



European Digital Media Observatory

EDMO IRELAND HUB BRIEFING REPORT

January 2023

The Journal

 NEWSWHIP



The
University
Of
Sheffield.



Institiúid DCU um Meáin Tírdhál,
Doimlethas agus Sochaí
DCU Institute of Future Media,
Democracy and Society



EDMO
Ireland



Co-financed by the Connecting Europe
Facility of the European Union

www.edmohub.ie

EDMO Ireland Hub Briefing Report

Disinformation in Ireland

January 2023



EDMO Ireland is one of fourteen hubs established as part of the European Digital Media Observatory (EDMO) to combat disinformation by bringing together fact-checkers, media literacy experts, technologists, and academic researchers.

EDMO Ireland comprises the Institute for Future Media, Democracy and Society at Dublin City University (coordinator), TheJournal, NewsWhip, and University of Sheffield.

Web: <https://edmohub.ie/>

Twitter: [@Ireland_EDMO](https://twitter.com/Ireland_EDMO)

Email: edmo@dcu.ie



TheJournal



**Co-financed by the Connecting Europe
Facility of the European Union**

EDMO has received funding from the European Union under Contract number: LC-01464044

Contents

| | |
|---|----|
| Introduction and summary | 4 |
| Developing situational awareness of disinformation in Ireland | 6 |
| How EDMO Ireland monitors disinformation | 6 |
| Assessing influence | 8 |
| Disinformation trends in Ireland | 11 |
| Actors | 11 |
| Narratives and themes | 12 |
| Tactics | 17 |
| Pre-bunking as a preemptive strategy | 20 |
| EDMO Ireland's pre-bunking campaign | 21 |
| Fostering a healthy information environment | 22 |
| Supporting fact-checking with AI-driven tools | 26 |
| Media literacy as resilience building | 28 |
| European Policy developments | 31 |
| Glossary of relevant terms | 35 |

Introduction and summary

As part of the EU strategy to limit the impact of harmful disinformation campaigns, the European Digital Media Observatory (EDMO) was established by the European Commission in 2020. It aims to counter disinformation by fostering national and multinational collaborations among researchers, fact-checkers, media literacy practitioners, technologists, and other stakeholders. EDMO Ireland is one of fourteen hubs operating across the EU.

EDMO Ireland is comprised of four partners:

- The *Institute for Future Media, Democracy, and Society* at Dublin City University (DCU FuJo) acts as coordinator, analyses disinformation, and develops media literacy campaigns;
- *TheJournal*, Ireland's only national fact-checker, investigates claims and collaborates on transnational investigations;
- *NewsWhip* deploys its commercial tools for analysing content trends and interactions; and
- The *University of Sheffield* develops new tools to enhance the detection of false claims and to support fact-checkers across the EU.

The formation of EDMO Ireland was supported by a wide range of stakeholders - including national authorities, media industry bodies, researchers, and civil society organisations - who share a common concern about the threats posed by disinformation and a common recognition of the need to build capacity to counteract these threats. EDMO Ireland aims to build that capacity through its core activities.

The briefing report provides insights into the work of EDMO Ireland across the following areas:

- **Developing situational awareness of disinformation in Ireland:** this section explains the methods and tools employed to monitor disinformation.
- **Disinformation trends in Ireland:** this section outlines major disinformation trends in Ireland in terms of actors, narratives, and tactics.
- **Pre-bunking as a preemptive strategy:** this section introduces pre-bunking as a counter-disinformation strategy and describes our current pre-bunking campaign.
- **Fostering a healthy information environment:** this section outlines how *TheJournal*'s fact-checking and outreach activities aim to improve the

information environment. It further highlights how AI tools can aid journalistic and fact-check investigations.

- **Building resilience through media literacy:** this section summarises the role of media literacy in countering disinformation and describes the range of media literacy activities undertaken by EDMO Ireland and its partners.
- **European policy developments:** this section briefly outlines EU policy developments as a context for developments at the Irish level. It further highlights the role of EDMO and previous work by DCU FuJo.

Developing situational awareness of disinformation in Ireland

A key goal in addressing online disinformation is developing situational awareness to anticipate potential threats and to develop appropriate responses including, for example, prebunking emerging narratives before they become viral. To develop situational awareness, it is necessary to maintain an up-to-date understanding of disinformation actors, narratives, and tactics across different platforms. It is also necessary to keep abreast of international trends as these have the potential to emerge, in some form, in the Irish context.

Maintaining situational awareness can be challenging given the volume of online media and the dynamic nature of the environment. For example:

- Disinformation is networked across a wide range of online platforms as well as mainstream and alternative media outlets;
- New actors, behaviours, tactics, and even platforms emerge in response to wider developments in society, politics, and media;
- Bad actors are opportunistic in exploiting current and unforeseen events; and
- Bad actors adapt their tactics in response to new developments and to evade disinformation countermeasures.

Consequently, it is necessary to monitor the disinformation environment on an ongoing basis using a variety of methods. This further entails training personnel with sufficient expertise to implement robust monitoring and employing relevant tools and technologies to complement human expertise.

How EDMO Ireland monitors disinformation

EDMO Ireland monitors disinformation trends across multiple online platforms through a mix of ethnographic and technological analyses. Put simply, ethnographic analysis involves spending regular time in online spaces to observe and assess their dynamics including the emergence of new platforms, groups, actors, or narratives that need to be monitored. This is complemented by a technological analysis of key platforms using NewsWhip Spike, a real-time media monitoring platform that allows users to track false stories, monitor questionable sources and narratives, and predict public engagement with a story.

In total, EDMO Ireland monitors 12 online platforms on a regular basis. These include well-known mainstream platforms (e.g., Twitter), less-well known mainstream platforms (e.g., Reddit), and alternative platforms with limited moderation policies (e.g., Telegram). The textbox below provides a brief description of each platform and how it is monitored. With the exception of Reddit, all the mainstream platforms listed

below are signatories of the EU Code of Practice on Disinformation. In contrast, none of the alternative platforms are signatories.

| Monitored platforms |
|---|
| <p>Mainstream platforms:</p> <ul style="list-style-type: none">● Facebook: Private Facebook Groups facilitate discussions among members with Group admins responsible for implementing moderation. DCU FuJo currently monitors 20 Irish-based Facebook Groups to track their disinformation activities. NewsWhip provides access to more than 200,000 public Facebook Pages.● Instagram: Instagram is a free photo and video sharing app that is especially popular among brands, celebrities, and influencers. NewsWhip has global coverage of content from “creator” and “business accounts. In addition, DCU FuJo monitors notable Irish influencers who have spread disinformation in the past.● Reddit: Reddit is a social news website and forum where content is socially curated and promoted by site members. NewsWhip provides access to all posts and comments across the 130,000 Reddit communities including “r/Ireland”. This community is not currently a major focus of disinformation activity, but it does provide an indication of the kinds of narratives that are gaining traction.● TikTok: TikTok is a short-format video platform that presents users with individualised video feeds based on their previous interactions. DCU FuJo monitors hashtags and engages with content to mimic the feed of those who engage with conspiracy theories and extremist content.● Twitter: Twitter is a micro-blogging social network that is often associated with breaking news. NewsWhip provides access to the real-time feeds of tweets from the most influential Twitter accounts. In addition, DCU FuJo monitors 63 individual accounts to examine how key actors interact and their following, replying, and liking behaviours.● YouTube: YouTube is a video sharing service where users can watch, like, share, comment and upload their own videos. NewsWhip Spike provides access to analytics about viral videos and influencers.● WhatsApp: WhatsApp is a messaging service that lets users send text, images, audio, or video to individuals or groups. TheJournal receives alerts from readers through a WhatsApp channel. These alerts are a means of identifying what kinds of disinformation are circulating in the encrypted groups of private individuals. |

Alternative platforms:

- **Bitchute:** see Oydsee
- **Odysee:** Both Odysee and Bitchute are video-hosting platforms that typically decline to implement content moderation policies. They are the go-to hosting sites for videos that breach the content standards of mainstream platforms. Both platforms host videos from RT (Russia Today), circumventing EU sanctions on the channel.
- **Telegram:** Telegram is a messaging platform that supports group communication through groups and channels. With only limited moderation, it is widely associated with extremist groups. DCU FuJo currently monitors 51 channels and 10 groups.
- **Gab:** see Gettr
- **Gettr:** Both Gab and Gettr operate as alternatives to Twitter. With lax moderation policies, they have attracted extremists and those banned, or likely to be banned, from other platforms. These platforms are not widely used in Ireland, but as some key actors maintain a presence on these platforms, DCU FuJo monitors their profiles.

In addition to the platforms listed above, EDMO Ireland uses technological and ethnographic analysis to track how certain narratives, topics, or terms are diffused across mainstream and alternative media including “alternative news” sites, blogs, and podcasts. As explained below, this facilitates an understanding of how certain false claims or ideas are mainstreamed and normalised overtime as they spread out from fringe online spaces.

