566639>
- Hart TC, Moneva A (2018) Conjunctive analysis of case configurations: an introduction to configural thinking. *REIC*. <https://doi.org/10.46381/reic.v16i0.148>
- Hillberg T, Hamilton-Giachritsis C, Dixon L (2011) Review of meta-analyses on the association between child sexual abuse and adult mental health difficulties: a systematic approach. *Trauma Violence Abuse* 12(1):38–49. <https://doi.org/10.1177/1524838010386812>
- Hosterman A, Johnson N, Stouffer R, Herring S (2018) Twitter, social support messages and the #MeToo movement. *J Soc Media Soc* 7(2):69–91
- Huffington Post (2021, March 24) #MeTooGay: Pourquoi les hommes gays sont particulièrement exposés aux violences sexuelles. *Le Huffington Post*. https://www.huffingtonpost.fr/entry/metoo-gay-pourquoi-les-hommes-gays-sont-particulierement-exposes-aux-violences-sexuelles_fr_6058b979c5b6bd95117ed5c8c
- Idoiaaga Mondragon N, Eiguren Munitis A, Belasko Txertudi M (2002) The breaking of secrecy: analysis of the hashtag #MeTooIncest regarding testimonies of sexual incest abuse in childhood. *Child Abuse Negl*. <https://doi.org/10.1016/j.chiabu.2021.105412>
- Ipsos (2020) Les Français face à l’inceste. Retrieved from <https://facealinceste.fr/upload/media/documents/0001/02/9b09b6479dc5cc783ff3ebfb07769e5911e173b8.pdf>
- Jahnke S, Imhoff R, Hoyer J (2015) Stigmatization of people with pedophilia: two comparative surveys. *Arch Sex Behav* 44(1):21–34. <https://doi.org/10.1007/s10508-014-0312-4>
- Kang K, Choo J, Kim Y (2020) Whose opinion matters? Analyzing relationships between bitcoin prices and user groups in online community. *Soc Sci Comput Rev* 38(6):686–702. <https://doi.org/10.1177/0894439319840716>
- Kar SK, Swain R (2020) Incest. In: Shackelford T, Weekes-Shackelford V (eds) *Encyclopedia of evolutionary psychological science*. Springer, Berlin. https://doi.org/10.1007/978-3-319-16999-6_1996-1
- Kearney MW (2019) rtweet: collecting and analyzing Twitter data. *J Open Source Softw*. <https://doi.org/10.21105/joss.01829>
- Kennedy AC, Prock KA (2018) “I still feel like I am not normal”: a review of the role of stigma and stigmatization among female survivors of child sexual abuse, sexual assault, and intimate partner violence. *Trauma Violence Abuse* 19(5):512–527. <https://doi.org/10.1177/1524838016673601>
- Khatua A, Cambria E, Khatua A (2018) Sounds of silence breakers: exploring sexual violence on Twitter. In: 2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp 397–400. <https://doi.org/10.1109/ASONAM.2018.8508576>
- Kouchner C (2021) *La familia grande*. Le Seuil, Paris
- Li P, Cho H, Qin Y, Chen A (2020) “#MeToo as a Connective Movement: Examining the Frames Adopted in the Anti-Sexual Harassment Movement in China”. *Social Science Computer Review* 0894439320956790. <https://doi.org/10.1177/0894439320956790>
- Lupton D (2015) *Digital sociology*. Routledge, London

- Lusky-Weisrose E, Fleishman T, Tener D (2022) “A little bit of light dispels a lot of darkness”: online disclosure of child sexual abuse by authority figures in the ultraorthodox Jewish community in Israel. *J Interpers Violence*. <https://doi.org/10.1177/08862605211028370>
- Manikonda L, Beigi G, Liu H, Kambhampati S (2018) Twitter for sparking a movement, Reddit for sharing the moment: #metoo through the lens of social media. *arXiv:1803.08022* [Cs]. <http://arxiv.org/abs/1803.08022>
- Méheut C (2021, January 18) Incest scandal sets off a new #MeToo movement in France. *The New York Times*. <https://www.nytimes.com/2021/01/18/world/europe/duhamel-france-incest.html>
