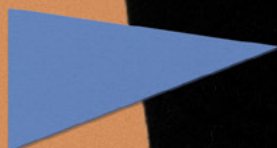
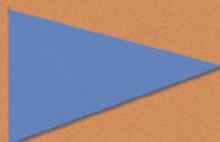



3.2

Big Data Case Studies

Carole Cadwalladr: The networked gaslighting of a high-impact investigative reporter

*Julie Posetti, Diana Maynard, Kalina Bontcheva,
and Nabeelah Shabbir*





“I have become, through evidence-based reporting, a cultural war hate figure and a ‘valid’ target.”

Carole Cadwalladr

Carole Cadwalladr is a multi-award winning British journalist whose investigative work exposed the Facebook-Cambridge Analytica scandal that ultimately led to the biggest fines in history being imposed on the social media giant by privacy regulators on both sides of the Atlantic. It also resulted in the collapse of the “sketchy” data analytics company closely associated with former US President Donald Trump’s successful 2016 election campaign (Lomas, 2019).

But from the moment *The Guardian* and *Observer*¹³¹ newspapers published the first story in Cadwalladr’s investigative series on the scandal that compromised the Facebook accounts of up to 87 million users, the journalist became the target of a malign, misogynistic, disinformation-laced campaign of online violence which grew increasingly threatening over time. This campaign has also created the enabling environment for her legal harassment by political actors.

Cadwalladr’s reportage linked the Cambridge Analytica scandal to both the election of former US President Donald Trump, and Brexit - the referendum which led to the withdrawal of the UK from the European Union. It suggested the widespread manipulation of Facebook users by political actors using microtargeting techniques, which fed into highly divisive politics in the US and the UK between 2016 and 2020. And it continues to resonate. Her journalism has led to criminal investigations, and parliamentary inquiries in multiple countries.

Despite reporting extensively on the global impacts of Facebook privacy breaches and disinformation on the platform, Cadwalladr barely uses Facebook, she said. Twitter has been the main vector for the cascading gendered abuse, threats and harassment that she endures. But the fuel is provided by pro-Brexit political actors and donors, whose abuse has been amplified by right-wing media, and even a prominent (now former) BBC Politics presenter. The objective, she said, is to hold her up to ridicule, discredit her, and thereby erode trust in her accountability journalism:

¹³¹ The *Observer* is The Guardian’s companion Sunday national newspaper. Both papers are published by Guardian News Media, along with the Guardian Weekly.

One of the most important things is the way that credentialled people become part of it. For me, it's right-wing journalists and 'trolling' MPs. That's when it becomes really hard to counteract. From the professional point of view, you're perceived as a sort of divisive, controversial figure [whereas] your male colleagues would not have attracted comment. And I think this is very effective.

In this interdisciplinary case study, the researchers conducted a big data analysis of nearly 2.1 million English language tweets directed at Carole Cadwalladr from December 1st, 2019 to January 14th, 2021. Applying the same Natural Language Processing (NLP) techniques that were used to analyse the data in the companion Maria Ressa big data case study¹³², 10,400 tweets directed at Cadwalladr were identified by Artificial Intelligence (AI) tools as clearly abusive, and subjected to granular analysis.

This quantitative research was conducted in parallel with longform qualitative interviews conducted with Cadwalladr, *Observer* editor Paul Webster, and Guardian Media Group's Director of Editorial Legal Services Gill Phillips to understand the impacts of the online violence Cadwalladr experiences, and the institutional responses to it. Then, a timeline of events was created to help us map her experience of sustained online violence over the past four years.

The nature of the abuse levelled at Cadwalladr is highly gendered, and at times misogynistic. Most typically, she is maligned and dismissed as a "mad cat lady," mocked for being middle-aged, childless and single; labelled "crazy", "hysterical" and a "conspiracy theorist". All this is "really humiliating", she said. The main goal of the abuse levelled at Cadwalladr is to discredit her professionally, thereby undermining trust in her critical reporting of the Cambridge Analytica scandal and its aftershocks, including questions about the accountability of the platforms regarding disinformation and hate speech.