Assessing influence

The existence of disinformation does not in itself indicate that it needs to be acted upon. An individual piece of disinformation may not have reached a wide audience or the audience may readily recognise it for what it is. Responding to disinformation in these scenarios may draw more attention to false claims than they would have ordinarily received. Moreover, there are different ways in which a piece of disinformation becomes influential and understanding these dynamics can play an important role in determining appropriate responses.

Sudden viral impact: In some scenarios, a false claim gains wide recognition in a short space of time as it spreads across online platforms. To assess this, NewsWhip Spike

provides insight into the traction and likely influence of stories across all major social platforms (see Figure 1). It also predicts the number of interactions a story is expected to earn 24 hours into the future, which facilitates the rapid detection of emerging disinformation stories and helps analysts and fact-checkers to prioritise where to focus their attention and efforts.

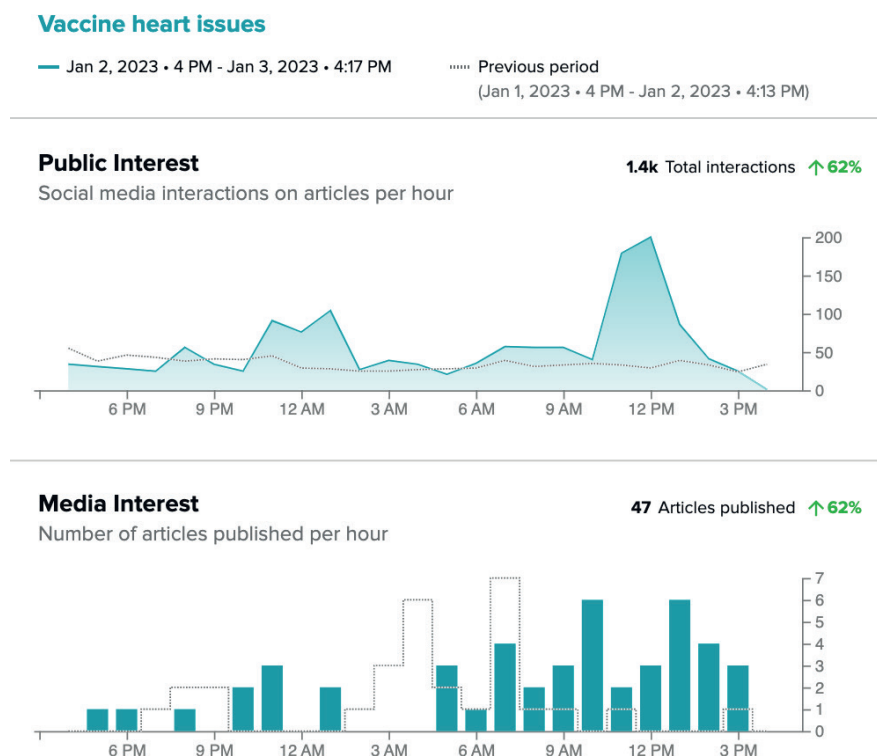


Figure 1: NewsWhip prediction charts for the topic “vaccine heart issues”. There has been a resurgence of Covid-19 disinformation in relation to sudden death and heart-related deaths.

Diffuse impact: In other scenarios, false claims have a more diffuse impact. Rather than a direct trajectory from publication to wide audience reach, false claims circulate in different online and offline media spaces and move from the fringes to the mainstream overtime (see “Plantation” textbox). Tracking this process is important because the normalisation of key claims and concepts is a goal of extremist actors, but the public at large may be unaware of the origins of these terms. This type of analysis typically requires in-depth human expertise. At the same time, a broader picture of changes can be obtained through NewsWhip Analytics as it enables the tracking of historical engagement trends around relevant content and among specific demographics.

“Plantation”:
Adapting the “great replacement” conspiracy theory for an Irish context

The popularisation of the term “plantation” is an example of diffuse impact whereby Irish extremists seek to adapt the “great replacement” conspiracy theory for an Irish context. The “great replacement” conspiracy theory is promoted by white nationalists and right-wing extremists. Put simply, it states that policies facilitating migration to Western countries are part of a plot to undermine or “replace” the political power and culture of white people.

- *Fringe origins:* The terms “plantation” and “great plantation” began gaining notable traction in Irish Telegram groups in September 2022 where they are used to describe an agenda to replace Irish culture, identity, and ethnicity. These terms specifically link the historical colonisation of Ireland to the “great replacement” conspiracy theory and to contemporary events.
- *Promotion:* By mid-November, this terminology was employed by alternative news outlets, including TheLiberal.ie, in coverage of anti-immigration protests and by members of the National Party and Irish Freedom Party. These accounts helped popularise the term across mainstream social media platforms including Facebook, Instagram, Twitter, and YouTube.
- *Mainstreaming:* The term is now becoming widespread. Anti-immigrant and anti-refugee protestors use the term in media interviews and it is even used by those attempting to counter the anti-refugee agenda. As such, it is an example of the successful normalisation of conspiracy theory concepts, which have the potential to draw people further towards extremist ideas.

Disinformation trends in Ireland

Below we provide an overview of disinformation trends in Ireland in terms of actors, narratives, and tactics. In light of recent and current events, there is a specific focus on activities by right-wing extremists.

Actors

Although disinformation can be produced or propagated by any actor, it is consistently promoted by those advocating conspiratorial, extremist, and anti-democratic views. In many respects, the Covid-19 pandemic marked a major turning point for disinformation in Ireland as conspiracy theorists, right-wing extremists, and anti-establishment protestors gained prominence on online platforms. Recently, right-wing extremists have become the most prominent group propagating disinformation across online platforms. This diverse group is discussed in more detail in the narratives section. The different types of actors operating in Irish online spaces are summarised below.

Fringe political parties with no elected representation such as the National Party and the Irish Freedom Party are particularly active across multiple platforms including Facebook, Instagram, Twitter, and Telegram. They operate party accounts at national level and, more recently, at local level (e.g., National Party Cork). Key figures from these parties also operate accounts such as the Irish Freedom Party leader Hermann Kelly.

Online influencers and personalities are those who have built up a following based on exposing or countering a mainstream media narrative. For example, Dave Cullen (AKA “Computing Forever”) operates a YouTube channel with more than 200,000 subscribers. It features cultural and media commentary as well as interviews with Irish conspiracy theorists and extremists. His Telegram channel has more than 10,000 followers. Such influencers often have a presence across mainstream and alternative platforms and tailor their content for the moderation policies of each platform type. In some instances, they sell products or take donations/subscriptions to fund their content.

Alternative media outlets are a noted feature of the disinformation environment in many countries. The network of alternative outlets in Ireland is growing with news websites, blogs, and podcasts. Typically, these outlets position themselves in direct opposition to mainstream news while presenting commentary as journalism. For example, The Irish Inquiry defines itself as “free speech in search of the truth” and operates a YouTube channel with more than 15,000 subscribers. It also posts videos across other platforms (see Figure 2).

Anonymous users are particularly active on Twitter. Although these accounts are not linked to any real person's identity, their affiliations are often made clear through their choice of name (e.g., IrishPatriot) and profile images (e.g., Irish flags). They play a role in amplifying content by engaging with accounts discussing major topics such as migration, politics, and vaccines.

Bots are automated programmes that generate and/or engage with content on a particular platform, most notably Twitter. Many high-profile individuals in Ireland, especially journalists, have reported a recent increase in bot followers on Twitter.



Figure 2: A video posted to multiple platforms by The Irish Inquiry in January 2023. The video claims Irish mainstream media are hiding evidence of vaccine deaths. The tone and style are part of a larger narrative that seeks to undermine trust in professional journalism and encourages a direct opposition to “mainstream media”. The video received over 1,148 views on Telegram; 15,000 views and 242 retweets on Twitter; and 131 shares, mostly to individual user feeds, on Facebook.

Narratives and themes

The narratives and themes that circulate in Ireland reflect broad international patterns with conspiracy theories and right-wing extremism flourishing in alternative platforms and increasingly visible on mainstream platforms. The narratives propagated by right-wing extremists in Ireland are broadly aligned to a nativist ideology. Nativism is a particular construction of nationalism that advocates protecting the interests of native or indigenous inhabitants over those of immigrants.

Nativist narratives have circulated for many years in Ireland, rising to prominence in tandem with wider events that are ripe for exploitation by extremists. The homelessness crisis; the crisis in accommodation for asylum seekers and refugees; the advance of diversity and inclusion; and the Covid-19 public health measures have all provided Irish right-wing extremists with an opportunity to promote their ideology to a wider audience. The current wave of protests against accommodation for refugees and asylum seekers follows protests and attacks against proposed direct provision centres in 2018 and 2019. It also follows explicit efforts to stoke fears of a “race war” through online campaigns. For example, following the shooting of George Nchenko by a member of the Garda Armed Support Unit in 2021, there was an organised effort to promote racist disinformation on Irish social media and to strengthen links with international right-wing extremists who were asked to participate in an online campaign promoting disinformation and hate speech.