- Miethe TD, Hart TC, Regoeczi WC (2008) The conjunctive analysis of case configurations: an exploratory method for discrete multivariate analyses of crime data. *J Quant Criminol* 24(2):227–241. <https://doi.org/10.1007/s10940-008-9044-8>
- Miró-Llinares F, Moneva A, Esteve M (2018) Hate is in the air! But where? Introducing an algorithm to detect hate speech in digital microenvironments. *Crime Sci*. <https://doi.org/10.1186/s40163-018-0089-1>
- Mojallal M, Simons RM, Simons JS (2021) Childhood maltreatment and adulthood proneness to shame and guilt: the mediating role of maladaptive schemas. *Motiv Emot*. <https://doi.org/10.1007/s11031-021-09866-6>
- Moneva A, Miro-Llinares F, Hart TC (2020) Hunter or prey? Exploring the situational profiles that define repeated online harassment victims and offenders. *Deviant Behav*. <https://doi.org/10.1080/01639625.2020.1746135>
- Nesi P, Pantaleo G, Paoli I, Zaza I (2018) Assessing the reTweet proneness of tweets: predictive models for retweeting. *Multimedia Tools Appl* 77(20):26371–26396. <https://doi.org/10.1007/s11042-018-5865-0>
- Ohlert J, Seidler C, Rau T, Fegert J, Allroggen M (2017) Comparison of psychopathological symptoms in adolescents who experienced sexual violence as a victim and/or as a perpetrator. *J Child Sex Abus* 26(4):373–387. <https://doi.org/10.1080/10538712.2017.1283652>
- Papacharissi Z (2002) The virtual sphere: the internet as a public sphere. *New Media Soc* 4(1):9–27. <https://doi.org/10.1177/14614440222226244>
- Peng H, Zhu J, Piao D, Yan R, Zhang Y (2011) Retweet modeling using conditional random fields. In: 2011 IEEE 11th international conference on data mining workshops, pp 336–343. <https://doi.org/10.1109/ICDMW.2011.146>
- Pereda N, Gallardo-Pujol D (2011) Revisión sistemática de las consecuencias neurobiológicas del abuso sexual infantil. *Gac Sanit* 25(3):233–239. <https://doi.org/10.1016/j.gaceta.2010.12.004>
- Pereda N, Guilera G, Forns M, Gómez-Benito J (2009) The prevalence of child sexual abuse in community and student samples: a meta-analysis. *Clin Psychol Rev* 29(4):328–338. <https://doi.org/10.1016/j.cpr.2009.02.007>
- Pereda N, Abad J, Guilera G (2016) Lifetime prevalence and characteristics of child sexual victimization in a community sample of Spanish adolescents. *J Child Sex Abus* 25(2):142–158. <https://doi.org/10.1080/10538712.2016.1123791>
- Pereda N, Greco AM, Hombrado J, Segura A, Gómez-Martín V (2018) ¿Qué factores inciden para romper el silencio de las víctimas de abuso sexual? *Revista Española De Investigación Criminológica* 16:1–27. <https://doi.org/10.46381/reic.v16i0.195>
- Pezzoni F, An J, Passarella A, Crowcroft J, Conti M (2013) Why do I retweet it? An information propagation model for microblogs. In: Jatowt A (ed) *Social Informatics*. SocInfo. Springer, Berlin, pp 360–369. <https://doi.org/10.1007/978-3-319-03260-3>
- Radio Télévision Suisse (2021, January 18) Sur les réseaux sociaux, déferlante de témoignages autour de l’inceste. *rts.ch*. <https://www.rts.ch/info/monde/11904369-sur-les-reseaux-sociaux-deferlante-de-temoignages-autour-de-linceste.html>
- R-Core-Team (2021) R: a language and environment for statistical computing. In: Version 3.6.1. R Foundation for Statistical Computing
- Ruchansky N, Seo S, Liu Y (2017) “CSI: A Hybrid Deep Model for Fake News Detection”. *Proceedings of the 2017 ACM Conference on Information and Knowledge Management*. <https://doi.org/10.1145/3132847.3132877>
- Schneider KT, Carpenter NJ (2019) Sharing #MeToo on Twitter: incidents, coping responses, and social reactions. *Equal Divers Incl: Int J* 39(1):87–100. <https://doi.org/10.1108/EDI-09-2018-0161>