The *Observer's* editor Paul Webster has witnessed the impacts on Cadwalladr: "Carole has been subjected to a fairly sustained and aggressive fusillade of online attacks in her reporting. She's a controversial writer who's chosen a series of very contested areas to report on, and she has been subjected to a great deal of online abuse." Cadwalladr's free-wheeling social media style and her refusal to be silent make her a bigger target. And when she posts controversial comments or corrects occasional errors on Twitter, the abuse becomes torrential. To Cadwalladr, this feels like the digital equivalent of a mob attack on an "unrespectable" woman. **"A few hundred years ago I would have been burned at the stake,"** she said.

The abuse has not been contained to the online realm. Cadwalladr was physically stalked in 2018 by a "shady" man with a military and cyber espionage background who first tried to befriend her, and then began sending threatening text messages (Cadwalladr, 2019a).

She has also been hit with multiple defamation claims by one of the subjects of her investigative reporting - a wealthy businessman who is also a prime instigator of the online harassment she experiences. That man is Leave.EU (the unofficial

10,400
TWEETS DIRECTED AT
CAROLE CADWALLADR
WERE "CLEARLY
ABUSIVE"

¹³² See the detailed methodology laid out in the full length Maria Ressa big data case study here: <https://www.icfj.org/our-work/maria-ressa-big-data-analysis>

pro-Brexit campaign) founder Arron Banks. Some of the tweets he directs at Cadwalladr have been received by her as menacing. The defamation cases filed by Banks against Cadwalladr pertain to her commentary on his relationship with the Russian Government. They have been labelled as Strategic Lawsuits Against Public Participation (SLAPPs) by a coalition of eight freedom of expression groups led by Reporters Without Borders (RSF) (UK Anti-SLAPP Coalition, 2021). They described the defamation actions as: “vexatious in nature and intended to silence Cadwalladr’s courageous investigative journalism” (RSF, 2019e).¹³³

The Guardian and *Observer* have also been threatened with legal action in the course of publishing Cadwalladr’s investigative reporting into Cambridge Analytica - including by Google and Facebook (Zerofsky, 2019). However Banks’ defamation claims target her as an individual - a typical characteristic of SLAPP suits. They respond to statements she made during a 2019 TED Talk and a tweet in which she shared the talk, rather than articles published by her employer on the same themes (Halton, 2019).

These factors, combined with her status as a freelance columnist at *The Observer* rather than a staff reporter, have left her personally exposed and responsible for her own legal defence. “The thing which I think is important, is that the lawsuit that I have is within this context of a four year campaign of harassment, intimidation, threats of violence... It was kind of like falling down the ‘rabbit hole.’” In her view, she is considered a threat to these very powerful actors.



Figure 15: Screenshot of a tweet from Arron Banks which quotes a tweet from the communications director (Andy Wigmore) of the unofficial Brexit campaign *Leave.EU*. The Wigmore Twitter account was suspended at the time of writing.

While the Guardian Media Group (which owns *The Guardian* and *Observer*) is not funding Cadwalladr’s defence in the SLAPP cases, her editors have publicly defended her reporting through statements like this (Bowcott and Hern, 2018):

Carole’s brave reporting has made waves around the world, and given the public much more insight into the secretive ways some powerful people and organisations have sought to influence our democracies... This case is a very worrying example of a wealthy person singling out an individual journalist and using the law to stifle legitimate debate and silence public interest journalism.

¹³³ In June 2022, Cadwalladr successfully defeated the remaining libel action brought against her by Banks in the UK High Court, relying on a public interest defence. See: <https://www.judiciary.uk/judgments/banks-v-cadwalladr/> and <https://www.theguardian.com/uk-news/commentisfree/2022/jun/19/arron-banks-set-out-to-crush-me-in-court-instead-my-quest-for-the-facts-was-vindicated>. However, Banks was subsequently granted the right to appeal the decision and the matter was ongoing at the time of writing.