Some of the key narrative themes that are prominent in the current protest include:

- **Threatening males:** References to “unvetted males” and “men of fighting age” are used to imply an imminent danger to Irish people, especially women and children, and their communities (see Figures 3 and 4). The National Party employs this rhetoric in its recruitment literature, which instructs Irish men to “make Ireland safe” (see Figure 5).
- **Inauthentic claims:** References to “economic migrants”, “fake asylum seekers”, and “fakeugees” imply that refugees and asylum seekers are lying about the dangers they face and attempting to scam the state.
- **Population replacement:** References to “an invasion” and “plantation” invoke fears about the decline of Irish identity and white ethnicity. The idea that Ireland is being “colonised” has been picked up by international conspiracy theory media including the notorious conspiracy theory website Infowars.
- **National chauvinism:** References to “Ireland first”, “Ireland belongs to the Irish” and “Ireland is full” advocate for the prioritisation of Irish people and the idea that Ireland is unable to accept any more refugees or asylum seekers. These terms are used as slogans and often trend on social media in tandem with discussions of the homelessness crisis. At the same time, a nativist ideology is often communicated symbolically through the prominent use of the Irish tricolour.

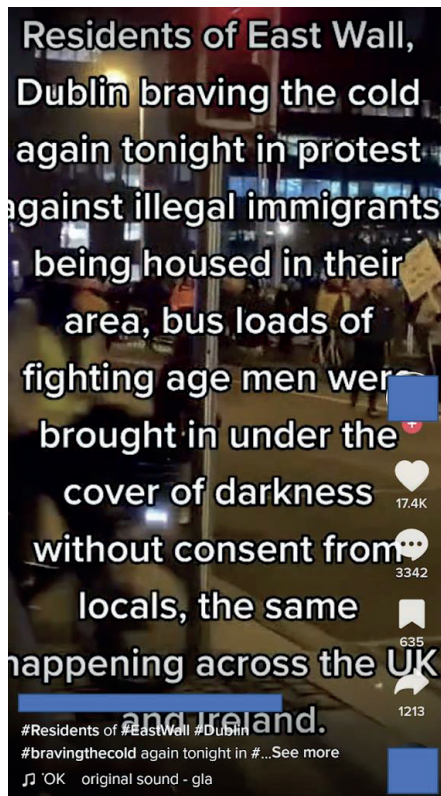


Figure 3: A TikTok post with over 17,000 views references “bus loads” of “fighting age men”.



Figure 4: A Telegram post depicts a refugee support poster which has been edited to imply that welcoming refugees is the same as welcoming rapists



Figure 5: National Party material shared in person and via social media in the first half of 2022

Other major narratives found in Ireland are presented below in order of prominence:

- **Conspiracy theories**, especially claims about an elite conspiratorial agenda, are a common factor across disinformation narratives. For example, in reference to an economic recovery plan drawn up by the World Economic Forum, conspiracy theorists invoke “the great reset” agenda to challenge election results, climate action, LGBTQ+ policies, the war in Ukraine, and public health measures. The concept of an elite conspiracy is also frequently invoked through discussions of a “NWO (new world order)”, “globalists”, “cabals”, and general opposition to “MSM (mainstream media)”. While such content can be localised (see Figure 6), much of what circulates in Irish spaces emanates from the US and US media personalities such as Tucker Carlson.
- **Health and wellness** disinformation narratives, especially in relation to the Covid-19 vaccine, are still very prevalent in Irish disinformation spaces. Much of this content is not Irish specific as it is shared from an international network of anti-vaccine actors. Recurring themes exploit news stories about celebrity deaths; sudden deaths, especially among young people, and excess deaths (see Figure 7); and illnesses such as RSV. The US documentary *Died Suddenly* (Peters 2022) was debunked by The Journal due to its prominence among anti-vaccine proponents in Ireland.
- **Gender and sexual identity** narratives emerge primarily from the UK and the US. Stories about drag queens and trans rights are prominent with recurring claims attempting to link LGBTQ people and LGBTQ education to child grooming (see Figure 8). These narratives are part of a wider “anti-woke” narrative that mocks social justice campaigns and efforts to promote diversity and inclusion.
- **Science and environment** disinformation narratives are not as prevalent as the preceding topics, but are nonetheless highly visible, especially on Facebook and Twitter. Climate change is often viewed through a conspiratorial lens that seeks to “expose” the ulterior motives behind climate action and downplay the evidence for climate change (see Figure 9). Greta Thunberg is a frequent target for abusive language and humour relating to climate change while Met Éireann weather warnings are viewed with suspicion. More generally, Irish environmental issues feed into wider anti-elite and “rural Ireland versus Dublin” narratives, which are very prominent.
- **Russia’s invasion of Ukraine** generated high volumes of disinformation in the first few months of the war, but it receives little discussion in Irish disinformation spaces now. Known claims from Russia, such as denying Russian atrocities and casting blame on Ukrainians, are sometimes circulated and primarily on Telegram.



Figure 6: Conspiracy theory narratives: An Irish Telegram channel posts about “The Great Reset” as an attempt to “centralise power over everyone” and violate “the people’s sovereignty”



Figure 7: Health and wellness narratives: A post by a Facebook Group user suggests that Ireland is experiencing excess deaths due to vaccine-induced sudden deaths.

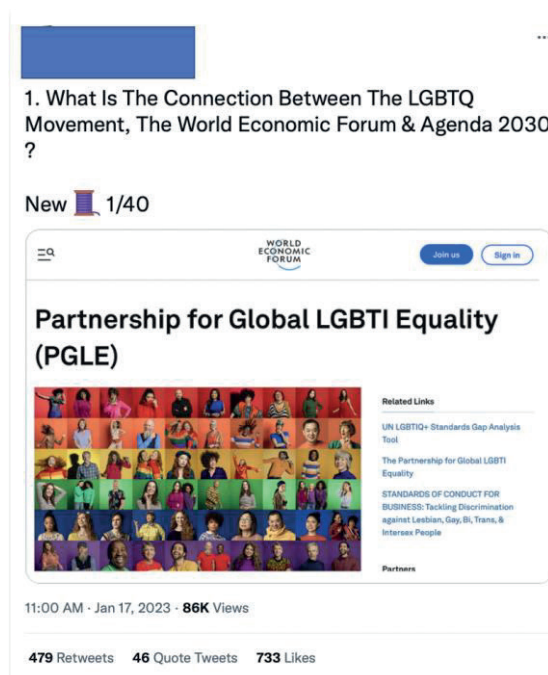


Figure 8: Gender and sexual identity narratives: A Twitter thread claims Ireland’s National LGBTI+ Inclusion Strategy is part of the Great Reset conspiracy theory. The thread garnered more than 86,000 views and was retweeted by 479 accounts.

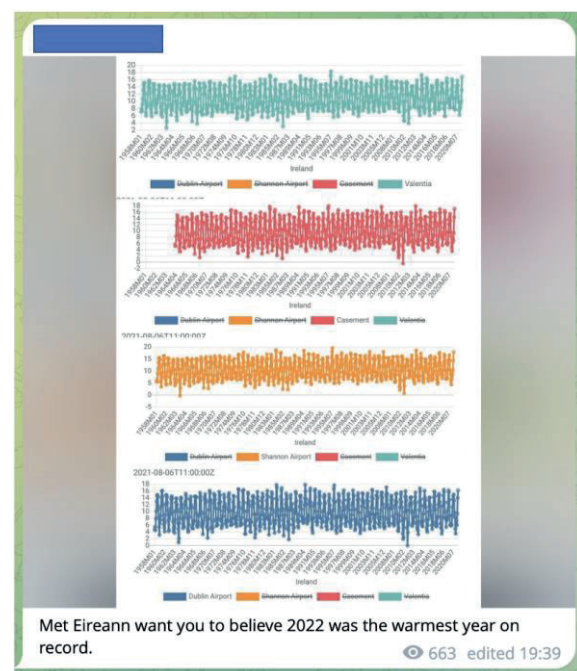


Figure 9: Science and environment narratives: A post in a Telegram channel rejects Met Eireann’s assertion that 2022 was the warmest year on record in Ireland

Tactics

In Ireland, as elsewhere, disinformation tactics have evolved considerably since concerns about “fake news” in 2016. Since the start of Covid-19 pandemic, there has been a clear effort to galvanise online support into offline organisation and action. Most recently, this is evident in the protests outside accommodation for refugees and asylum seekers. The online dimensions of these protests rely on the circulation of disinformation and misinformation, filmed footage of protest speeches and actions, a network of “alternative” investigations and commentary, and the production of memes and slogans that can be easily shared and disseminated by supporters. For example, video footage of men arriving on a bus in the middle of the night became widely circulated on social media. Such videos are often unrelated to the discussion or topic, but they act as antagonising talking points to drive online discussions.