- Shin J, Jian L, Driscoll K, Bar F (2018) The diffusion of misinformation on social media: temporal pattern, message, and source. *Comput Hum Behav* 83:278–287. <https://doi.org/10.1016/j.chb.2018.02.008>
- Silge J, Robinson D (2016) tidytext: text mining and analysis using tidy data principles in R. *J Open Source Softw*. <https://doi.org/10.21105/joss.00037>
- Silge J, Robinson D (2017) *Text mining with R: a tidy approach*. O’Reilly Media, Sebastopol
- Suk J, Abhishek A, Zhang Y, Ahn SY, Correa T, Garlough C, Shah DV (2021) #MeToo, networked acknowledgment, and connective action: how “empowerment through empathy” launched a social movement. *Soc Sci Comput Rev* 39(2):276–294. <https://doi.org/10.1177/0894439319864882>
- Tener D, Murphy SB (2015) Adult disclosure of child sexual abuse: a literature review. *Trauma Violence Abuse* 16(4):391–400. <https://doi.org/10.1177/1524838014537906>
- The Guardian (2017, December 1) Alyssa Milano on the #MeToo movement: ‘We’re not going to stand for it any more’. <http://www.theguardian.com/culture/2017/dec/01/alyssa-milano-mee-too-sexual-harassment-abuse>
- Twenge JM, Martin GN, Spitzberg BH (2019) Trends in U.S. adolescents’ media use, 1976–2016: the rise of digital media, the decline of TV, and the (near) demise of print. *Psychol Pop Media Cult* 8(4):329–345. <https://doi.org/10.1037/ppm0000203>
- Ullman D (2007) Relationship to perpetrator, disclosure, social reactions, and PTSD symptoms in child sexual abuse survivors. *J Child Sex Abus* 16(1):19–36. https://doi.org/10.1300/J070v16n01_02
- Vollman BK (2021) When did I become a victim? Exploring narratives of male childhood sexual abuse. *J Child Adolesc Trauma*. <https://doi.org/10.1007/s40653-021-00346-0>
- Wang Y, Ma F, Jin Z, Yuan Y, Xun G, Jha K, Su L, Gao J (2018) EANN: event adversarial neural networks for multi-modal fake news detection. In: *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pp 849–857. Association for Computing Machinery, New York. <https://doi.org/10.1145/3219819.3219903>
- Watanabe K, Zhou Y (2020) Theory-driven analysis of large corpora: semisupervised topic classification of the UN speeches. *Soc Sci Comput Rev*. <https://doi.org/10.1177/0894439320907027>
- Watson.ch (1 March 2021) Six questions que vous vous posez sûrement sur #MeTooInceste. <https://www.watson.ch/fr/1984721063>. Retrieved 14 April 2021
- Wolf MR, Pruitt DK (2019) Grooming hurts too: the effects of types of perpetrator grooming on trauma symptoms in adult survivors of child sexual abuse. *J Child Sex Abus* 28(3):345–359. <https://doi.org/10.1080/10538712.2019.1579292>
- World Health Organization (2003) Guidelines for medico-legal care for victims of sexual violence. Sexual child abuse. Retrieved

- from <https://www.who.int/gender-equity-rights/knowledge/924154628x/en/>
- Xiong Y, Cho M, Boatwright B (2019) Hashtag activism and message frames among social movement organisations: semantic network analysis and thematic analysis of Twitter during the #MeToo movement. *Public Relat Rev* 45(1):10–23. <https://doi.org/10.1016/j.pubrev.2018.10.014>
- Zalberg S (2017) The place of culture and religion in patterns of disclosure and reporting sexual abuse of males: a case study of ultra orthodox male victims. *J Child Sex Abus* 26(5):590–607. <https://doi.org/10.1080/10538712.2017.1316335>
- Zaman T, Fox E, Bradlow E (2014) A bayesian approach for predicting the popularity of tweets. *Ann Appl Stat* 8(3):1583–1611. <https://doi.org/10.1214/14-aos741>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.