In the past three years, Cadwalladr's experience of online abuse and harassment has also been influenced by the coronavirus pandemic. She said she has become a "national punching bag". Her fightback strategy in the face of sustained online violence has involved countering disinformation through investigative journalism, forming support networks with other women journalists experiencing online violence, and launching civil society collaborations designed to hold social media companies and political actors to account. It is a "solutions focus," Cadwalladr said.

These collaborations include a network of scientists critical of the UK's response to COVID-19 and the "Real Facebook Oversight Board" - both initiatives of grassroots advocacy group, The Citizens.¹³⁴

I'm just trying to process what I've gone through for the last four years... I'm trying to look structurally at the news and information ecosystem and work out how we can counteract disinformation [and] help others who are experiencing this. I really have had enough of being the national punching bag.

Cadwalladr recruited Maria Ressa - the subject of the other big data case study presented as part of this study - to sit on the Real Facebook Oversight Board (Halpern, 2020). And while their cases are not directly comparable, there are some noteworthy similarities in the patterns of attack, and the role of gendered online violence against a journalist in creating the enabling environment for legal harassment.

In both cases, 'patriotic trolling', niche bloggers/influencers and partisan media figures fuel the attacks, but domestic political actors, not foreign States, are the biggest source of the online violence they experience. As Cadwalladr points out, there is a "misconception" that online attacks come from "faceless foreign bots". She points instead to a "right wing information system that is all powerful in Britain, working at all of these different levels" - which she described as "pernicious and troublesome".

40%

OF THE ABUSE
LEVELLED AT
CADWALLADR
WAS DESIGNED TO
UNDERMINE HER
CREDIBILITY & TRUST
IN HER JOURNALISM

12 Key Findings

Types of abuse levelled against Cadwalladr:

- **55%** of obvious abuse detected by the researchers targeted Cadwalladr at the personal level. It was highly gendered and designed to hold her up to ridicule, humiliate, belittle and discredit.

¹³⁴ In October 2021 the non-profit activist group The Citizens (a group of journalists, filmmakers, advertising creatives, data scientists, artists, students and lawyers) [<https://the-citizens.com/about-us/>] tweeted about various successes, including winning a legal battle to challenge government transparency in the UK High Court <https://twitter.com/allthecitizens/status/1448039316010250241>

- **40%** of the abuse was categorised as harassment designed to undermine Cadwalladr's professional credibility and trust in her journalism.
- **5%** of the abuse was politically-based.
- **21%** of all obvious abuse (across the personal and professional categories) levelled at her was sexist, misogynistic or sexually explicit.

Characteristics of abuse against Cadwalladr:

- The online violence Carole Cadwalladr experiences is a feature of the enabling environment for her offline legal harassment.
- The abuse was constant and sustained, with several peaks per month delivering intense abuse.
- The cumulative impacts of the sustained online abuse, harassment and attacks over a four year period have created a gaslighting effect, chilling Cadwalladr's investigations and delivering deep personal impacts.

Abuse tactics and triggers:

- Disinformation tactics are clearly in evidence.
- There is obvious evidence of trolling behaviours among abusers, with key identifiable political instigators leading the pack.
- The online violence experienced by Cadwalladr has been instigated by subjects of her investigative reporting, amplified by fringe right-wing media, worsened by conservative columnists, and inflamed by the sexist behaviour of high profile establishment journalists.
- Cadwalladr's internationally consequential, multi-award winning journalism focused upon powerful individuals and entities are what made her a prime target for online violence.
- Cadwalladr is disproportionately attacked for making minor corrections and contentious comments, as the spikes of abuse in the Twitter data demonstrate.