Some notable characteristics of disinformation campaigns in Ireland are summarised below:

- **Cross-platform mobilisation** facilitates coordination and radicalisation as mainstream audiences are targeted on popular platforms such as Instagram and Twitter while more extreme audiences are cultivated on less moderated platforms such as Telegram. On these platforms, extreme content appears to bolster community cohesion as participants discuss plans to disseminate their ideas and make them appealing for wider audiences.
- **Participatory memes and content templates** make it easy to roll out content and messages at speed. As noted above, slogans such as “Ireland First” or “East Wall says no” can be turned into posters and hashtags (see Figure 10). These can be replicated by anyone, which encourages participation in the campaign.
- **Niche influencers** play an important role in amplifying disinformation to their followers. Niche influencers often operate under the radar, which means they are often not moderated as strictly as high-profile influencers or as other non-newsworthy accounts. Given the small size of Ireland, figures can gain significant influence over a specific demographic with a relatively small following.
- **Online groups**, such as Facebook Groups and private WhatsApp groups, support the authentic and bottom-up spread of disinformation within communities. These produce highly salient disinformation narratives because they emanate from authentic users, rather than fake, inauthentic accounts. Extremists and conspiracy theorists frequently highlight the goal of spreading ideas to community groups.
- **Difficult-to-verify content** is content that evades content moderation policies and fact-checking. This may include broad claims that are not of a factual

nature or it may include long-form videos (e.g., online talk shows, game reviews) in which false claims are embedded in other content.

- **Online abuse and incivility** are a persistent backdrop to disinformation. Certain figures, such as the former Lord Mayor of Dublin Hazel Chu, are consistently targeted for abuse (see Figure 11). The most extreme kinds of abuse are found in alternative platforms while the same sentiments are expressed in milder forms on mainstream platforms.
- **Threats of physical violence and doxing** are evident on some small-scale Telegram groups (i.e., less than 500 subscribers) and in the comments on Bitchute and Odysee. Doxing is the act of publishing private or identifying information about an individual without the permission of that individual and often with malicious intent. Some Telegram channels focus on identifying those who are associated with anti-fascist activities (see Figure 12). Details, such as their place of work, are shared alongside threats and, in some instances, video footage is shared of individuals being confronted by channel members (see Figure 13).
- **Attacks on the credibility of mainstream media and journalism** are a consistent feature of disinformation campaigns by all actors. News outlets are accused of being biased, incompetent, and in the service of elite agendas. RTÉ as the public broadcaster and, to a lesser extent, TheJournal, as the only registered fact-checking entity in the Republic of Ireland, are frequent targets of abuse.



Figure 10: Memes and content templates: Twitter post which highlights the incorporation of Irish identity into anti immigrant protests, with strong use of the tricolour.



Figure 11: Example of nativism in a Telegram post shared by a popular figure at the protests. It implies that the Irish politician, Hazel Chu, is not really Irish due to her Chinese ethnicity. The post is also a

word play on the slogan “Ireland belongs to the Irish”.

She really shouldn't be giving it large on Twitter being from [redacted] as there's people there that'll “deal with her” appropriately if she's not careful.



Anyone in [redacted] should give her a shout in the [redacted] and if she doesn't behave herself we'll have to pay her a visit.  355 10:47 

Figure 12: Excerpt from a Telegram post which suggests intimidating an anonymous Twitter user who's photograph, name and place of work was shared within the channel.

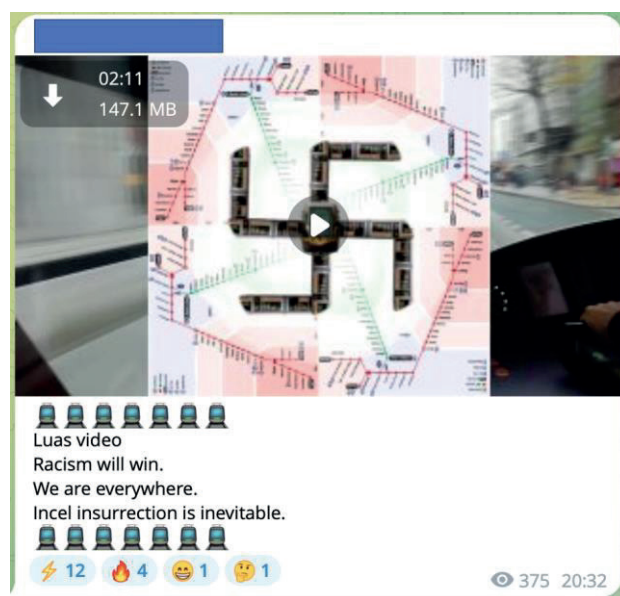


Figure 13: Video posted in a Telegram channel which celebrates news coverage of racist violence and slurs on the Luas.

Pre-bunking as a preemptive strategy

Preemptive approaches to tackling misinformation are commonly referred to as preemptive debunking or pre-bunking. Rather than attempt to counter disinformation after it has spread, pre-bunking aims to make audiences resilient to disinformation before they encounter it. Research indicates that pre-bunking is generally effective at increasing attitudinal resistance against manipulation; although more research is needed to better understand the dynamics.

Writing in the *European Psychologist* journal, Jon Roozenbeek (University of Cambridge), Eileen Culloty and Jane Suiter (DCU FuJo) explained pre-bunking as follows: Although several approaches to pre-bunking exist, the most common framework is inoculation theory. Medical inoculations are (usually) weakened and harmless pathogens that, upon introduction, prompt the body to create antibodies. Inoculation theory posits that the same can be achieved with unwanted attempts at persuasion. People can build attitudinal resistance against future persuasion attempts by pre-emptively exposing them to a “weakened” dose of the unwanted persuasive argument.

The major components of pre-bunking, as currently practised, are outlined below:

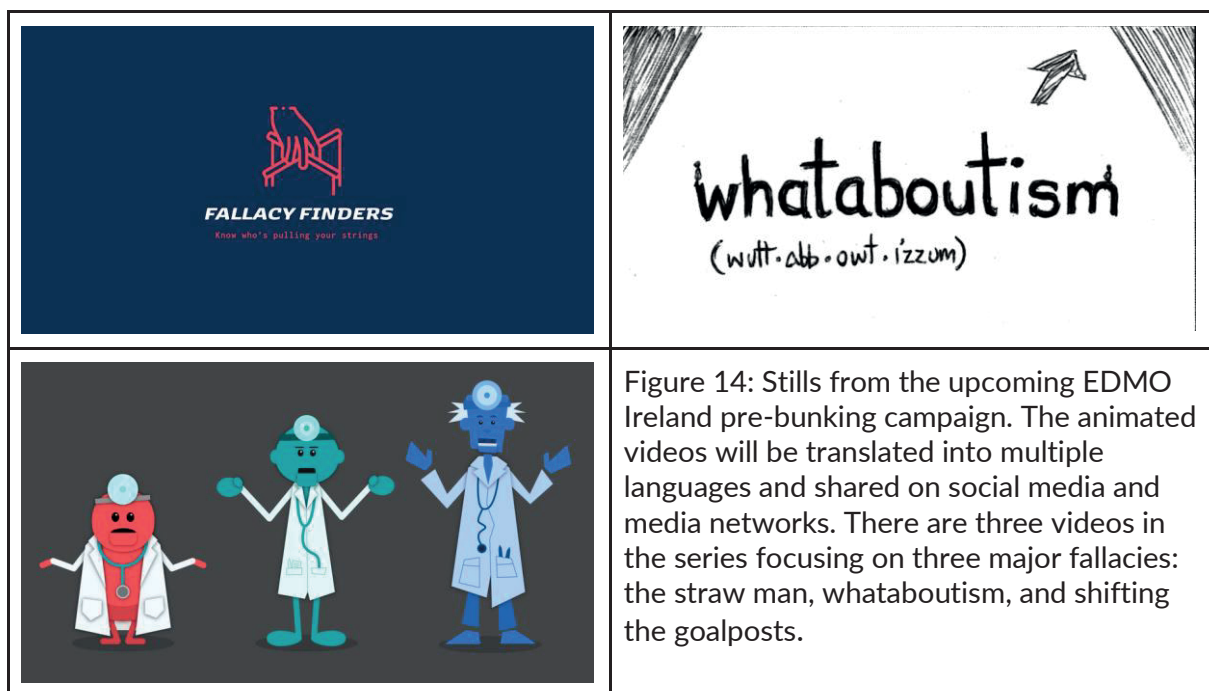
- **Inoculation treatments:** Inoculation treatments consist of two core components: 1) a warning of an impending attack on one’s beliefs or attitudes (i.e., a forewarning of impending manipulation), and 2) a pre-emptive refutation of this upcoming manipulation attempt.
- **Active and passive inoculations:** Inoculations may be active or passive. With passive inoculation, people are provided with counter-arguments against the unwanted persuasion attempt, usually in the form of a short piece of text or a video. Conversely, with active inoculation, people generate their own counterarguments through internal rehearsal (rather than simply being provided with them), for example by playing a game. There is currently not enough evidence to assess whether active or passive inoculations are more effective overall.
- **Issue-based and technique-based inoculations:** Inoculations may be issue-based or technique-based. Issue-based inoculations seek to inoculate people against individual persuasive attacks or specific examples of disinformation such as disinformation about climate change. In contrast, technique-based inoculations confer resistance against manipulation strategies or tactics such as logical fallacies, emotional manipulation, or conspiracy theories.

Both approaches have their advantages: issue-based inoculations may be more effective than technique-based ones when it is known what disinformation people are likely to be exposed to in the near future. Technique-based inoculations, on the other hand, have the benefit of applying to a wider range of misinformation, at the expense of specificity. As prebunking is preventative in nature (i.e., it occurs prior to exposure to misinformation), it is not usually possible to know exactly what disinformation to prebunk (as you cannot know what disinformation people will be exposed to in the future, or in what form).

EDMO Ireland's pre-bunking campaign

In 2022, EDMO Ireland partnered with the Department of Foreign Affairs and the University of Cambridge to produce a series of pre-bunking videos to counter common disinformation claims about the Russian war on Ukraine.

These animated videos are examples of passive, technique-based inoculation as they provide viewers with a counter-argument to resist common manipulation strategies: specifically, the straw man fallacy, whataboutism, and the shifting the goalposts strategy. Once produced in English, the animated videos will be presented to the network of EDMO hubs and translated into multiple languages for international distribution (see Figure 14).



Fostering a healthy information environment

TheJournal is the only national fact-checking entity in the Republic of Ireland. As the fact-checking partner in EDMO Ireland, TheJournal conducts investigations, publishes fact-checks, and collaborates with European fact-checkers.

The position of news outlets as trusted sources of information has been under siege for some time. The Digital News Report Ireland 2022 noted that Irish audiences registered a higher average level of trust than their European counterparts in mainstream news sources (52% vs 42%) but that level has fluctuated since tracking of this sentiment by the Reuters report in Ireland began in 2015. It took noticeable hits in periods where talk of “fake news” became amplified by specific events, eg, in Donald Trump’s first full year as US president in 2017, or in the pandemic-hit year 2020 when conspiracy and disinformation around Covid19 and vaccines were rife.

Combine that with a declining interest in “the news” itself, especially among younger cohorts, and journalists are left asking: If not from us, where are people getting their information? Online users have endless opportunities to find information from a myriad of sources, to build up preferred networks, and find ways to speak to each other that does not require a mainstream middle (news) person. The digital sphere offers the freedom to “do your own research” - a phrase celebrated and maligned in equal measure, depending on who is employing it and to what end.

This is where journalism still has a vital role to play, once media outlets enter into a contract of trust with their audience to be transparent about how they are getting their information and how they decide what to publish, or not publish. It is the process of combining traditional skills of fact-checking, data research and objective interviewing with the maxim of “show, don’t tell” to the foreground. Audiences want more than ever to know how newsrooms “do their research”.

This idea of bringing audiences deeper into newsroom processes, so they feel more ownership over and trust in the information being presented was the basis of an 18-month project The Journal concluded in June 2022 called The Good Information Project. Its chief objective was to equip news-interested citizens with solutions and the skills to find information on big issues and combat misinformation in digital spaces that lack media literacy.

The Journal newsroom and FactCheck teams did this through three stages:

- **organised listening sessions** - on site, in webinars, in surveys - to surface the key questions and knowledge gaps that need to be answered around a particular topic that was impacting their lives;
- **empowering/informing the audience** - using journalism to create content that catered to those gaps, answering their questions and creating chances for the audience to speak to experts directly, eg, live panels held in cities around Ireland;
- **creating information ambassadors** - equipping the audience with ways to return what they had learned to their own networks, eg, multimedia content to be distributed for specific social media platforms, quizzes, information packs for schools.

Journalists can be powerful allies in the media literacy effort. One clear role for journalism in the face of misinformation and disinformation is not to tell audiences what is false, but also to create and share the tools that lead us to good information, the type that ideally should inform good quality journalism.

Our interaction with users surfaced more clearly what their specific perceptions - or misperceptions were on important issues, and how that impacts public decision-making. This conversation is also being held among fact-checkers themselves. The Journal contributes to a monthly report by EDMO fact-checkers that tracks disinformation trends across the continent. This sharing of information can also serve as an early warning, as in the case of a conspiracy theory video that revived long-disproven disinformation about Covid-19 vaccines, leading to a concerted effort by fact-checking units to locate and debunk this particular push in their own territory, Ireland included.

Within the Ireland EDMO hub more specifically, our day-to-day fact-checking experience is informing the shape of new tools being researched and developed by our academic and tech partners in the hope of finding ever more efficient ways to surface misinformation that needs to be tackled because of the severity of the disinformation or a clear and immediate impact on public opinion.

Understanding what is swaying public opinion is also where journalism can greatly increase the power of specific fact-checks. Sitting within a daily newsroom, The Journal FactCheck has the advantage of accessing editorial meetings, noting the stories and topics beginning to dominate the national conversation. Journalists practice verification every day in their work; specific fact-checking efforts around an ongoing story can add context and a reference point for readers.

TheJournal's Fact-Checks in 2022

In 2022, The Journal FactCheck produced **67 fact-checks** and **12 analysis articles** which explored disinformation-related issues. It created a monthly newsletter and contributed to international fact-checking work on initiatives such as **EDMO Ukraine Taskforce**. The topics covered by the fact-checks were varied and included Covid-19, vaccines, climate change, the war in Ukraine, migration, monkeypox, conspiracy theories and transgender issues as well as political claims relating to housing, government policy, and the Northern Ireland protocol.

A news story on repeat in 2022, in particular the second half of the year, featured protests outside refugee centres around various parts of the country. One such protest caught public attention when videos emerged of loudhailers aimed at a convent in Fermoy, Co Cork where women and children were among the asylum-seekers and refugees being housed. Analysis from The Journal, assisted by the FactCheck team, made the connection between this and other protests, how they were organised and by whom, explaining to the audience how this is not an isolated incident, and why such protests have common motivations, motifs and impacts.

This is a form of living media literacy; a reader may now recognise these elements if and when a similar incident pops up on their doorstep, when it might be difficult to resist attempts to influence coming through local social networks. The trust the media wishes online audiences would have in them has not disappeared; it has just gone elsewhere, and this frontline is where journalism can reinsert itself.

To do that though, journalism is having to increasingly come to terms with getting on board with many objective measures of trust - transparency about funding, processes, and being open to independent scrutiny of their work. The Journal FactCheck undergoes an annual verification audit with the International Fact-Checking Network, will undertake a similar process with the newly-established European Fact-Checking Standards Network in 2023 and is subject to complaints submitted to the Press Council of Ireland.

But most important is that The Journal take as many opportunities as possible to explain how they work, especially when it comes to making choices about what misinformation to tackle when there is sadly so much to choose from.

A stark example of decoding such a decision came in the depths of the cold snap that descended on the country in early December, and a seemingly well-meaning

Facebook post about the alleged death of a young homeless woman in a tent went viral and fed into a much larger and more harmful narrative about refugees being prioritised over Irish people in need.

The Journal Editor Sineád O'Carroll took to Twitter to unspool a thread on the editorial decisions that led to the publication of a piece from our FactCheck editorial lead Stephen McDermott tracking the trajectory of a viral post which claimed a young girl had been found dead in a tent during the coldest spell of this winter.

Making the decision to “fact-check” such a post was difficult because, in the midst of a housing crisis and cost-of-living crisis, it is conceivable that such a tragedy could happen. “And by calling it out as fake, it might seem we are saying there ‘nothing to see here’ about housing and homelessness,” wrote O'Carroll.

“But it’s important to know how these ‘reports’ are jumped on and used to further completely separate narratives and agendas. Which is exactly what happened here regarding asylum seekers and refugees.”

McDermott, in the course of his article, noted that “this situation is not unique to Ireland: the European Digital Media Observatory (EDMO) reported last month that false narratives and disinformation targeting refugees has been on the rise across the continent recently”.

Many of these claims, he wrote, suggest that refugees seeking asylum in European countries receive favourable treatment from authorities in the countries they enter.

“And although a claim about the homeless girl is not directly about refugees, it has been used to further an agenda whose narrative falls into that category.”

The original post on a charity’s Facebook page was deleted, as were stories carried on three news media websites, and a TD had to clarify comments he made about the “death” in Dáil Éireann, updating the parliament record to reflect that the story was unconfirmed.

Nonetheless, as McDermot notes, while the original post has been deleted, “its traces may linger for some time”.