OF ALL OBVIOUS
ABUSE AGAINST
CAROLE CADWALLADR

21%

WAS SEXIST,
MISOGYNISTIC OR
SEXUALLY EXPLICIT

The gaslighting effect

Cadwalladr has been subjected to five years of deeply sexist and misogynistic online violence that takes the form of a constant wave of abuse, with several crests each month. This pattern of abuse - building gradually over time - appears designed to destroy the confidence of the target and undermine their credibility without appearing overtly offensive on a tweet-by-tweet basis - a classic case of networked gaslighting. This method typically operates via an organised, or semi-organised, network of abusers for greatest effect (the 'pile-on', 'dogpiling', or 'brigading' approach), and can lead to echo chambers of abuse, where the same abusive message is retweeted many times. These methods are clearly in play in the case of Cadwalladr. Her abusers are frequently interlinked - organisationally, or through association, with pro-Brexit rhetoric.

i. Analysis: Hysterical hag, stupid bitch, crazy cat lady, shut up!

10,400 separate instances of obvious abuse against Cadwalladr were identified by the researchers in a Twitter dataset consisting of 2.1 million English language tweets collected between December 2019 and January 2021. Sexist, misogynistic and sexually explicit abuse represented 40% of obvious abuse in the 'personal attacks' category and 21% of abuse detected overall. It is notable, too, that the pile-on of abuse Cadwalladr experienced included significant elements of anti-journalism and anti-mainstream news media rhetoric, reflecting the demonisation of the press on the global stage, and the weaponisation of terms like "fake news" to chill critical reporting.



Figure 16: The word cloud above shows the most frequently occurring abusive terms tweeted at Carole Cadwalladr (occurring at least 20 times), normalised by case.

The dominant abusive phrases and terms illustrate the gendered nature of the online violence Cadwalladr experienced during the period studied (e.g., variations on "stupid woman", "crazy cat lady", and "witch"). The abuse is clearly designed

to belittle and humiliate her, while also eroding trust in her reporting. The main objective appears to be to intimidate her into silence.

Next, after analysing the clearly identifiable abuse in the ‘personal’ category (55% of all abuse), the researchers classified it into two subcategories:

- **Sexist, misogynistic or explicit abuse (40%)**
- **Other types of personal abuse (60%)**

40% of the abuse identified is aimed at undermining Cadwalladr’s journalistic credibility, especially her alleged “stupidity” as well as her integrity. Typical slurs of this kind include labelling her a “liar” or claiming she is talking “crap”, “bullshit”, or “bollocks”. The objective of this abuse is to scare her off critical reporting. She is frequently told to “fuck off”. She is often referred to as “Carole Codswallop” or “Codswallop Cadwalladr”, again insinuating that she ‘talks rubbish’ and her journalism is not to be trusted. But it is the sustained, low-intensity, high-frequency, high-volume nature of the attacks Carole Cadwalladr experiences, not only the content of the abuse, that can be understood to be so cumulatively damaging.

Looking at breakdown by frequency, we established that over 25% of the credibility-related abuse involves abusive terms implying that Cadwalladr is stupid or mentally ill in some form (e.g., “idiot”, “twat”, “moron”, “cretin”, “silly”, “crazy”, “fool”, “nutter”), with the term “liar” also appearing over 300 times, alongside terms like “shut up”, “STFU” and “fuck off”. This language is specifically designed to undermine Cadwalladr’s journalistic reputation and professional credibility while also seeking to silence her. ‘Sexist, misogynistic, and explicit’ terms include examples such as “dickhead”, “fuck off”, “cunt”, “witch”, and “hag”, along with terms involving sexual acts and intimate body parts.