This is where journalism can find its place in the fight against misinformation, by highlighting the relationship between false claims and broader false narratives as well as the gap that can be exploited when there are unaddressed societal issues which people are frustrated and concerned about.

Supporting fact-checking with AI-driven tools

The University of Sheffield is developing a suite of tools to support fact-checkers. These tools include automatic verification, stance classification, MEME OCR analysis, disinformation narratives classification, hyperpartisan detection, among others. The Sheffield team focus on textual, multimodal (e.g. memes) and network analysis approaches to develop cutting-edge tools that can provide further information to journalists, supporting their round the clock work on debunking false narratives. The team also supports the development of the InVID-WeVerify plugin (Fake News Debunker toolkit), a "swiss army knife" for journalists and fact-checkers with multiple tools, including image and video analysis, search on the new Database of Known Fakes and a network analysis tool for visualising the network of disinformation narratives on Twitter.

Stance classification tool: In particular, the team is working with TheJournal to improve its stance classification tool. This work is motivated by cases where nothing or very little information is available about a rumour on social media, and it is then worth monitoring what others are talking about it. For example, as shown in Figure 15, a reply to a factual tweet can be supporting the statement, denying the statement, asking for more information about the statement (querying), or just adding information that is not relevant to debunking (commenting). In practice, this type of classification may help fact-checkers decide where to focus their attention by identifying issues that are causing disagreement and controversy.

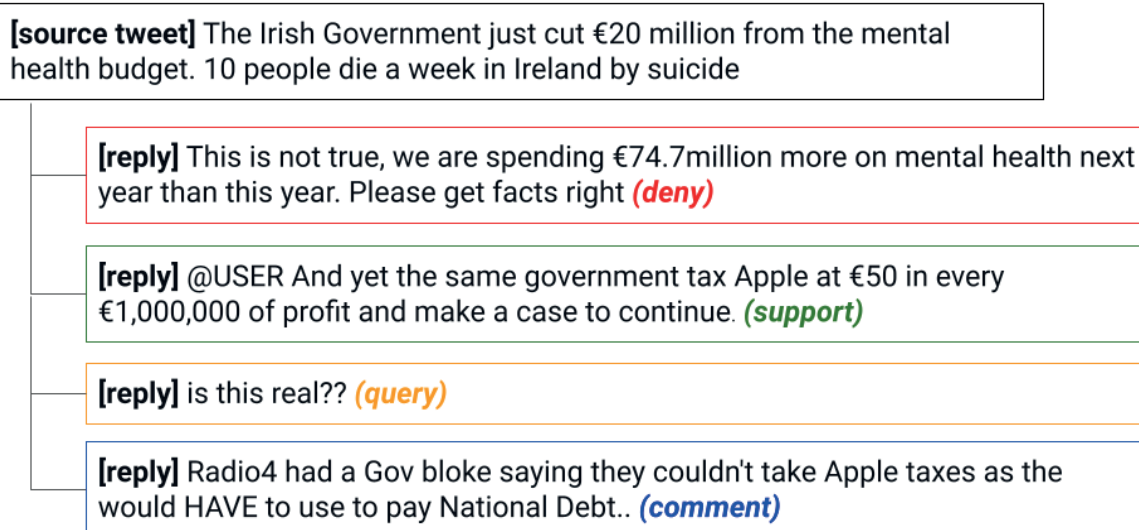
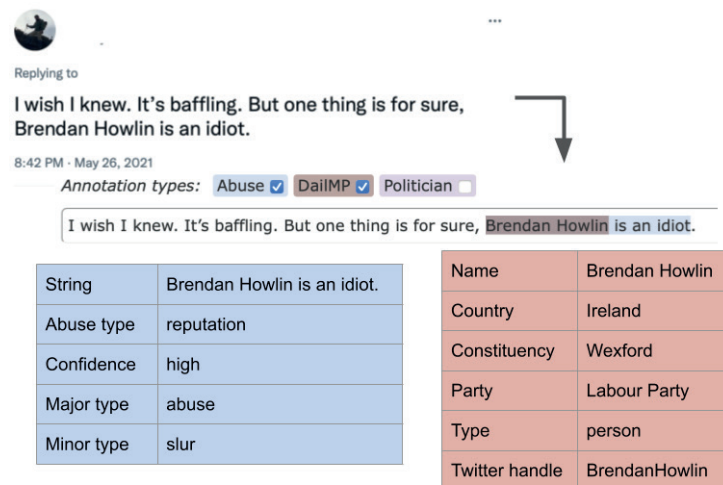


Figure 15: Hypothetical example of stance classification in a Twitter thread about a rumour.

Political Abuse Monitor: Disinformation often appears in tandem with abuse and harassment. Another tool being developed within EDMO Ireland is the Political Abuse Monitor. This tool identifies abusive comments in textual form and is tailored also to

tag politicians' names. The first version of this tool included UK members of parliament for the 2015, 2017 and 2019 General Elections and candidates for the 2017 and 2019 elections. This was then extended to include members of the Irish Dáil Éireann. Figure 16 shows an example of an abusive tweet directed to Irish MP Brendan Howlin that was processed by our tool. The tool can identify the abusive text and classifies the type of abuse as a "slur" and directed to the "reputation" of a person. It also retrieves information about the Irish TD, including their constituency, party and Twitter handle.



The screenshot shows a tweet from a user replying to another. The tweet text is: "I wish I knew. It's baffling. But one thing is for sure, Brendan Howlin is an idiot." The tweet is dated 8:42 PM · May 26, 2021. Below the tweet, the tool displays "Annotation types" with three checkboxes: "Abuse" (checked), "DailMP" (checked), and "Politician" (unchecked). A search bar below this shows the text "I wish I knew. It's baffling. But one thing is for sure, Brendan Howlin is an idiot." with "Brendan Howlin is an idiot." highlighted. Below the search bar, there are two tables. The first table, with a blue header, shows the analysis results for the highlighted text. The second table, with a red header, shows the retrieved information about the politician mentioned in the text.

| | |
|------------|-----------------------------|
| String | Brendan Howlin is an idiot. |
| Abuse type | reputation |
| Confidence | high |
| Major type | abuse |
| Minor type | slur |

| | |
|----------------|----------------|
| Name | Brendan Howlin |
| Country | Ireland |
| Constituency | Wexford |
| Party | Labour Party |
| Type | person |
| Twitter handle | BrendanHowlin |

Figure 16: Example of use of the Political Abuse Monitor tool using a real abusive tweet targeting an Irish TD.

Media literacy as resilience building

Media literacy is increasingly recognised for its important role in building resilience to disinformation and related threats to democratic societies. Media literacy refers to the lifelong process of acquiring the knowledge, skills and practices that are necessary to be a consumer and producer of media content in a critical, creative and responsible manner. Fundamentally, media literacy is about empowering citizens to make well-informed decisions about the content and information they consume. In this sense, media literacy underpins fundamental concepts and values such as democracy, freedom, equity, justice, and tolerance.

In relation to disinformation specifically, media literacy can equip people with the knowledge and skills they need to critically evaluate the information they encounter, learn how to identify trustworthy sources, analyse multimodal messages and use information in a responsible way.

Media Literacy in Ireland: In recent years, Ireland has developed a significant infrastructure to support media literacy. The country is well-placed to play an influential role in the upscaling of media literacy activities across Europe. Much of this is due to the consolidation of media literacy expertise and the development of Media Literacy Ireland (MLI) in 2017.

MLI is an independent association facilitated by the Broadcasting Authority of Ireland/Media Commission. It is chaired by Prof. Brian O'Neill, a member of Council of Europe's Expert Group on Digital Citizenship Education. The Vice-Chair, Dr Eileen Culloty, is the coordinator of EDMO Ireland and the coordinator, Martina Chapman, is author of a 2016 Council of Europe report on media literacy in the EU. MLI's 250+ members represent a broad range of sectors - education, libraries, civil society, media, technology platforms - and collaborate to develop new partnerships and sustainable media literacy projects. Launched in 2019, the national Be Media Smart campaign established a campaign infrastructure drawing on the expertise and resources of members. The concept and resources were freely shared with European partners resulting in the replication of the campaign in four regions (North Macedonia, the Czech Republic, Bosnia and Herzegovina, and Norway).

In relation to education, some elements of media literacy are spread across the national curriculum. There are also opportunities to explore media literacy in Transition Year programmes and a short course on Digital Media Literacy was developed and is available to Junior Cycle teachers across the country. Unfortunately, there is no data available on how many schools have implemented it. However, Ireland follows most European countries in the sense that it lacks a national strategy to implement media literacy in schools and media literacy is often conceived as an

ICT-related subject. Beyond formal education, a wide range of actors have a media literacy remit. For example, charities and NGOs working towards digital and social inclusion recognise that media literacy is an essential skill for the 21st Century. However, these organisations may lack the background knowledge to provide media literacy guidance to their members and target audiences.

EDMO Ireland

Led by DCU FuJo, EDMO Ireland has three overarching goals to build capacity in media literacy at national and European levels.

Assessing needs and opportunities for media literacy initiatives: This work involves mapping out the opportunities for media literacy initiatives in different sectors of society, such as schools, universities, and libraries. For instance, EDMO Ireland has delivered workshops on various topics around media literacy for both children and adults in public libraries. We have also designed and delivered webinars on disinformation to secondary teachers in partnership with Arts in Junior Cycle, an educational programme that offers training and support for secondary teachers, including the ones teaching the Digital Media Literacy short course. The hub was also involved in the development of classes on topics related to media literacy for students of the Professional Masters of Education (PME) in Trinity College Dublin, a compulsory pre-service education programme for secondary teachers in the country. In 2022, we delivered the first class on critical media literacy for PME students in Ireland.

In 2023, DCU FuJo will be part of an international research project involving five countries that will complement the work carried out by EDMO Ireland. The project will investigate current teacher training initiatives across the continent, assess the opportunities available, and provide solutions in the form of digital resources and professional networks.

Implementing media literacy campaigns: In partnership with Media Literacy Ireland (MLI) and the Local Government Management Agency (LGMA), EDMO Ireland is developing a media literacy campaign for public libraries across the country. For the pilot study in Spring-Summer 2023, DCU FuJo will design and deliver in-person training programmes and online webinars. It will also create an online course that will be available for both librarians and members of the public who want to take a self-paced online training session on media literacy.

Ultimately, if sufficient funding is identified, the aim is to train around 800 librarians covering all public library jurisdictions. Public librarians across the country will then be equipped with the necessary knowledge and skills to provide assistance and education to the public on an ongoing basis. Media Literacy Ireland may then run a

national Be Media Smart campaign to highlight the availability of media literacy support at local libraries. EDMO Ireland aims to use this project as a case study that can be replicated in different countries and contexts.

Depository of training and resources: Media Literacy Ireland aims to become the first port of call for anyone interested in learning more about media literacy. EDMO Ireland was responsible for organising a new section on the new MLI website launched in December 2022 called “Training and Development”, which provides the public with resources and training courses on different topics related to media literacy, such as disinformation, online safety, news media and data privacy. The resources include websites, videos, reports, lesson plans, academic papers and others, and they are segmented by topic, age, and formats (see Figure 17).

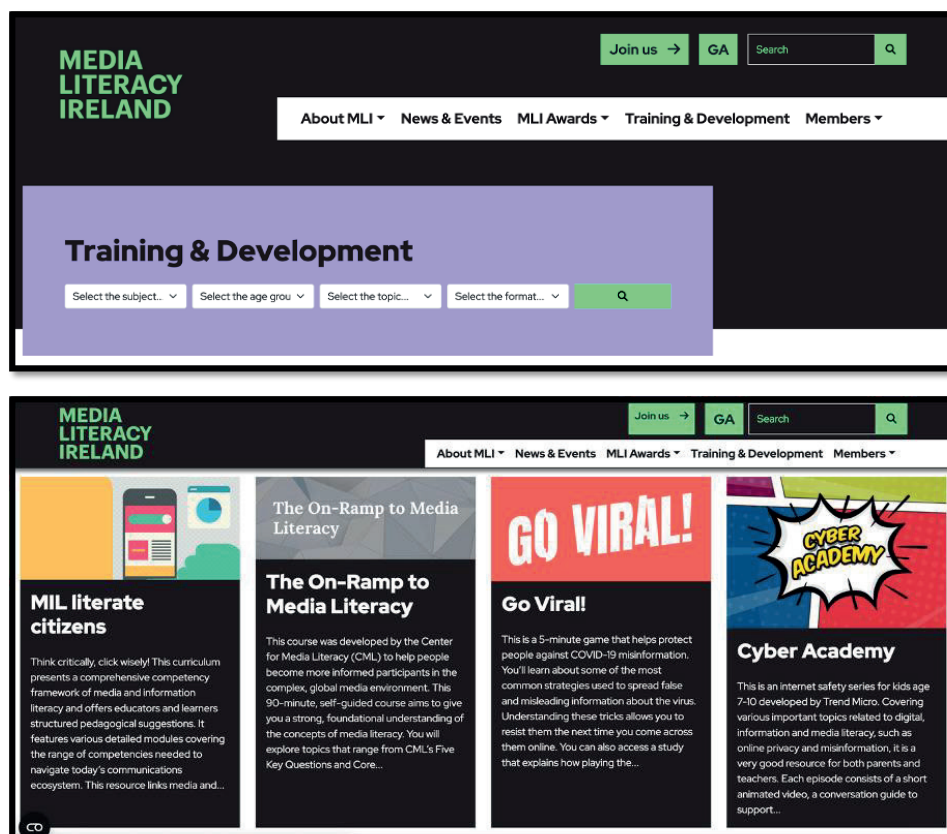


Figure 17: The Training and Development page on Media Literacy Ireland website.

European Policy developments

From a legal and regulatory perspective, online disinformation operates in a grey area. Although it may cause significant public harm, it is not illegal in most democratic countries. This is further complicated by the fact that the regulatory structures that govern the media-information environment need to be updated to address the new dynamics of digital media. In Ireland, these issues are being addressed through a suite of new policy instruments including the Online Safety and Media Regulation Act; the Electoral Reform Act; and the National Counter Disinformation Strategy. Here we provide a brief overview of the major EU policies that establish an essential context for Irish developments and the potential role of EDMO and EDMO Ireland.

The Digital Services Act (DSA) is the nucleus of EU actions to address digital media and digital markets as it establishes an EU-wide regulatory framework for digital platforms. Wide-ranging in scope, it specifically addresses a number of issues associated with disinformation including algorithmic accountability and online political advertising. For example, it introduces measures to increase algorithmic accountability regarding how information is prioritised and targeted and it aims to introduce a common framework for political advertising.

The DSA was formally agreed in October 2022, although its implementation process will not be complete until early 2024. The DSA applies to Online Intermediary Service Providers (OIPs), who offer services within the EU. This includes a wide range of services from internet service providers to social media platforms. However, the level of obligation that OIPs have under the DSA is determined by their role, size and impact. Very Large Online Platforms (VLOPs) or Search Engines (VLOSEs) with over 45 million users, which is 10% of the EU population, are subject to the highest level of obligation.

A broad principle underpinning the DSA is that what is illegal offline should also be illegal online, but it also includes requirements for transparency in areas like algorithmic recommendation systems and targeted advertising. Platforms will be responsible for assessing and mitigating systemic risks associated with their usage and VLOPs will have to conduct a systemic risk assessment once a year. The categories of systemic risk outlined within the DSA include a provision which covers disinformation or the “intentional manipulation” of digital services in ways that may have “an actual or foreseeable negative effect on the protection of public health, minors, civic discourse, or actual or foreseeable effects related to electoral processes and public security.”

The DSA also requires that vetted researchers should have access to data in order to investigate systemic risks. Access to data held by online platforms has proven a

contentious issue. In 2022, an EDMO working group on Access to Data Held by Digital Platforms for the Purposes of Social Scientific Research produced a framework in conjunction with platforms that outlines a process by which researchers can access data in a GDPR-compliant way. It includes a draft code of conduct on how platforms can share data with independent researchers while protecting users' rights.

Article 45 of the DSA makes provision for codes of conduct, which will be used to help enforce certain aspects of the regulation relating to systemic risks. It is expected that the EU Code of Practice on Disinformation will essentially act as a code of conduct under the DSA and pave the way for penalties if a Code signatory and VLOP fails to keep their commitments with the Code. Penalties under the DSA can include a fine of up to 6% of a company's global turnover.

The EU Code of Practice on Disinformation is a self-regulatory mechanism for platforms, and other stakeholders, to report their actions to counteract disinformation. It has been in place since 2018. As a self-regulatory mechanism, platforms voluntarily agreed to participate in the Code and reported on their compliance. The European Commission assigned the European Regulators Group for Audiovisual Media Services (ERGA) with the task of monitoring the effectiveness of the Code and its implementation.

To date, the Broadcasting Authority of Ireland has commissioned DCU FuJo to produce three reports assessing the implementation of the Code; ElectCheck (2019), CodeCheck (2020) and CovidCheck (2021). In each report, FuJo identified significant inconsistencies in the reporting standards and the quality of efforts made by the platforms. The most recent report, CovidCheck, found that the reports submitted by the platforms were highly repetitive, often irrelevant and generally failed to provide the data that was requested. The report recommended the establishment of more robust monitoring and reporting procedures. After presenting the key findings to the ERGA monitoring group, the FuJo recommendations were endorsed and incorporated into ERGA's own recommendations for the revision of the Code.