Instances of political abuse frequently involve anti-Brexit sentiment such as calling her a “remoaner” (a pejorative term for people who supported the campaign for the UK to remain in the EU). Other insults included political abuse such as “Common Purpose globalist whores”, “Fuck off you commie twat” and frequent use of “Libtard” (a pejorative term for a liberal thinker) which categorised Cadwalladr as left of centre and lacking objectivity - another way to discredit a journalist in the UK context. Some of the worst abuse levelled against Carole Cadwalladr in this category during the period of this study involved people wishing she would die, as illustrated by this tweet:

@carolecadwalla The only failure is scum like you. You [sic] still throwing your toys out the pram that we left the EU. Hope you get COVID and die from it¹³⁵

Many tweets using such hashtags combined misogynistic language with pejorative descriptions of the mainstream press. Below is an example of a tweet sent to Cadwalladr after she tweeted a link to a story about this UNESCO-supported

¹³⁵ This tweet, from a now suspended Twitter account, was published November 20th, 2020.

research, ahead of her appearance on a 2020 World Press Freedom Conference panel (Cadwalladr, 2020; Posetti, Harrison and Waisbord, 2020):

@carolecadwalla @julieposetti Female? Male? All #scummedia should be treated accordingly - there should be no discrimination when it comes to destroying all you Common Purpose globalist whores. Is there anything lower than you and your ilk? #presstitute #evil

This now deleted tweet, which also tagged this study's lead author (i.e. it radiated to capture affiliated targets) features the familiar hashtag associated internationally with gender-based online violence against women journalists - #presstitute.

According to the analysis of the Twitter dataset studied, the main themes associated with abuse against Carole Cadwalladr were: democracy; public health and COVID-19; Brexit, Europe and immigration; and foreign affairs. These are also the issues that Cadwalladr most frequently wrote about and commented on during the period in focus.

But one of the most striking features of the abuse against her was not the emphasis on these issues as such, but the deployment of misogynistic tropes. She is frequently labelled "mad" and "hysterical" by critics who emphasised her age, relationship status and childlessness, and questioned her desirability. This sexist 'othering' is a way to treat her as a deviation from the norm. It has also been amplified by some very high-profile figures in the UK press, in addition to popular right-wing blogs. To Cadwalladr, this has all the hallmarks of a witch hunt.

"Codswallop", the "crazy cat lady" and the former BBC presenter

2,921 mentions of the term "cat lady" and its various manifestations (e.g., "cat woman") were detected in the dataset of obvious abuse against Cadwalladr. It is one of the most pernicious and pervasive forms of abuse in evidence, and it was usually used in conjunction with "mad" or "crazy." "I would still say that to this day Andrew Neil is largely responsible for the ongoing misogynistic abuse and threats that I get," Cadwalladr said. She is referring to a now deleted tweet from the then BBC Politics presenter and chair of the parent company of the conservative magazine *The Spectator*, Andrew Neil. In November 2018, Neil tweeted about Cadwalladr, branding her a "mad cat woman" and ridiculing her as "Karol Kodswallop", insinuating that she resembled a character from the satirical cartoon *The Simpsons*.

According to this research, the Leave.EU donor and founder Arron Banks - who, as described above, pursued SLAPP-style defamation claims against Cadwalladr - was the first Twitter user to disparage the journalist as a "crazy cat lady". He also frequently labelled her "Carole Codswallop". This language was used to denigrate her as she pursued stories probing possible links between the Cambridge Analytica scandal and Leave.EU, the unofficial pro-Brexit campaign that Banks funded. Ten days before Neil called Cadwalladr a "mad cat woman", she wrote in *The Guardian*

that since her investigative reporting began to focus on Leave.EU and Banks, he had called her: “...hysterical, insane, a lunatic, a mad woman, a conspirator, a loony, a mad cat lady, a nasty piece of work, a criminal, a bully, a mad cat lady, a loony, a tinfoil hat nutter, a hacker, a mad cat lady, a loony, a bitter Remoaner, a lone conspiracy theorist, an enemy of the people” (Cadwalladr, 2018c).



Figure 17: Business Insider's Senior Politics Editor Adam Bienkov¹³⁶ defends Cadwalladr after she receives sexist abuse on Twitter from Andrew Neil.