A 'strengthened' Code was released in 2022 with 34 signatories agreeing to 44 commitments and 128 measures. The Code addresses the following areas:

- **Demonetisation: cutting financial incentives for purveyors of disinformation:** The strengthened Code aims to ensure that purveyors of disinformation do not benefit from advertising revenues. Signatories commit to stronger measures avoiding the placement of advertising next to disinformation, as well as the dissemination of advertising containing disinformation. The Code also sets up a more effective cooperation among the players of the advertising sector, allowing stronger joint action.

- **Transparency of political advertising:** Recognising the importance of political advertising in shaping public life, the strengthened Code commits signatories to put in place stronger transparency measures, allowing users to easily recognise political ads by providing more efficient labelling, committing to reveal the sponsor, ad spend and display period. Moreover, signatories commit to putting in place efficient and searchable ad libraries for political advertising.
- **Ensuring the integrity of services:** The Code will strengthen the measures to reduce manipulative behaviour used to spread disinformation (e.g. fake accounts, bot-driven amplification, impersonation, malicious deep fakes), and establishes a stronger cooperation among signatories to fight the challenges related to such techniques. A cross-service understanding of unpermitted manipulative behaviours and practices to spread disinformation will be agreed among signatories. They will be also required to periodically review the list of tactics, techniques and procedures (TTPs) employed by malicious actors, and will implement clear policies, covering the range of behaviours and practices identified.
- **Empowering users:** Users will be better protected from disinformation through enhanced tools to recognise, understand and flag disinformation, to access authoritative sources, and through media literacy initiatives. In particular, the Code will ensure that safe design practices are put in place to limit the spread of disinformation and ensure more transparency of their recommender systems, adapting them to limit the propagation of disinformation.
- **Empowering researchers:** The Code foresees that online platforms provide better support to research on disinformation. Researchers will have a better and wider access to platforms' data. This means ensuring automated access to non-personal, anonymised, aggregated or manifestly made public data, and working towards putting in place a governance structure to simplify access to data requiring additional scrutiny.
- **Empowering the fact-checking community:** The new Code will extend fact-checking coverage across all EU Member States and languages and ensure that platforms will make a more consistent use of fact-checking on their services. Moreover, the Code works towards ensuring fair financial contributions for fact-checkers' work and better access to fact-checkers to information facilitating their daily work.
- **Transparency centre and Task-force:** The Transparency Centre, accessible to all citizens, will allow for an easy overview of the implementation of the Code's

measures, providing transparency and regular updates of relevant data. The permanent Task-force will keep the Code future-proof and fit-for-purpose, by establishing a forum – inter alia - to review and adapt the commitments in view of technological, societal, market and legislative developments. The Task-force is composed of representatives of signatories, the European Regulators' Group for Audiovisual Media Services, the European Digital Media Observatory and the European External Action Service, and is chaired by the Commission.

- **Strengthened Monitoring framework:** The Code comes with a strong monitoring framework, including Service Level Indicators to measure the Code's implementation throughout the EU and at the Member State level. By the beginning of 2023, signatories will provide to the Commission the first baseline reports on their implementation of the Code. Subsequently, Very Large Online Platforms, as defined in the Digital Services Act (DSA), will report every six-months while other Signatories will report on a yearly basis. The strengthened Code also contains a clear commitment to work towards establishing structural indicators, allowing to measure the overall impact of the Code on Disinformation.

EDMO hubs are expected to contribute expertise to the monitoring framework. As noted, it is expected that the EU Code of Practice on Disinformation will essentially act as a code of conduct under the DSA and pave the way for penalties if a Code signatory and VLOP fails to keep their commitments with the Code. Building on previous work assessing the implementation of the Code in Ireland and across the EU, DCU FuJo has applied to join the EDMO taskforce that will guide how EDMO hubs contribute to the monitoring framework.

Glossary of relevant terms

Algorithms: An algorithm is a set of instructions that a computer follows to complete a task. Social media platforms use different algorithms to make automated decisions about the content shown or recommended to an end-user.

Algorithmic transparency: Algorithmic transparency refers to the extent to which the factors influencing automated decisions are made visible to those who use and regulate systems that employ algorithmic decision-making.

Algorithmic accountability: Algorithmic accountability concerns the capacity to assess the fairness of algorithms and to hold entities responsible for the algorithms they employ.

Amplification: In the context of disinformation, amplification refers to the process through which disinformation spreads to a wide audience. *Algorithmic amplification* refers to the bias of social media algorithms towards certain kinds of content and the capacity of such algorithms to be gamed or manipulated by bad actors. Amplification can also occur when disinformation is reported by news media or influential figures.

API: An API, or application programming interface, is a means by which data from one web tool or application can be exchanged with, or received by another.

Artificial intelligence (AI) and machine learning: Artificial intelligence (AI) is the application of computer systems to perform tasks that normally require human intelligence. Advances in AI are driven by machine learning whereby computers develop the ability to complete or learn a task without being explicitly programmed to do that task. In the context of disinformation, AI is employed as an effective way to detect and moderate content at speed and scale. However, AI advances also enable disinformation through more efficient targeting, personalization, and the creation of synthetic content (see deep fakes).

Automation: Automation is the process of designing a machine or software application to complete a task with little or no human direction (see algorithms and artificial intelligence).

Bad actors: Bad actors is a catchall term for those who intentionally create and propagate disinformation. Bad actors may be states, corporations, social movements, or individuals and their motivations span a spectrum of political, ideological, and financial interests. Bad actors vary considerably in terms of the groups they target and the levels of coordination involved in their activities.

Bot: A bot is a software application that is programmed to automate certain tasks. Social media bots are programmed to generate and/or engage with content on a particular platform. In the context of disinformation, bots are often used to amplify false information to create the illusion of public interest.

Clickbait: Clickbait is sensationalised online content, often in the form of a headline, that is designed to attract attention and entice users to click the content link. Clickbait often aims to generate revenue from advertising based on the clicks the content receives (see content farm).

Code of Practice on Disinformation: The Code of Practice on Disinformation was developed by the European Commission in 2018 as a self-regulatory framework that encouraged online platforms to commit to voluntary measures countering disinformation. Following criticisms and review, a strengthened Code was agreed in 2022. It details 44 commitments and 128 specific measures and includes an expanded list of (voluntary) signatories including representatives of relevant trade associations, fact-checkers, and civil society groups. Signatories are asked to report on their efforts in four key areas: demonetising disinformation; increasing the transparency of political advertising; reducing manipulative behaviour; and safeguarding users.

Content farm: A content farm is a website that creates a large volume of low-quality content to generate money from advertising based on the clicks the content receives (see clickbait).

Dark ads: Dark ads are advertisements that are only visible to the publisher and their target audience.

Deep fakes: Deep fakes are media (images, videos, sound recordings) in which the words or actions of an individual have been fabricated in a highly convincing way. Deep fakes are produced using artificial intelligence (see artificial intelligence).

Digital literacy: Digital literacy is the ability to find, evaluate, organise, and use information in all its various formats. In the context of disinformation, it is closely related to media literacy and digital literacy (see media literacy and digital literacy).

Disinformation: Disinformation is false or misleading information that is created or spread with the intention to deceive or cause harm. Disinformation can take many forms including, for example, conspiracy theories, coordinated campaigns, rumours, and propaganda. Those who create disinformation typically have political, financial, psychological, or social motivations. Disinformation may be assessed for its degree of

truthfulness or accuracy, the nature of the malicious intent behind its creation or distribution, and the harms it poses. Disinformation is sometimes distinguished from misinformation based on the intention to deceive (see misinformation).

Doxing: Doxing is the act of publishing private or identifying information about an individual without the permission of that individual and often with malicious intent.

Fact-checking: Fact-checking is the process of determining the truthfulness and accuracy of published information or claims.

Influence operation: An influence operation or coordinated influence operation is a coordinated effort by domestic or foreign actors to influence a target state or audience. using a range of deceptive means, including suppressing independent information sources in combination with disinformation

Media literacy: Media Literacy is the ability to access, analyse, evaluate and create media in a variety of forms. In relation to disinformation, media literacy can equip people with the knowledge and skills they need to critically evaluate messages, and use information in a responsible way. When applied to disinformation, media literacy is related to *information literacy* (the ability to find, evaluate, organise, and use information) and *digital literacy* (the competencies individuals require to fully participate in a digital world).

Misinformation: Misinformation is sometimes used as a catchall term for various kinds of false or misleading content (see disinformation). However, in some contexts misinformation refers to false or misleading information that is created or spread without the intention to cause harm. In such contexts, it is distinguished from disinformation based on the absence of malicious intent.

Post-truth: Post-truth refers to a social condition in which facts and evidence are largely deemed irrelevant for the formation of public opinions in favour of personal beliefs and affiliations.

Trolling: Trolling refers to the deliberate provocation or harassment of individuals or online communities. Online trolls can operate individually or as part of a coordinated troll farm.

Verification: Verification, in the context of online media, is the process of assessing the authenticity or origin of media content.

Notes

[illegible]

Notes

[illegible]

Notes

[illegible]