Figure 18: Cadwalladr tweets about the abuse she received following Andrew Neil's 2018 tweet.¹³⁷

This abuse mirrors a pattern of online violence against women journalists that is now recognisable internationally. It involves a process of instigating, amplifying and legitimising gendered online harassment, abuse and attacks. In Cadwalladr's case, the sexist and misogynistic tropes deployed by Banks and his supporters were amplified by the popular right-wing blog Guido Fawkes, and then legitimised by sympathetic voices in the mainstream news media.

Within this right-wing information ecosystem, we see abusive content about Cadwalladr being cross-pollinated through retweeting and quote-tweeting by various actors affiliated with, or sympathetic to, the campaign for Britain's exit from the EU, and/or disparaging towards Cadwalladr's reporting. Notable among them during 2017/2018 are Arron Banks, Leave.EU's Twitter account, Leave.EU's Communications Director Andy Wigmore (who remained suspended from Twitter

¹³⁶ <https://twitter.com/AdamBienkov/status/1062295554946932736?s=20>.

¹³⁷ <https://twitter.com/carolecadwalla/status/1064074315828744192?s=20>

in mid 2022), the Guido Fawkes blog, a foreign State actor, *The Spectator*, and its holding company chair, Andrew Neil (Murray, 2020; Cadwalladr, 2019b).

Neil deleted his “mad cat woman” tweet following a backlash, but he did not apologise (Mayhew, 2018a). The BBC’s corporate communications team did, however, issue a tweet in which they stated that Neil recognised the tweet was “inappropriate” (BBC News Press Team, 2018). The following day, the then BBC Director General Tony Hall told a conference (BBC Media Centre, 2018):

On Twitter there are constant anonymous threats to journalists simply reporting on opinions that some people might not want to hear. Some of the material that journalists have had to face is quite frankly disgraceful. It is an attempt to intimidate people and stop them doing their jobs. For the sake of all journalists - we need to defend our role - seeking out the facts, no matter how inconvenient they may be for others. Because journalism matters - whether you’re in broadcasting, in the press or working online.

The next week, *The Guardian* reported that a number of senior women journalists at the BBC had complained to executives about Neil’s tweet, suggesting inequitable treatment regarding social media policy enforcement (Waterson, 2018; BBC, 2018e). A BBC spokesperson was quoted as saying: “The tweet was sent from Andrew’s personal account, however the BBC has social media guidelines which it expects all staff to follow and these have been discussed with Andrew.”

Meanwhile, the misogynistic insult continues to reverberate: “I am now called a ‘mad cat lady’ dozens of times a day on social media, every day,” Cadwalladr said (Cadwalladr, 2018d). In addition, “Codswallop”, and Neil’s variant “Kodswallop”, are also pejorative nicknames still in regular use in abusive tweets against Cadwalladr according to this research.

ii. Who is sending the abusive tweets targeting Carole Cadwalladr?

Of the 10,400 obviously abusive tweets identified in the dataset, there are 7,744 unique authors. Of the serial abuse senders, the most prolific sent 38 obviously abusive tweets, while 155 tweeters sent five or more that the NLP tool categorised as abusive. Only two of the top eight most prolific abuse senders in the dataset still have live Twitter accounts. Of the other six, three have been deleted and three were suspended at the time of writing. Leave.EU’s Director of Communications Andy Wigmore was the Twitter user who abused Cadwalladr most prolifically during the period. He sent 38 highly abusive tweets before his account was suspended on November 8th, 2020. Here is one of the abusive tweets Wigmore sent before his suspension, in which he alludes to SLAPP suits and tags both Andrew Neil (who triggered the “mad cat woman” pile-on of 2018) and Arron Banks, who sued Cadwalladr for libel:



*Pointing out facts by @afneil (a proper journalist) must be like a SLAPP round the chops or scraping your nails down a chalk board
Codswallop @carolecadwalla get used to it more facts coming your way. @Arron_banks*

The second most abusive account was an anonymous Twitter account which overlapped with another anonymous account using the same profile photo and an almost identical Twitter handle. Both accounts were suspended in late 2020. Out of the 7,744 authors of obviously abusive tweets in the dataset, as of March 15th, 2021, 943 of these had deleted their accounts, and 1,021 had their accounts suspended. In total, this means that just over 25% of the accounts are no longer active. This suggests that those accounts abusing Cadwalladr most prolifically are people whose misconduct on the platform has been relatively extreme, generating a high removal response rate from Twitter. Authors of abusive tweets also: have more recently established accounts than tweet authors in the non-abusive set; have fewer followers; follow fewer users; and post slightly fewer tweets. Viewed together, this data pertaining to the behaviour of the Twitter users in the dataset who were obviously abusive towards Cadwalladr during the period under examination indicates clear patterns of trolling behaviour with identifiable instigators and ‘ring leaders’.



Figure 19: Cadwalladr says she is a target of a toxic right-wing media ecosystem in the UK.

iii. Abuse spikes and triggers

Cadwalladr is an outspoken and reactive Twitter user, reflecting her status as a high-profile columnist at the *Observer*, and she is relatively unfettered in her social media conduct compared to journalists bound by corporate social media policies because she is not a staff writer. At the same time, she is also a woman and a journalist who feels battered after years of exposure to unrelenting online violence. “It’s changed my life,” Cadwalladr said.



Early on, she was hacked; she has been stalked; she believes that the trolling has seriously damaged her reputation within the UK news industry; she acknowledges her reporting has been chilled; she has experienced the trend of victim blaming - judged for “having answered back on Twitter” - which is slowly muting her social media use; and two of the whistleblowers at the heart of her Cambridge Analytica investigations were also targeted in coordinated online and news media attacks designed to undermine their credibility, she said.

All of this was going on as Cadwalladr was winning a cache of premier British journalism awards, such as the 2018 Orwell political journalism prize, and she was recognised as a Pulitzer Prize finalist with the *New York Times*. Then, she faced possible bankruptcy, as defamation and libel actions brought by one of the powerful subjects of her award-winning investigative reporting wended their way through the courts. “I’m processing actually what it’s like being under attack for four years. At various points, I think I have overreacted to things as well, you know. You’re so used to being attacked that you respond like you are on a bit of a trigger.”

Her realtime reflections are weaponised against her, as are any realtime errors. Cadwalladr corrects the few mistakes she does make but she is offered no forgiveness, nor shown any grace on Twitter by her dedicated detractors and trolls. The dataset demonstrates a continuous ebb and flow pattern of abuse, frequently associated with a backlash against her commentary, with two to three attack spikes per month. This sort of digital gaslighting - a long-range persistent pattern of abuse - is designed to wear the target down.

As a woman covering the global political, social and human rights ramifications of big technology and viral disinformation, and historically divisive political movements (on both sides of the Atlantic); who dared to break major investigative stories as a features writer and columnist, Cadwalladr could be seen as a natural lightning rod for online abuse in the toxic communications ecosystem of the early 21st century. But while her reporting and tweets are often cast as “controversial”, the abuse she receives is an entirely disproportionate response to her occasional errors and her engaging, reflexive, opinionated and consequential writing.

iv. The role of The Guardian and Observer

The online violence Carole Cadwalladr is subjected to is designed to silence her, chill her critical journalism and discredit those who publish it, namely: *The Guardian* and *Observer*. In August 2020, Guardian Media Group took legal action to shutdown a website generating fake *Guardian* headlines and byline profiles which were being shared with the trending Twitter hashtag #TrollingTheGuardian. Carole Cadwalladr was one of the journalists targeted by the now defunct guardianmeme.com website. But the action was triggered by a tweet from a UK member of parliament who was similarly attacked after she wrote a *Guardian* column (Mayhew, 2020b). The MP tweeted that the site put her in danger, along with her family and staff.

Guardian Media Group’s decision not to fund Cadwalladr’s defence in the ongoing defamation cases - the enabling environment which includes the online violence she experiences - is a source of grief and frustration for Cadwalladr, who regularly pointed out that she could “lose her house” as a result of the litigation. The

company's position was that because the defamation and libel claims were made against Cadwalladr as an individual, in relation to speeches (reviewed by her employer prior to delivery) and a tweet, not reports carrying similar claims and imputations published by the outlets, they are not a party to the legal action, and they are unwilling to set a precedent by funding Cadwalladr's defence.

"If she had been sued individually over something she wrote for us, we would be all over it. But if we do this for her, where do we stop?" Editorial Legal Director Gill Phillips asked. "So we've had to say, really regrettably, we cannot financially support this, but we will give you whatever support we can morally, ethically, including every time we do anything on SLAPPs. And we will continue to defend the journalism she does for us."

However, Cadwalladr said that her concerns about her employer's response to the online violence and associated legal harassment she experiences, are not purely financial. She also described what she perceived to be a lack of organisational support. Her response has been to crowdfund her legal defence (GoFundMe, 2019). She raised several hundred thousand pounds from supporters. She is also working proactively to advance protocols for dealing with online violence against women journalists at *The Guardian* and *Observer* - emphasising the function of disinformation campaigns in such attacks, along with the experiences of more isolated freelancers, and the need to publicly defend women journalists in the midst of attacks.

v. The role of the platforms

Cadwalladr said she is very frustrated with the big tech companies that act as vectors for online violence, viral disinformation and privacy erosion. Facebook and Google have been the focus of her investigative journalism while the online attacks she experiences are facilitated by Twitter. "I report stuff all the time on Twitter and they never ever do anything. They never take it down," she said. "There's these endless spoof accounts...I'm always reporting stuff." She was particularly aggrieved by Twitter's failure to remove a Leave.EU deep fake video in which she was depicted being repeatedly slapped. The video, which featured one man with a gun and one with a hammer, was clearly "incitement to violence," in Cadwalladr's view. In the end, it was *The Guardian's* intervention in contacting the source of the post that resulted in its removal, 48 hours later, she said.

Believing that reporting, blocking, deleting and muting are acts of futility against self-regenerating troll armies, Cadwalladr has turned to grassroots campaigning and the development of loose support networks through her work at The Citizens. This has also involved working with data scientists she has met on Twitter to map networks of abuse and develop response mechanisms to support other users under attacks. Concerning Facebook and the impunity with which the company has been able to act, Cadwalladr said representatives repeatedly lied to her in the course of the Cambridge Analytica investigation, and she believes that the company is just too big to hold accountable:

The FTC fined them a record US\$5 billion but it had no impact whatsoever. Their share price actually went up...because there is no mechanism to hold them to account. The only accountability structure we have has no effect because the platforms are too big. You have got to find other ways, and more creative ways essentially.

Her alternative approach has been to launch the “Real Facebook Oversight Board” - a collection of academics, civil society experts, social media critics and prominent journalists who do not have faith that the Facebook-funded official Oversight Board and its limited remit will do the kind of urgent accountability work needed, with appropriate transparency, and at the scale required.

One thing that needs to change within the social media companies is resourcing to deal with online abuse and harassment, according to *Observer* Editor Paul Webster. “They need far more people,” he said. Adding: “They need to recognise their roles and responsibility as publishers and act accordingly. So curate that material in the way that we do - take legal responsibility for it. So if somebody is libelled in tweets and in online posts, then you have recourse to the people who carry the material as well as the people who make the libels.”

While *The Guardian* has had some limited success in getting the platforms to remove abusive content in certain cases, and deplatform individual offenders in others, the scale of the problem is not able to be managed by present systems, according to Guardian Media Group’s Director of Editorial Legal Services Gill Phillips: “Look, of course, we can write letters and ask people to stop it, and point out it’s horrible. [But] the only thing that can stop it is the people who are giving them the platforms that allow them to make these attacks.”

